Modeling the Emergence of an Exemplar Verb in Construction Learning

Libby Barak, **Afsaneh Fazly**, and **Suzanne Stevenson**
Department of Computer Science
University of Toronto
{libbyb,afsaneh,suzanne}@cs.toronto.edu

Abstract

Using a computational model of verb argument structure learning, we study a key assumption of the usage-based theory: that the acquisition of a construction relies heavily on the existence of a high-frequency exemplar verb that accounts for a large proportion of usages of that construction in the input. Importantly, unlike the psycholinguistic experiments that focus on the learning of an artificial novel construction using novel verbs, here we examine the acquisition of the English sentential complement construction from naturalistic input. Our results provide new insights into exemplar-based learning in the context of naturalistic input with multiple semantic classes, and a diverse set of constructions for the verbs.

Introduction

Verb argument structure acquisition is a challenging task that children face early in their life. In order to correctly use a verb, children must learn the syntactic structures that the verb appears in, as well as the semantic relations among the arguments of the verb. Nonetheless, children learn the correct usages of many verbs at a young age. Usage-based theories of language acquisition suggest that children learn the argument structure regularities mainly from the input they receive. These theories are supported by observing that children initially learn verb argument structures on an item-by-item basis, and only later generalize their verb-specific knowledge into abstract *constructions* that map a particular syntactic form to certain semantic properties (Tomasello, 2000).

The distributional properties of verb usages in child-directed input highly affect the developmental path of the acquisition of argument structure constructions. For example, several studies have shown that children tend to learn high-frequency verbs earlier (Naigles & Hoff-Ginsberg, 1998; Matthews, Lieven, Theakston, & Tomasello, 2005), and that they are more likely to detect grammatical anomalies in sentences containing such verbs (Theakston, 2004; Ambridge et al., 2008). Moreover, the relative frequency of a verb with a particular syntactic construction has been shown to positively correlate with the ability of young children to recall sentences containing that verb (Kidd, Lieven, & Tomasello, 2006, 2010). Most importantly, there is evidence that the acquisition of a construction is connected to a high-frequency exemplar verb; e.g., *give* is the exemplar for the ditransitive construction (Goldberg, 1999; Kidd et al., 2006, 2010). In fact, several studies have shown that it is not just the amount of overall exposure to a construction that affects its acquisition, but instead learning seems to be facilitated by a high number of usages of a particular exemplar verb (Casenhiser & Goldberg, 2005; Wonnacott et al., 2008).

The above psycholinguistic studies perform experiments on children, and hence are often limited in the number of items they can investigate, and in how much they can tease apart the various interacting factors that might play a role on the results. For example, the sentence recall tasks performed by Kidd et al. (2010) examine only eight complement-taking verbs (CTVs). Moreover, due to their choice of verbs, they cannot separate the effects of overall frequency and relative construction frequency on their results. Using a computational model of argument structure learning, we extend these investigations into a larger set of CTVs, and also manipulate input in such a way that we can tease apart the effects of the various frequency factors. Our results are consistent with the findings of Kidd et al. (2006, 2010), that the relative frequency of a verb with a sentential complement is positively correlated with the ability of young children to recall sentences containing the verb in that construction. However, through computational modeling, we are further able to provide evidence on the interaction of verb frequency and relative construction frequency in accounting for their findings.

Studies examining the effect of a high-frequency exemplar verb in the acquisition of novel constructions often do so in the context of an artificial language learning task, where children are introduced to a novel verb mapped to a novel (or familiar) event semantics (Casenhiser & Goldberg, 2005; Wonnacott et al., 2008). We use our computational model to investigate the existence and role of an exemplar verb in the acquisition of the English finite sentential complement (SC) syntax — a complex structure that has received less attention in the experimental studies (though see Kidd et al., 2006, 2010). Importantly, the use of a model enables us to vary distributional properties of the input in a way not easily achieved in a human experimental setting. Inspired by the work of Casenhiser and Goldberg (2005), we study the role of a high-frequency exemplar verb (*think*) for the acquisition of SC, but we do so in the context of a diverse set of verbs and constructions, as is the case in the naturalistic input that children receive. Our results suggest that the acquisition of a construction is facilitated by the relative frequency with which a class of semantically-related verbs appear with the syntactic form associated with the construction.

The Computational Model

We use an extended version of the verb argument structure acquisition model of Alishahi and Stevenson (2008), which we have used in studying the acquisition of mental state verbs (Barak, Fazly, & Stevenson, 2012). This model has appropriate characteristics for our study: (i) it focuses on argument structure learning, and the interplay between syntax and semantics; (ii) it is probabilistic and hence can naturally reflect

Input Utterance: *He thinks Mom made pancakes.*

Extracted Frame:

main predicate	**think**
other predicate	make
event primitives	{ *state*, *consider*, *cogitate*, *action* }
event participants	{ *experiencer*, *perceiver*, *considerer* } { *agent*, *animate* } { *theme*, *changed* }
syntactic pattern	arg1 *verb* arg2 *verb* arg3
argument count	3
complement type	finite

Table 1: An example of an input frame based on the utterance *He thinks Mom made pancakes*, and the semantic information assumed to be available from the scene (not shown).

the role of the statistical properties of the input in the formation of constructions; and (iii) it is incremental, which allows us to investigate changes in behaviour over time.

The input to our model is a sequence of *frames*, where each frame is a collection of *features* that resemble what children can extract from the utterances they hear and the typical learning scenes they preceive from their environments. We use features that include both semantic properties (i.e., *event primitives* and *event participants*), and syntactic properties (i.e., *syntactic pattern*, *argument count*, and *complement type*). Table 1 presents an example of an input frame given a child-directed utterance in a typical learning scene.

The model incrementally clusters the input frames into *constructions* that reflect probabilistic associations of semantic and syntactic features across similar verb usages. Note that a cluster is not simply a set of similar frames, but instead an abstraction over these frames represented as probability distributions over the possible values of each feature.

Algorithm for Learning Constructions

The model clusters input frames into constructions on the basis of their overall similarity in the values of their features. Importantly, the model learns these constructions incrementally, considering the creation of a new construction for a given frame if the frame is not sufficiently similar to any of the existing constructions. Formally, the model finds the best construction (including a new one) for a given frame F as in:

$$\text{BestConstruction}(F) = \underset{k \in Constructions}{\text{argmax}} \; P(k|F) \qquad (1)$$

where k ranges over all existing constructions and a new one. Using Bayes rule:

$$P(k|F) = \frac{P(k)P(F|k)}{P(F)} \propto P(k)P(F|k) \qquad (2)$$

The prior probability of a construction $P(k)$ is estimated as the proportion of observed frames in k, assigning a higher prior to constructions that are more entrenched (i.e., observed more frequently). The likelihood $P(F|k)$ is estimated based on the values of features in F and the frames in k:

$$P(F|k) = \prod_{i \in frameFeatures} P_i(j|k) \qquad (3)$$

Semantic class	Verb	Overall frequency	Frequency with finite-SC
Belief	think	13829	100%
	bet	391	100%
	guess	278	76%
	know	7189	61%
	believe	78	21%
Desire	wish	132	94%
	hope	290	86%
Communication	tell	2953	64%
	say	8622	60%
	ask	818	29%
Perception	hear	1370	21%
	see	9717	14%
	look	5856	9%

Table 2: The Overall frequency of the 13 CTVs in our data, along with their relative frequency with finite-SC. Verbs are grouped by semantic class, and only the 13 verbs that appear with this construction are listed.

where i refers to the i^{th} feature of F and j refers to its value. The conditional probability of a feature to have value j in construction k, is calculated using a smoothed version of:

$$P_i(j|k) = \frac{\text{count}_i(j,k)}{n_k} \qquad (4)$$

where $\text{count}_i(j,k)$ is the number of times feature i has the value j in construction k, and n_k is the number of frames in k.

Generation of the Input Corpora

We generate artificial corpora for our simulations, since we do not have access to sufficient data of actual utterances paired with scene representations. To create naturalistic data that resembles what children are exposed to, we build an input-generation lexicon that is based on the distributional properties of actual child-directed speech (CDS). We extracted our verbs and their distributional properties from the CDS to 8 children from CHILDES (MacWhinney, 2000).[1] We selected 31 verbs from different semantic classes and different frequency ranges, including 11 Physical Action (*come, go, fall, eat, play, get, give, take, make, put, sit*), 5 Perception (*hear, listen, look, see, watch*), 5 Communication (*ask, say, speak, talk, tell*), 5 Belief (*think, know, guess, bet, believe*), and 5 Desire (*want, wish, like, mind, need*). For each verb, we manually analyzed a random sample of 100 CDS usages (or all usages if fewer than 100) to extract distributional information about argument structures. Many of these verbs can take a (finite or infinitival) SC. Our focus in this work is on the finite-SC construction, and so we use the term Complement-Taking Verb (CTV) to refer to verbs that appear with the finite SC, following Kidd et al., 2010. Table 2 lists the 13 CTVs in our data, along with their semantic class, their overall frequency, and their relative frequency with the finite SC).

[1]Corpora of (Brown, 1973; Suppes, 1974; Kuczaj, 1977; Bloom, Hood, & Lightbowny, 1974; Sachs, 1983; Lieven, Salomo, & Tomasello, 2009).

We construct the input-generation lexicon by listing each of the 31 verbs (i.e. the *main predicate*), along with its overall frequency, as well as the frequency with which it appears with each argument structure. Each entry contains values for the syntactic and semantic features (see Table 1 for examples). By including these features, we assume that a learner is capable of (i) understanding basic syntactic properties of an utterance, such as syntactic categories (e.g., noun and verb) and word order, and (ii) perceiving and conceptualizing the general semantic properties of events — including mental, perceptual, communicative, and physical actions — as well as those of the event participants. Values for the semantic features (the *event primitives* and *event participants*) are taken from several resources, including Alishahi and Stevenson (2008), VerbNet (Kipper et al., 2008), and Dowty (1991). For each simulation in our experiments (explained below), we use the input-generation lexicon to automatically generate an input corpus of frames that reflects the observed frequency distribution in CDS. We perform 100 simulations, each on $20,000$ frames, and examine the behaviour of our model over the course of learning.

Experiment 1: The Imitation Task

Our goal here is to examine the role of verbs' overall frequency and their frequency with finite-SC, and the interaction of these frequencies, in the acquisition of argument structure constructions. Our simulations are inspired by the *imitation task* in which participants are asked to repeat a recently-heard utterance. Kidd et al. (2006, 2010) use this approach to examine the effect of verb frequency with finite-SC on how well children repeat utterances, in particular focusing on the relation between frequency of a verb with finite-SC and its likelihood of being correctly repeated, or substituted by another verb.

Experimental Setup

Following Kidd et al. (2006, 2010), we focus on whether our model correctly repeats the verb of a sentence in an imitation task involving CTVs with sentential complements. We present the model with a full frame representing a complete utterance plus its corresponding scene, analogous to the presentation of a sentence with an accompanying picture, as in the psycholinguistic experiments. We then ask the model to predict the best verb in response to that frame, essentially asking it to repeat the just-presented verb.

To consider the responses of the model over a developmental trajectory, we train it on the full corpus, and at periodic fixed points during training, we present it with a test frame for each of the 13 CTVs in our lexicon, to see how it responds to each CTV. All the test frames have the same syntactic features (i.e., *syntactic pattern*, *argument count*, and *complement type*) corresponding to a finite SC that contains a transitive action verb, paired with the appropriate semantic features for the given CTV (see Table 1). For consistency, we use the same physical action verb for the embedded verb (*other predicate*) in all 13 test frames, but randomly vary this verb across each of 100 simulations.

As in Kidd et al. (2006, 2010), we focus on the patterns of *verb repetition* and *verb substitution* among the model's responses. We record for each of the 13 test frames (at each point of testing) which verb the model predicts as its best response to that frame. To do this, we calculate the likelihood of each of our 31 verbs v given a test frame F_{test}, as in:

$$P(v|F_{\text{test}}) \quad = \sum_{k \in Clusters} P_{\text{main}}(v|k)P(k|F_{\text{test}}) \qquad (5)$$

where $P_{\text{main}}(v|k)$ is the probability of the main predicate feature having the value v in cluster k, calculated as in Eqn. (4), and $P(k|F_{\text{test}})$ is calculated as in Eqn. (2) (see Section for details). The model's response is taken to be the verb with the highest likelihood; this resembles the single choice of a verb made by the participants in the psycholinguistic experiments.

Results: Verb Repetition

Kidd et al. (2006, 2010) observe a positive correlation between the frequency of a verb with finite-SC and the proportion of its correct repetitions. We focus on how frequency with finite-SC impacts our model's correct repetition of a verb. Figure 1 presents the proportion of times that each of the 13 CTVs are correctly repeated, which we refer to as the repetition accuracy. According to these results, a high frequency with finite-SC is neither a necessary nor a sufficient condition for a verb to be correctly repeated by our model. For example, although the two verbs with the highest repetition accuracy (i.e., *think* and *say*) have high frequencies with finite-SC, other verbs with high frequency with finite-SC (i.e., *bet*, *guess*, *know*, *wish*, and *hope*) are not easy for our model to repeat. In addition, *see* is among the top four verbs to be correctly repeated, although it has relatively low frequency with finite-SC (see Table 2). These results suggest that other factors beyond the frequency with finite-SC examined by Kidd et al. (2006, 2010) may play a role here.

Our model enables us to explore some of the possible factors, and to make predictions that could be verified through experiments with children. For example, the overall frequency of the verb affects the model's responses: Out of the four highest-frequency verbs (*think*, *see*, *say*, *know*), three also have a higher repetition accuracy compared with the other CTVs. However, like frequency with finite-SC, overall frequency alone does not predict the responses: The repetition rate is not in frequency order, and *know* is high frequency but has a low repetition rate. In fact, we note that, except for the verb *think*, the model rarely repeats Belief verbs correctly, regardless of their frequencies. These results point to another factor that might affect the performance of our model in repeating a verb: the frequency with which semantically-related verbs appear with the same syntactic pattern as the verb to be repeated. To illustrate, when given a test frame that represents the semantic properties of a Belief verb with SC syntax, the model predicts the Belief verb with the highest frequency since it will have more occurrences in the clusters

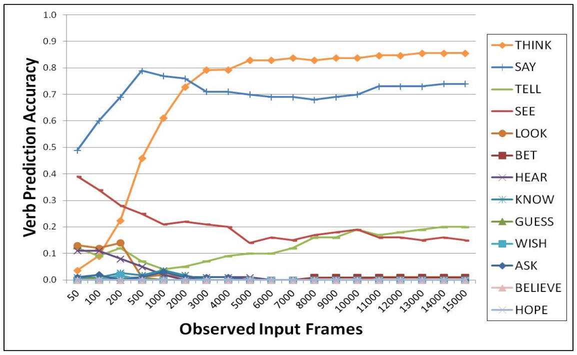

Figure 1: Verb repetition accuracy for each of the 13 CTVs.

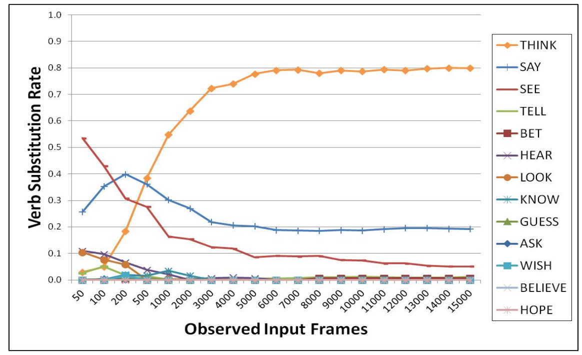

Figure 2: Verb substitution rate for each of the 13 CTVs.

that the model bases its predictions on (see Eqn. 5). E.g., the verb *know* will have usages in clusters with many more occurrences of *think*, and hence the model will mainly produce *think* in response to test frames containing the semantics of *know* with the SC syntax. To further understand the interaction of overall frequency and frequency with finite-SC, and distribution over semantic classes, we next look at the patterns of verb substitution by our model.

Results: Verb Substitution

Interestingly, Kidd et al. (2006, 2010) found that in a large number of cases, children specifically substituted the verb *think* in place of the verb they heard. They thus suggest that *think* is an 'exemplar' for the finite-SC construction. Figure 2 presents the proportion of times each of the 13 CTVs is produced by our model in place of the other 12 verbs, which we refer to as the substitution rate. That is, for each verb v, its substitution rate reflects the proportion of times that our model incorrectly produces v in response to the test frames for the other 12 CTVs (out of 100 simulations). In line with the findings of Kidd et al. (2006, 2010), the model substitutes the verb *think* for the other verbs with a very high likelihood from an early stage (See Figure 2)

Kidd et al. (2006, 2010) attribute their finding to the high frequency of the verb *think* with finite-SC. However, we have observed that *think* also has the highest overall frequency among the 13 CTVs (see Table 2). In addition, *think* is a Belief verb, and it is known that people form a strong association between Belief verbs and the finite-SC syntax (Gleitman et al., 2005). It is thus not clear whether the status of *think* as an exemplar for the finite-SC construction is solely due to its high frequency with finite-SC, or if it is also affected by these other factors: (a) the high overall frequency of *think*, and/or (b) the overall strong connection of Belief verbs to the construction. We explore these factors in the next set of experiments.

Interaction of the Different Frequency

One of the advantages of using a computational model is that we can manipulate the input to study the effects and interactions of the different frequency factors. Here, we manipulate the input such that we can examine the effects on the substitution patterns in our model of: overall frequency, frequency with finite-SC, as well as the frequency with finite-SC of a verb class as a whole. We perform three new experiments, in each of which we switch the overall frequency of *think* with one of the following three verbs: *guess*, *believe*, and *tell*. The goal is to change the input such that it is not the case that the verb with the highest overall frequency is also the verb with a frequency of 100% with finite-SC (as is the case with *think*)—that is, we want to tease apart the effect of these two frequencies.

The first interesting finding is comparing the results of making *guess* vs. *believe* (other Belief verbs) the highest-frequency verb (in place of *think*). This explores the impact of a relatively high (but not 100%) frequency with finite-SC (for *guess*, of 76%) and a low frequency with finite-SC (for *believe*, of 21%), in the context of a very high overall frequency. We find that, as in the original results with *think*, *guess* is substituted for other verbs a very high proportion of the time (75%). However this does not hold for *believe*; when it is the highest-frequency verb, the Belief verb with next highest overall frequency and relatively high frequency with finite-SC (*know*) becomes the verb most often substituted for others, with a substitution rate of 43%. This behaviour predicts that both a high overall frequency *and* a relatively high frequency with finite-SC are required for a verb to be treated as an 'exemplar' of the finite-SC construction.

We also examined the result of making *tell*, which is not a Belief verb, the highest frequency verb with fintie-SC (again, in place of *think*). Interestingly, although *tell* is a verb with a relatively high frequency with finite-SC (like *guess* above), *tell* does not become the verb the model most frequently substitutes for other verbs (in contrast to *guess*). In this case, *know* — a Belief verb — is the verb most frequently substituted for others. This suggests that the semantics of the verb also plays an important role in determining the substitution behaviour. The strong association of particular (Belief-verb) semantics with the finite-SC syntactic pattern are necessary to the verb substitution behaviour.

In summary, our findings suggest a somewhat different view from that of Kidd et al. (2010), who suggested that *think* was an exemplar verb in their experiments mainly because of its 100% frequency with finite-SC. The results of the input manipulation with *guess* and *believe* predict that for a verb to be the exemplar for a construction (here the finite-SC), it has to have a sufficiently high overall frequency and also appear with the construction with a relatively high frequency. In addition, although a semantically diverse group of verbs appears with finite-SC, the input manipulation involving *tell* suggests that the exemplar verb will come from the Belief class, since Belief verbs as a whole have an overall high frequency of appearance with the SC syntax.

Experiment 2: Generalization

Experiment 2 further examines the role of verb, construction, and semantic verb class frequencies in the acquisition of the finite SC. Given the noted strong association between the finite-SC syntax and Belief semantics, we focus here on the emergence of a 'Belief–SC' construction.

Experimental Setup

Following Casenhiser and Goldberg (2005), we focus here on the effect of the distributional pattern of verb usages with a particular construction on the acquisition of that construction. Casenhiser and Goldberg (2005) introduce five novel verbs appearing in a novel construction (a novel syntactic pattern paired with a novel meaning), to 5-to 7-year-old children in two input conditions: The skewed condition where one verb accounts for half of the occurrences of the construction, and the balanced condition with roughly equal number of usages of each verb. The study used a preferential-looking paradigm to show that participants in the skewed condition were significantly better at generalizing the newly-learned construction to a new novel verb (by looking at the scene with the appropriate semantics), compared to the balanced condition.

Our results in Experiment 1 imply that, in addition to the frequency with finite-SC of the individual verbs, their semantic class also influences the learning and use of verbs in a construction. This interaction of semantic classes is not addressed by the artificial language experiment of Casenhiser and Goldberg (2005), since it includes only a single class. Using a computational model enables us to explore the impact of a skewed vs. balanced distribution in a naturalistic setting, in which verbs from different semantic classes occur in the same syntactic frame under investigation (here, the finite SC), and verbs occur with multiple constructions (not just the one under investigation). Specifically, we examine how strongly our model learns the Belief–SC construction given the skewed input of our CDS-based data, compared to a balanced input, both with the same total exposure to CTVs. We form the balanced input by re-distributing the overall number of occurrences of CTVs so that each CTV would have an equal number of occurrences with finite-SC. Note that, all CTVs have an equal number of occurrences with finite-SC in the balanced input. However, because there is a different number of CTVs in each semantic class, the total number of finite-SC usages still slightly differs across classes.

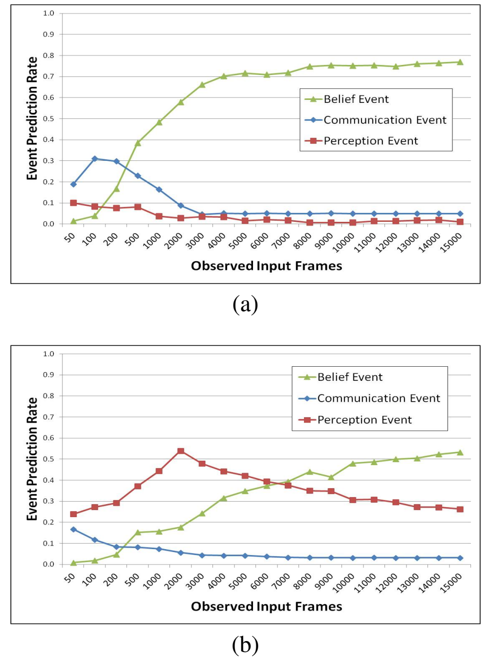

Figure 3: The 3 highest likelihood values of semantic properties of the event (a) given the CDS-based distribution, (b) given artificially balanced frequencies with finite-SC.

We need to evaluate the ability of the model to generalize its knowledge of the Belief–SC construction in response to a novel verb when training on these two types of input. However, the model is incapable of engaging in preferential looking, as in Casenhiser and Goldberg (2005). Instead, we simulate preferential looking in our model as a choice between possible sets of *event primitives*, given a test frame. Following the psycholinguistic settings, we construct the test frame with a novel verb in place of the *main predicate*, where event participants are associated with a belief event, but the semantics of the predicate is missing. In other words, the test frame represents the belief construction used to test the children, and each set of event primitives represents one of the test scenes in a preferential looking task. At each point of testing, over 100 simulations, we record the set of *event primitives* that the model predicts as its best response to the partial test frame. This prediction corresponds to the selection of the scene with the appropriate action, given the arguments and syntax of the construction (as in Casenhiser and Goldberg).

Results

Figure 3(a) and (b) report the proportion of times each of the three most likely sets of *event primitives* is chosen by our model as the most appropriate one, which we refer to as the event prediction rate.[2] Figure 3(a) shows that the semantics of Belief events is highly associated with the arguments and syntax of novel Belief verbs from an early stage, given

[2]Other sets of *event primitives* have lower likelihoods than the likelihoods presented here.

the skewed condition. That is, the Belief–SC construction is strongly entrenched given the naturalistically-skewed input.

However, in the balanced condition, as shown in Figure 3(b), only much later in training is the Belief event semantics predicted with the highest rate for the test frames. As in the results of Experiment 1, there is an effect of overall frequency in addition to frequency with finite-SC, of both verbs and classes. In the balanced input, each CTV has the same number of occurrences with finite-SC; hence there is only a small difference in the total number of occurrences of the different classes with this pattern. Recall that, to balance the input in terms of the CTV usages, we had to change the overall frequencies of the verbs and classes. In particular, we note that the overall frequency of the Belief class in the balanced input is much lower than that of the Perception class. The model is thus exposed to many more usages of Perception verbs with the finite-SC compared to the usages of Belief verbs with the same syntax, causing the observed delay in the formation of a strong Belief–SC construction.

Summary

We have used a computational model to examine the effect of various distributional properties of the input on the acquisition of argument structure constructions. Specifically, we have examined the interaction of several factors in the emergence of an exemplar verb for the finite SC construction. Our results suggest that exemplar-based learning of a construction (such as the finite SC) is sensitive to several properties of the input, including overall verb frequency, frequency of each verb with the construction, and the frequency of each semantic verb class with the construction. These results are in line with the psycholinguistic findings (e.g., Naigles & Hoff-Ginsberg, 1998; Casenhiser & Goldberg, 2005; Wonnacott et al., 2008; Kidd et al., 2006, 2010). Moreover, they further our understanding of the exemplar-based learning mechanism by providing a broader investigation of the role of each of the above factors in the context of naturalistic input that contains multiple classes of verbs each appearing with multiple constructions. Our findings signify the importance of considering the interaction of the various distributional factors in the design of psycholinguistic experiments.

References

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

Ambridge, B., Pine, J., Rowland, C., & Young, C. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, *106*(1), 87–129.

Barak, L., Fazly, A., & Stevenson, S. (2012). Modeling the acquisition of mental state verbs. In *Proc. of the 3rd Wkshp. on Cognitive Modeling and Computational Linguistics.*

Bloom, L., Hood, L., & Lightbowny, P. (1974). Imitation in language development: If, when, and why. *Cognitive Psychology*, *6*(3), 380–420.

Brown, R. (1973). *A first language: The early stages.* Harvard U. Press.

Casenhiser, D., & Goldberg, A. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, *8*(6), 500–508.

Dowty, D. (1991). Thematic Proto-Roles and Argument Selection. *Language*, *67*(3), 547–619.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64.

Goldberg, A. (1999). The emergence of the semantics of argument structure constructions. *The emergence of language*, 197–212.

Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, *21*(2), 93–107.

Kidd, E., Lieven, E. V., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences*, *32*(1), 132–142.

Kipper, K., Korhonen, A., Ryant, N., & Palmer, M. (2008). A large-scale classification of English verbs. *Language Resources and Evaluation*, *42*(1), 21–40.

Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 589–600.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*(3).

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (Vol. 2). Psychology Press.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, *20*(1), 121–136.

Naigles, L., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*(1), 95–120.

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. *Children's language*, *4*.

Suppes, P. (1974). The semantics of children's language. *American psychologist*, *29*(2), 103.

Theakston, A. (2004). The role of entrenchment in children's and adults' performance on grammaticality judgment tasks. *Cognitive Development*, *19*(1), 15–34.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253.

Wonnacott, E., Newport, E., & Tanenhaus, M. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, *56*(3), 165–209.