The role of verb polysemy in constructional profiling
A cross-linguistic study of GIVE in the dative alternation

Karolina Krawczak
Faculty of English, Adam Mickiewicz University in Poznań
karolina@wa.amu.edu.pl

Abstract: This study employs corpus-based quantitative methods to investigate the interaction between the semasiological structure of a single verb and the onomasiological structuring of the dative alternation in English and Polish. More precisely, it zooms in on the TRANSFER construction, as instantiated with the lexical category GIVE in the two languages. The primary objective is to examine the relationship between morpho-syntactic variation and lexical semantic variation. More specifically, the study addresses the importance of accounting for semasiological lexical structure in modeling morpho-syntactic structure. It is argued here that the polysemous nature of lexemes that are licensed by constructions has an impact on the choice of alternate constructional variants. In other words, some meanings of a given lexeme are likely to be more distinctly associated with one construction than the other. The results obtained in the study for both English and Polish provide supporting evidence for this claim.

Keywords: dative alternation, GIVE-verbs, semasiology, onomasiology, multivariate statistics, cross-linguistic

1. Introduction

The present study addresses the importance of semasiological variation of a lexeme licensed by a given construction[1] to explaining onomasiological variation observed in morpho-syntactic structure. More precisely, the question we are posing here is whether the polysemy of a lexeme can offer a key to understanding speaker's choice of alternative constructions. The importance of lexical semantics in constructional inquiries has been pointed out before (e.g., Boas 2003; Glynn 2004), but it is a question that is yet to be explored empirically in relation to the otherwise extensively studied dative alternation. The alternation, obtaining between two dative constructions, is here investigated in association with the verb *give* in English and with its perfective and imperfective equivalents in Polish, i.e., *dać*/*dawać*. The two variants are illustrated in (1) and (2):

[1] Construction is here understood in the Goldbergian sense of a "form and meaning pairing" (Goldberg 2006).

(1) Cx A: ['give' + RECIPIENT + THEME]
a. *She gave* [*Peter* $_{\text{NP RECIP}}$] [*the keys* $_{\text{NP THEME}}$].
b. *Dała* [*Piotrowi* $_{\text{DAT NP RECIP}}$] [*klucze* $_{\text{ACC NP THEME}}$].

(2) Cx B: ['give' + THEME + RECIPIENT]
a. *She gave* [*the keys* $_{\text{NP THEME}}$ [*to* [*Peter* $_{\text{NP RECIP}}$]$_{\text{PP}}$].
b. *Dała* [*klucze* $_{\text{ACC NP THEM}}$] [*Piotrowi* $_{\text{DAT NP RECIP}}$].

The formal difference between the two constructions lies in their word order. Construction A, as exemplified in (1a) for English, where the proper noun designating the recipient precedes the theme, is referred to as the double object construction. The other construction, Cx B, illustrated for English in (2a), where the order of participants is reversed, is known as the prepositional dative. In Polish, both variants are double object constructions and case marking is used to differentiate between the two objects in the argument structure, with the Dative marking the RECIPIENT, and the Accusative indicating the THEME. In addition, in Polish, the verb slot can be instantiated by either the perfective or imperfective form of the lexical category GIVE.

Overall, at the level of function, there is a subtle difference in meaning between the two constructional variants, which can be understood in terms of a shift in focus and topicality: while Cx A can be described as profiling the new possessor and the state of possession or control which is transferred, Cx B can be said to foreground the possessed object and its change of location (cf. e.g., Newman 1996: 62f.; Goldberg 2002). As Bresnan *et al*. (2007) point out, the function of the former constructional choice focuses on "causing a change of state (possession)", while the latter on "causing a change of place (movement)". This realization can be attributed, among others, to Green (1974) and Pinker (1989).

Prior research on the dative alternation is extensive, particularly in English, where the construction has received considerable attention from different theoretical

paradigms (e.g., Arnold *et al.* 2000; Gries 2003b; Bresnan *et al*. 2007; Bresnan & Ford 2010; Bresnan & Hay 2008; Bresnan & Nikitina 2009; Theijssen 2012; Wolk *et al.* 2013). This extensive research has identified a range of formal and discursive variables that were demonstrated to explain the grammatical variation adequately. These explanatory variables include: (a) Animacy of Recipient, (b) Definiteness of Recipient / Theme, (c) Pronominality of Recipient / Theme, (d) Givenness of Recipient / Theme, (e) Length of Recipient / Theme, and (f) Person of Recipient. In this context, Collins (1995) puts forward the so-called "Receiver / Theme Differentiation" principle, which Bresnan *et al*. (2007) link to what in Optimality Theory is known as Harmonic Alignment. This principle states that objects that are given, definite, shorter and pronominal come before those that are non-given, indefinite, longer and nominal. Put differently, objects that are less demanding in terms of cognitive processing are more likely to be aligned with the "immediately post-verbal position", whereas those that require expending more mental energy to comprehend are more likely to be relegated to the sentence-final position (Bresnan *et al*. 2007). This principle, which is further motivated by the "end-weight" rule (Behaghel 1910; Arnold *et al.* 2000; Wasow 2002), is claimed to override any other constraints.

Another related proposal that also explains the variation by focusing on the clause-final position of a given element is referred to as a "principle of dominance" (Erteschik-Shir 1979: 451, as quoted in Williams 1994: 40). It holds that those elements that the speaker intends to bring to the interlocutor's attention and that are discursively dominant will come at the end of a sentence. In a similar vein, Thompson (1988), with regard to the post-verbal position of the object, posits that it is most likely to be allocated to "topicworthy" objects, which, in addition to being

characterized by the features listed above for this slot of the constructions, also tend to designate animate things and be instantiated by proper nouns (Williams 1994: 41).

As already mentioned above, none of the prior studies examining the dative alternation has included among its predictors the semasiological structure of the verbs pertaining to this constructional variation. What is more, in one of the most influential works in the field, Bresnan *et al.* (2007) treat the variable as random in their final logistic regression analysis, thus excluding it from the predictive modeling of this linguistic phenomenon. One reason why the semantic variation of the verbs partaking in the alternation has not been taken into account may be the overwhelming scope of the task. To account for the multiple meanings of even only the most frequent verbs associated with the alternation would present a considerable challenge. To reduce the complexity, this study focuses on the semasiological structure of a single verbal category, i.e., GIVE, which is the basic-level exponent of the TRANSFER construction and its most prototypical member (cf. Newman 1996).

The **polysemy of GIVE** has been discussed in detail by Newman (1996), who draws a map of the various semantic extensions radiating from the central literal sense of physical transfer of an object between animate agents. It is this network of meanings that has served as the basis for establishing the values for which to annotate the uses of GIVE attested in the dataset (see Table 2 for the list of values). Let us, therefore, discuss the internal semantic distinctions within the GIVE category, as proposed by Newman (1996). The central and literal sense of GIVE involves the movement of a certain entity not only from one physical entity to another, but also from one control zone to another (Newman 1996: 144). In addition to this focal meaning, which provides the conceptual source for figurative mappings, Newman (1996: e.g., 77, 233) identifies eight such metaphorical extensions, noting that even

though the list is not in itself exhaustive, it represents accurately the extended uses that are possible. Among the senses thus enumerated that are also relevant to the present study we find the following (after Newman 1996: 77): (i) interpersonal communication; (ii) emergence; (iii) causation; (iv) enablement; and (v) schematic interaction. Before we briefly discuss each of the senses,[2] it is noteworthy that the common denominator of all the figurative extensions is a sense of "abstract motion" of an entity from one point to another, where the origin of the motion is conceived of as the trajector, while the thing in motion and its destination (if present) constitute two distinct landmarks (Newman 1996: 138, 224).

Now, with regard to the first sense, i.e., interpersonal communication, it concerns interactive events between animate agents, with the latter understood literally or metonymically. This usage can be illustrated by such interpersonal acts as *giving advice*, *giving one's word* or *giving an order*, all of which have their respective equivalents in Polish – *dać radę*, *dać słowo*, *dać rozkaz*. This extension, as noted by Newman (1996: 137f.) relies heavily on the conceptualization of communication in terms of the CONDUIT metaphor, first introduced by Reddy (1979). The next metaphorical extension is subsumed under the umbrella term of emergence. In uses that fall into this category, one entity (a landmark) comes *out of* another (a trajector), being thus produced or caused. Some pertinent examples in English that instantiate this usage include *giving milk* (about a cow), *giving fruit* (about a tree), *giving shade* (about a tree), *giving warmth* (about the sun or fire) or *giving a sound*. The Polish exponent of GIVE can be used in the same manner, as evidenced by *dać mleko* ('give milk'), *dać owoce* ('give fruit'), *dać cień* ('give shade'), *dać ciepło* ('give warmth') or *wydać dźwięk* ('out-give a sound'). In the last case the verb *dać* is prefixed by *wy-*,

[2] The discussion of the subsences along with the examples is based on Newman (1996: 136ff.).

whose function additionally intensifies the sense of emergence. Causation and enablement are the next two senses identified by Newman (1996), where the occurrence of one thing is engendered or made possible by another, as in *give sb a job / a promotion* or *give sb the right to do sth* and the respective equivalents in Polish, i.e., *dać komuś pracę / awans*, *dać komuś prawo do czegoś*. Another clear example of causation would be English *give sb a headache* and Polish *dać komuś coś do zrobienia* ('have sb do sth'), while enablement could be illustrated by *dać komuś coś zrobić* ('let sb do sth'). The last subsense of GIVE that is relevant here has to do with schematic interaction between entities, as in *give sth a try*, *give sth a wash* or *give sb a kiss,* only the last of which finds its correspondence in Polish – *dać komuś całusa*. In such constructions, as noted by Newman (1996: 202), it is the nominal element that elaborates on the otherwise "schematic" semantic structure of GIVE. It should be pointed out, however, that the scaffolding provided by GIVE in such uses also adds an element of intentional and telic behavior on the part of the instigator (Newman 1996: 202). As we will see in Section 2, the senses discussed above are further refined in light of the usage nuances found in the data.

The present study has a number of descriptive **goals**. In the most general terms, the objective is to test the findings of Bresnan *et al*. (2007) for the dative alternation in English. However, there are three important differences between their study and the present inquiry. Firstly and most importantly, based on the assumption that semasiological variation (polysemy) of the verb contributes crucially to the onomasiological structuring of constructions, this study includes lexical semantic contribution of to the constructional profiling. As already mentioned, Bresnan *et al*. (2007) exclude lexical effects from their logistic regression model by treating verb sense as a random variable. To make this inclusion feasible, the analysis here is

limited to only one lexeme. In so doing, we also avoid the problem pointed out, for example, by Gries & Stefanowitsch (2004), namely that different verbs that are licensed by the dative constructions will often have their own preferences for one constructional variant or the other. Secondly, in order to verify the results for English, the study employs a different type of data: spontaneous, dialogic and blog-based. Such data can be said to lie between written and spoken registers, as used by Bresnan *et al.* (2007), where the former is more typically well thought through and carefully edited. Finally, the analytical tools will also be applied to another language, Polish, to test the relevance of the predictors to explaining the variation in a language that has not been analyzed in this respect before.

There are three corresponding **hypotheses** that will be tested in the study. Firstly, it is expected that the results obtained in Bresnan *et al*. (2007) will be confirmed for the new dataset for English. Secondly, it is assumed that the integration of the semasiological variation of the verb into the model will improve the descriptive and predictive accuracy of the analysis. Finally, it is also hypothesized that the results will extend to Polish, thus explaining in a statistically significant and predictively accurate model the choice between the two constructions in this language.

2. Method and Data

The method employed in the present study is known as the **Profile-based Approach** (Gries 2003a, 2006) or the **Multifactorial Usage-Feature Analysis** (Glynn 2009, 2010a, 2010b, 2014). It has been developed within the framework of Cognitive Linguistics in the work of Geeraerts *et al*. (1994, 1999), Gries (1999, 2003a, 2003b, 2006), Heylen (2005), Gries & Stefanowitsch (2006), Divjak (2006), Glynn (2009, 2010a, 2010b, 2014), Speelman & Geeraerts (2009), Glynn & Fischer (2010) or

Glynn & Robinson (2014). It aims at identifying frequency-based behavioral profiles of the linguistic phenomenon under investigation, which is achieved in a two-step procedure. Firstly, all the contextualized examples are annotated manually for a range of usage characteristics, which may include purely morpho-syntactic features, but which may also incorporate semantic and sociolinguistic values. Depending on whether these variables are directly observable or operationalizable in such terms, the process of data annotation can be automatized to varying degrees. This procedure of data annotation results in a complex matrix of multifactorial interactions, whose processing and subsequent interpretation calls for dedicated analytical tools. The metadata are therefore submitted to multivariate statistical modeling, which makes pattern identification possible. Multivariate methods, as the name suggests, allow us to account *simultaneously* for the impact of all the variables that we deem crucial to explaining the linguistic behavior in question. Importantly, such methods, in addition to revealing the frequency-based behavioral profiles, allow us to test our hypotheses in a rigorous manner.

The data in this study amount to over 600 occurrences of the two constructions in Polish and American English. The summary of the data is provided in Table 1.

Table 1: Data summary

Construction	**American English**	**Polish**	**Total**
Cx A: S GIVE RECIPIENT THEME	160	146	306
Cx B: S GIVE THEME RECIPIENT	153	153	306
Total	313	299	612

The observations were extracted from the blog-based components of the *TenTen* corpus for the two languages (SketchEngine, Kilgarriff *et al*. 2014). The extraction was based on regular expressions, which was followed by manual cleaning of the data. In the cleaning process, any observations that did not contain all three arguments

(i.e., Subject, Theme, Recipient) or that were highly idiomatic were excluded from the analysis.[3] All the contextualized examples were then manually annotated for a clearly defined set of usage-features. In addition to the variables found significant in Bresnan *et al*. (2007), i.e., variables (i)-(viii) in Table 2, the data were also tagged for verb sense, which is crucial here as it relates to the central claim of the paper, i.e., the importance of verb polysemy to the onomasiological choice between constructional variants. The complete annotation schema is presented in Table 2, where all the variables and their respective features (or values) are enumerated.

Table 2: Annotation Schema

	Variable	**Feature**
(i)	Recipient Animacy	Animate, Inanimate
(ii)	Definiteness	Definite, Indefinite
(iii)	Pronominality	Pronominal, Nominal
(iv)	Givenness	Given, Non-Given
(v)	Theme Concreteness	Concrete, Abstract
(vi)	Person of Recipient	First, Second, Third
(vii)	Number	Singular, Plural
(viii)	Length	Calculated as the natural log of the difference in the number of words btw. Theme & Recipient
(ix)	Lexical Sense	Causation, Change of State, Communication, Emergence, Enablement, Physical Contact, Render Available, Transfer of Possession

We will now consider examples for the variables that require some explanation with regard to the decision process that was followed in the annotation. The variables that are more directly observable, i.e., Animacy, Definiteness, Pronominality, Givenness, and Concreteness are illustrated in examples (3) and (4) and then discussed in greater detail below.

(3) *So everything we have in our home is very new, 8 years old or younger so I have been looking for older things to* ***give the home a warmer, aged feeling****.*
(Definite, Inanimate & Given Recipient; Indefinite, Abstract & Non-Given Theme)

(4) *I'll* ***give you an example****.*
(Pronominal, Animate & Given Recipient; Nominal, Non-Given & Abstract Theme)

[3] Bresnan *et al*. (2007) also exclude such observations from their analysis.

With regard to the first variable of Recipient Animacy, both human and animal objects were tagged as animate. This is in line with what Bresnan *et al.* (2007) propose in their analysis, having thus simplified the more complex schema adopted by Garreston *et al*. (2004). Metonymic referents in this position were also annotated as animate, as illustrated in (5).

(5) *Our planning reforms will put local communities in the driving seat by **giving** new powers to **neighborhoods** to write their own plans*.

Definiteness was a feature ascribed to objects that were personal pronouns, proper nouns or that were accompanied by a possessive pronoun, a demonstrative determiner or definite article. The last usage characteristic is absent in Polish, where articles do not exist, which is why the task was slightly more complex for this language. However, reliance on the immediate context normally sufficed to address the problem, as illustrated in (6), where the previous sentence makes it clear that the Recipient, even though not preceded by any determiner, is specified and definite. Example (3), in turn, provides an illustration in English of a clearly definite object occupying the Recipient position and an explicitly indefinite object designating the Theme.

(6) *Zaproponowałam więc, żeby troszkę konie rozluźnić i zakłusowałyśmy, a chwilę później pozwoliłyśmy koniom wyciągnąć nogi w galopie. ... Zwolniłyśmy po jakimś czasie,*

dałyśmy	***koniom***	*chwilę*	*wytchnienia*	*w stępie*
gave-1PL	horses-DAT	moment-ACC	rest-GEN	in walk-LOC

'So I suggested relaxing the horses a bit and we trotted, and a moment later we let the horses stretch out their legs in a gallop. It was wonderful … We slowed down after a while, we gave the horses a moment to rest in the walk and Max used it to play with the waves attacking his legs.'

The next variable for which the data were annotated is Pronominality, where objects "headed by pronouns (personal, demonstrative, and indefinite)" (Bresnan *et al*. 2007) were assigned the feature Pronominal. All other objects were classified as Nominal. An example of a pronominal object in the Recipient position is given in (7),

where the pronoun *ktoś* (‘someone’) is used:

(7) *Ja jeśli już **daję komuś** +*
I if at all give-1SG someone-DAT +
to za to, że udzielił najlepszej odpowiedzi ...
‘If I give someone a plus at all, it is because they have provided the right answer.'

The discourse status of the object, i.e., its newness vs. givenness, also referred to as “accessibility in discourse” (Bresnan *et al*. 2007), was established on the basis of whether or not it was possible to identify its co-referent in the preceding few sentences. In addition, personal pronouns referring to the first and second person singular and plural were also treated as discursively accessible to the addressee. The same practice is followed in Bresnan *et al*. (2007), where, in turn, the authors adopt the procedure employed by Prince (1981) and Michaelis & Hartwell (2007). Examples (3) and (4) given above are a clear illustration of both values.

The last variable that was also demonstrated to be a significant predictor in Bresnan *et al*. (2007) and that calls for some clarification is Theme Concreteness. The feature <Concrete> was assigned to objects that were spatially defined and could be described as experienceable through perception (cf. Krawczak *et al*. 2016), whereas objects that had no perceptible physical form were annotated as Abstract. These two values are illustrated in examples (8) and (3), respectively.

(8) *The fever itself rose very slowly throughout the week, despite our **giving** her **intravenous antibiotics** three times a day at home, and it looked like we might have to hospitalize her, but Thursday evening the fever dropped, and was gone by Friday*.

Finally, let us turn to the different senses that were identified for the verbal category GIVE in the uses that were attested in the data. As already indicated, these values are largely based on the semasiological network proposed by Newman (1996) and discussed in Section 1. Sentences (9)-(16) illustrate the individual values of <Lexical Sense>, which were decided on the basis of contextual clues. We will here only discuss those senses that have not been explained above, i.e., <Change of State>,

<Physical Contact> and <Render Available>, as illustrated in (10), (14) and (15), respectively.

(9) *Duke had somehow banged the leg on something and **given himself a contusion***. (Causation)

(10) *It **gives me great comfort** to belong to this huge family.* (Change of State)

(11) *Today, I was trying to drown out the noise do something on my computer as Nick **was giving me a running commentary** on a television show he was watching*. (Communication)

(12) *Ich zieloną i kwitnącą oprawę stanowią drzewa, tak często rosnące przy kapliczkach. **Dają cień wędrowcom**, odwiedzającym nadsańskie okolice*. (Emergence)
'Their green and blossoming frame is provided by trees, so often growing near chapels. They give shade to travellers visiting the area.'

(13) *Mityng Weltklasse **dał okazję kilku światowym gwiazdom do rewanżu***. (Enablement)
'Mityng Weltklasse gave an opportunity to a few world starts for revenge.'

(14) *The results seemed promising and so today after I picked him from the ground and **gave him a big hug***. (Physical contact)

(15) *Na szczęście ruszała się, usiadła. Jedyna myśl to szybko zadzwonić po pogotowie. **Dałam mamie moją komórkę***. (Render available)
'Fortunately, she moved, sat up. My only thought was to quickly call an ambulance. I gave mom my mobile phone.'

(16) *The complete run was marked at $10, but Retailer Tim **gave it to me for five bucks***. (Transfer of possession)

<Change of State> refers to situations in which GIVE is used to indicate that a given entity, which is most likely to be animate, moves metaphorically from one state to another, as in (10). The change may also concern the intensification of an already existing state experienced by a subject, as illustrated in (17).

(17) *For a while there I was trying to be a poet and reading a lot of poetry, which **gave my hatred for the contemporary literary scene bite and drive***.

The next semantic value ascribed to GIVE designates <Physical Contact> between the grammatical subject of the verb and the Recipient. The type of contact between the two entities is determined by the Theme, as can be witnessed in (14). This feature is related to what Newman (1996) refers to as <Schematic Interaction>, but it is more specific in that it only involves events of physical interaction.

The last sense that has been identified here and that differs from the list proposed by Newman (1996) is <Render Available>, where an entity passes from one zone of control to another, but this is likely to be a temporary state or to involve

situations where a thing or service are provided voluntarily. The former case is exemplified in (15), where the speaker passes her phone to her mother so that she can make a phone call.

All the observations of the two constructional variants were manually annotated for the variables presented in Table 2 and discussed above. The annotation was performed methodically for the two languages, resulting in a complex grid, where each observation was accompanied by nine tags specifying its usage characteristics. In order to identify the contextual environments that determine the choice between the two constructions and to test the hypotheses put forward at the end of Section 1, the metadata were submitted to multivariate statistical modeling in the form of logistic regression analysis.

3. Results and Discussion

Four logistic regression models were fitted, two for English (see Tables 3 & 4) and two for Polish (see Tables 5 & 6). All the models were checked for multicolinearity and none of the factors had variance inflation of more than 3.28. This highest value was found for Polish Model 2 (Table 6). With regard to the two models for English, one of them (Table 2) included only the factors accounted for in Bresnan *et al*. (2007), the other (Table 3) also included the lexical senses of the verb. The reason behind having two models was to check whether the integration of the semantic variable into the analysis would improve the predictive power of the model, as hypothesized here. With respect to the two models for Polish, the same procedure was followed, i.e., Model 2, unlike Model 1, includes <Lexical Sense> in the predictors. Let us turn now to Model 1 for English, presented in Table 3, to see which of the variables considered by Bresnan *et al*. (2007) prove significant in our analysis.

Before we consider the results, let us explain briefly how to interpret the table. The first column of Table 3 lists all the variables (or predictors) and their respective values that were found to be statistically significant in this logistic regression analysis. The other two columns of the table specify the effect size and level of significance of the correlation identified between a given value of the predictor and one of the two constructional variants. The correlations that we observe here provide only partial support for the findings of Bresnan *et al*. (2007), indicating that the dimensions of Pronominality, Giveness, Definiteness and the Length Difference between the Recipient and the Theme are indeed significant in distinguishing between the use of the two constructions in English. Interestingly, however, no confirmation is obtained for the importance of Recipient Animacy and Theme Concreteness.

Table 3: English Model 1. Fixed-Effects Binary Logistic Regression
Dative Alternation ~ Bresnan Features

		Effect Size / Coefficient & Significance	
Predictors		**Cx A**: Recipient – Theme	**Cx B**: Theme – Recipient
Recipient Pronominality:	Pronominal	1.4787 (***)	————
Recipient Givenness:	Non-Given	————	1.8281 (***)
Theme Definiteness:	Indefinite	1.1360 (**)	————
Theme Pronominality:	Pronominal	————	1.4394 (***)
Length (log scale):	Longer Recipient	————	1.8984 (***)
Model Statistics			
C statistic: 0.94			
Nagelkerke R^2: 0.70			
AIC: 211.10			

Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1

Looking more closely at the correlations revealed here, we can see that Pronominal Recipients are significantly associated with the double object construction (Cx A), while Pronominal Themes are significantly correlated with the prepositional dative (Cx B). Similarly, Indefinite Themes are important predictors for the double object construction (Cx A), whereas Non-Given and longer Recipients are significantly linked to the prepositional dative (Cx B). This confirms the Harmonic Alignment

principle mentioned in Section 1. No multicolinearity was identified in this model, with the variance inflation factor of no more than 1.25. The overall performance of the model can be evaluated as exceptionally good with the C statistic at 0.94 and the Nagelkerke R^2 at 0.70, where normally "acceptable discrimination" is achieved with the C statistic measure of 0.70 (Hosmer & Lemeshow 2000: 162). Let us now see what happens when the semasiological variation of the verb is added to the model. Table 4 shows the results of this analysis.

Analysis of variance (anova), which enables us to compare models, reveals that the difference between English Model 1 (Table 3) and English Model 2 (Table 4) is statistically significant with p=0.01057(*). This means that adding the variable of <Lexical Sense> represents an important improvement in the predictive modeling procedure. The findings obtained here confirm the importance of the same explanatory variables that were identified as significant predictors of the linguistic choice in the simpler model presented in Table 3. The Harmonic Alignment rule is thus again fully corroborated here. More importantly, the results clearly show that, in line with our hypothesis, the descriptive accuracy and predictive power of the model is indeed improved when semasiological variation of the verb is added to the analysis. This can be evaluated on the basis of the C statistic and Nagelkerke R^2, both of which are higher in Model 2. At the same time, even though the complexity of the model is increased, its parsimony is comparable, if not improved, as can be observed on the basis of the AIC score, which drops slightly in Model 2.

Table 4: English Model 2. Fixed-Effects Binary Logistic Regression
Dative Alternation ~ Bresnan Features + Lexical Sense

		Effect Size / Coefficient & Significance	
Predictors		**Cx A**: Recipient – Theme	**Cx B**: Theme – Recipient
Recipient Pronominality:	Pronominal	1.5649 (***)	————
Recipient Givenness:	Non-Given	————	1.6835 (***)
Theme Definiteness:	Indefinite	0.9924 (*)	————
Theme Pronominality:	Pronominal	————	1.7380 (***)
Length (log scale):	Longer Recipient	————	2.0519 (***)
Lexical Sense:	Causation	4.1879 (*)	————
Lexical Sense:	Communication	————	1.8092 (.)
Lexical Sense:	Render Available	————	1.3706 (*)
Model Statistics			
C statistic: 0.95			
Nagelkerke R^2: 0.742			
AIC: 207.16			

Signif. Codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1

What is more, as indicated by the coefficients, it is the semantic variable that is the most important predictor in this model, with the effect size of 4.769 for the value <Lexical Sense: Causation>. The correlation identified in this respect predicts that when GIVE is used in the TRANSFER construction to designate "Causation", it is Cx A, where the Recipient precedes the Theme, that will be chosen by the speaker. This tendency identified here is not unmotivated. It is only natural that in uses such as those observed in (9) or (18), encoding causation, the speaker should opt for word order that finds its reflection in experience, i.e., the instigator of the caused process (causer) impacts upon the receiver (cause) so that the end-result is engendered. In example (9), the instigator and the recipient coincide and the result is clearly undesired and unintentional, while (18) illustrates a more typical case of causation with the causer and the □ause as two independent agents and the caused event representing a desired and intended result.

(18) *The beings of the Heap go into your house at night and touch objects in your house to collect good memories that they combine to* ***give you good dreams****.*

This observed correlation between Cx A and the meaning of causation provides

supporting evidence for a more general association proposed in prior research (e.g., Green 1974; Pinker 1989; Bresnan *et al.* 2007) and discussed in Section 1, whereby Cx A is said to express the schematic meaning of "change of state". The two other values of <Lexical Sense> that are significant predictors, i.e., the senses "Communication" and "Render Available", both are correlated with Cx B, where the Recipient occurs in the clause-final position and is preceded by the Theme, as in examples (19) and (20), respectively:

(19) *We got the mainsail down, and with Twinkle Toes on the helm **giving instructions to Santa Claus** on the throttle, we pulled into the slip in triumph*.

(20) *She is very active in the local chapter of Amnesty International, she goes once a week to the train station to **give medical care to homeless people**, she took a second job at the major's office.*

The word order predicted for these two senses could be said to be a more natural reflection of the perceived directionality of events in reality, where the giver provides something that s/he has and is in control of to the Recipient. This correlation also supports the claim discussed in Section 1 that Cx B encodes the abstract meaning of "change of location". Finally, it should be noted that in the analysis presented in this model, the highest variance inflation factor score is 3.7 for the variable <Lexical Sense> and, more precisely, the usage feature <Enable>, which is not returned as a significant value.

Let us now turn to the logistic regression model in Table 5, which presents the results for Polish. We can see that four of the factors that Bresnan *et al*. (2007) find to be significant predictors of the dative alternation in English emerge as such for Polish, i.e., <Recipient Pronominality>, <Recipient Givenness>, <Theme Givenness>, and <Length Difference>. Compared to the independent variables that were revealed to be significant in explaining the alternation in English (Tables 3 and 4), <Definiteness> of either the Recipient or the Theme is not a significant predictor for Polish. Why this should be so is not immediately clear, but one reason could have to do with the fact

that Polish does not have an explicit marker of definiteness or indefiniteness in the form of the respective articles. We do find, however, that <Theme Givenness>, which was not found to be significant for English in our analyses above, is a significant predictor here.

Table 5: Polish Model 1. Fixed-Effects Binary Logistic Regression
Dative Alternation ~ Bresnan Features

		Effect Size / Coefficient & Significance	
Predictors		**Cx A**: Recipient – Theme	**Cx B**: Theme – Recipient
Recipient Pronominality:	Pronominal	0.9873 (**)	————
Recipient Givenness:	Non-Given	————	0.6314 (*)
Theme Givenness:	Non-Given	0.7515 (.)	
Length (log scale):	Longer Recipient	————	1.7151 (***)
Model Statistics			
C statistic: 0.851			
Nagelkerke R^2: 0.467			
AIC: 295.40			

Signif. Codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.10

With regard to the correlations that are identified here as significant in predicting the choice between the two constructional variants, we can see that, similarly to the findings obtained for English, the results for Polish also support the claim that objects that are nominal, inaccessible in previous discourse and longer occur in the final position in the dative construction. The highest variance inflation factor (vif) here measures 1.23, which is indicative of there being no multicolinearity between the independent variables. The overall performance of the model is excellent, as evidenced by the C statistic measure or the Nagelkerke R^2. Let us see now what happens when we add the <Lexical Sense> to the analysis (Table 6).

Firstly, analysis of variance (anova) performed on the two models, i.e., Polish Model 1 and Polish Model 2, demonstrates that the difference between them is statistically significant with p=0.006226 (**). Similarly to what we have observed for English, this means that adding <Lexical Sense> to the list of predictors improves the

performance of the explanatory model in a statistically significant manner.

Table 6: Polish Model 2. Fixed-Effects Binary Logistic Regression
Dative Alternation ~ Bresnan Features + Lexical Sense

		Effect Size / Coefficient & Significance	
Predictors		**Cx A**: Recipient – Theme	**Cx B**: Theme – Recipient
Recipient Pronominality:	Pronominal	1.1234 (**)	————
Recipient Givenness:	Non-Given	————	0.6494 (*)
Theme Givenness:	Non-Given	0.8112 (.)	
Length (log scale):	Longer Recipient	————	1.7766 (***)
Lexical Sense:	Causation	2.5242 (*)	————
Lexical Sense:	Render Available	2.9575 (**)	————
Model Statistics			
C statistic: 0.870			
Nagelkerke $R^2 = 0.527$			
AIC = 290.03			

Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1

Looking at the correlations, we find that the same variables that we have seen as statistically significant in Polish Model 1, are also returned as such here. The most important finding in Model 2 is that, in line with our hypothesis, the semasiological variation of the verb is again significant in differentiating between the two constructions. More importantly still, and similarly to what we have observed in the analysis for English, presented in Table 4, <Lexical Sense> is also the strongest predictor of the constructional choice, as demonstrated by the effect sizes of its two levels that are significant, i.e., <Lexical Sense: Causation> and <Lexical Sense: Render Available>. Interestingly, these are the two levels of the predictor that we have also found to be statistically significant for English. In the Polish model presented in Table 6, however, both these values of <Lexical Sense> are predictors of Cx A, where the Recipient precedes the Theme, as in examples (15), for "Render Available" and in (21), for "Causation".

(21) C*hyba nie pójdę na plastykę we wtorek.*
Dałam ***mamie*** ***rysunek***
Gave-1SG mom-DAT picture-ACC
(od dwóch tygodni go robi). Rysowała grubym ołówkiem a potem stwierdziła ze ona nie da rady.

> 'I don't think I will attend my arts and crafts class on Tuesday. I gave mom a picture (she has now been working on it for two weeks). She was drawing it with a thick pencil and then decided she wouldn't be able to do it.'

With regard to <Lexical Sense: Causation>, it should be noted that in (21), the caused event is encoded elliptically and it is the context that specifies the meaning of "Causation". The full causative construction would be complemented by *do zrobienia* ('to do'). We have already discussed the motivation for the correlation between Cx A and this sense of GIVE for English, where the same observation was made. With respect to <Lexical Sense: Render Available>, this correlation for Polish is predicted as a significant contextual clue for Cx A, and not Cx B, as was the case for English. This divergence between the two languages is an interesting finding that should be further explored.

The performance of Polish Model 2 and its goodness of fit are improved when compared to Polish Model 1. This can be assessed on the basis of the C statistic and Nagelkerke R^2. With the C statistic score at 0.87, the model can be evaluated as descriptively and predictively accurate with an excellent degree of discrimination between the two constructional variants. Comparing the AIC scores of the two Polish models shows that their parsimony is comparable, despite the addition of <Lexical Sense> in Model 2. We should also note that there is no risk of multicolinearity in the model, where the highest variance inflation factor is at 3.28, a value observed for <Lexical Sense: Enable>.

It should also be noted that for both English and Polish when <Lexical Sense> was entered as a random variable, the results of the models presented in Table 3 and Table 5, respectively, did not change at all. The mixed-effects logistic regression analyses performed equally well, the only difference being that the effect sizes of the correlations identified therein were lower.

4. Conclusion

The present study pursued two main objectives and sought to test corresponding hypotheses. Firstly, the goal was to test the findings of Bresnan *et al.* (2007) on a new dataset for English and on another language, i.e., Polish, with a focus on just one verb participating in the alternation under analysis. It was expected that the results obtained in Bresnan *et al*. (2007) would be confirmed here for both languages. As we have seen, in our analysis, we gained support for the importance of only some of the explanatory variables. Importantly, however, the variables that were found to be statistically significant do demonstrate that the "end-weight" principle operates in both English and Polish behind the choice of the constructional variant of the dative alternation.

Secondly and more importantly, the other objective in this inquiry was to investigate the interaction between semasiological lexical structure and constructional profiling in determining the use of alternate constructions. The central claim here was that lexical semantics should not be excluded from the analysis, as it is not random, but rather constitutes one of the determining variables that condition constructional choice. Our findings clearly show that the semasiological variation contributed by the lexeme is indeed an important predictor of the use of the two variants in the dative alternation. This is true for both English and Polish, where <Lexical Sense> was identified as the highest rank predictor. What is more, in both languages, the models that contained <Lexical Sense> among the predictors outperformed the models that disregarded lexical semantics. In addition, the correlations revealed here, especially in English, provide some support for the claim that Cx A, where the Recipient precedes the Theme, is significantly linked to the schematic meaning of "change of state", whereas Cx B, where the Recipient follows the Theme, encodes the abstract sense of

"change of location" (cf. Green 1974; Pinker 1989; Bresnan *et al*. 2007).

The fact that lexical sense should be the strongest predictor of a constructional choice is an important finding both in descriptive and theoretical terms. Descriptively, it informs the rich body of research on the dative alternation and on GIVE. Theoretically, it sheds light on the interaction between lexicon and grammar. More precisely, it shows that, in the investigation of constructional choices, apart from accounting for the impact of structural and discursive factors, we should also incorporate the effects of lexical semantics. As Goldberg (2002: 349f.) notes, "the meaning of a clause is more than the meaning of the argument structure construction used to express it", which is why in trying to identify the behavioral profiles of alternations or grammatical choices, we should not disregard "individual verbs". We have here seen a clear example of how the semasiological variation of a single verbal category, such as GIVE, can help us map more accurately the structure of onomasiological variation in grammar, such as the dative alternation.

REFERENCES

Arnold, J., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of complexity and information structure on constituent ordering. *Language*, 76, 28–55.

Behaghel, O. (1910). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25, 110–142.

Boas, H. (2003). *A constructional approach to resultatives*. Stanford: CSLI.

Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In S. Featherston & W. Sternefeld (Eds.), *Roots: Linguistics in search of its evidential base*, (pp. 76–96). Berlin: Mouton de

Gruyter.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Kraemer, & Zwarts, J. (Eds.), *Cognitive foundations of interpretation*, (pp. 69–94). Amsterdam: Royal Netherlands Academy of Sciences.

Bresnan, J., & Ford, M. . (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* ,86, 186–213.

Bresnan, J., & Hay, J. (2008). Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua*, 118, 245–259.

Bresnan, J., & Nikitina, T. (2009). The gradience of the dative alternation. In L. Uyechi, & L. H. Wee (Eds.), *Reality exploration and discovery: Pattern interaction in language and life*, (pp. 161–184). Stanford: CSLI Publications.

Collins, P. (1995). The indirect object construction in English: An informational approach. *Linguistics*, 33, 35–49.

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

Erteschik-Shir, N. (1979). Discourse constraints on dative movement. In T. Givon (Ed.) *Discourse and syntax.* New York: Academic Press.

Garretson, G., O'Connor, M. C., Skarabela, B., & Hogan, M. (2004). Coding practices used in the project Optimal Typology of Determiner Phrases. http://npcorpus.bu.edu/ documentation/index.html.

Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). *The structure of lexical variation: Meaning, naming, and context*. Berlin: Mouton de Gruyter.

Geeraerts, D., Grondelaers, S., & Speelman, D. (1999). *Convergentie en divergentie in de Nederlandse woordenschat*. Amsterdam: Meertens Instituut.

Glynn, D. (2004). Constructions at the crossroads. The place of construction grammar between field and frame. *Annual Review of Cognitive Linguistics*, 2, 197–233.

Glynn, D. (2009). Polysemy, syntax, and variation: A usage-based method for Cognitive Semantics. In V. Evans, & S. Pourcel (Eds.), *New directions in Cognitive Linguistics*, (pp. 77– 106). Amsterdam: John Benjamins.

Glynn, D. (2010a). Synonymy, lexical fields, and grammatical constructions: A study in usage- based Cognitive Semantics. In H.-J. Schmid, & S. Handl (Eds.), *Cognitive foundations of linguistic usage-patterns: Empirical studies*, (pp. 89–118). Berlin: Mouton de Gruyter.

Glynn, D. (2010b). Testing the hypothesis: Objectivity and verification in usage-based Cognitive Semantics. In D. Glynn, & K. Fischer (Eds.), *Quantitative Cognitive Semantics: Corpus-driven approaches*, (pp. 239–270). Berlin: Mouton de Gruyter.

Glynn, D. 2014. The many uses of *run*: Corpus methods and Socio-Cognitive Semantics. In D. Glynn & J. Robinson (Eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, (pp. 117–144). Berlin: Mouton de Gruyter.

Glynn, D., & Fischer, K. (Eds). 2010. *Quantitative Cognitive Semantics: Corpus-driven approaches*. Berlin: Mouton de Gruyter.

Glynn, D., & Robinson, J. (Eds.) 2014. *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.

Goldberg, A. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics*, 13, 327–356.

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

Green, G. (1974). *Semantics and syntactic regularity*. Bloomington: Indiana University Press.

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

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

Gries, St. Th. (2003b). Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics*, 1, 1–27.

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

Gries, St. Th., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics*, 9, 97–129.

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

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

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: John Wiley & Sons.

Kilgarriff, A., Baisa, V., Buta, J., Jakubek, M., Kov, V., Michelfeit, J., Rychly, P.,

Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography* ,1, 7–36.

Krawczak, K., Fabiszak, M., & Hilpert, M. (2016). A corpus-based, cross-linguistic approach to mental predicates and their complementation: Performativity and descriptivity vis-à-vis boundedness and picturability. *Folia Linguistica*, 50, 475–506.

Michaelis, L. A., & Hartwell, S. F. (2007). Lexical subjects and the conflation strategy. In N. Hedberg, & R. Zacharski (Eds.), *Topics in the grammar-pragmatics interface: Papers in honor of Jeanette K. Gundel*, (pp. 19–48). Amsterdam: Benjamins.

Newman, J. (1996). *Give: A cognitive linguistic study.* Berlin: Mouton de Gruyter.

Pinker, S. (1989). *Learnability and cognition. The acquisition of argument structure*. Cambridge, MA: MIT Press.

Prince, E. (1981). Toward a new taxonomy for given-new information. In P. Cole (Ed), *Radical pragmatics*, (pp. 223–255). New York: Academic Press.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought*, (pp. 284–310). Cambridge: Cambridge University Press.

Speelman, D., & Geeraerts, D. (2009). Causes for causatives. The case of Dutch *doen* and *laten*. In T. Sanders, & E. Sweetser (Eds.), *Causal categories in discourse and cognition*, (pp. 173–204). Berlin: Mouton de Gruyter.

Theijssen, D. L. (2012). Making choices. Modeling the English dative alternation. PhD dissertation, Radboud Universiteit Nijmegen.

Wasow, T. (2002). *Postverbal behavior*. Stanford: CSLI.

Williams, R.S. (1994). A statistical analysis of English double object alternation. *Issues in Applied Linguistics*, 5, 37–58.

Wolk, C., Bresnan, J., Rosenbach, A., & Szmrecsányi, B. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica*, 30, 382–419.