# Acquisition of Desires before Beliefs: A Computational Investigation

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

## Abstract

The acquisition of Belief verbs lags behind the acquisition of Desire verbs in children. Some psycholinguistic theories attribute this lag to conceptual differences between the two classes, while others suggest that syntactic differences are responsible. Through computational experiments, we show that a probabilistic verb learning model exhibits the pattern of acquisition, even though there is no difference in the model in the difficulty of the semantic or syntactic properties of Belief vs. Desire verbs. Our results point to the distributional properties of various verb classes as a potentially important, and heretofore unexplored, factor in the observed developmental lag of Belief verbs.

## 1 Introduction

Psycholinguistic studies have shown great interest in the learning of Mental State Verbs (MSVs), such as *think* and *want*, given the various cognitive and linguistic challenges in their acquisition. MSVs refer to an entity's inner states, such as thoughts and wishes, which the language learner must be able to perceive and conceptualize appropriately. Moreover, such verbs often appear in a Sentential Complement (SC) construction, which is complex for children because of the embedded clause.

Despite some shared properties, MSVs are a heterogeneous group, with different types of verbs exhibiting different developmental patterns. Specifically, a wealth of research shows that children produce Desire verbs, such as *want* and *wish*, earlier than Belief verbs, such as *think* and *know* (Shatz et al., 1983; Bartsch and Wellman, 1995; Asplin, 2002; Perner et al., 2003; de Villiers, 2005; Papafragou et al., 2007; Pascual et al., 2008). Some explanations for this pattern posit that differences in the syntactic usages of Desire and Belief verbs underlie the observed developmental lag of the latter (de Villiers, 2005; Pascual et al., 2008). In particular, Desire verbs occur mostly with an infinitival SC (as in *I want (her) to leave*), while Belief verbs occur mostly with a finite SC (a full tensed embedded clause, as in *I think (that) she left*). Notably, infinitivals appear earlier than finite SCs in the speech of young children (Bloom et al., 1984, 1989). Others suggest that Desire verbs are conceptually simpler (Bartsch and Wellman, 1995) or pragmatically/communicatively more salient (Perner, 1988; Fodor, 1992; Perner et al., 2003). Proponents of the conceptual and pragmatic accounts argue that syntax alone cannot explain the delay in the acquisition of Belief verbs, because children use finite SCs with verbs of Communication (e.g., *say*) and Perception (e.g., *see*) long before they use them with Belief verbs (Bartsch and Wellman, 1995).

We use a computational model of verb argument structure acquisition to shed light on the factors that might be responsible for the developmental gap between Desire and Belief verbs. Importantly, our model exhibits the observed pattern of learning Desire before Belief verbs, without having to encode any differences in difficulty between the two classes in terms of their syntactic or conceptual/pragmatic requirements. The behaviour of the model can thus be attributed to its probabilistic learning mechanisms in conjunction with the distributional properties of the input. In particular, we investigate how the model's learning mechanism interacts with the distributions of several classes of verbs — including Belief, Desire, Perception, Communication, and Action — in the finite and infinitival SC syntax to produce the observed pattern of acquisition of Desire and Belief verbs. Using a computational model can reveal the poten-



*Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pages 231–240,
Sofia, Bulgaria, August 8-9 2013. 

tial effects of interactions of verb classes in human language acquisition which would be difficult to investigate experimentally. Our results suggest that the distributional properties of relevant verb classes are a potentially important, and heretofore unexplored, factor in experimental studies of the developmental lag of Belief verbs.

## 2 The Computational Model

We require an incremental model in which we can examine developmental patterns as it gradually learns relevant aspects of argument structures. This task calls for an ability to represent the semantic and syntactic properties of verb usages, including those containing MSVs and other kinds of verbs taking sentential complements (SCs). Most computational models of verb argument structure acquisition have largely focused on physical action verbs (Alishahi and Stevenson, 2008; Chang, 2009; Perfors et al., 2010; Parisien and Stevenson, 2011). Recently, Barak et al. (2012) extended the incremental Bayesian model of Alishahi and Stevenson (2008) to include the syntactic and semantic features required for the processing of MSVs and other verbs that take SCs. While Barak et al. (2012) modeled some developmental patterns of MSVs overall, their work did not account for the difference between Desire and Belief verbs. In this section, we present their model, which we adopt for our experiments. In Section 3, we describe how we modify the representation of the input in Barak et al. (2012) to enable our investigation of the differences among the MSV classes.

### 2.1 Overview of the Model

The input to the Barak et al. (2012) model is a sequence of *frames*, where each frame is a collection of syntactic and semantic *features* representing what the learner might extract from an utterance s/he has heard paired with a scene s/he has perceived. In particular, we consider syntactic properties, including *syntactic pattern*, *argument count*, and *complement type*, as well as semantic properties, including *event primitives* and *event participants*. Table 1 presents a sample frame illustrating possible values for these features.

The model incrementally groups the input frames into *clusters* that reflect probabilistic associations of the syntactic and semantic features across similar verb usages. Each learned cluster is a probabilistic (and possibly noisy) representation of an argument structure construction: e.g., a cluster containing frames corresponding to usages such as *I eat apples*, *She took the ball*, and *He got a book*, etc., represents a Transitive Action construction.[1] Note that a cluster operates as more than simply a set of similar frames: The model can use the probabilistic associations among the various features of the frames in a cluster to generalize over the individual verb usages that it has seen. For example, if the model is presented with a frame corresponding to a transitive utterance using a verb it has not observed before, such as *She gorped the ball*, the example cluster above would lead the model to predict that *gorp* has semantic event primitives in common with other Action verbs like *eat*, *take*, and *get*. Such probabilistic reasoning is especially powerful because clusters involve complex interactions of features, and the model reasons across all such clusters to make suitable generalizations over its learned knowledge.

| | |
|---|---|
| head predicate | *think* |
| other predicate | *make* |
| **Syntactic Features:** | |
| syntactic pattern | arg1 *verb* arg2 *verb* arg3 |
| argument count | 3 |
| complement type | SC-fin |
| **Semantic Features:** | |
| event primitives | $\{$ $state, consider, cogitate, action$ $\}$ |
| event participants | $\{$ $experiencer, perceiver, considerer$ $\}$<br>$\{$ $agent, animate$ $\}$<br>$\{$ $theme, changed$ $\}$ |

Table 1: An example input frame. The Syntactic features reflect an utterance such as *He thinks Mom made pancakes*: i.e., syntactic pattern 'arg1 *verb* arg2 *verb* arg3', 3 arguments, and finite SC. The Semantic features reflect a corresponding conceptualized belief event with a physical action described in the SC ($\{state, consider, cogitate, action\}$) whose 'arg1' participant ($\{experiencer, perceiver, considerer\}$) perceives the 'arg2' ($\{agent, animate\}$) acting on the 'arg3' ($\{theme, changed\}$).

### 2.2 Algorithm for Learning Clusters

The model groups input frames into clusters on the basis of the overall similarity in the values of their syntactic and semantic features. Importantly, the model learns these clusters incrementally; the number and type of clusters is not predetermined. The model considers the creation of a new cluster for a given frame if the frame is not sufficiently similar to any of the existing clusters. Formally, the model finds the best cluster for a given input

[1]Note that, because the associations are probabilistic, a construction may be represented by more than one cluster.

frame $F$ as in:

$$\text{BestCluster}(F) = \underset{k \in Clusters}{\text{argmax}} \; P(k|F) \quad (1)$$

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

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

The prior probability of a cluster $P(k)$ is estimated as the proportion of frames that are in $k$ out of all observed input frames, thus assigning a higher prior to larger clusters, representing more frequent constructions. The likelihood $P(F|k)$ is estimated based on the match of feature values in $F$ and in the frames of $k$ (assuming independence of the features):

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

where $i$ refers to the $i^{th}$ feature of $F$ and $j$ refers to its value, and $P_i(j|k)$ is calculated using a smoothed version of:

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

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

### 2.3 Attention to Mental Content

One factor proposed to play an important role in the acquisition of MSVs is the difficulty children have in being aware of (or perceiving the salience of) the mental content of a scene that an utterance may be describing (Papafragou et al., 2007). This difficulty arises because the aspects of a scene associated with an MSV — the "believing" or the "wanting" — are not directly observable, as they involve the inner states of an event participant. Instead, younger children tend to focus on the physical (observable) parts of the scene, which generally correspond to the event described in the embedded clause of an MSV utterance. For instance, young children may focus on the "making" action in *He thinks Mom made pancakes*, rather than on the "thinking".

A key component of the model of Barak et al. (2012) is a mechanism that simulates the gradually-developing ability in children to attend to the mental content rather than solely to the (embedded) physical action. This mechanism basically entails that the model may "misinterpret" an input frame containing an MSV as focusing on the semantics of the action in the sentential complement. Specifically, when receiving an input frame with an MSV, as in Table 1, there is a probability $p$ that the frame is perceived with attention to the semantics corresponding to the physical action verb (here, *make*). In this case, the model correctly includes the syntactic features as in Table 1, on the assumption that the child can accurately note the number and pattern of arguments. However, the model replaces the semantic features with those that correspond to the physical action event and its participants. At very early stages, $p$ is very high (close to 1), simulating the much greater saliency of physical actions compared to mental events for younger children. As the model "ages" (i.e., receives more input), $p$ decreases, giving more and more attention to the mental content, gradually approaching adult-like abilities.

## 3 Experimental Setup

### 3.1 Generation of the Input Corpora

Because there are no readily available large corpora of actual child-directed speech (CDS) associated with appropriate semantic representations, we generate artificial corpora for our simulations that mimic the relevant syntactic properties of CDS along with automatically-produced semantic properties. Importantly, these artificial corpora have the distributional properties of the argument structures for the verbs under investigation based on an analysis of verb usages in CDS. To accomplish this, we adopt and extend the *input-generation lexicon* of Barak et al. (2012), which is used to automatically generate the syntactic and semantic features of the frames that serve as input to the model. Using this lexicon, each simulation corpus is created through a probabilistic generation of argument structure frames according to their relative frequencies of occurrence in CDS. Since the corpora are probabilistically generated, all experimental results are averaged over simulations on 100 different input corpora, to ensure the results are not dependent on idiosyncratic properties of a single generated corpus.

Our input-generation lexicon contains 31 verbs from various semantic classes and different frequency ranges; these verbs appear in a variety

| Semantic class | Verb | Frequency | % Relative frequency with SC-fin | SC-inf |
|---|---|---|---|---|
| Belief | think | 13829 | 100 | - |
| | bet | 391 | 100 | - |
| | guess | 278 | 76 | - |
| | know | 7189 | 61 | - |
| | believe | 78 | 21 | - |
| Desire | wish | 132 | 94 | - |
| | hope | 290 | 86 | - |
| | want | 8425 | - | 76 |
| | like | 6944 | - | 51 |
| | need | 1690 | - | 60 |
| Communication | tell | 2953 | 64 | - |
| | say | 8622 | 60 | - |
| | ask | 818 | 29 | 10 |
| | speak | 62 | - | - |
| | talk | 1322 | - | - |
| Perception | hear | 1370 | 21 | 25 |
| | see | 9717 | 14 | - |
| | look | 5856 | 9 | - |
| | watch | 1045 | - | 27 |
| | listen | 413 | 33 | 2 |
| Action | go | 20364 | - | 5 |
| | get | 16493 | - | 14 |
| | make | 4165 | - | 10 |
| | put | 8794 | - | - |
| | come | 6083 | - | - |
| | eat | 3894 | - | - |
| | take | 3239 | - | - |
| | play | 2565 | - | - |
| | sit | 2462 | - | - |
| | give | 2341 | - | - |
| | fall | 1555 | - | - |

Table 2: The list of our 31 verbs from the five semantic classes, along with their overall frequency, and their relative frequency with the finite SC (SC-fin) or the infinitival SC (SC-inf).

of syntactic patterns including the sentential complement (SC) construction. Our focus here is on learning the Belief and Desire classes; however, we include verbs from other classes to have a realistic context of MSV acquisition in the presence of other types of verbs. In particular, we include (physical) Action verbs because of their frequent usage in CDS, and we include Communication and Perception groups because of their suggested role in the acquisition of MSVs (Bloom et al., 1989; de Villiers, 2005). Table 2 lists the verbs of each semantic class, along with their overall frequency and their relative frequency with the finite (SC-fin) and infinitival SC (SC-inf) in our data.

For each of these 31 verbs, the distributional information about its argument structure was manually extracted from a random sample of 100 CDS usages (or all usages if fewer than 100) from eight corpora from CHILDES (MacWhinney, 2000).[2] The input-generation lexicon then contains the overall frequency of each verb, as well as the relative frequency with which it appears with each of its argument structures. Each argument structure entry for a verb also contains the values for all the syntactic and semantic features in a frame (see Table 1 for an example), which are determined from the manual inspection of the usages.

The values for syntactic features are based on simple observation of the order and number of verbs and arguments in the usage, and, if an argument is an SC, whether it is finite or infinitival. We add this latter feature (the type of the SC) to the syntactic representation used by Barak et al. (2012) to allow distinguishing the syntactic properties associated with Desire and Belief verbs. Note that this feature does not incorporate any potential level of difficulty in processing an infinitival vs. finite SC; the feature simply records that there are three different types of embedded arguments: SC-inf, SC-fin, or none. Thus, while Desire and Belief verbs that typically occur with an SC-inf or SC-fin have a distinguishing feature, there is nothing in this representation that makes Desire verbs inherently easier to process. This syntactic representation reflects our assumptions that a learner: (i) understands basic syntactic properties of an utterance, such as syntactic categories (e.g., noun and verb) and word order; and (ii) distinguishes between a finite complement, as in *He thinks that Mom left*, and an infinitival, as in *He wants Mom to leave*.

The values for the semantic features of a verb and its arguments are based on a simple taxonomy of event and participant role properties adapted from several resources, including Alishahi and Stevenson (2008), Kipper et al. (2008), and Dowty (1991). In particular, we assume that the learner is able to perceive and conceptualize the general semantic properties of different kinds of events (e.g., *state* and *action*), as well as those of the event participants (e.g., *agent*, *experiencer*, and *theme*). In an adaptation of the lexicon of Barak et al., we make minimal assumptions about shared semantics across verb classes. Specifically, to encode suitable semantic distinctions among MSVs, and between MSVs and other verbs, we aimed for a representation that would capture reasonable as-

[2] Brown (1973); Suppes (1974); Kuczaj (1977); Bloom et al. (1974); Sachs (1983); Lieven et al. (2009).

sumptions about high-level similarities and differences among the verb classes. As with the syntactic features, we ensured that we did not simply encode the result we are investigating (that children have facility with Desire verbs before Belief verbs) by making the representation for Desire verbs easier to learn.

In the results presented in Section 4, "our model" refers to the computational model of Barak et al. (2012) together with our modifications to the input representation.

### 3.2 Simulations and Verb Prediction

Psycholinguistic studies have used variations of a *novel verb prediction* task to examine how strongly children (or adults) have learned to associate the various syntactic and semantic properties of a typical MSV usage. In particular, the typical Desire verb usage combines desire semantics with an infinitival SC syntax, while the typical Belief verb usage combines belief semantics with a finite SC syntax. In investigating the salience of these associations in human experiments, participants are presented with an utterance containing a nonce verb with an SC (e.g., *He gorped that his grandmother was in the bed*), sometimes paired with a corresponding scene representing a mental event (e.g., a picture or a silent video depicting a *thinking* event with heightened saliency). An experimenter then asks each participant what the nonce verb (*gorp*) "means" — i.e., what existing English verb does it correspond to (see, e.g., Asplin, 2002; Papafragou et al., 2007). The expectation is that, e.g., if a participant has a well-entrenched Belief construction, then they should have a strong association between the finite-SC syntax and belief semantics, and hence should produce more Belief verbs as the meaning of a novel verb in an finite-SC utterance (and analogously for infinitival SCs and Desire verbs).

We perform simulations that are based on such psycholinguistic experiments. After training the model on some number of input frames, we then present it with a test frame in which the main verb (*head predicate*) is replaced by a nonce verb like *gorp* (a verb that doesn't occur in our lexicon). Analogously to the human experiments, in order to study the differences in the strength of association between the syntax and semantics of Desire and Belief verbs, we present the model with two types of test frames: (i) a typical desire test frame, with syntactic features corresponding to the infinitival SC syntax, optionally paired (depending on the experiment) with semantic features associated with a Desire verb in our lexicon; and (ii) a typical belief test frame, with syntactic features corresponding to the finite SC syntax, optionally paired with semantic features from a Belief verb.[3]

Given a test frame $F_{\text{test}}$, we use the clusters learned by the model to calculate the likelihood of each of the 31 verbs $v$ as the response of the model indicating the meaning of the novel verb, as in:

$$\begin{aligned} & P(v|F_{\text{test}}) \\ & \quad = \sum_{k \in Clusters} P_{\text{head}}(v|k) P(k|F_{\text{test}}) \\ & \quad \propto \sum_{k \in Clusters} P_{\text{head}}(v|k) P(F_{\text{test}}|k) P(k) \end{aligned} \tag{5}$$

where $P_{\text{head}}(v|k)$ is the probability of the $\text{head}$ feature having the value $v$ in cluster $k$, calculated as in Eqn. (4); $P(F_{\text{test}}|k)$ is the probability of the test frame $F_{\text{test}}$ given cluster $k$, calculated as in Eqn. (3); and $P(k)$ is the prior probability of cluster $k$, calculated as explained in Section 2.2.

What we really want to know is the likelihood of the model producing a verb from each of the semantic classes, rather than the likelihood of any particular verb. For each test frame, we calculate the likelihood of each semantic class by summing the likelihoods of the verbs in that class:

$$P(\text{Class}|F_{\text{test}}) = \sum_{v_c \in \text{Class}} P(v_c|F_{\text{test}})$$

where $v_c$ is one of the verbs in Class, and Class ranges over the 5 classes in Table 2. We average the verb class likelihoods across the 100 simulations.

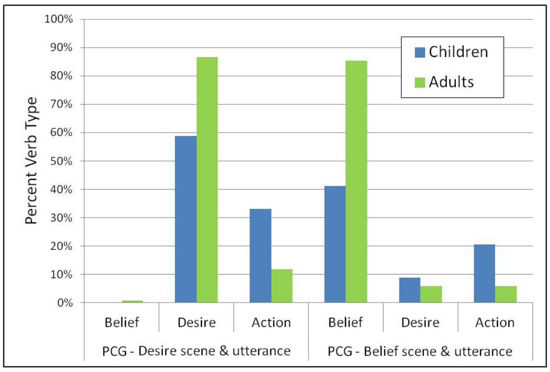
## 4 Experimental Results

The novel verb prediction experiments described above have found differences in the performance of children across the two MSV classes (e.g., Asplin, 2002; Papafragou et al., 2007). For example, children performed better at predicting that a novel verb is a Desire verb in a typical desire context (infinitival-SC utterance paired with a desire scene), compared to their performance at identifying a novel verb as a Belief verb in a typical belief

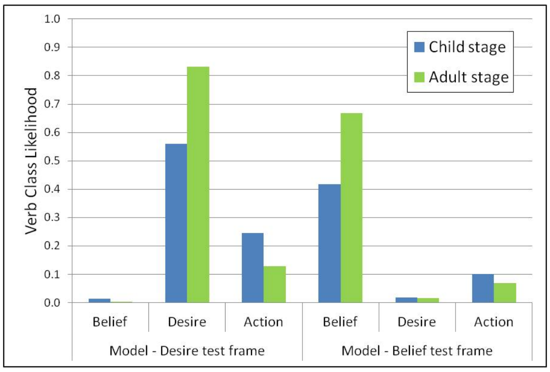
[3]Table 2 shows that, in our data, Belief verbs occur exclusively with finite clauses in an SC usage. Although Desire verbs occur in both SC-inf and SC-fin usages, the former outnumber the latter by almost 30 to 1 over all Desire verbs.

context (finite-SC utterance accompanied by a belief scene). In Section 4.1, we examine whether the model exhibits this behaviour in our verb class prediction task, thereby mimicking children's lag in facility with Belief verbs compared to Desire verbs.

Recall that some researchers attribute the above-mentioned developmental gap to the conceptual and pragmatic differences between the two MSV classes, whereas others suggest it is due to a difference in the syntactic requirements of the two classes. As noted in Section 3.1, we have tailored our representation of Desire and Belief verbs to not build in any differences in the ease or difficulty of acquiring their syntactic or semantic properties. Moreover, the possibility in the model for "misinterpretation" of mental content as action semantics (see Section 2.3) also applies equally to both types of verbs. Thus, any observed performance gap in the model reflects an interaction between its processing approach and the distributional properties of CDS. To better understand the role of the input, in Section 4.2 we examine how the distributional pattern of appearances of various semantic classes of verbs (including Belief, Desire, Communication, Perception and Action verbs) with the finite and infinitival SC constructions affects the learning of the two types of MSVs.




(a) Human participants in Papafragou et al. (2007)




(b) The model

Figure 1: (a) Percent verb types produced by adult and child participants given a desire or belief utterance and scene. (b) The model's verb class likelihoods given a desire or belief test frame. Child stage is represented by 500 input frames compared to the $10,000$ input frames for Adult stage.

### 4.1 Verb Prediction Simulations

Here we compare the verb prediction responses of the participants in the experiments of Papafragou et al. (2007) (PCG), with those of the model when presented with a novel verb in a typical desire or belief test frame. (See Section 3.2 for how we construct these frames.) PCG report verb responses for the novel verb meaning as desire, belief, or action, where the latter category contains all other verb responses. Looking closely at the latter category in PCG, we find that most verbs are what we have termed (physical) Action verbs. We thus report the verb class likelihoods of the model for the Belief, Desire, and Action verbs in our lexicon. To compare the model's responses with those of the children and adults in PCG, we report the responses of the model to the test frames at two test points: after training the model with 500 input frames, resembling the "Child stage", and after presenting the model with $10,000$ input frames, representing the "Adult stage".

Figure 1(a) gives the percent verb types from PCG;[4] Figure 1(b) presents the results of the model. Similarly to the children in PCG, the model at earlier stages of learning ("Child stage") is better at predicting Desire verbs for a desire test frame (**.56**) than it is at predicting Belief verbs for a belief test frame (**.42**) — cf. $59\%$ Desire vs. $41\%$ Belief prediction for PCG. In addition, as for both the children and adult participants of PCG, the model produces more Action verbs in a desire context than in a belief context at both stages.

We note that although the adult participants of PCG perform well at identifying both Desire and Belief verbs, the model does not identify Belief verbs with the same accuracy as it does Desire verbs, even after processing $10,000$ input frames (i.e., the "Adult stage"). In Section 4.2, we will see that this is due to the model forming strong associations between the Communication and Perception verbs and the SC-fin usage (the typical syntax of Belief verbs). These associations might be

[4]Based on results presented in Table 4, Page 149 in Papafragou et al. (2007), for the utterance and scene condition.

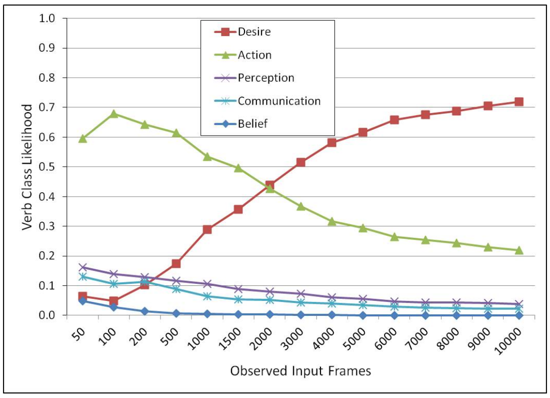
overly strong in our model because of the limited number of verbs and verb classes — an issue we will need to address in the future. We also note that, unlike the results of PCG, the model only rarely produces Desire verbs in a Belief context. This also may be due to our choice of Desire verbs, which have extremely few SC-fin usages overall.

To summarize, similarly to children (Asplin, 2002; Papafragou et al., 2007), the model performs better at identifying Desire verbs compared to Belief verbs. Moreover, we replicate the experimental results of PCG without encoding any conceptual or syntactic differences in difficulty between the two types of verbs. Specifically, because the representation of Desire and Belief classes in our experiments does not build in a bias due to the ease of processing Desire verbs, the differential results in the model must be due to the interaction of the different distributional patterns in CDS (see Table 2) and the processing approach of the model. Although this finding does not rule out the role of conceptual or syntactic differences between Desire and Belief verbs in delayed acquisition of the latter, it points to the importance of the distributional patterns as a potentially important and relevant factor worth further study in human experiments. We further investigate this hypothesis in the following section.

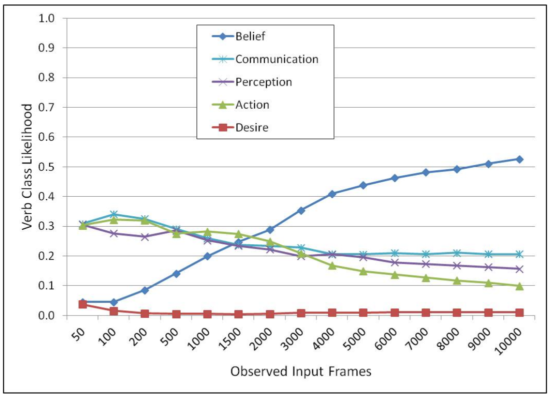
### 4.2 A Closer Look at the Role of Syntax

The goal of the experiments presented here is to understand how an interaction among the 5 different semantic classes of verbs, in terms of their distribution of appearance with the two types of SC constructions, coupled with the probabilistic "misinterpretation" of MSVs in the model, might play a role in the acquisition of Desire before Belief verbs. Because our focus is on the syntactic properties of the verbs, we present the model with partial test frames containing a novel verb and syntactic features that correspond to either a finite SC usage (the typical use of a Belief verb) or an infinitival SC usage (the typical use of a Desire verb).[5] We refer to the partial test frames as SC-fin or SC-inf test frames. We test the model periodically, over the course of $10,000$ input frames, in order to examine the progression of the verb class likelihoods over time.

[5]Verb prediction given an isolated utterance has been performed with adult participants (e.g., Gleitman et al., 2005; Papafragou et al., 2007). Here we simulate the settings of such experiments, but do not compare our results with the experimental data, since they have not included children.




(a) Model's likelihoods given SC-inf test frame




(b) Model's likelihoods given SC-fin test frame

Figure 2: The model's verb class likelihoods for the individual semantic classes.

First, we examine the verb class prediction likelihoods, given an SC-inf test frame; see Figure 2(a). We can see that all through training, the likelihoods are mainly divided between Desire and Action verbs, with the Desire likelihood improving over time. Looking at Table 2, we note that the Desire and Action verbs have the highest frequency of occurrence with SC-inf (taking into account both the overall frequency of verbs, and their relative frequency with SC-inf), contributing to their strength of association with the infinitival-SC syntax. Note that the very high likelihood of Action verbs given an SC-inf test frame, especially at the earlier stages of training, cannot be solely due to their occurrence with SC-inf, since these verbs mostly occur with other syntactic patterns. Recall that the model incorporates a mechanism that simulates a higher probability of erroneously attending to the physical action (as opposed to the mental event) at earlier stages, simulating what has been observed in young children (see Section 2.3 for details). We believe that this mechanism is responsible for some of the Action verb responses of the model for an SC-inf test frame.

Next, we look at the pattern of verb class likelihoods given an SC-fin test frame; see Figure 2(b). We can see that the likelihoods here are divided across a larger number of classes — namely, Action, Communication, and Perception — compared with Figure 2(a) for the SC-inf test frame. Since Action verbs do not occur in our data with SC-fin (see Table 2), their likelihood here comes from the misinterpretation of mental events (accompanied with SC-fin) as action. The initially high likelihoods of Communication and Perception verbs results from their high frequency of occurrence with SC-fin. Because at this stage Belief verbs are not always correctly associated with SC-fin due to the high probability of misinterpreting them as action, we see a lower likelihood of predicting Belief verbs. Eventually, the model produces more Belief responses than any other verb class, since Beliefs have the highest frequency of occurrence with the finite-SC syntax.

To summarize, our results here confirm our hypothesis that the distributional properties of the verb classes with the finite and infinitival SC patterns, coupled with the learning mechanisms of the model, account for the observed developmental pattern of MSV acquisition in our model.

## 5 Discussion

We use a computational model of verb argument structure learning to shed light on the factors that might underlie the earlier acquisition of Desire verbs (e.g., *wish* and *want*) than Belief verbs (e.g., *think* and *know*). Although this developmental gap has been noted by many researchers, there are at least two competing theories as to what might be the important factors: differences in the conceptual/pragmatic requirements (e.g., Fodor, 1992; Bartsch and Wellman, 1995; Perner et al., 2003), or differences in the syntactic properties (e.g., de Villiers, 2005; Pascual et al., 2008). Using a computational model, we suggest other factors that may play a role in an explanation of the observed gap, and should be taken into account in experimental studies on human subjects.

First, we show that the model exhibits a similar pattern to children, in that it performs better at predicting Desire verbs compared to Belief verbs, given a novel verb paired with typical Desire or Belief syntax and semantics, respectively. This difference in performance suggests that the model forms a strong association between the desire semantics and the infinitival-SC syntax — one that is formed earlier and is stronger than the association it forms between the belief semantics and the finite-SC syntax. Importantly, the replication of this behaviour in the model does not require an explicit encoding of conceptual/pragmatic differences between Desire and Belief verbs, nor of a difference between the two types of SC syntax (finite and infinitival) with respect to their ease of acquisition. Instead, we find that what is responsible for the model's behaviour is the distribution of the semantic verb classes (Desire, Belief, Perception, Communication, and Action) with the finite and infinitival SC syntactic patterns in the input.

Children are also found to produce semantically-concrete verbs, such as Communication (e.g., *say*) and Perception verbs (e.g., *see*), with the finite SC before they produce (more abstract) Belief verbs with the same syntax. Psycholinguistic theories have different views on what this observation tells us about the delay in the acquisition of Belief verbs. For example, Bartsch and Wellman (1995) suggest that the earlier production of Communication verbs shows that even when children have learned the finite-SC syntax (and use it with more concrete verbs), they lack the required conceptual development to talk about the beliefs of others. Our results suggest a different take on these same findings: because Communication (and Perception) verbs also frequently appear with the finite-SC syntax in the input, the model learns a relatively strong association between each of these semantic classes and the finite SC. This in turn causes a delay in the formation of a sufficiently-strong association between the Belief verbs and that same syntax, compared with the association between the Desire verbs and the infinitival SC.

de Villiers (2005) suggests that associating Communication verbs with the finite-SC syntax has a facilitating effect on the acquisition of Belief verbs. In our model, we observe a competition between Communication and Belief verbs, in terms of their association with the finite-SC syntax. To further explore the hypothesis of de Villiers (2005) will require expanding our model with enriched semantic representations that enable us to investigate the bootstrapping role of Communication verbs in the acquisition of Beliefs.

## References


Afra Alishahi and Suzanne Stevenson. 2008. A computational model of early argument structure acquisition. *Cognitive Science*, 32(5):789–834.

Kristen N. Asplin. 2002. *Can complement frames help children learn the meaning of abstract verbs?* Ph.D. thesis, UMass Amherst.

Libby Barak, Afsaneh Fazly, and Suzanne Stevenson. 2012. Modeling the acquisition of mental state verbs. *NAACL-HLT 2012*.

Karen Bartsch and Henry M. Wellman. 1995. *Children talk about the mind*. New York: Oxford Univ. Press.

Lois Bloom, Lois Hood, and Patsy Lightbown. 1974. Imitation in language development: If, when, and why. *Cognitive Psychology*, 6(3):380–420.

Lois Bloom, Matthew Rispoli, Barbara Gartner, and Jeremie Hafitz. 1989. Acquisition of complementation. *Journal of Child Language*, 16(01):101–120.

Lois Bloom, Jo Tackeff, and Margaret Lahey. 1984. Learning *to* in complement constructions. *Journal of Child Language*, 11(02):391–406.

Roger Brown. 1973. *A first language: The early stages*. Harvard Univ. Press.

Nancy Chih-Lin Chang. 2009. *Constructing grammar: A computational model of the emergence of early constructions*. Ph.D. thesis, University of California, Berkeley.

Jill G. de Villiers. 2005. Can language acquisition give children a point of view. In *Why Language Matters for Theory of Mind*, pages 199–232. Oxford Univ. Press.

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

Jerry A Fodor. 1992. A theory of the child's theory of mind. *Cognition*, 44(3):283–296.

Lila R. Gleitman, Kimberly Cassidy, Rebecca Nappa, Anna Papafragou, and John C. Trueswell. 2005. Hard words. *Language Learning and Development*, 1(1):23–64.

Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2008. A large-scale classification of English verbs. *Language Resources and Evaluation*, 42(1):21–40.

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

Elena Lieven, Dorothé Salomo, and Michael Tomasello. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20(3):481–507.

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

Anna Papafragou, Kimberly Cassidy, and Lila Gleitman. 2007. When we think about thinking: The acquisition of belief verbs. *Cognition*, 105(1):125–165.

Christopher Parisien and Suzanne Stevenson. 2011. Generalizing between form and meaning using learned verb classes. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*.

Belén Pascual, Gerardo Aguado, María Sotillo, and Jose C Masdeu. 2008. Acquisition of mental state language in Spanish children: a longitudinal study of the relationship between the production of mental verbs and linguistic development. *Developmental Science*, 11(4):454–466.

Amy Perfors, Joshua B. Tenenbaum, and Elizabeth Wonnacott. 2010. Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language*, 37(03):607–642.

Josef Perner. 1988. Developing semantics for theories of mind: From propositional attitudes to mental representation. *Developing theories of mind*, pages 141–172.

Josef Perner, Manuel Sprung, Petra Zauner, and Hubert Haider. 2003. Want That is understood well before Say That, Think That, and False Belief: A test of de Villiers's linguistic determinism on German–speaking children. *Child development*, 74(1):179–188.

Jacqueline Sachs. 1983. Talking about the There and Then: The emergence of displaced reference in parent–child discourse. *Children's language*, 4.

Marilyn Shatz, Henry M. Wellman, and Sharon Silber. 1983. The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14(3):301–321.

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