Distributional determinants of learning argument structure constructions in first and second language

Yevgen Matusevych (Y.Matusevych@uvt.nl)
Department of Culture Studies & Tilburg Center for Cognition and Communication (TiCC), Tilburg University

Afra Alishahi (A.Alishahi@uvt.nl)
Tilburg Center for Cognition and Communication (TiCC), Tilburg University

Ad Backus (A.M.Backus@uvt.nl)
Department of Culture Studies, Tilburg University
PO Box 90153, 5000 LE Tilburg, the Netherlands

Abstract

Learning argument structure constructions is believed to depend on input properties. In particular, in a cued production task, verb production within each construction has been shown to depend on three input factors: frequency of a verb in a construction, contingency of verb–construction mapping, and verb semantic prototypicality. Earlier studies have estimated these values from a language corpus, without accounting for variation in the input to individual learners. We use a computational model to control for such variation, and our results replicate those reported for human learners. The second issue that we address relates to different ways of representing constructions: while the earlier studies employ form-only representations, we run an additional analysis for form–meaning representations. Again, the results show the impact of all three input properties on the verb production, but their relative impact depends on the representations used.

Keywords: language learning; argument structure constructions; learning factors; frequency effects

Introduction

Cognitive, or usage-based, approaches to language acquisition claim that our linguistic knowledge is based on our individual experiences with the language (e.g., Kemmer & Barlow, 2000). This is why a crucial role within such theories is assigned to the language input that learners are exposed to. In this view, cognitive skills, such as pattern recognition, categorization, induction, etc., enable learners to notice various statistical regularities in the input and build up the structured knowledge of language (Tomasello, 2003). Usage-based theories are often associated with constructionist accounts. Such accounts claim that our linguistic knowledge can be represented as a structured inventory of constructions, or form–meaning pairings, which vary from fully specific to fully schematic. Schematic, or abstract, constructions emerge as generalizations from individual instances—a particular example of this process is described in Goldberg's (1995) account of clausal-level argument structure (AS) constructions. The nature of such constructions is well captured by the term "argument structure generalizations", suggested by Goldberg, Casenhiser, and Sethuraman (2004).

Although the input is believed to play a crucial role in language learning, it is yet unclear how exactly the input shapes our lingustic knowledge—that is, which properties of the input determine how well we know specific language units and how readily we use them. One particular proposal in this respect, which addresses the learning of AS constructions, has been made by Ellis and colleagues in a series of studies (Ellis & O'Donnell, 2012; Ellis, O'Donnell, & Römer, 2014a, 2014b). They have analysed how the production of verbs in AS constructions is affected by certain properties of these verbs, related to the distribution of their forms and meanings in the constructions. In particular, Ellis et al. (2014a, 2014b) have carried out a series of experiments with L1 and L2 speakers. In these experiments the frequency of verb production in a given construction has been shown to be predictable from three characteristics of the verb, which quantitatively describe its use in the construction, as reflected in language corpora:

1. **Frequency**, or the token frequency of occurrence of the verb within this construction. This is believed to reflect the degree of accessibility of the verb, or its entrenchment: more entrenched verbs are produced more often.
2. **Contingency**, or the reliability of the verb–construction mapping. This is related to the selectivity of verbs within each construction: verbs that are strongly associated with the construction are more likely to be produced first.
3. **Prototypicality**, or the centrality of the verb meaning: learners would more readily, or more frequently produce verbs whose meanings are most central to the construction.

Methodological issues

The values of frequency, contingency, and prototypicality have been calculated by Ellis et al. (2014a, 2014b) based on the occurrence of each verb in the British National Corpus. This implies that all learners are exposed to language input with exactly the same distributional properties. On the contrary, the usage-based approach claims that language learning is driven by usage, therefore different language experiences of individual learners lead to different language representations in learners (Dąbrowska, 2012; Misyak & Christiansen, 2012). The variation is even higher among L2 learners, who may be exposed to very different kinds of L2 input (DeKeyser, 2013). In short, it is important to account for the individual variation in the langage input. This is a challenging task for studies with human subjects, because it is nearly impossible to account for the whole learning history of each individual learner. On the other hand, computational modeling can tackle the problem by providing us with a maximum control over input data (Poibeau, Villavicencio, Korhonen, & Alishahi, 2013). In an earlier study (Matusevych, Alishahi, & Backus, 2014) we have shown how various input variables can be systematically manipulated in a computational model of AS construction learning (Alishahi & Stevenson, 2008). Similarly, if it is important to keep the naturalistic input without manipulating its characteristics, we can also simply monitor such characteristics. In the current study, we exploit this advantage of the model and use it for simulating the earlier experimental studies (Ellis et al., 2014a, 2014b).

These earlier studies also give rise to a more theoretical question. In theory, they adopt the definition of constructions as form–meaning pairings. In practice, however, they analyse constructions as grammar patterns, informed by the COBUILD dictionary project (Hunston, Francis, & Manning, 1996). Each pattern contains slots for the verb and its arguments, together with a specific preposition (if any), e.g. *V of N*. Therefore, the patterns are form-based and contain little semantic information. This may result in grouping semantically divergent verbs together—e.g. a cognition verb *think*, a communication verb *speak*, and a perception verb *smell* are analysed within the same pattern *V of N* (example from Ellis et al., 2014a). Ellis and O'Donnell (2012) suggest that the same analysis should be carried out for construction representations informed by other theories of construction grammar.

The current study addresses these two issues. First, we intend to investigate whether the findings of Ellis et al. (2014a, 2014b) can be obtained in controlled computational simulations of L1 and L2 learning, in which we monitor the actual characteristics of the input to each individual learner. Second, we carry out an additional analysis, in which constructions are represented as form–meaning pairings rather than form-based patterns, and the distributional characteristics are monitored for these pairings. We compare the results of the two analyses, to investigate whether the findings are consistent for the two types of construction representations.

Computational model

We use a computational model of AS construction learning from exposure (Alishahi & Stevenson, 2008). The exposure consists of individual argument structure usages (AS usages). Each usage is represented as a set of features that correspond to linguistic and perceptual cues normally available to human learners in an utterance they hear and the accompanying scene. The features include the verb, its lexical meaning, its arguments, their lexical meanings, their thematic roles and linguistic cases, the syntactic pattern and prepositions. Values of some features, such as the verb and arguments, are represented as a single symbol (e.g., *acquire*), while values of other features, in particular lexical meanings and thematic roles, are formed by sets of semantic primitives: e.g., the lexical meaning of *acquire* is a set {GET, HAS POSSESSION, TRANSFER, CAUSE, COST}. These primitives have been automatically extracted from WordNet (Miller, 1995) and VerbNet (Schuler, 2006). As for syntactic patterns, we define them as strings describing the word order, e.g. ARG1 VERB *in* ARG2.

The learning task consists in learning AS generalizations by abstracting from individual verb usages (cf. Goldberg et al., 2004). Such generalizations are represented as groups of AS usages: each group combines the properties of all the participating usages, but does not store the information about each AS usage individually. During the learning, the model receives one AS usage at a time and adds it either to one of the existing groups, or to a new one. The decision for each usage is made based on two factors: the similarity between the AS usage and each group, and the number of usages in each group (which reflects its degree of entrenchment). In case of bilingual learning, L1 and L2 usages are treated equally. The model is not explicitly informed about the language of each AS usage, therefore L1 and L2 usages can potentially be mixed within groups. In other words, the learning of both languages is grounded in their use, which is reflected in the input. A more detailed description of the model can be found in our earlier study (Matusevych et al., 2014).

Eliciting language use

Ellis et al. (2014a, 2014b) in their experiments use an elicited production task. Participants are given a number of phrases, from which the actual verb is removed (e.g., *it ___ about the...*), and they have to produce verb(s) that fit the empty slot. In a similar way, we elicit the verb production by simulated computational learners. At the end of the learning process, the model is provided with a test set containing a number of AS usages, in which the value of the verb feature is masked. For each test AS usage, the model is asked to provide a number of most likely verbs that fit the usage, given the values of other features. The probability of using each verb is estimated considering (1) how likely it is that the test usage belongs to a specific group, and (2) how frequently each verb occurs within this group.

Input data

The input data sets come from two monolingual corpora of German (the SALSA corpus, see Burchardt et al., 2006) and English (the PropBank, see Palmer, Gildea, & Kingsbury, 2005). From each corpus, a number of AS usages were extracted (see Table 1), each of which had a FrameNet frame type associated with it (e.g., RECEIVING, DEPARTING, see Ruppenhofer, Ellsworth, Petruck, Johnson, & Scheffczyk, 2006). For the German data the frame types were contained in the SALSA corpus itself, while for the English data we used the existing mapping between PropBank propositions and FrameNet frames (Palmer, 2009). Frame types were not used as input features available to simulated learners, but only as a tool for carrying out our analysis of form–meaning mappings. Additionally, the semantic information was automatically obtained from computational resources such as VerbNet and WordNet (via the existing FrameNet–WordNet map-

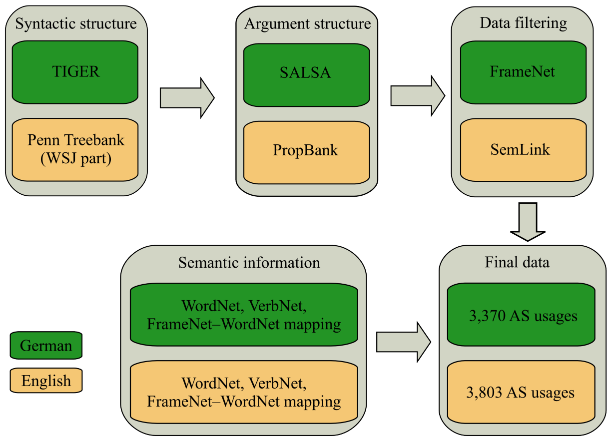

Figure 1: Schematic description of preparing the input data.

ping, see Bryl, Tonelli, Giuliano, & Serafini, 2012). Figure 1 schematically illustrates the procedure.

A quantitative descripton of the English and the German data sets is provided in Table 1. It contains information about the total number of AS usages in the data, the number of verb types, frame types, and syntactic patterns.

Table 1: Characteristics of the two data sets.

Data set	AS usages	Verb types	Pattern types	Frame types
German	3,370	301	97	179
English	3,803	351	51	151

Experimental setup

We run a computational simulation of the experiment 2 described by Ellis et al. (2014a). Their experiment investigated only L1 learning, however a similar experiment for L2 learning was reported by Ellis et al. (2014b). Therefore, we run our simulations for both L1 and L2 learning, both in English and in German (hence four sets of simulations), and compare our results to those of the two earlier studies.

The actual input to the computational model is randomly sampled from the appropriate data set(s), so that the model's actual language experience varies among simulations. Thereby, running a number of simulations corresponds to collecting data from a population of language learners. In each set of simulations, 30 learners are exposed to a total number of AS usages N. Given the size of our data sets and the sampling procedure, we have found that simulated learners achieve a high proficiency in each language after they are exposed to $6,000$ AS usages in this language. Therefore, L1 learners start with no linguistic knowledge and receive $N = 6,000$ L1 usages, following which they perform the elicited production task in L1 (see Figure 2a). Bilingual learners, on the other hand, receive in total $N = 12,000$ AS usages. First they are exposed to $6,000$ L1 usages, which results in the emergence of groups consisting of L1 usages.

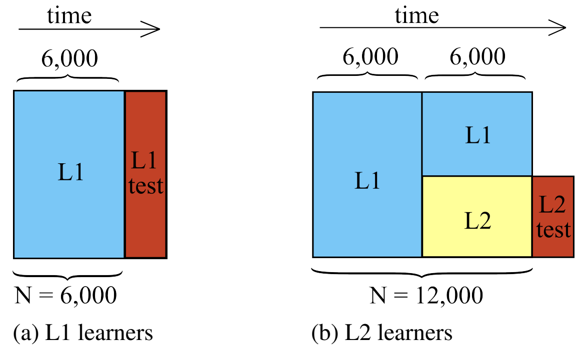

Figure 2: Schemes of the experimental setup.

Following this, learners receive $6,000$ more usages of bilingual input, in which L1 and L2 are mixed in a proportion $1:1$ (see Figure 2b). Note that the input in this latter case contains only $\frac{1}{2} \times 6,000 = 3,000$ L2 AS usages, but this is done to simulate L2 learners whose exposure to L2 is limited, and whose L2 proficiency is lower than L1 proficiency.[1] Following the acquisition process, L2 learners perform the elicited production task in L2.

For the production task (both in L1 and L2), a set of 300 stimuli is randomly generated from the respective data set (English or German). The head verb is always masked, so that the production is yielded by the values of other features: syntactic pattern, prepositions, number of arguments, argument cases and thematic roles. The arguments and all the lexical meanings are also masked, as they would associate the stimuli only with particular verbs. To test each learner on a variety of constructions, we ensure that the test set does not contain more than one usage of the same verb in the same construction. For each test stimulus, learners produce a number of verbs that they find most likely to occur in the respective slot. Each verb V in the data is accompanied by a probability value P_V, which shows how likely V fits the stimulus. Upon obtaining the production outcomes from each learner, we filter out verbs with $P_V \leq 1\%$, and those that never occured in the target construction in the input to this specific learner, in order to reduce the unwanted noise in the data. Following this, we look at how P_V can be explained by each of the three determinants.

Learning determinants

For each individual learner, we record the information about the actual frequency of each verb in each construction (F_{VC}), the frequency of each construction (F_C), and the overall frequency of each verb (F_V) in the input. Besides, we also know which verbs participate in every construction in the input to each learner. This allows us to calculate the values of the three learning determinants for each learner individually.

[1] In an earlier study (Matusevych et al., 2014) we manipulated parameters of exposure such as N and the L1:L2 proportion, and found that the simulated learning process was robust to the variation of these parameters.

Frequency. The token frequency of each verb in each construction is directly accessible via F_{VC}.

Contingency. Following the earlier experiments (Ellis et al., 2014a, 2014b), contingency is measured using a one-way dependency statistic ΔP (construction $\rightarrow$ verb), or ΔP_{CV} (Ellis et al., 2014a), which in this case reflects the degree to which a given construction selects a particular verb in the input. This measure can be calculated as follows:

$$\Delta P_{CV} = P(V|C) - P(V|\neg C) = \frac{F_{VC}}{F_C} - \frac{F_V - F_{VC}}{N - F_C} \tag{1}$$

Prototypicality. Since each verb meaning in AS usages is represented as a set of semantic primitives, the most prototypical verb V for a construction C is the one whose meaning M_V shares most primitives with the meanings M_i of all the other verbs i occurring in C ($i \in C$). In this terms, the prototypicality of a verb V can be calculated as follows:

$$Prt_V = \frac{\sum_{i \in C} \frac{|M_i \cap M_V|}{|M_V|}}{|C|} \tag{2}$$

Note that the value of Prt_V may also vary among different learners, because some of them may have not encountered in their input a certain verb occuring in a certain construction.

As it is clear from the definitions of the learning determinants, their values are calculated for each verb in each construction. In other words, their values depend on the type of construction representations used in the analysis. In our replication of Ellis et al.'s (2014a, 2014b) experiments, the values of the learning determinants are associated with form-based patterns, while in the second analysis that we carry out the values are calculated for form–meaning pairings.

Statistical methods

In each analysis, we use linear mixed effects models to predict P_V values by the three main effects—F_{VC}, ΔP_{CV} and Prt_V. Before fitting the model, we log-transform the outcome and the predictors (following Ellis et al., 2014a), center the predictors (to reduce their collinearity), and standardize the outcome and the predictors (to make all the β coefficients directly comparable). The random effect structure of each model is kept as maximal as it is justified by the data sample (Barr, Levy, Scheepers, & Tily, 2013). In each case, random effects account for the variation between constructions (CONSTRUCTION factor) that the test stimuli belong to, and—within each construction—between verbs that are masked in the test stimuli (CONSTRUCTION:PREDICATE factor). For each model we calculate the goodness of fit using a marginal and a conditional R^2, which quantify the amount of variance explained by the fixed factors and the full model, respectively (Johnson, 2014). Due to the difficulties related to the inference of p values in mixed effects models, we do not report such values, but only the respective 95% confidence intervals estimated via parametric bootstrap with 100 resamples (Bates, Maechler, Bolker, & Walker, 2015).

Table 2: Summary of mixed effects models predicting the use of verbs within form-based patterns.

Language	**Goodness of fit**		**Predictor**	**Statistics**		
	R^2_m	R^2_c		β	*SE*	95% *CI*
L1 English	.17	.47	F_{VC}	0.44	0.06	$[0.33, 0.56]$
			ΔP_{CV}	0.06	0.01	$[0.04, 0.07]$
			Prt_V	0.17	0.04	$[0.09, 0.25]$
L1 German	.07	.50	F_{VC}	0.21	0.01	$[0.18, 0.23]$
			ΔP_{CV}	0.09	0.02	$[0.05, 0.14]$
			Prt_V	0.19	0.04	$[0.11, 0.26]$
L2 English	.15	.55	F_{VC}	0.47	0.09	$[0.29, 0.66]$
			ΔP_{CV}	0.04	0.03	$[-0.02, 0.11]$
			Prt_V	0.13	0.04	$[0.05, 0.19]$
L2 German	.06	.47	F_{VC}	0.20	0.02	$[0.15, 0.25]$
			ΔP_{CV}	0.12	0.01	$[0.09, 0.14]$
			Prt_V	0.12	0.04	$[0.05, 0.19]$

Results

Constructions as form-based patterns

In the results reported here, the values of frequency, contingency, and prototypicality are computed for form-based patterns. Table 2 provides the summary of mixed effects models predicting the learners' verb use from the values of the three determinants. First, we can observe that the goodness of fit (both R^2_m and R^2_c) for L1 English and L1 German is generally higher than that for L2 English and L2 German, respectively. We explain this by the fact that the language input during the bilingual learning consists of both L1 and L2 usages, therefore, it is characterized by a higher variability than the L1-only input.

As for the actual predictors, we can see that for most predictors in all four data sets (in L1 and L2 English and German) the 95% confidence intervals do not contain zero, with the only exception of ΔP_{CV} in L2 English. This is evidence of the independent contribution of each variable—F_{VC}, ΔP_{CV}, and Prt_V—to predicting the probability of the verb production in our simulated data. The lack of the evidence for the effect of ΔP_{CV} in L2 English may be explained by the interaction of this variable with F_{VC}, which we discuss in the concluding section. Overall, however, the results are in line with the findings reported by Ellis et al. (2014a, 2014b). It could be interesting to compare the magnitude of the effect of each predictor in our simulations vs. their studies, but they report non-standardized regression coefficients, which makes the comparison difficult.

Another factor we should mention is the overall fit of the models reported in Table 2. Although the conditional R^2_c values are similar for all four data sets (between .47 and .55), the marginal R^2_m values are lower for German than for English data (.07 vs. .17 for L1, and .06 vs. .15 for L2). In other words, random effects explain more variance in the German data sets, compared to the English data. The higher amount of the random variance in the German data may be simply due to the larger number of groups per random factor (CONSTRUCTION and CONSTRUCTION:PREDICATE) in German data sets: 89 and 703 in L1 German vs. 47 and 536 in L1 English; 89 and 688 in L2 German vs. 45 and 531 in L2 English, respectively.

This, however, does not explain why the R^2_m values are generally low. We discuss this in the concluding section, but before that we run the second analysis, to investigate whether the results hold for different construction representations.

Constructions as form–meaning pairings

This time we compute the values of frequency, contingency, and prototypicality for constructions which comprise form (syntactic pattern) and meaning (frame type), e.g. PLACING : ARG1 VERB ARG2 *in* ARG3. The statistical results are summarized in Table 3.

Table 3: Summary of mixed effects models predicting the use of verbs within form–meaning pairings.

Language	Goodness of fit		Predictor	Statistics		
	R^2_m	R^2_c		β	*SE*	95% *CI*
L1 English	.21	.61	F_{VC}	0.13	0.04	[0.06, 0.20]
			ΔP_{CV}	0.27	0.04	[0.20, 0.34]
			Prt_V	0.44	0.02	[0.39, 0.48]
L1 German	.21	.65	F_{VC}	0.13	0.02	[0.09, 0.18]
			ΔP_{CV}	0.14	0.03	[0.08, 0.20]
			Prt_V	0.46	0.02	[0.41, 0.51]
L2 English	.18	.56	F_{VC}	0.12	0.04	[0.04, 0.21]
			ΔP_{CV}	0.24	0.03	[0.17, 0.30]
			Prt_V	0.38	0.02	[0.34, 0.42]
L2 German	.25	.65	F_{VC}	0.10	0.03	[0.04, 0.14]
			ΔP_{CV}	0.24	0.03	[0.18, 0.31]
			Prt_V	0.50	0.02	[0.46, 0.54]

Again, we see that the confidence intervals for F_{VC}, ΔP_{CV}, or Prt_V do not include zero in all four data sets, this time with no exception. Therefore, all three predictors contribute to explaining the probability of verb use to a certain extent. The most interesting aspect, however, is to compare the explanatory power of each variable in the two types of analysis. For this, we plot the sizes of the respective standardized β coefficients—see Figure 3. Overall, the pattern for the four sets of the form–meaning data is more consistent than for the form-based data sets: prototypicality explains most variation, followed by contingency, followed by frequency. There is no clear pattern in this respect for the form-based data, in particular due to the differences between the German and English data mentioned in the previous section.

The most consistent difference between the results for the form-based and the form–meaning data is the higher impact of prototypicality for the latter representations. Evidently, form–meaning pairings are more semantically coherent than form-based patterns, which explains the higher predictive power of verb semantic prototypicality in the second analysis. In short, the overall findings hold both for form-based and for form–meaning data, but the magnitudes of the individual effects depend on the representations used.

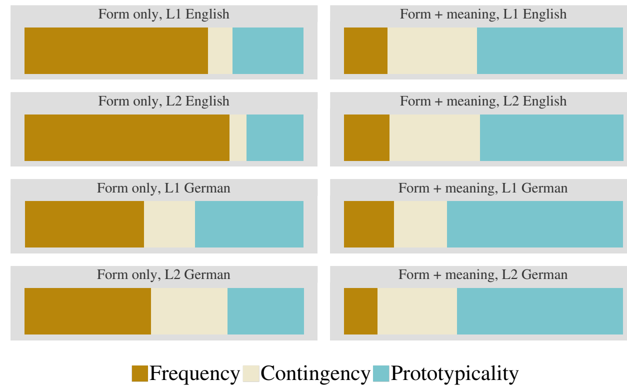

Figure 3: Contributions of the predictors across the analyses.

Discussion

The present paper has addressed two methodological issues related to the two earlier studies, in which the learning of AS constructions is explained by three input-related distributional factors. The first issue is whether individual differences between participants in Ellis et al.'s (2014a, 2014b) studies might have affected their results. In our computational setup the measures of frequency, contingency, and prototypicality reflected the actual characteristics of the input to each individual learner, while in the earlier experiments the measures were calculated for a language corpus, and did not account for each learner's individual language experience. In our simulation, all three variables had significant individual contributions (with one exception) to predicting the verb production. This is in line with the results reported by Ellis et al. (2014a, 2014b).

The second question was whether the results for form-based patterns would be generalizable to form–meaning pairings, which is a more common definition of constructions within the usage-based framework. Overall, the results were consistent between the two types of analysis: all the variables showed significant individual contributions. However, the magnitude of the individual effects differed: for form–meaning pairings, the impact of semantic prototypicality was found to be substantially higher than for form-based patterns. Thus, the use of form-based patterns in the analysis may result in underestimating the power of prototypicality in explaining the process of AS construction learning.

Overall, this study confirms the proposal by Ellis et al. (2014a, 2014b) that verb token frequency, contingency of verb–construction mapping, and verb semantic prototypicality have individual impact on the verb production. An important issue, however, relates to whether this three-factor model of explaining the verb production is optimal and complete. In our simulations, the model fit (R^2_c and R^2_m) was rather low in some cases, which suggests the model can be further improved. In particular, the current model includes verb token frequency *within a construction* (that is, the joint frequency of a verb and a construction) and verb–construction contingency. Both variables, in fact, relate to the associations between verbs and constructions, so including them both to the model may be redundant. This may also explain why the effect of contingency did not even reach the significance level for L2 English form-based analysis. On the other hand, the total (marginal) verb frequency is absent from the model, but there is experimental evidence in psychology that the total item frequency affects cued recall (Clark & Burchett, 1994). To conclude, the three-factor model of verb production may need to be refined in further research.

References

Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, *32*(5), 789–834.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*.

Bryl, V., Tonelli, S., Giuliano, C., & Serafini, L. (2012). A novel Framenet-based resource for the semantic web. In S. Ossowski & P. Lecca (Eds.), *Proceedings of the 27th Annual ACM Symposium on Applied Computing* (pp. 360–365). New York: Association for Computing Machinery.

Burchardt, A., Erk, K., Frank, A., Kowalski, A., Pado, S., & Pinkal, M. (2006). The SALSA corpus: a German corpus resource for lexical semantics. In N. Calzolari et al. (Eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC-2006)* (pp. 969–974). European Language Resources Association (ELRA).

Clark, S. E., & Burchett, R. E. (1994). Word frequency and list composition effects in associative recognition and recall. *Memory & Cognition*, *22*(1), 55–62.

DeKeyser, R. M. (2013). Age effects in second language learning: Stepping stones toward better understanding. *Language Learning*, *63*, 52–67.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*(3), 219–253.

Ellis, N. C., & O'Donnell, M. B. (2012). Statistical construction learning: Does a zipfian problem space ensure robust language learning. In P. Rebuschat & J. N. Williams (Eds.), *Statistical Learning and Language Acquisition.* Boston: De Gruyter Mouton.

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014a). The processing of verb-argument constructions is sensitive to form, function, frequency, contingency and prototypicality. *Cognitive Linguistics*, *25*(1), 55–98.

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014b). Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism*, *4*(4), 405–431.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, *15*(3), 289–316.

Hunston, S., Francis, G., & Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs.* London: HarperCollins Publishers.

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946.

Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In S. Kemmer & M. Barlow (Eds.), *Usage-Based Models of Language.* Stanford: CSLI Publications.

Matusevych, Y., Alishahi, A., & Backus, A. (2014). Isolating second language learning factors in a computational study of bilingual construction acquisition. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 988–994). Austin: Cognitive Science Society.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38*(11), 39–41.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302–331.

Palmer, M. (2009). SemLink: Linking PropBank, VerbNet and FrameNet. In A. Rumshisky & N. Calzolari (Eds.), *Proceedings of the 5th International Conference on Generative Approaches to the Lexicon* (pp. 9–15). Stroudsburg: Association for Computational Linguistics.

Palmer, M., Gildea, D., & Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, *31*(1), 71–106.

Poibeau, T., Villavicencio, A., Korhonen, A., & Alishahi, A. (2013). Computational modeling as a methodology for studying human language learning. In A. Villavicencio, T. Poibeau, A. Korhonen, & A. Alishahi (Eds.), *Cognitive Aspects of Computational Language Acquisition.* Heidelberg: Springer.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R., Johnson, C. R., & Scheffczyk, J. (2006). *FrameNet II: Extended Theory and Practice.*

Schuler, K. K. (2006). *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Unpublished doctoral dissertation, University of Pennsylvania.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.