Karolina Krawczak*, Małgorzata Fabiszak and Martin Hilpert

A corpus-based, cross-linguistic approach to mental predicates and their complementation: Performativity and descriptivity vis-à-vis boundedness and picturability

DOI 10.1515/flin-2016-0018
Submitted August 4, 2015; Revision invited November 17, 2015;
Revision received March 17, 2016; Accepted May 31, 2016

Abstract: This corpus-based study investigates the complementation patterns of mental predicates in a cross-linguistic context. More precisely, it examines five equivalent mental verbs from English, German, and Polish and analyzes whether their complements are cognitively construed in different ways in first-person uses of those verbs as opposed to third-person uses. Two types of complementation are considered: we contrast nominal complements with clausal complements. Based on the results of prior studies into Polish *myśleć* 'think' and *wierzyć* 'believe', we hypothesize that first-person singular occurrences of mental predicates will be more readily associated with clausal complements designating non-bounded and non-picturable objects. Conversely, third-person uses of the verbs are expected to be linked to nominal complements that denote bounded and picturable objects. The hypotheses are tested with bivariate and multivariate quantitative techniques. Our results have both descriptive and theoretical implications. Descriptively, we aim to identify the differences in construing the complement of mental predicates, depending on the grammatical person of the syntactic subject. Theoretically, we provide empirical evidence that is relevant for the long-recognized distinction between performativity and descriptivity of mental verbs.

Keywords: mental predicates, complementation, constructions, corpus-based, bivariate and multivariate statistics

***Corresponding author: Karolina Krawczak,** Faculty of English, Adam Mickiewicz University, Poznań, al. Niepodległości 4, 61-874 Poznań, Poland, E-mail: karolina@wa.amu.edu.pl
Małgorzata Fabiszak, Faculty of English, Adam Mickiewicz University, Poznań, al. Niepodległości 4, 61-874 Poznań, Poland, E-mail: fagosia@wa.amu.edu.pl
Martin Hilpert, Institut de langue et littérature anglaises, Université de Neuchâtel, Espace Louis-Agassiz 1, CH-2000 Neuchâtel, Switzerland, E-mail: martin.hilpert@unine.ch

1 Introduction

Language provides speakers with ample opportunities to express the same idea in different ways. For instance, mental predicates such as English *believe* or *forget* can occur with a range of different complement types. The two pairs of sentences in (1) and (2) illustrate how speakers may choose between a nominal complement and a syntactically more complex clausal structure. In both sets, the second variant is a constructed version of the first one, which has been taken from COCA (Davies 2008–2013):

(1) a. *I have always* ***believed in fate and destiny***.
b. *I have always* ***believed that fate and destiny exist***.
(2) a. *She even* ***forgot his name***.
b. *She even* ***forgot what his name was***.

Why does English provide these two variants, and when do speakers choose one or the other? The present study addresses this kind of variation and aims to identify the formal and semantic–pragmatic constraints that determine speakers' choices between alternate constructions. In order to model this use in a predictively and descriptively accurate manner and to identify the variables that are at stake, we apply bivariate (Pearson's Chi-square test of independence) and multivariate statistical methods (multiple correspondence analysis and logistic regression). Taking our cues from the framework of usage-based Construction Grammar, we assume that "general cognitive, pragmatic, and processing constraints" (Goldberg 2006: 3) shape speakers' use and knowledge of language. Ultimately, therefore, empirical inquiries into the determinants of constructional choices are important both descriptively and theoretically.

In the most general terms, the two pairs of sentences in (1) and (2) are instantiations of an alternation between a nominal and a clausal complement construction: (SUBJECT + PREDICATE + NOMINAL COMPLEMENT) and (SUBJECT + PREDICATE + CLAUSAL COMPLEMENT). In these examples, the predicate position is filled by present-perfect- and simple-past-tense forms of the verbs of cognition *believe* and *forget*. In addition to these two mental predicates, the present study considers three others, i. e., *think*, *remember*, and *understand* for English and their respective translational equivalents in German and Polish (see Table 2 for a list of verbs). In the former utterance of each sentence pair in (1) and (2), the complement is formally realized as a noun phrase, whereby the mental object is reified. In the latter sentence, on the other hand, the object is encoded in a clause, which renders the construal more dynamic or processual. Another important difference

between the examples presented in (1) and (2) is the grammatical person of the subject of the mental predicate. Based on prior research into the complementation of mental predicates in Polish (Fabiszak et al. 2012; Krawczak and Kokorniak 2012; Fabiszak et al. 2014), this difference is here hypothesized to be a significant variable determining the alternation, arguably irrespective of the language or the type of the mental predicate.

On the basis of the perfective and imperfective verbal realizations of BELIEVE in Polish, Fabiszak et al. (2012) demonstrate that complements designating a private experience, which is less readily accessible interpersonally, are correlated with the perfective aspect of the verb and expressed through clausal complementation. In turn, an experience that is interpersonally shared has been shown to be associated with the imperfective aspect and expressed as a nominal complement; here, the objects thus encoded undergo reification and acquire a more intersubjectively accessible and hence graspable character. The correlation of an experience that is publicly available with the imperfective aspect can be interpreted to indicate that this experience is conceptualized as more stable and lasting (Fabiszak et al. 2012). In other words, opinions and beliefs that are held intersubjectively, regardless of what they concern, are likely to be indefinable in terms of temporal boundaries. It is noteworthy that the association of clausal complementation with the perfective aspect of the verb and of nominal complements with the imperfective aspect has also been identified as statistically significant in Fabiszak et al. (2014), where the analysis focused on the prefixed forms of *myśleć* 'think'.

The results obtained in Krawczak and Kokorniak (2012), in turn, indicate that first-person singular occurrences of the prefixed and unprefixed forms of *myśleć* 'think' combine more readily with complements that designate more complex and abstract objects of thought of lesser tangibility. Such instances of use afford direct access to the mental object and can be referred to as "performative" (Nuyts 2001). This is because the speaker, when formulating a given statement regarding mental states, "subscribes to and accepts responsibility for the epistemic evaluation underlying it" (Nuyts 2001: 385). Third-person singular occurrences of cognition verbs, on the other hand, which instantiate "descriptive" attribution-based uses (Nuyts 2001), are comparatively more readily associable with concrete objects. Such objects can be more easily ascribed to third-person subjects on the basis of observation, where first-hand experience is unavailable. The study presented in Krawczak and Kokorniak (2012) also reveals a correlation of first-person subjects with the perfective aspect of the predicate and third-person subjects with the imperfective. This association may be taken to mean that the descriptive third-person uses of cognition verbs favor non-bounded forms of mental

experience, while first-person performative uses are more definite and bounded in character.[1] This preference for unbounded imperfective uses of verbs like *think* or *believe* with third-person subjects may be motivated conceptually by the elusive and ungraspable character of mental states experienced by other subjects. A consequence of this inherent inaccessibility and, therefore, obscurity of other people's mental states is that when attributing such states to others, the speaker is likely to opt for means of expression that will reflect the uncertainty that marks any such attribution. Additional support in Krawczak and Kokorniak (2012) for this relatively lower degree of certainty enjoyed by the speaker in attribution-based uses of mental predicates shows that such uses manifest a statistically significant correlation with hypothetical adverbial modification, as exemplified by *perhaps*, *probably*, or *maybe*.

The findings of the above studies lead us to formulate two sets of hypotheses, one concerning the formal characteristics of constructions involving mental predicates, the other dealing with their semantic dimension. The formal hypothesis states that while first-person singular uses of cognition verbs will be associated more significantly with clausal complements across the three languages, their third-person singular occurrences will be linked more canonically to nominal complements. Regarding the semantic attributes of the complements of mental predicates, it is hypothesized that more concrete objects of greater picturability will correspond to third-person singular uses of mental predicates, while complements designating objects that are relatively more abstract and hence less picturable will be linked more immediately to first-person subjects. The hypotheses are summarized in Table 1.

Table 1: Hypotheses.

Hypotheses	**Grammatical person and complement**	
	First person singular	**Third person singular**
Formal dimension	Clausal complements	Nominal complements
Semantic dimension	Abstract/ephemeral objects	Concrete/tangible objects

1 The present study, while drawing on these findings, will not test the correlation between the grammatical person and the grammatical aspect of the mental predicate. Rather, it will focus exclusively on the lexical aspect of the complement associated with the predicates under investigation (for further discussion, see Section 3).

While the formal distinction between nominal and clausal complements is straightforward and does not need to be further operationalized, the semantic dimension clearly requires a measurable definition. How the abstractness and tangibility of the object designated by the complement of the mental predicates have been operationalized will be explained in Section 3. Importantly, these hypotheses do not concern purely descriptive dimensions. Rather, testing their accuracy will produce theory-informing results. This will be attained with the use of bivariate and multivariate statistics. Such methods make possible not only the identification of "multidimensional and socio-conceptually realistic profiles" (Glynn 2014c: 311) of the linguistic phenomena under investigation, but they also provide information about the predictive strength of any patterns thus revealed.

2 Methodology

The methodology employed in this study is known as Profile-Based Analysis or Multifactorial Usage-Feature Analysis. It has been developed in the work of Geeraerts et al. (1994, 1999), Gries (1999, 2003), Heylen (2005), Gries and Stefanowitsch (2006), Divjak (2006, 2010), Glynn (2007, 2010a, 2010b, 2014b), Gries and Divjak (2009), Glynn and Fischer (2010), or Glynn and Robinson (2014). The fundamental assumption here is that contextualized language use can give us an insight into the structure of language, whether within a single linguistic community or across different speech communities. This methodology is designed to model usage or, more precisely, linguistic choices made by speakers, in the form of frequency-based generalizations across many usage events.

Two specific steps are followed in any study employing this methodology. The first stage is a fine-grained qualitative analysis, the second one a quantitative modeling of the annotated data. More precisely, the first phase involves detailed manual annotation of all the occurrences of the phenomenon under investigation for a range of usage features. These features can concern linguistic form, in which case the process of tagging the examples can be semiautomated to various degrees. They can also concern the semantic–pragmatic dimension of language use, which, in turn, necessitates close reading of the contextualized examples. The features for which the data are annotated will depend on the research questions posed in a given study, possible hypotheses to be tested as well as prior empirical and theoretical investigations in the relevant domain. This procedure of qualitative analysis of hundreds and commonly thousands of observations results in a very complex multidimensional grid of usage features.

Given such a multifaceted nature of the resultant data structure, recourse must be taken to quantitative analytical tools that can help us reveal the emergent tendencies in the behavior of the linguistic phenomenon under investigation. This is the second step in the Multifactorial Usage-Feature Analysis. There are different techniques that can be implemented at this stage. In the most general terms, they divide into exploratory and confirmatory methods such as hierarchical cluster analysis (e. g., Divjak and Fieller 2014) and correspondence analysis (e. g., Glynn 2014a), in the former case, or logistic regression modeling (e. g., Speelman 2014), in the latter. Exploratory statistics are used to find orderly patterns in the data, identified on the basis of frequency-based positive and negative correlations of usage characteristics. Confirmatory statistics, in turn, enable us to establish which correlations are statistically significant and important in their predictive power and how accurate they are descriptively and predictively. Both types of analysis are employed in the present study. In addition, we will also employ bivariate statistics in the form of Pearson's Chi-square test of independence, which will be applied in order to test the first hypothesis concerning the formal dimension of clausal and nominal complementation.

3 Data and analysis

3.1 Data

The data for this study were extracted from the newspaper and magazine sections of the *Corpus of Contemporary American English* (Davies 2008–2013) and the *National Corpus of Polish* (Pęzik 2012). For German, the examples were taken from a comparable source, the daily newspaper *Mannheimer Morgen* (http://www.morgenweb.de/). The choice of the journalistic genre was dictated by its cross-linguistic comparability and the high likelihood of finding a large number of the constructions under investigation, where the speaker either presents his or her own mental state or ascribes such a state to another subject.

The dataset (see Table 5 for the data summary) amounts altogether to over 5,000 occurrences of the mental predicates used in the first- and third person singular with nominal and clausal complements. The exact proportions of the complement types were established relative to their distribution in the respective corpora. Since in this study we were only interested in nominal and clausal complements, any examples that diverged from this profile were excluded, e. g., parenthetical uses in which the complement was elided. The

extraction was based on the following four search strings (FIRST PERSON SG. + MENTAL PREDICATE + CLAUSE), (THIRD PERSON SG. + MENTAL PREDICATE + CLAUSE), (FIRST PERSON SG. + MENTAL PREDICATE + NOUN PHRASE), and (THIRD PERSON SG. + MENTAL PREDICATE + NOUN PHRASE). The keywords in the three languages that filled the mental predicate slot in these constructions are provided in Table 2.

Table 2: Keywords.

Language	Predicate 1	Predicate 2	Predicate 3	Predicate 4	Predicate 5
English	*think*	*believe*	*understand*	*remember*	*forget*
German	*denken*	*glauben*	*verstehen*	*erinnern*	*vergessen*
Polish	*myśleć*	*wierzyć*	*rozumieć*	*pamiętać*	*zapomnieć*

The verbs were selected as the most representative instantiations of predicates designating mental states that permit both nominal and clausal complementation across the speech communities analyzed here. Equal numbers of the past-tense forms of the five verbs were extracted for each of the three languages. The past tense was chosen because at least for some of the predicates such as *remember* or *forget* past-tense uses seem more frequent, which was the only reason why the variable of the grammatical tense was controlled for. In English and German, both simple-past- and perfect-tense forms were extracted. In Polish, imperfective aspect forms of the verbs were retrieved, except *zapomnieć* 'forget', which is much more frequent in its perfective usage in the sampled component of the corpus (2,316 vs. 135 occurrences).

The exact number of examples involving the two complementation types for each person was determined proportionally to their relative representation in the respective corpora. Controlling for the distribution of the complement types relative to the grammatical person of the predicates was important in the present inquiry because one of the hypotheses to be tested focuses on this formal characteristic of usage. The hypothesis in question, as put forward in Section 1 (see Table 1), states that clausal complements are expected to be associated more significantly with first-person singular uses of the verbs, while nominal complementation is hypothesized to be more typical of the third-person occurrences of the predicates. Taking equal numbers of such observations or determining the proportions arbitrarily would skew the results.

Another point that should be made with regard to data extraction concerns clausal complementation alone. For both Polish and German, the extraction of the mental predicates associated with clausal complements was highly automatic in that in both languages the dependent clause is introduced by a

subordinating conjunction, which is *że* in Polish and *dass* in German. The situation was different in English, where the complementizer *that* is not an obligatory element in a dependent clause. For that reason, and insofar as we did not want to exclude the zero complementizer clauses, the procedure of selecting examples with clausal complements was more manual here.

3.2 Analysis

All the contextualized examples were annotated for four variables (or factors), each involving binary distinctions (or levels). The variables were selected as informative with respect to the choice of complement and as a means to operationalize the problem and ultimately test the hypotheses formulated in Section 1. Two of these variables are formal, the two others semantic. The variables and the respective levels they subsume are presented in Table 3 and then discussed and exemplified below.

Table 3: Annotation schema.

Variables		Levels
Formal	Grammatical person	First person, third person
	Complement type	Clausal, nominal
Semantic	Boundedness	Bounded, non-bounded
	Picturability	Picturable, non-picturable

The formal variables have already been illustrated in examples (1) and (2) in Section 1, and they are straightforward enough not to require any further explanation here. The semantic variables of boundedness and picturability were introduced in the study as a way of operationalizing the degree of abstractness or, conversely, concreteness of the object of thought and its resultant tangibility. This is important for the second hypothesis to be tested in the present study, which concerns the semantic dimension of the constructions under investigation (see Table 1). In this regard, it was assumed that objects that are bounded, either spatially or temporally, and that can be easily pictured or imagined are relatively more concrete and graspable, whereas those that are unbounded and non-picturable can be understood to be comparatively more abstract and considerably less tangible.

Before discussing some actual examples of the two variables, let us define the notions of boundedness and picturability in conceptual terms. The category

of boundedness,[2] as pointed out by Langacker (1999: 223; cf. 2009: 65, 148–149), is prototypically realized by perfective verbs (e. g., *She thought **that he had finished***) and count nouns (e. g., *She thought **about him***), while imperfective verbs (e. g., *She thought **that they were flying***) and mass nouns (e. g., *She thought **about space travel***) typically, but not necessarily, make "no intrinsic reference to bounding". With respect to the category of aspect, as indicated by Radden and Dirven (2007: 177–178), the perfective imposes a "maximal viewing frame", while the imperfective, whether applied to events or states, is associated with a "restricted viewing frame" and "limited duration".

In a similar vein, Talmy (2000: 50) defines boundedness, which for him falls within the "system of configurational structure", in terms of indefinite continuation for unbounded events, and as a matter of demarcation and individuation into a single "unit entity" for bounded quantities. Naturally, this understanding, as recognized by Talmy (2000), translates directly into the distinction between mass as opposed to count nouns and imperfective as opposed to perfective verbs.

However, the distinction is not always that simple, especially in the case of verbs. For example, if we take the construction [SUBJECT + VERB + *in* + AMOUNT OF TIME], it becomes apparent that such predicates as *sleep* or *read*, even though used in the perfective aspect, are in fact unbounded because they cannot fill in the predicate slot in the above construction, which specifies boundedness (cf. Talmy 2000: 51):

(3) a. **He read in three hours*.
b. *He read for three hours*.

To bound the activity, the grammatical operation of "portion excerpting"[3] (Talmy 2000: 50) needs to be performed, as illustrated in sentence (3b).[4] Here, this period-of-time or "durational" (Radden and Dirven 2007: 180) verb, whose lexical semantics, irrespective of the grammatical aspect, renders it a prototypically unbounded activity, is delimited and thus becomes bounded.

2 Boundedness could be investigated in relation to both the mental predicates and their clausal or nominal complements. However, we will only focus on the latter aspect. The former was examined in Fabiszak et al. (2014) and in Krawczak and Kokorniak (2012), where perfective and imperfective forms of Polish *myśleć* 'think' were studied.

3 The same process applies to nouns: compare *She thought about music* (unbounded) with *She thought about the music of Mozart* (bounded).

4 Note that the examples given in (3) and (4) could well be clausal complements of the mental predicates under analysis in this study.

On the other hand, processes, events, and activities that are typically point-in-time or, put differently, brief or "punctual" (Radden and Dirven 2007: 180), and so bounded in their temporal scope, such as *wake up* or *look through*, can be "debounded" (Talmy 2000: 52) by lexical and/or grammatical means, as illustrated in the constructed sentences in (4b) and (4d) below.

(4) a. *She woke up at 7 this morning/in two minutes this morning*.
b. *She was waking up without end, never getting any sleep*.
c. *She was waking up for two hours*.
d. *She was waking up for hours and hours*.

The utterance in (4a) is a clear case of a bounded action, indicating the exact temporal scope within the boundaries of which it is circumscribed. Here, there is no conflict between the grammatical and lexical aspect of the verb. In (4b), on the other hand, the activity is stretched over time without any demarcation. This is achieved grammatically through the use of the imperfective progressive aspect and lexically through the phrase *without end*, employed by Talmy to the same effects (2000: 53). In (4c), on the other hand, the lexical specification *for two hours* bounds the activity despite its imperfective grammatical aspect, just as in (4d) the phrase *for hours and hours*, which is not specific enough to be bounding, keeps the unbounded character of the verb.

The above distinctions concerning boundedness as ascribed to verbs can be presented in a tabular format, based on the classification proposed by Radden and Dirven (2007: 177–182). Accordingly, Table 4 outlines the conceptual characteristics of states and events in relation to boundedness and provides adequate examples illustrating the distinctions.

As we can see, states remain unbounded irrespective of the viewing frame, i. e., regardless of whether we zoom in on a particular temporal excerpt or consider it holistically. The situation is different with events, where characterization in terms of boundedness is much more complex. The complexity is evidenced by the further subdivision proceeding along the parameters of duration and telicity. This categorization draws on Vendler's (1957) classification of verbs into four mains groups: achievements (telic, instantaneous), accomplishments (telic, durational), activities (atelic, durational), and states (atelic, durational). Overall, events that are bounded are describable in terms of temporal boundaries, which can be complete and enclose the event into a holistic unit (Figure 1a), or partial, i. e., involving only the beginning or the end, as visualized in Figure 1 (b) and (c), respectively.

It is not necessarily self-evident whether in the case of partial boundaries, an event should be treated as bounded or not. One solution is to allow for

Table 4: Boundedness and verbs.

Object type and conceptual properties	Exemplification	
Unbounded permanent states		
– Stability, homogeneity, and infiniteness	*She lives in Europe*	
– Maximal viewing frame	*She loves her job*	
Unbounded nonpermanent states		
– Stability, homogeneity but implied boundaries	*She is living in Europe*	
– Restricted viewing frame	*She is loving her new job*	
Bounded events	**Durational and telic**	**Punctual and telic**
– Internal dynamism and heterogeneity	*She wrote an article*	*She woke up*
– Holistic and external conceptualization	**Durational and atelic**	**Punctual and atelic**
– Maximal viewing frame	*She swam in a pool*	*She sneezed*
– Definable in terms of duration and telicity		
Unbounded events	**Durational and telic**	**Durational and telic**
– Internal dynamism and heterogeneity	*She's writing an article.*	*She is waking up*
– Non-holistic and internal conceptualization	**Durational and atelic**	**Durational and atelic**
– Restricted viewing frame	*She is swimming in a pool*	*She is sneezing*
– Definable in terms of duration and telicity		

Figure 1: Boundedness of processes.

different degrees of granularity (e. g., unbounded – semi-bounded – bounded) and then see which of these levels explains the data best. In the present inquiry, however, we have opted for a maximally elegant and thus the simplest possible set of explanatory variables, i. e., we will draw a two-way distinction and treat half-bounded processes as bounded.

The other explanatory variable for which our data were manually annotated is picturability. Its positive value, i. e., picturable, can be assigned to those things, relations, processes, or events that can be perceived visually. In the case of things, this will normally mean that mass nouns will be more typically, though not necessarily, non-picturable, while count nouns should, by definition, be picturable. This is because they normally constitute integral wholes that are easily evoked as such. In the most prototypical cases of picturability, therefore, this will ultimately mean that an object can be graphically represented, e. g., in a drawing. In less representative instances, it will mean that if we were to be

presented with a photo, e. g., of an event such as a conference, a birthday party, or a wedding reception, we would be able to recognize it on the basis of the image thus captured.

Now that the two factors have been defined, we can consider examples from our dataset that will provide specific illustrations of the distinctions drawn here. Let us first look at the positive values of the two semantic factors, i. e., boundedness and picturability, starting with the former. Examples (5) and (6) both illustrate bounded objects, the difference being the type of complementation and hence the type of boundaries that operate in each case.

(5) *The fuzzy hand pocket lining saved the day when* ***I forgot my gloves***. (Bounded, nominal)

(6) *We went on two chic, fabulous dates. And then he never called again. Every single day, men don't call women back. But to say that I was utterly devastated by it does not properly describe the deep pain this rejection unleashed in me or how clearly it brought into focus my chaotic, disordered life.* ***I believed I had lost the prince*** *to grant all my deeply buried wishes*. (Bounded, clausal)

In sentence (5), the object is encoded by a nominal complement of *forget* and it is a clear instance of a thing that is spatially delimited. We might say, following Langacker (1999: 59), that it constitutes "a specific bounded region within" the category of clothes. A case of temporal demarcation is instantiated in example (6), where the subject expresses an emotionally charged belief she held regarding the loss of a man she had been interested in. An important indicator of the boundedness of the verb *lose* is the fact that it is used in the perfective aspect, which clearly defines its temporal scope. We view the event from an external and maximal or general perspective, rather than zooming in on any particular event. Finally, it is noteworthy that in Vendler's verbal classification, it would constitute an instance of achievement, as *losing something* or *somebody* suggests a telic and punctual event.

With regard to picturability, then, its positive value is exemplified in sentences (7) and (8). In the former case, the object that is categorized as picturable is encoded as a noun phrase; in the latter, it is rendered in a clause.

(7) *As* ***I so vividly remembered Sister Mary's bedside prayers***, ***her kind smiling face, her white habit with its centerpiece of her pectoral cross****...*. (Picturable, nominal)

(8) *The moment my chin hit the bottom of the pool*, ***I understood what had happened***. *It was the shallow end. I had shattered two vertebrae in my spine*. (Picturable, clausal)

The utterance in (7) enumerates a number of objects in the complement of *remember*, all of which can be described as visually perceivable. First, we have the specific activity performed by a specific person at her bed, which must have been witnessed by the speaker and so it can be easily brought back in memory and vividly so. Then, the focus shifts to the physical characteristics of the person, her face and clothes, both of which can be perceived effortlessly. At the same time, all of these objects are also bounded either temporally, in the case of the *bedside prayers*, or spatially, in the case of the *smiling face* and *the white habit*. The example in (8) is an instance of a picturable object expressed in a dependent clause of the verb *understand*. Here, the speaker describes a diving accident, as a result of which he was left paralyzed. The verb *happen*, used in the past perfect tense, refers not only to the moment of impact, but also to the whole situation, including the jump and its tragic consequences, all of which can be visually experienced.

Let us now consider the negative values in the two attributes, i. e., non-bounded and non-picturable. The former quality is exemplified in sentences (9) and (10), the latter in (10) and (11). As before, we will consider examples of such objects encoded as nominal and clausal complements. Sentence (9) is a case of a non-bounded clausal object, while in (10), the object is an abstract noun:

(9) *At times* ***I forgot that I was watching*** *television, which inevitably means that I was being sold something*. (Non-bounded, clausal)

(10) ***She thought*** *and read a lot* ***about happiness****, but mostly she achieved that delicious state by setting herself specific, tangible tasks for different areas of her life*. (Non-bounded, non-picturable, nominal)

In the former sentence, the activity described in the dependent clause of *forget* is expressed in the imperfective aspect and, additionally, the mental predicate is modified by the adverb *at times*. Both these features point to the continuous character of the process of *watching television*. We should also note that the lexical semantics of the predicate *watch* lends itself easily to such an unbounded temporally extended reading. Despite its non-bounded character, the activity is clearly picturable. In the utterance in (10), the object of *think* is also unbounded, but here it is an abstract noun describing a generic emotional state of *happiness*. This sentence simultaneously provides an example of an object that is clearly not picturable. While one might be able to visually perceive an instance of happiness, the state in general is too abstract to visualize. Example (11) illustrates a non-picturable clausal object:

(11) *In my career as a family therapist*, ***I earnestly believed environment played the greatest role in shaping who we become as people***. (Non-picturable, clausal)

Here, the complement of *believe* encodes a proposition which is too abstract and too complex to be amenable to visual perception.

The distinction between bounded and non-bounded objects is not always straightforward. For example, in (12), the predicate in the dependent clause is used in the nonprogressive aspect and so, on such formal grounds, might be regarded as bounded.

(12) *Up until day three, in fact*, ***I'd believed that any successful backcountry trek required bluebird skies and bright sunshine***.

However, there are two aspects to be considered here. First, the predicate designates a certain state of affairs, rather than an event, and states, as noted earlier, are always unbounded. This is true irrespective of whether they are permanent or not and regardless of whether they are expressed through the perfective or imperfective aspect. Second, the process designated by the verb *require* is additionally de-bounded by its subject *any successful trek*, which generalizes the statement and takes it beyond any specific boundaries.

Example (13), in turn, involves nominal complementation, where the noun *darkness*, typically an uncountable abstract noun, would normally be classified as unbounded.

(13) ***I remembered a darkness*** *barely illuminated by too few streetlights*.

In this sentence, however, *darkness* is preceded by the indefinite article and is further post-modified phrasally. This clearly "establishes a selective window of attention", to use Talmy's (2000: 284) wording, on some specific instance of darkness that is thus bounded both in space and time.

Another potentially problematic example involving nominal complementation and the possibility for either a bounded or unbounded reading is given in (14a) and (14b).

(14) a. *Zawsze* ***wierzyłem w zwycięstwo****, choćby nawet nie było ku temu podstaw (...)*
'I always believed in victory, even if there was no reason for that.'
b. *Teraz* ***wierzyłem w szybkie zwycięstwo***.
'Now I believed in a quick victory.'

In (14a), victory is treated as a general abstract phenomenon and is conceptualized from a maximal viewing frame. In other words, we are not focusing on any specific instance of winning some specific event. This is further reinforced by the generalizing adverbial modification of the predicate through *zawsze* 'always', which implies the continuous homogeneous nature of the mental experience and, hence, by extension, of the object of belief. In (14b), on the other hand, the speaker is clearly referring to a specific event and a specific instance of victory, *szybkie zwycięstwo* 'a quick victory'. Here, as was the case in (13), our attention zooms in on a particular type of the phenomenon in question. In addition, the predicate is modified by the specifying adverb *teraz* 'now', which simultaneously narrows down the object of belief even further.

The annotation was performed by the three authors of the present paper, with each taking the data for one language only. Once the annotation had been completed, the Polish and English data were submitted to secondary annotation for the two semantic factors. Thus, the Polish data were annotated by the person previously responsible for the English data, and vice versa. This procedure was motivated by the fact that some inconsistencies were identified in the understanding of the two categories between the two annotators. The objective here was to eradicate any mistakes in the annotation. In order to achieve this goal, we first re-annotated a small sample from each dataset and then discussed any differences in order to fine-tune our operational definition of the two semantic categories. The definition and some examples were then consulted with the third author, responsible for the German dataset. Following this input, the definition was further refined and then applied methodically to all the examples in Polish and English. Any observations, where the results of the primary and secondary annotation diverged, were closely reconsidered and a final decision was made as to their status. The summary of the data is presented in Table 5.

Table 5: Data summary.

Variable	Levels						Total
Lexeme	*believe*	291	*glauben*	477	*wierzyć*	340	1,108
	forget	303	*vergessen*	455	*zapomnieć*	300	1,058
	remember	298	*erinnern*	287	*pamiętać*	297	882
	think	344	*denken*	379	*myśleć*	300	1,023
	understand	307	*verstehen*	339	*rozumieć*	300	946
Complement type	Clausal: 2,572		Nominal: 2,445				5,017
Grammatical person	First person: 1,955		Third person: 3,062				5,017
Boundedness	Bounded: 2,689		Non-bounded: 2,328				5,017
Picturability	Picturable: 1,860		Non-picturable: 3,157				5,017

It specifies the exact distribution of the semantic and formal variables as well as the number of predicates per language.

4 Results

This section presents the quantitative results of the study. This will be done in two steps. First, in Section 4.1, we will test the formal hypothesis concerning the distribution of clausal and nominal complements. Given the simple binary nature of the hypothesis, this will be attained through the use of the bivariate Chi-square test for independence. Second, in Section 4.2, we will test our hypothesis regarding the semantic characteristics of the complement. To that purpose, we will employ the exploratory method of binary correspondence analysis and the confirmatory method of mixed-effects logistic regression modeling.

4.1 Formal hypothesis testing

In this section, we test the first hypothesis put forward in Section 1 concerning the relation between the grammatical person of the mental predicate and the form of the complement. On the basis of prior findings for Polish *wierzyć* 'believe' and *myśleć* 'think', it was proposed that the first-person occurrences of verbs of cognition will be significantly associated with clausal complements, while third-person uses of such verbs will be positively correlated with nominal complementation. It was also assumed that such results will be obtained irrespective of both the type of mental predicate involved and the language that is sampled. In other words, we expect that for both grammatical persons, a clear and consistent pattern in complementation preferences will emerge across the three languages and the five verbs, a pattern that will conform to the hypothesized dichotomy. To see whether this is indeed the case, we subdivided the data into two groups, depending on the grammatical person of the predicate: one group contained all the occurrences of the first-person uses of the mental predicates across the three languages, while the other consisted solely of the third-person uses of the verbs in the three linguistic communities. These two sets were then each submitted separately to the bivariate Chi-square test for independence to establish whether the first-person uses indeed correlate significantly with clausal complements and third-person occurrences with nominal complementation, as assumed in hypothesis 1. The difference in complement preferences was found to be

statistically significant in both instances, with $p < 0.001$. Table 6 shows the observed (O) and expected (E) frequencies for all the predicates, demonstrating a clear discrepancy between the two.

Table 6: Expected and observed frequencies.

Mental predicate	Clausal complementation		Nominal complementation	
	Expected	Observed	Expected	Observed
BELIEVE ENG	73.87	134	76.13	16
BELIEVE GER	186.65	269	192.35	110
BELIEVE PL	84.2	111	86.78	60
FORGET ENG	72.39	36	74.60	111
FORGET GER	167.94	128	173.06	213
FORGET PL	73.87	70	76.13	80
REMEMBER ENG	74.86	34	77.14	118
REMEMBER GER	122.14	78	125.86	170
REMEMBER PL	72.89	57	75.11	91
THINK ENG	93.08	123	95.92	66
THINK GER	140.85	187	145.15	99
THINK PL	73.87	101	76.13	49
UNDERSTAND ENG	75.84	83	78.16	71
UNDERSTAND GER	121.64	49	125.35	198
UNDERSTAND PL	73.87	48	76.13	102

ENG: English; GER: German; PL: Polish.

The expected frequencies would be observed if the null hypothesis were to be true. In other words, they would hold true if there was no difference in preference for either type of complementation. As the results indicate, this is clearly not the case here. The Pearson residuals obtained for both types of complements are presented graphically in Figure 2. They are calculated on the basis of the following formula (from Gries 2014: 370):

$$\frac{\text{Observed} - \text{Expected}}{\sqrt{\text{Expected}}}$$

The dot plot on the left visualizes the behavioral patterns of the verbs relative to clausal complementation; the plot located on the right shows the usage tendencies of the predicates with respect to nominal complementation. The predicates in relation to language are listed on the left along the vertical axis of the first map, but they are relevant to both plots. The horizontal axes, in turn, provide the scale for the Pearson residuals, against which to evaluate the degree of association of the predicates with either type of complementation or their disassociation from either. Positive and negative values are indicative of

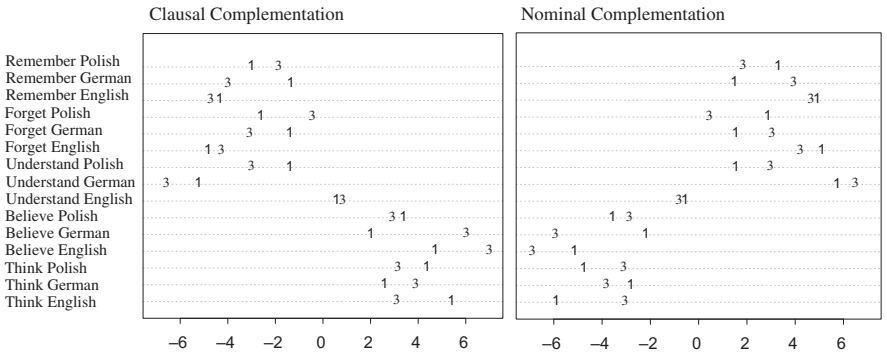

Figure 2: Pearson residuals for clausal and nominal complementation across first- and third-person uses of mental predicates in German, English, and Polish.

positive and negative correlation, respectively, in this regard. Finally, points 1 and 3 scattered in the dot plots correspond to the first- and third-person occurrences of the verbs.

Before we consider the specific results, two general observations can be made. First, as evidenced in Figure 2, our hypothesis appears to be only partially confirmed. The supporting evidence that we find for it can be described as verb-specific. This, in turn, takes us to the second observation, namely, that the results reveal the importance of the verb to the structuring of the data in terms of complementation patterns. More specifically, the findings suggest that the lexical semantics of the specific cognition verbs overrides the hypothesized relation between grammatical person and syntactic complexity. More importantly, the corresponding verbs across the three languages exhibit the same tendencies: regardless of the grammatical person, the verbs are consistently positively or negatively correlated with one of the two complement types. Overall, therefore, we can say that even though we seem to have found partial evidence for our hypothesis, it is, in fact, an epiphenomenal result of the specific verbs. This is evident from the fact that the two dot plots are mirror reflections of each other, where the correlation of each verb with one type of complementation is the exact inverse of its correlation with the other type. Let us now turn to the specific results for clausal complementation, visualized in the first dot plot in Figure 2.

Here, it was hypothesized that the preference for clausal complementation would be predominantly observed for the first-person uses of the mental predicates. The values of the Pearson residuals can be used to establish the relative rank or importance of the given positive or negative correlation. Interestingly, when we look at the bottom half of the dot plot, where all the positive

correlations are found, we can see that the strongest association with clausal complements obtains for the third-person uses of *believe* in English, immediately followed by the third-person occurrences of its German cognate. This association of third-person uses of BELIEVE in the two languages with clausal complementation goes directly against our claim. The next strongest correlates for clausal complementation are first-person occurrences of *think* and *believe* in English and *myśleć* 'think' in Polish, followed by third-person uses of *denken* 'think' in German and first-person uses of *wierzyć* 'believe' in Polish. The weakest association holds for first-person uses of *understand* in English. Interestingly, *verstehen* 'understand' in German and *rozumieć* 'understand' in Polish are negatively correlated with clausal complementation, which (given the values of the Pearson residuals) is particularly salient for the former language. Along with its relatively weak positive correlation for English, this might suggest that the category UNDERSTAND inherently prefers reified objects expressed nominally. The verbs *forget* and *remember* and their equivalents in German and Polish manifest a clear dissociation from clausal complementation.

When we consider the visualization for nominal complementation in the second dot plot in Figure 2, we can observe an overall pattern that mirrors that found in the first dot plot. It shows a clear transition from predicates favoring one type of the complement to those associated more distinctly with the other. Overall, the same predicates manifest the same complementation preferences across their first- and third-person uses. This pattern emerges as independent of the language. What it means is that, rather than being constrained by the grammatical person or cross-linguistic variation, complement choice appears to be determined by the lexical semantics of the mental predicate. Admittedly, however, there is a weak effect of the grammatical person on the strength of the observed association, which might ultimately imply some influence of this factor on the behavior of the predicates, both intra- and interlinguistically. Whether such an interpretation is correct, and if so, to what extent, would require further investigation.

As the results in this section demonstrate, our hypothesis regarding the influence of grammatical person on the choice of the complement type finds only partial support. Whatever evidence we identify is weakened by overall lexical effects. We therefore need to conclude that the lexical semantics of the mental predicates under investigation here overrides the expected constructional profiling. A regular pattern has emerged across the three linguistic communities, where the verbal instantiations of BELIEVE and THINK are strongly associated with clausal complementation, while those of REMEMBER and FORGET are distinctly associated with nominal complements. The usage patterns manifested by the exponents of UNDERSTAND are slightly more language-specific in

that in German and Polish, they are aligned with the predicates leaning toward nominal complementation, while *understand* in English remains rather non-distinct in its complement preferences. We need to conclude, therefore, that we do not find conclusive evidence for our first hypothesis.

4.2 Semantic hypothesis testing

In this section, the second hypothesis, regarding the relation between the grammatical person and the conceptual properties of the complement, is tested through multivariate statistical modeling. This will be done in two steps. First, the exploratory method of binary correspondence analysis will be employed to identify the behavioral tendencies of the predicates in their functional context of use (i. e., language, grammatical person, and the conceptual characteristics of the complement). Next, the confirmatory technique of mixed-effects logistic regression analysis will be used to see whether any of the patterns thus identified are statistically significant and accurate in both predictive and descriptive terms.

4.2.1 Exploratory results

The exploratory results are visualized in Figure 3. The plot is a binary correspondence analysis showing the correlations obtaining between the lexemes as used in the first and third person singular in their respective linguistic communities and the conceptual characteristics manifested by the complement. For the purposes of this analysis, the two semantic variables have been combined, resulting in a four-way distinction instead of two binary factors. Before we consider the specific correspondences that emerge in the plot, let us briefly introduce the method employed here.

Correspondence analysis is a "space-reduction technique" (Glynn 2014a: 443); it yields a biplot of the metadata on the basis of the "Chi-squared distance matrices" (Baayen 2008: 129) calculated for the variables subjected to the mapping. The relative proximity of the respective data points in the plot can be taken as an indication of the strength of their correlation (Glynn 2014a). The size of the data points reflects the contribution of a given feature to the structuring of the data in the visualization. Finally, the overall accuracy of the representation in the two-dimensional map can be established on the basis of the percentages provided for each axis. This tells us how much variation in the data is explained in these

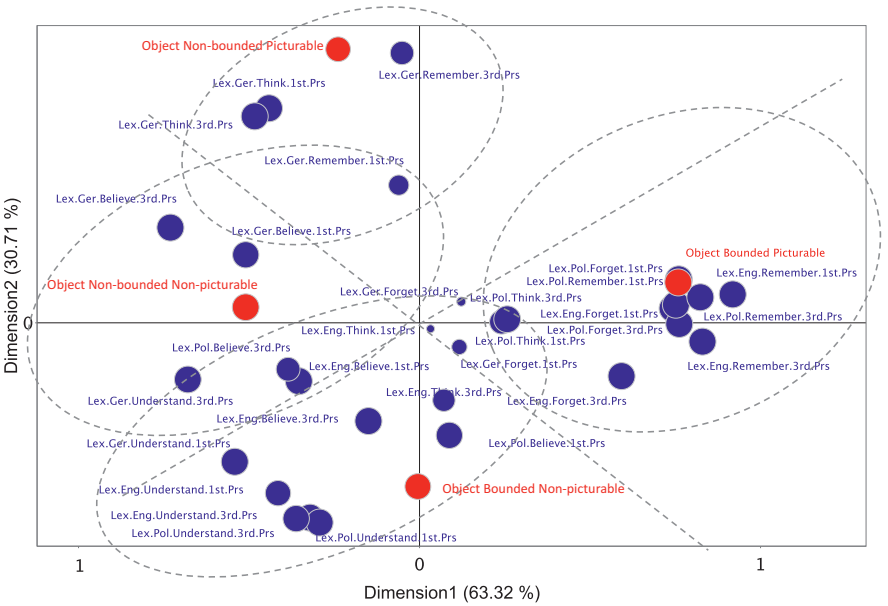

Figure 3: Binary Correspondence Analysis: Lexeme–Language–Person and Object Type.

two visualized dimensions (Glynn 2014a). In the case of the plot presented in Figure 3, we can see that we account for over 60 % of variation in the first dimension and for another 30 % in the second dimension. This is an excellent result. Let us now see what patterns are identified here.

In the most general terms, the plot reveals four main groupings of usage features, each organized around a different semantic value of the complement. For clarity, these four individual sets have been encircled in the plot. They are distributed across the four quadrants of the map. The location of some of the data points for the predicates shows that they are attracted by two or even three clusters simultaneously. Another macro-level observation that can be made here is that the plot further divides diagonally into two parts, as indicated in the figure. The common denominators of the two parts are constituted by the two values of boundedness and picturability. Finally, it should be noted that the lexemes which are positioned in the middle of the plot, i. e., third-person singular uses of *vergessen* 'forget' in German and first-person singular occurrences of *think* in English, clearly lack distinct correlations. This is further corroborated by the relatively small size of these data points. Their location discloses their apparent association with all four clusters. Let us now consider more closely the four groupings that emerge in Figure 3 and examine how the usage patterns relate

to our second hypothesis. To facilitate interpretation, the results are presented in tabular format in Table 7.

Table 7: Clusters revealed in the binary correspondence analysis.

	Non-bounded Non-picturable	Non-bounded Picturable	Bounded Non-picturable	Bounded Picturable
First person	BELIEVE (Ger.) BELIEVE (Eng.)	REMEMBER (Ger.) THINK (Ger.)	BELIEVE (Pol.) UNDERSTAND (Pol.) UNDERSTAND (Eng.) UNDERSTAND (Ger.)	REMEMBER (Pol.) REMEMBER (Eng.) FORGET (Eng.) FORGET (Pol.) THINK (Pol.)
Third person	UNDERSTAND (Ger.) BELIEVE (Ger.) BELIEVE (Pol.)	REMEMBER (Ger.) THINK (Ger.)	UNDERSTAND (Pol.) UNDERSTAND (Eng.) THINK (Eng.) BELIEVE (Eng.)	FORGET (Eng.) FORGET (Pol.) REMEMBER (Pol.) THINK (Pol.)

Eng.: English; Ger.: German; Pol.: Polish.

Evidence that would support our hypothesis should indicate positive correlations between the first-person occurrences of the predicates and objects that are non-bounded and non-picturable. Conversely, for the third-person observations of the verbs of cognition, we should see distinct associations with bounded and picturable objects. Objects that can be described as bounded but non-picturable or non-bounded but picturable, being transient categories between the other two cases, should ideally correlate with the smallest number of uses. As can be seen in Table 7, the actually observed correspondences are not consonant with what was expected. More precisely, it will be seen that the specific correspondences obtaining for the four clusters are determined by lexical semantics and language rather than by constructional preferences.

Let us consider the first grouping for non-bounded and non-picturable objects. In Figure 3, this cluster is located at the intersection of the left-hand bottom and upper quadrants. It is associated with five predicates. The first among these is *glauben* 'believe' in German, for both grammatical persons. The position of these two data points in the plot is indicative of the positive correlation of *glauben* in the first and third person with non-bounded and non-picturable objects. In addition, this type of objects is also correlated with *wierzyć* 'believe' in Polish and *verstehen* 'understand' in German for the third person singular. Such associations go directly against our hypothesis. In the first person singular, *believe* in English is also part of this grouping, which, in

turn, is in line with the hypothesis. It should be noted, however, that first-person uses of *believe*, third-person uses of *wierzyć* 'believe' as well as, to a lesser degree, third-person occurrences of *verstehen* 'understand' are all simultaneously attracted by bounded and non-picturable objects. A cluster for this type of objects is centered on the *y*-axis, being thus spread across the bottom half of the plot. When we look at Figure 3, the partial overlap between these two clusters can be taken as an indication of this. Some of these correspondences are illustrated in sentences (15)–(19).

(15) *Auch dass sie krank sei* ***glaubte*** *er nicht*.
'Also he did not believe that she was sick.'

(16) *Er* ***verstand*** *nichts von den mathematischen Gesetzen der Bewegung*.
'He did not understand anything of the mathematical laws of motion.'

(17) *Ich habe nur in den Spiegel gestarrt und nicht* ***geglaubt****, was ich hörte*.
'I just stared in the mirror and did not believe what I heard.'

(18) *I* ***believed*** *that my personality and sense of self were too well-defined for me to harm myself or others*.

(19) *Sam Gallup* ***wierzył*** *w zbiorową mądrość prostych ludzi, a nie ufał politykom*.
'Sam Gallup believed in the collective wisdom of simple people, and didn't trust politicians.'

The second group of verbs, located in the left-hand upper quadrant of the plot, is associated with non-bounded and picturable objects. This is a very consistent class both along lexical and linguistic lines. It brings together the verbal exponents of THINK and REMEMBER in German for both grammatical persons. Given that it is only German occurrences of the verbs that we find here, this cluster is more informative with respect to the usage tendencies of these two mental predicates in German than it is with regard to the constructional profiling of mental predicates. One more point that deserves our attention is that the first-person singular uses of *erinnern* 'remember' are located in the intersection of this cluster and the one discussed above, which suggests that these uses of *erinnern* are also linked to non-bounded and non-picturable objects. This, in turn, can be interpreted to mean that the complements associated with *erinnern* in the first person tend to designate non-bounded objects. The correlations found in this cluster are exemplified in (20)–(23).

(20) *Der Kühlschrank war leer*, ***erinnerte*** *ich mich heute nicht zum ersten Mal*.
'The fridge was empty, as I remembered not for the first time today.'

(21) *Als Kasimir das Wort süß hörte, **erinnerte** er sich sofort an süße Schnecken...*
'As Kasimir heard the word *sweet*, he immediately remembered sweet rolls...'

(22) *Ich **dachte**, ich würde noch länger am Krankenbett bleiben*
'I thought I would stay longer still at the sickbed.'

(23) *Sie **dachte** an die Traurigkeit, die sie so oft bei ihm beobachtet hatte.*
'She thought about the sadness that she had observed with him so often.'

In the bottom half of the plot, we find the cluster associated with bounded and non-picturable objects, which for the first person accommodates the verbal instantiations of UNDERSTAND across the three linguistic communities and *wierzyć* 'believe' in Polish. This type of objects is also correlated strongly with the third-person uses of *understand* in English and the relevant equivalent in Polish and, to a lesser degree, with the third-person occurrences of *think* and *believe* in English. However, the verbs that are clearly most distinctly associated with bounded and non-picturable objects are English *understand* and Polish *rozumieć* 'understand', used in both grammatical persons, as is evidenced by their location in the plot. Polish exponents of BELIEVE in the first person and English uses of *think* and *believe* in the third person are less prominent members of this class, being simultaneously attracted to bounded and picturable objects in the case of the two former predicates and to non-bounded and non-picturable objects for English *believe* in the first person singular. Given these associations, we can see that irrespective of the language, the verbal uses of UNDERSTAND tend to be associated with objects that are non-picturable. In both English and Polish, they also lean toward bounded objects, as illustrated in (24) and (25). In German, in turn, we can observe a tendency toward non-bounded objects, especially for the third-person occurrences of *verstehen* 'understand', as exemplified in (26).

(24) *Mąż nie **rozumiał** potrzeb żony.*
'The husband didn't understand his wife's needs.'

(25) *I **understood** and understand today and respect the decision that the president made.*

(26) *der ›Coldman‹ sprach unaufhörlich in einer Sprache, die sie nicht **verstand**.*
'He said something inaudibly in a language, which she did not understand.'

Finally, the largest cluster of mental predicates is organized around objects that can be defined as bounded and picturable. For the first person singular, this type of objects is associated with the instantiations of REMEMBER and FORGET in

Polish and English and with *myśleć* 'think' in Polish. On the basis of the location of these data points, we can see that the first four are distinctly associated with bounded and picturable objects, while the last one appears to be attracted, to a smaller extent, also by objects that are non-bounded and non-picturable. With regard to third-person occurrences, which in light of our hypothesis should be particularly numerous here, we only have four associations, i. e., *pamiętać* 'remember' and its English equivalent, both of which are close correspondences, and the Polish and English exponents of FORGET. In the case of FORGET, the association is considerably less distinct for English, for which we can observe a simultaneous usage tendency toward bounded and non-picturable objects. It is evident that this cluster is uniform both lexically and linguistically. These regularities demonstrate that the common conceptual denominator here is not the grammatical person, but lexeme and language. Some examples for this cluster are given below.

(27) *Rodzina bardzo szybko* ***zapomniała*** *o staruszce*.
'The family forgot about the old lady very quickly.'

(28) ***Zapomniałam*** *o magnoliach, które stały w wodzie na stole*.
'I forgot about the magnolias that were standing in water on the table.'

(29) *Kobieta jednak* ***pamiętała*** *co się wydarzyło i zgłosiła sprawę na policję*.
'The woman, however, remembered what had happened and reported the case to the police.'

(30) ***Pamiętałem*** *ich z twarzy*.
'I remembered them by face.'

What can we conclude based on this exploratory investigation into the conceptual effects observed for the objects of mental predicates? Undoubtedly, the distribution of the predicates relative to grammatical person fails to provide supportive evidence for the patterns that were expected in light of our semantic hypothesis. The discrepancies across the clusters in terms of grammatical person as well as lexical and linguistic effects show yet again that it is lexical semantics and, in this case, also variation across the three languages that override possible constructional profiling. Among the most conspicuous tendencies that are observed are the cross-linguistic preferences of the verbal instantiations of REMEMBER and FORGET for picturable objects and of UNDERSTAND for non-picturable objects. Boundedness appears to be less lexically determined. Indeed, there are some indications that the correlation of boundedness with the objects of mental predicates might be the result of the construal imposed by the grammatical person. This is the case for English *believe*, which is associated with non-bounded

objects in the first person, but bounded objects in the third person. Similarly, German *verstehen* 'understand' and Polish *wierzyć* 'believe' are both associated with non-bounded objects in the third person, but with bounded objects in the first person. Could this be taken to mean that boundedness might give us a better insight into construal differences in mental predicates relative to the grammatical person? To see if this is the explanatory variable that can indeed help us account for the choice of construal in cognition verbs, we will now turn to the confirmatory method of logistic regression analysis.

4.2.2 Confirmatory results

In this section, we turn to confirmatory statistical modeling of the choice between the first-person and third-person construal with mental predicates relative to the conceptual properties of the complement. More specifically, we will use mixed-effects binary logistic regression modeling with the lexical category and language treated as random variables. The reason for treating these factors as random and thus excluding their contribution is the variation that they introduce, which has been identified already at the exploratory stage in Section 4.2.1. Adding these two variables as random variables to our model enables us to account for this variation while filtering out the statistical noise that they would introduce. There is also another reason for treating the lexeme as a random variable, which is the constructional, rather than lexical, focus of the present study.

Two models were fitted: one with the predictor Object Type, characterized by a four-way distinction, as already employed in the correspondence analysis in Section 4.2.1, the other with two separate predictors boundedness and picturability. In the latter model, only the former variable, i. e., boundedness, turned out to be statistically significant. This fact might be taken as a tentatively affirmative answer to the question posed at the end of the previous section. Given that in all the other respects, the two models are nearly identical, we will here look at the model where Object Type incorporating a four-way distinction was used as an explanatory variable. Table 8 presents the results of this analysis.

The accuracy of the model can be evaluated on the basis of the *C* statistic score provided at the bottom, which is at 0.65. This diverges by at least 0.05 from a value that would indicate "acceptable discrimination" (Hosmer and Lemeshow 2000: 162), which means that the model is not predictively strong. Before we consider the levels of the predictor, let us note that the variance values for the random effects, neither of which is zero, suggest that if the

Table 8: Mixed-effects logistic regression analysis for first-person vs. third-person uses of mental predicates: Person ~ Object Type + (1|Lexeme) + (1|Language).

Fixed effects	Estimate	Std. error	*t* Value	pMCMC
Intercept	1.58645	0.07180	22.097	0.472
Bounded and picturable	−0.00836	0.01877	−0.445	0.792
Non-bounded and non-picturable	0.03182	0.01446	2.201	0.038*
Non-bounded and picturable	0.01038	0.02679	0.387	0.578
C statistic: 0.65				**Random effects**
				Lexeme (variance: 0.0002)
				Language (variance. 0.01476)

Significance codes: *** < 0.001, ** < 0.01, * < 0.05.

variables were not treated as random here, they would be likely to affect the results to a greater or lesser extent. To assess whether the independent variable can be regarded as a statistically significant and important predictor determining the use of the first- over third-person uses of the mental predicates, we need to look at the last and second columns of Table 8, respectively. The last column provides the *p* value, which has here been calculated through Markov chain Monte Carlo (MCMC) sampling.

As we can see, the only level that proves statistically significant is that which denotes non-bounded and non-picturable objects. This feature has a positive estimate, which means that it is associated with the third-person uses of the mental predicates.[5] The effect size of this association, which can be evaluated on the basis of the estimate value provided in the second column of Table 8 and which serves to establish the relative importance of a given feature in predicting the outcome, is rather negligible. This means that despite the statistical significance of this type of complements, the variable Object Type, operationalized in terms of boundedness and picturability, does not provide the ultimate key to understanding and predicting the choice of complement type with the two grammatical persons. Even if its relative importance were to be higher, it would still contradict the correlation that was expected in our second hypothesis. Not only is the type of complement associated with the third person non-bounded, but it is also non-picturable.

5 Incidentally, in the model where boundedness and picturability were treated separately and where the former variable alone turned out to be significant, the results were parallel.

The reason for this unexpected result may be twofold. One explanation is that the simple binary operationalization differentiating between objects that are picturable and non-picturable, on the one hand, and bounded and non-bounded, on the other, is too schematic and that finer distinctions need to be made. It might well be that if more leeway had been allowed in the classification, as was the case in the original studies (Fabiszak et al. 2012; Krawczak and Kokorniak 2012; Fabiszak et al. 2014), the results would have had different implications. Of course, the other possibility and explanation for the results, which contradict our two hypotheses, might well be that the potential for picturing or imagining a given object, irrespective of the granularity of the distinctions that we should allow for, is not what determines the choice of the complement of mental predicates relative to the grammatical person of their syntactic subject. To see which of these two options is the case, we would need further investigation, involving the re-operationalization of the semantic properties of mental objects.

5 Conclusions

The present paper has addressed the question of the constructional profile or construal imposed by the grammatical person on the complement of mental predicates. In examining this question, we had two objectives, one descriptive, the other theoretical. First, the goal was to identify the differential effect on complement choice of first- and third-person uses of five mental predicates across three languages. More precisely, the aim was to provide further empirical evidence for the findings obtained in prior research into mental predicates in Polish by testing two hypotheses, one concerning a syntactic alternation between two complement types, the other focusing on the semantic properties of the complement. This, in turn, was to be directly informative with respect to the long-standing theoretical distinction between performative and descriptive uses of mental predicates, a distinction dating back to Benveniste (1971) and taken up by Nuyts (2001) and Verhagen (2005) in more contemporary research into epistemicity and evidentiality.

We hypothesized that complements encoded as nouns and representing concrete mental objects will be more commonly associated with third-person subjects. This is because the speaker attributes a given mental state to another "nonlocal" (Bresnan et al. 2007) person, and so it may be assumed that such reified and concrete objects will be more easily attributable to others. On the other hand, in first-person occurrences of mental predicate constructions, where the speaker has direct access to his/her own thoughts, a more likely choice may

be the more dynamic construal imposed by clausal complements, for one thing, and, for another, a preference for more abstract and intangible objects. As the results in Sections 4.1 and 4.2 demonstrate, neither of the hypotheses has been confirmed.

The findings that we obtain here, however, are revealing and provide valuable feedback, both on the methodological and theoretical plane. Methodologically, an important implication of our study is that when dealing with semantic dimensions in language, one should probably allow for more fine-grained distinctions. Indisputably, elegance and simplicity of explanations, which we were striving for in the present inquiry, are of great significance in empirical research. However, binary distinctions in semantics, which only allow for the presence or absence of a given usage characteristic, are likely to fail to account for any subtle variation, which normally cannot be expressed in dichotomous terms. This may well be why our experiment attempting to capture conceptual variation along binary parameters did not answer our research questions. The theoretical insight that we gain is that the performative and descriptive uses of mental predicates do not seem to exhibit statistically significant syntactic differences. More specifically, they do not have distinct preferences in terms of their complementation pattern. The differences that we did find were all motivated by lexical variation rather than by the construal imposed by the grammatical person. This, in turn, may be interpreted to suggest that clausal complements, despite their inherent dynamism, are no less likely to occur with descriptive attribution-based uses than nominal complements, which reify the mental object, arguably rendering it more manageable in the intersubjective space. This in itself is an interesting finding, even if negative, as it implies that the answer to the difference in construal between descriptive and performative uses of mental predicates is more likely to lie in their semantic and pragmatic features than in their syntactic patterning.

Acknowledgments: We would like to express our gratitude to the two anonymous reviewers for their comments and suggestions. Our sincere thanks also extend to Hubert Cuyckens, the editor of *Folia Linguistica*, for his most constructive help. Any remaining shortcomings are our own.

References

Baayen, Harald. 2008. *Analyzing linguistic data*: *A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Benveniste, Emile. 1971. Subjectivity in language. In Emile Benveniste, *Problems in general linguistics*, 223–230. Coral Gables: University of Miami Press.

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

Davies, Mark. 2008–2013. *The Corpus of Contemporary American English: 450 million words, 1990–present*. http://corpus.byu.edu/coca/.

Divjak, Dagmar. 2006. Ways of intending: A corpus-based Cognitive Linguistic approach to near-synonyms in Russian. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis*, 19–56. Berlin: Mouton de Gruyter.

Divjak, Dagmar. 2010. *Structuring the lexicon: A clustered model for near-synonymy*. Berlin: De Gruyter Mouton.

Divjak, Dagmar & Nick Fieller. 2014. Cluster analysis: Finding structure in linguistic data. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 405–443. Amsterdam & Philadelphia: John Benjamins.

Divjak, Dagmar & Stefan Th. Gries. 2006. Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2(1). 23–60.

Divjak, Dagmar & Stefan Th. Gries. 2009. Corpus-based Cognitive Semantics: A contrastive study of phrasal verbs in English and Russian. In Barbara Lewandowska-Tomaszczyk & Katarzyna Dziwirek (eds.), *Studies in cognitive corpus linguistics*, 273–296. Frankfurt am Main: Peter Lang.

Fabiszak, Małgorzata, Anna Hebda, Iwona Kokorniak & Karolina Krawczak. 2014. The semasiological structure of Polish *myśleć* "to think": A study in verb-prefix semantics. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 223–251. Amsterdam: John Benjamins.

Fabiszak, Małgorzata, Anna Hebda & Barbara Konat. 2012. Dichotomy between private and public experience: The case of Polish *wierzyć* "believe". In Christopher Hart (ed.), *Online proceedings of UK-CLA meetings* 1, 164–176. Hertfordshire: The UK Cognitive Linguistics Association. http://www.uk-cla.org.uk/proceedings/volume_1.

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

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

Glynn, Dylan. 2007. *Mapping meaning: Toward a usage-based methodology in Cognitive Semantics*. Leuven: University of Leuven (KU Leuven) dissertation.

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

Glynn, Dylan. 2010b. Testing the hypothesis: Objectivity and verification in usage-based Cognitive Semantics. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative Cognitive Semantics: Corpus-driven approaches*, 239–270. Berlin: De Gruyter Mouton.

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

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

Glynn, Dylan. 2014c. Techniques and tools: Corpus methods and statistics for semantics. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 307–341. Amsterdam: John Benjamins.

Glynn, Dylan & Kerstin Fischer (eds.). 2010. *Quantitative methods in Cognitive Semantics: Corpus-driven approaches*. Berlin: De Gruyter Mouton.

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

Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Gries, Stefan Th. 1999. Particle movement: A cognitive and functional approach. *Cognitive Linguistics* 10(2). 105–145.

Gries, Stefan Th. 2003. *Multifactorial analysis in Corpus Linguistics: A study of particle placement*. London: Continuum Press.

Gries, Stefan Th. 2006. Corpus-based methods and Cognitive Semantics: The many senses of *to run*. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis*, 57–99. Berlin: Mouton de Gruyter.

Gries, Stefan Th. 2014. Frequency tables: Tests, effect sizes, and explorations. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 365–390. Amsterdam: John Benjamins.

Gries, Stefan Th. & Dagmar Divjak. 2009. Behavioral profiles: A corpus-based approach to cognitive semantic analysis. In Vyvyan Evans & Stephanie Pourcel (eds.), *New directions in Cognitive Linguistics*, 57–75. Amsterdam: John Benjamins.

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

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

Hosmer, David W. & Stanley Lemeshow. 2000. *Applied logistic regression*. New York: John Wiley & Sons.

Krawczak, Karolina & Iwona Kokorniak. 2012. A corpus-driven quantitative approach to the construal of Polish "think". *Poznań Studies in Contemproary Linguistics* 48(3). 439–472.

Langacker, Ronald W. 1999. *Grammar and conceptualization*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. 2009. *Investigations in Cognitive Grammar*. Berlin: Mouton de Gruyter.

Nuyts, Jan. 2001. *Epistemic modality, language, and conceptualization*. Amsterdam: John Benjamins.

Pęzik, Piotr. 2012. Wyszukiwarka PELCRA dla danych NKJP [PELCRA search engine for the NKJP data]. In Adam Przepiórkowski, Mirosław Bańko, Rafał Górski & Barbara Lewandowska-Tomaszczyk (eds.), Narodowy Korpus Języka Polskiego [The National Corpus of the Polish language], 253–273. Warszawa: PWN.

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

Speelman, Dirk. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 487–533. Amsterdam: John Benjamins.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Vol 1: *Concept structuring systems*. Cambridge, MA: MIT Press.
Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160.
Verhagen, Arie. 2005. Constructions of intersubjectivity: Discourse, syntax, and cognition. Oxford: Oxford University Press.