

# ANGLICA

An International Journal of English Studies

25/2 2016

## EDITOR

prof. dr hab. Grażyna Bystydzieńska [g.bystydzienka@uw.edu.pl]

## ASSOCIATE EDITORS

dr hab. Marzena Sokołowska-Paryż [m.a.sokolowska-paryz@uw.edu.pl]

dr Anna Wojtyś [a.wojtys@uw.edu.pl]

## ASSISTANT EDITORS

dr Katarzyna Kociołek [kkociolek@uw.edu.pl]

dr Magdalena Kizeweter [m.kizeweter@uw.edu.pl]

## ADVISORY BOARD

Michael Bilynsky, University of Lviv  
Andrzej Bogusławski, University of Warsaw  
Mirosława Buchholtz, Nicolaus Copernicus University, Toruń  
Xavier Dekeyser, University of Antwerp / KU Leuven  
Bernhard Diensberg, University of Bonn  
Edwin Duncan, Towson University, Towson, MD  
Jacek Fabiszak, Adam Mickiewicz University, Poznań  
Jacek Fisiak, Adam Mickiewicz University, Poznań  
Elżbieta Foeller-Pituch, Northwestern University, Evanston-Chicago  
Piotr Gąsiorowski, Adam Mickiewicz University, Poznań  
Keith Hanley, Lancaster University  
Christopher Knight, University of Montana, Missoula, MT  
Marcin Krygier, Adam Mickiewicz University, Poznań  
Krystyna Kujawińska-Courtney, University of Łódź  
Zbigniew Mazur, Maria Curie-Skłodowska University, Lublin  
Rafał Molenccki, University of Silesia, Sosnowiec  
John G. Newman, University of Texas at Brownsville  
Michał Jan Rozbicki, St. Louis University  
Jerzy Rubach, University of Iowa, Iowa City  
Piotr Ruskiewicz, Pedagogical University, Cracow  
Hans Sauer, University of Munich  
Merja Stenroos, University of Stavanger  
Krystyna Stamirowska, Jagiellonian University, Cracow  
Jeremy Tambling, University of Manchester  
Peter de Voogd, University of Utrecht  
Anna Walczuk, Jagiellonian University, Cracow  
Jean Ward, University of Gdańsk  
Jerzy Wehna, University of Warsaw

## GUEST REVIEWERS

Magdalena Bator, University of Social Sciences, Warsaw  
Magdalena Charzyńska-Wójcik, John Paul II Catholic University of Lublin  
Joanna Esquibel, SWPS University  
Anna Fornalezyk-Lipska, University of Warsaw  
Barry Keane, University of Warsaw  
Marta Kisielewska-Krysiuk, University of Warsaw  
Jarosław Krajka, Maria Curie-Skłodowska University in Lublin  
Jadwiga Linde-Usiekiewicz, University of Warsaw  
Sylwester Łodej, Jan Kochanowski University, Kielce  
Zbigniew Możejko, University of Warsaw  
Mieczysław Nasiadka, University of Warsaw  
Jerzy Nykiel, University of Bergen  
Monika Opalińska, University of Warsaw  
Agnieszka Piskorska, University of Warsaw  
Dorota Rut-Kluz, University of Rzeszów  
Marta Sylwanowicz, University of Social Sciences, Warsaw  
Piotr Szymczak, University of Warsaw  
Anna Warso, SWPS University  
Iwona Witczak-Plisiecka, University of Łódź



UNIVERSITY  
OF WARSAW

## Object Length and Other Factors in Prepositional Variation: The Case of *Speak, Talk and Write*

---

### Abstract

The existence of prepositional variation in constructions such as *speaking about/of something* can be motivated semantically in that *about* and *of* invite different construals of the topic they introduce. This article reports on a study carried out to test the hypothesis that the semantic distinction is reflected in the length of the object which complements the preposition. The results suggest that object length might indeed play a part in the choice between *about* and *of*. The study also reveals additional factors which might influence speakers' choice of the topic-introducing preposition.

---

### 1. Introduction

One characteristic feature of all natural languages is the presence of variation, which might be defined as “differences in linguistic form without (apparent) changes in meaning” (Walker 2010, 16). In English, an example of such formal differences is the complementation of lexemes with a prepositional phrase, as in *tolerance for/of*. The present study is concerned with this type of variation as found in three common communication verbs: *speaking, talking* and *writing*.

Each of the predicates can be complemented with *about* or *of* and whichever of these prepositions is used, the meanings of the resulting combinations, e.g. *talk about* and *talk of*, are largely similar in that they both introduce the topic of speech. Thus, these phrases can be used in similar linguistic contexts, as the following examples from the *British National Corpus* demonstrate:

- (1a) Eva and Charlie and I had a drink and **talked about** other things.
- (1b) During the meal, as if by mutual consent, they **talked of** other things, but it was difficult.

The actual choice of the variant preposition might depend on a variety of factors, which include nuanced semantic differences between the prepositions. This possibility constitutes the starting point for the hypothesis tested in this study. It

is assumed that the preference for *about* or *of* might be correlated with the length of the complement which follows the preposition. The hypothesized link between the semantics of *about* and *of* and the length of their objects is addressed in greater detail in Section 2.

Since there are various other factors which might play a part in the choice of the topic-introducing preposition, a number of such characteristics are investigated here, too. They are discussed in Section 3, which also focuses on other methodological aspects of the study, such as data collection. Section 4 presents the results of the investigation, which are then discussed in the closing Section 5.

In this article, the prepositions which introduce the communication topic are termed *topic markers*, following Lindstromberg (2010, 141). The kind of variation under scrutiny is called the *about/of variation* or the *topic marker variation*, while the investigated phrases (*speak about*, *speak of*, etc.) are referred to as *communication topic constructions*.

## 2. Semantics of *about* and *of* in the context of communication verbs

A well-known principle in cognitive linguistics (cf. Lemmens 1998, 14–15) states that a difference in linguistic form necessarily signifies a difference in meaning. In other words, it is assumed that two structurally different constructions cannot be identical semantically. As Dik (1997, 18) advises:

Whenever there is some overt difference between two constructions X and Y, start out on the assumption that this difference has some kind of functionality in the linguistic system. Rather than pressing X into the preconceived mould of Y, try to find out why X and Y are different, on the working assumption that such a difference would not be in the language unless it had some kind of task to perform.

In the light of such an approach to variation, it is reasonable to expect that the *about* constructions, i.e. *speak about*, *talk about* and *write about*, perform different functions from their *of* counterparts. Various authors, presented in the remainder of this section, have provided clues that this assumption might be correct, as they note certain semantic distinctions between the topic markers.

Vorlat (1982, 27) observes that the choice between *of* and *about* to complement *speak* is not dictated by the kind of topic introduced by the predicate but by how the speaker approaches this topic. Vorlat states that when the speaker's intention is to indicate that the subject matter is being elaborated on, he or she might prefer *speak about*. Conversely, when the speaker wishes to suggest a more cursory or focused treatment of the topic, i.e. a simple mention, he or she will opt for *speak of*. Dirven (1982) offers similar observations concerning the complementation of *talk*. The topic-introducing *about* might be selected to intimate that the speaker

is considering every facet of the subject matter, while *of* to convey the sense of focusing on some specific aspects of the topic with the implication that “one could always take more and more elements from it” (Dirven 1982, 62).

A comparison of Vorlat’s and Dirven’s remarks demonstrates semantic similarities between *speak about* and *talk about*, on the one hand, and between *speak of* and *talk of*, on the other. The former pair of constructions conveys a sense of a broader perspective on the topic, while the latter a sense of a more restricted view. This indicates that the *about/of* variation in *speak* and *talk* (and, by extension, in *write*) might be attributed to the meaning of the preposition which constitutes the head of the object.

Other analyses provide evidence for this semantic interpretation: *of*, for a long time treated as devoid of any meaning and serving purely grammatical functions (cf. Langacker 1992, 483–484), might be described as generally having an integrative sense (Lindstromberg 2010, 205–206). In a phrase such as *the peak of the mountain*, the preposition indicates that the subject, *the peak*, is somehow integrated with the object, *the mountain*; here the integration takes the form of a part and whole relation. Other such forms, as noted and exemplified by Langacker (2008, 18), include identity, e.g. *the month of January*, and intrinsic functions, e.g. *the father of the bride*. The sense of integration associated with *of* means that one entity is an “*inherent and restricted subpart*” (Langacker 1992, 484; original emphasis) of another.

These integrative and restricting properties which characterize *of* stand in contrast to the semantics of *about*. In its most immediate, physical sense, the latter preposition indicates dispersion: “location or motion in vicinity but in no particular direction” (Rudzka-Ostyn 2003, 180). This basic meaning is then metaphorically extended “from motion on a surface to mental motion on topics” (Rudzka-Ostyn 2003, 181). As noted by Lindstromberg (2010, 255), in such senses *about* “must once have evoked images of one’s interest and attention enveloping” the subject. In current English, the function of *about* as a topic marker has become so prominent that the word is now the default preposition used to present its object as a topic (Lindstromberg 2010, 141). In the same vein, *Oxford English Dictionary* states in its entry for *about* that the word is currently “the regular preposition employed to define the subject matter of verbal or mental activity in a large number of verbal phrases such as *to speak, think, ask, dream, hear about*”.

The “dispersive” meaning which underlies *about* explains why *speak about* and *talk about* might be used to indicate that the speaker’s attention is directed towards various aspects of a topic. Given this tendency, the topic-restricting meaning of *speak/talk/write of* might have arisen in opposition to the metaphorical meaning of *speak/talk/write about*. If the latter constructions point to dispersion, the *of* variants have come to take on a contrasting meaning of concentration, i.e. one in which the speaker focuses on just selected aspects of the topic. The sense of focus fits well with the integrative behaviour displayed by *of*.

Although the above remarks might seem to suggest that the semantic distinction between the *about* and *of* constructions is almost clear-cut, the difference is in fact rather nuanced. Furthermore, the meanings of the prepositions are only one of various factors potentially influencing the choice of the variant constructions, as was briefly noted in section 1. These possible factors include the variety of English, idiolectal preferences and syntactic priming.

Still, the semantic generalizations described in this section should surface as certain tendencies in the complementation of *speak*, *talk* and *write*. It is hypothesized here that this surfacing is detectable in the length of the object which follows the communication topic constructions. The reasoning behind this assumption is as follows. In communication-related meanings, the use of *about* implies that the speaker is not only focusing on the topic itself, but is also considering its various aspects. Such an extended treatment might result in a longer linguistic expression used to describe the subject matter. From this follows that the prepositional complements used with *speak/talk/write about* are on the whole more likely to be of greater length than those used after *speak/talk/write of*, especially as the *of* constructions imply a more cursory, or concentrated, treatment of the topic.

Put differently, the hypothesis proposed here states that the semantic effects brought about by the prepositions translate into the length of their objects. Such a formal feature, however, might not be the only marker of whether the treatment of a topic is presented as dispersed or concrete. As an example, one study (Krawczak and Glynn in press) investigates the possibility that these semantic effects surface in the abstractness of the topic. The authors start from the same observations concerning the differing construals connected with *about* and *of*. They assume that the sense of indeterminacy inherent in *about* is more likely to attract abstract nouns, while *of*, which invites a sense of focus, is associated with concrete nouns. The present investigation takes account of this hypothesis, too, by considering the abstract-concrete distinction in the data analysis discussed in the following section.

### 3. Methodology

This article reports on a study conducted in order to test the object length hypothesis described above, as well as to uncover other variables which might influence the choice between the topic-introducing *about* and *of*. These goals were pursued by means of collecting and analysing a large number of naturally occurring linguistic data and subjecting the results to two quantitative techniques, the chi-squared test and binary logistic regression.

The first step in data extraction consisted in specifying the source of the linguistic material. The data were taken from the *TIME Magazine Corpus* (Davies 2007); a roughly 100-million word collection of more than 275 thousand texts published in the U.S. American weekly *Time* between 1923 and 2006. The study

does not focus on the differences between the individual constructs in terms of their frequencies, but in terms of the circumstances under which speakers prefer one topic marker over the other. For this reason, the corpus was sampled to generate a dataset in which each of the examined constructions is represented by an equal number of text fragments. This was achieved by selecting twenty random examples of a construction for each of the nine decades covered in the corpus. In other words, the dataset contains 180 examples of each of the six constructions (*speaking about, speaking of* etc.), which amounts to a total of 1,080. This particular sampling procedure was adopted to satisfy the assumptions of the statistical techniques used in the study.

The actual data selection process consisted in querying the corpus interface for fragments in which the verb of interest, in any of its word forms, is followed by *about* or *of*. The verb and the preposition did not need to be adjacent; there might have been up to five intervening words between them, so as to cover expressions such as *speaking in private about*. This maximum number of intervening words was a limit imposed by the corpus's online query interface. The search focused only on a specific ordering of tokens, namely that in which the verb precedes the preposition. In other words, examples such as the one in 2 were disregarded. (This and other quotations from the *TIME Magazine Corpus* are cited with the article title and year of publication.)

- (2) (...) Powers seemed unmindful of history, and the faraway cities **of** which he **talked** were apparently little more than dots on the map to him. (The Boy from Virginia 1960)

As the next step, random text fragments were retrieved from the set of search results. To be included in the study, those examples needed to fulfil a number of selection criteria. Among others, one *Time* article could be the source of at most one example of a specific combination of the verb and topic marker. Furthermore, the construction in each analysed example was required to have a fully compositional meaning and function, which means that fragments such as 3, where *speaking of* serves to indicate some kind of scarcity, were outside the purview of the study. The reason for the exclusion of such cases is that when communication topic constructions are used idiomatically, variation in the choice of the topic marker, if present at all, might depend on factors other than those operating in the case of compositional meaning.

- (3) With no government to **speak of**, even the most powerful warlords have limited influence over their satraps elsewhere and no hope at all of exercising control over free-lance bandits. (Warlord Country 1993)

The collection of text fragments assembled in this way served as the basis for feature analysis (cf. e.g. Rudzka-Ostyn 1989), a method of preparing linguistic data for quantitative investigation. This method consists in examining and annotating

examples of language use with respect to a number of characteristics relevant to the phenomenon under scrutiny. Each such feature, e.g. Tense, typically has a specific set of values to choose from, e.g. past and non-past.

The choice of characteristics to annotate for is driven by the goals of a study. Here, the fundamental annotation feature expresses the length of the object which complements the topic marker; its importance derives from the fact that it is directly related to the hypothesis tested here. The length might be operationalized using various measurements, e.g. words, syllables or phrases. This study makes use of the first of these options; a word is operationally defined here as a sequence of characters separated by spaces. Such a definition makes it possible to easily and confidently perform automated length counts.

As stated in Section 1, the present investigation also examines a range of other factors which might play a part in the selection of the topic marker. First, it tests the hypothesis proposed in Krawczak and Glynn (in press) by recording, in a feature called Abstractness, whether the object of the topic-introducing preposition represents an abstract or concrete entity. A general annotation principle adopted was that if the referent of the object is tactile, it was marked as concrete, otherwise as abstract. However, even though such a guideline might help in rendering the individual annotation decisions more objective, the actual classification of a fragment is still largely dependent on subjective evaluation.

Both Object length and Abstractness refer to the object of a topic-introducing preposition. It is not inconceivable that the *about/of* variation could also be connected with other elements of the communication topic construction, for example the subject. Such a relation has been proposed by Vorlat (1982, 29), who, discussing the verb *speak*, observes that first-person subjects tend to co-occur with *about*. The reason she proposes for this association is that “when reporting about one’s own speech act, one may well tend to refer to a topic as discussed”. To test this observation, the present study records whether the agent of communication in an analysed text fragment is expressed as the first, second or third person. When the agent is not explicitly given, as in the example below, the value “unknown” was used in the annotation.

- (4) It is common to **speak of** the newspapers of today as purely commercial enterprises managed with an eye single to profits (...) (Publishing Morals 1941)

The analysis also focused on the communication verb itself in that it included the features Finiteness, Voice, Polarity and Verb form. The first three are binary: they record, respectively, whether the form of the verb phrase is finite or non-finite, whether that phrase is in the active or passive voice and whether it is negated or not. Verb form can take one of the following values: base, third person singular, past, past participle and present participle. To illustrate, the form of the verb in 5 is annotated as past participle.

- (5) Engman, who has vigorously pushed antitrust actions in his two years at the FTC, has **talked** before **about** “conspiracies of silence” concerning prices in other professions. (Toward Open Pricing 1975)

As mentioned previously, the data search included cases when the verb and the topic marker are separated by up to five words. It is speculated here that the choice between *about* and *of* might be conditioned by the presence or absence of this kind of discontinuity, encoded in a feature called Intervening words. Discontinuity also underlies another feature, Stranded form, which specifies whether the topic marker and its object are adjacent or separated, for instance as in 6.

- (6) The death of such a grammarian as Browning **wrote of** was reported last week from London (...) (Death of a Grammarian 1943)

The features included in the analysis described above were treated as variables to be examined in the statistical tests employed in the study. These variables are summarized in Table 1; note the presence of one not yet discussed: Binned object length. This is a variant of Object length converted from a numerical to a categorical variable in that the range of identified object lengths is divided into three groups: short, medium and long. The transformation was carried out for the purposes of the chi-squared test, which operates on categorical data only.

**Table 1.** Set of variables resulting from feature analysis (for factors, the values in parentheses refer to the number of occurrences of each level)

Variable	Possible values
Topic marker:	<i>about, of</i> (540 each)
Construct:	<i>speak about, speak of, talk about, talk of, write about, write of</i> (180 each)
Object length:	1 .. 41
Binned object length:	short (374), medium (489), long (217)
Topic abstractness:	abstract (823), concrete (257)
Finiteness:	finite (594), non-finite (486)
Polarity:	negative (26), positive (1054)
Voice:	active (1066), passive (14)
Verb form:	base (297), 3rd person (137), past (378), past participle (67), present participle (201)
Stranded form:	stranded (1020), non-stranded (60)
Subject person:	1st (68), 2nd (24), 3rd (943), unknown (45)
Intervening words:	present (285), absent (795)

The outcome of the feature analysis was a so-called wide data table, in which a row represents one observation (a text fragment in this study), while a column represents one variable. The annotated dataset was subsequently inspected with confirmatory statistical techniques, whose results are presented and discussed in the remainder of this article.

#### 4. Results

The first step taken in the quantitative analysis consisted in running chi-squared tests on pairs of factors, where one variable was either Construction or Topic marker, while the other was selected from among the remaining factors. The results are presented in Table 2.

**Table 2.** Results of chi-squared tests ( $p$ -value and effect size), each of which was run on the pairs of variables specified in the row and column headers (numbers in grey signify results which are not reliable because of the data failing to meet all the assumptions of the test)

	Construction		Topic marker	
	$p$ -value	Cramer's V	$p$ -value	Cramer's V
Finiteness	< 0.001	0.327	< 0.001	0.279
Intervening words	< 0.001	0.365	0.003	0.090
Polarity	< 0.001	0.153	0.001	0.097
Stranded form	< 0.001	0.219	< 0.001	0.172
Subject person	< 0.001	0.131	< 0.001	0.196
Topic abstractness	< 0.001	0.209	0.943	0.002
Binned object length	< 0.001	0.153	< 0.001	0.190
Verb form	< 0.001	0.220	< 0.001	0.292
Voice	0.072	0.097	1	0

The  $p$ -values resulting from the analysis indicate a statistically significant result for almost all of the examined pairs of variables. This is to be expected, since the chi-squared test is dependent on the sample size: the larger the number of observations, the more likely it is to obtain  $p$ -values below the conventional level of 0.05. This is why, among other reasons, Table 2 additionally reports Cramer's V values. These metrics are independent of the sample size and represent so-called association measures, or effect sizes. While the  $p$ -values indicate how likely it is that there exists an association between the examined variables, the effect sizes express how large this association is. Cramer's V scores range from 0 to 1, i.e. from a non-existent to a perfect association. The intermediate values are interpreted vari-

ously, the convention adopted here is as stated by Janda (2013, 10–11): scores of up to 0.3 indicate a small effect size, up to 0.5 a moderate effect size, while those including and above 0.5 can be treated as large.

As Table 2 demonstrates, the associations identified by the chi-squared test are mostly of little or moderate strength. The best results, with Cramer's V exceeding 0.3, have been identified for Intervening words and Verb finiteness, either in conjunction with Construction. To establish which combinations of levels are primarily responsible for the relationship, Pearson's residuals can be consulted. These indicate how the observed frequency in a table cell differs from the frequency which would be expected if there was no association between the variables. The results for the present dataset are presented in Table 3.

**Table 3.** Pearson's residuals for the Construction and Verb finiteness and Construction and Intervening words

Construction	Verb finiteness		Intervening words	
	finite	non-finite	present	absent
<i>Speak about</i>	-0.603	0.667	-5.343	8.923
<i>Speak of</i>	2.312	-2.556	-0.217	0.363
<i>Talk about</i>	-5.327	5.889	2.476	-4.135
<i>Talk of</i>	1.709	-1.889	1.086	-1.814
<i>Write about</i>	-1.608	1.778	0.999	-1.669
<i>Write of</i>	3.518	-3.889	0.999	-1.669

Positive and negative values in the table indicate, respectively, that a construction appears in the data more and less often than might be expected. Thus, the Pearson's residuals for Construction and Verb finiteness show a clear difference between the topic markers: the *about* phrases are associated with finite verb forms, while the *of* phrases with non-finite. The table furthermore demonstrates that *talk about* is the phrase primarily responsible for the relationship between the two variables, i.e. between Construction and Verb finiteness.

On the other hand, the lack of independence between Construction and Intervening words is apparently driven not by the preposition, but by the verb, with *Speak* connected with the absence of such words. Further factors revealed by Table 2 to have a noteworthy effect size are Object head, Stranded form and Verb form. An inspection of Pearson's residuals in these cases does not demonstrate any clear preposition- or verb-based patterns such as those in Table 3.

The chi-squared tests have allowed for the identification of differences between pairs of categorical variables. In order to inspect the *about/of* variation by considering all of these variables at once, binary logistic regression is used as a final step in the investigation. This technique involves modelling the behaviour of one binary variable, such as Topic marker, in the simultaneous presence of multiple other

variables, called *predictors* or *explanatory variables*. (For a discussion on the use of logistic regression in linguistics, see e.g. Spielman 2014.) Table 4 presents the results of an analysis carried out on the *Time* corpus data; positive estimates point to an increased probability of encountering *of*, while negative ones are related to *about*. The model was built as parsimonious, i.e. it contains only such predictors which are statistically significant.

**Table 4.** Binary logistic regression model of the *about/of* variation; the success outcome is taken to be *of*

<b>Regressor</b>	<b>Estimate</b>	<b>Probability</b>	
(Intercept)	1.251	0.007	**
Binned object length:			
medium	-0.028	0.878	
short	-0.711	< 0.001	***
Finiteness: non-finite	-0.904	< 0.001	***
Intervening words: present	-0.743	< 0.001	***
Subject person:			
1st person	-1.970	< 0.001	***
2nd person	-0.718	0.219	
3rd person	-0.390	0.299	
Stranded form: stranded	-1.854	< 0.001	***
Verb form:			
3rd person singular	0.729	0.019	*
past	0.144	0.546	
past participle	0.953	0.002	**
present participle	-0.282	0.203	
<b>Diagnostics</b>			
AIC:	1316.7		
ROC:	0.738		
VIF (maximal value):	3.989		

The results presented in Table 4 are consistent with those originating from the chi-squared tests because all of the variables found there to relate to Topic marker are also included in the logistic regression model. A notable predictor is Stranded form, for which the logit of -1.854 corresponds to the estimated odds of 0.157, i.e. when there is stranding, the odds for *of* are approximately one sixth of those when the preposition is directly followed by the object. Interestingly, the model also includes Intervening words, a variable found previously to be connected mostly with the communication verb itself rather than its complementing preposition. However, its effect is not prominent, as might be expected from the results so far.

The model indicates that the *about/of* variation also depends on Verb form. When compared against the referential level of the base form, the use of the present participle increases the probability of encountering *of* as the topic marker, while the remaining forms point to *about*. Another verb-related factor, Finiteness, has proved to be a significant predictor, too, with the level “finite” being related to *about*. It should be pointed out that both of these variables are interconnected in that some of their levels correspond to each other directly, e.g. 3rd person singular forms are always finite. Thus, the two factors contribute to multicollinearity, an undesirable effect in which two or more predictors in a model are strongly associated. Nevertheless, Verb form and Finiteness have been kept in the model because of the still marginally acceptable Variance Inflation Factor (VIF) values, which are measures of the degree of multicollinearity. Values above 4 can be considered problematic (cf. Speelman 2014, 528). The presence of Finiteness and Verb form in Table 4 testifies to the role of the verb in the topic marker variation and so this element of the communication topic construction merits further investigation, so that its influence can be explained.

The remaining two predictors identified as significant are Subject person and Object length. As regards the former factor, first-person subjects are associated with the *about* constructions, and this relationship is among the strongest identified in the model. The estimated odds for *about* are roughly seven times higher when the subject is in the first person than when it is unknown. The presence of Object length in the proposed model supports the hypothesis that the length of the object which complements the topic marker is a predictor of that marker. However, unlike initially assumed, short objects are associated with *about*, not *of*. The strength of that association is not great, as evidenced by the estimates. Incidentally, note that the model makes use of Binned object length, not its numeric variant: this is because the categorical variable leads to a better model in terms of diagnostics.

## 5. Discussion

The main hypothesis tested in this study assumed that the length of the object which complements the topic-introducing preposition is related to the topic marker variation in that *about* and *of* would attract, respectively, longer and shorter phrases. While the logistic regression model presented in Table 4 provides support for the existence of a relation between Topic marker and Object length, the direction of this relation is in reality the opposite of what was expected. It has been found that longer objects are more likely to be used with *of*. A number of possible explanations might be offered for this observed effect.

First, the relationship between the topic marker and the length of the object might arise due to the operation of another, confounding variable. To illustrate, it is not inconceivable that the choice between *about* and *of* relates to register in that

e.g. *of* appears more often in formal contexts. At the same time, increased formality could be concomitant with an increased use of longer phrases. As a result of such supposed tendencies, there would exist a relationship between object length and the choice of the topic marker. Additionally, in the case of short, monosyllabic objects, another reason for preferring *about* can be prosodic, as an utterance such as *They speak about it* might be preferred, in terms of sentence stress, to *They speak of it*.

Theoretically, the role of Object length might also be explained in a more straightforward way. As mentioned in Section 2, *about* is currently the default topic marker in English. Given this position, the word might be expected to be the more frequent choice to complement a communication verb. At the same time, as shown by the data, the phrases which complement topic markers are mostly short. In this investigation, for example, nearly 46% of objects in the examined text fragments have a length of fewer than three words. On this view, the relation between Object length and Topic marker would only be due to a mere co-occurrence of two frequent elements of the language.

While such an effect might indeed exist, it certainly is not the only reason behind the association between Topic marker and Object length. This is because *about* and *of* are represented in the dataset by the same number of text fragments. Thus, if the association was only due to frequency effects as described above, it would not be observed in this study. Furthermore, a preliminary search in the *TIME Magazine Corpus* indicates that *about* is not always the dominant topic marker in communication verbs, as for one of the examined lexemes, *speak*, the more frequent option seems to be *of*.

Whatever the reason for the observed pattern might be, the results demonstrate that if the choice between the topic-introducing *about* and *of* is influenced semantically, then this tendency does not seem to be reflected in the length of the object, at least not in the way assumed in the study. A similar finding concerns the abstractness of the topic. This variable has been tested by Krawczak and Glynn (in press), who found it to have only a limited impact on the selection of the topic marker. The present study uses a somewhat different, more coarse-grained operationalization of abstractness, yet has likewise failed to identify any relation between that variable and the topic marker. Admittedly, the results of the chi-squared test indicate that there is a statistically significant lack of independence between Construction and Topic Abstractness, yet this relationship is weak, as shown by the effect size.

Possibly, the semantic effects of the topic-introducing *about* and *of* might come to the fore not in the object of the verb, but in other elements of the communication topic construction. These could be, for example, adverbials of manner which modify the communication verb. They might be investigated in a feature analysis which attempts to distinguish between adverbials semantically compatible with the dispersion indicated by *about* and those compatible with the integrative sense inherent in *of*. Yet another possible approach to examining the role which the semantics of *about* and *of* play in topic marker variation consists in attempting to directly specify

whether the object of the preposition is presented as elaborated upon or not. For a feature analysis, the topic markers should best be removed from the text fragments under scrutiny (e.g. “He spoke . . . his experiences”), so as not to suggest any interpretation. Preferably, such an analysis would involve the work of two or more annotators, whose consistency can be specified with inter-rater agreement measures.

As regards the remaining factors investigated in the study, one found to play a part in the variation is Intervening words, whose presence is related to *about*. As discussed above, this variable seems to be associated first and foremost with the verb itself. Still, logistic regression demonstrates that Intervening words is a significant predictor of the topic marker, even if its effect is not strong. An explanation of this relation can be proposed which pertains to linguistic processing: the further away the topic-introducing *of* is from the communication verb, the less transparent its role as a topic marker might become, due to the various functions fulfilled by this word. Consequently, in situations when speakers place any phrases between the preposition and its object, they might opt for *about* as potentially less ambiguous in its topic-introducing role.

The relation between Topic marker and Intervening words could also be an expression of the underlying semantics of the topic-introducing *of*. As discussed, the preposition is characterized by an integrative meaning; these semantic properties might result in a tendency for this word to be used in the close vicinity of the phrases whose referents it integrates. An advantage of this explanation is that it could also be used in relation to Stranded form. The quantitative analysis revealed that when the topic marker is non-contiguous with its object, the preferred preposition is *about*. The reason why *of* appears less often in such settings might again be connected with the “binding” nature of that preposition. Thus, the preference for *about* as a stranded topic-introducing preposition could result from the different meanings of the two topic markers: an expression of semantic tendencies which was sought in the object length hypothesis.

## References

### (a) Database

*The British National Corpus, Version 3 (BNC XML Edition)*. 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. <http://www.natcorp.ox.ac.uk/>

Davies, Mark. 2007. *TIME Magazine Corpus: 100 million words, 1920s–2000s*. <http://corpus.byu.edu/time/>

### (b) Dictionaries

*Oxford English Dictionary*. <http://www.oed.com/>

**(c) Special studies**

- Dik, Simon C. 1997. *The Theory of Functional Grammar: The Structure of the Clause*. Berlin: Walter de Gruyter.
- Dirven, René. 1982. "Talk : Linguistic Action Perspectivized as Discourse." *The Scene of Linguistic Action and its Perspectivization by Speak, Talk, Say and Tell*. Eds. René Dirven, Louis Goossens, Yvan Putseys, and Emma Vorlat. Amsterdam, Philadelphia: John Benjamins. 37–85.
- Janda, Laura A. 2013. "Quantitative Methods in Cognitive Linguistics: An Introduction." *Cognitive Linguistics – The Quantitative Turn: The Essential Reader*. Ed. Laura A. Janda. Berlin, Boston: Walter de Gruyter. 1–31.
- Krawczak, Karolina, and Dylan Glynn. In press. "Operationalising Construal. *Of About* Prepositional Profiling for Cognitive and Communicative predicates." *Construals in Language and Thought: What Shapes What?* Ed. Carmen M. Bretones Callejas. Amsterdam: John Benjamins.
- Langacker, Ronald W. 1992. "The Symbolic Nature of Cognitive Grammar: The Meaning of *Of* and of *Of*-periphrasis." *Thirty Years of Linguistic Evolution: Studies in Honour of René Dirven on the Occasion of his Sixtieth Birthday*. Ed. Martin Pütz. Philadelphia, Amsterdam: John Benjamins.
- Langacker, Ronald W. 2008. "The Relevance of Cognitive Grammar for Language Pedagogy." *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven*. Ed. Sabine De Knop, and Teun D. Rycker. Berlin, New York: Mouton de Gruyter. 7–36.
- Lemmens, Maarten. 1998. *Lexical Perspectives on Transitivity and Ergativity: Causative Constructions in English*. Amsterdam, Philadelphia: John Benjamins.
- Lindstromberg, Seth. 2010. *English Prepositions Explained* (2nd ed.). Amsterdam: John Benjamins Publishing Company.
- Rudzka-Ostyn, Brygida. 1989. "Prototypes, Schemas, and Cross-category Correspondences: the Case of *Ask*." *Linguistics* 27.4: 613–662.
- Rudzka-Ostyn, Brygida. 2003. *Word Power: Phrasal Verbs and Compounds, a Cognitive Approach*. Berlin, New York: Mouton de Gruyter.
- Speelman, Dirk. 2014. "Logistic Regression: A Confirmatory Technique for Comparisons in Corpus Linguistics." *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Ed. Dylan Glynn, and Justyna A. Robinson. Amsterdam, Philadelphia: John Benjamins. 487–533.
- Vorlat, Emma. 1982. "Framing the Scene of Linguistic Action by Means of *Speak*." *The Scene of Linguistic Action and its Perspectivization by Speak, Talk, Say and Tell*. Ed. René Dirven, Louis Goossens, Yvan Putseys, and Emma Vorlat. Amsterdam, Philadelphia: John Benjamins. 9–36.
- Walker, J. A. 2010. *Variation in Linguistic Systems*. New York: Routledge.