Modeling the Acquisition of Words with Multiple Meanings

Libby Barak, Sammy Floyd, and Adele Goldberg
Psychology Department
Princeton Univesrity
{lbarak,sfloyd,adele}@princeton.edu

Abstract

Learning vocabulary is essential to successful communication. Complicating this task is the underappreciated fact that most common words are associated with multiple senses (are **polysemous**) (e.g., baseball *cap* vs. *cap* of a bottle), while other words are **homonymous**, evoking meanings that are unrelated to one another (e.g., baseball *bat* vs. flying *bat*). Models of human word learning have thus far failed to represent this level of naturalistic complexity. We extend a feature-based computational model to allow for multiple meanings, while capturing the gradient distinction between polysemy and homonymy by using structured sets of features. Results confirm that the present model correlates better with human data on novel word learning tasks than the existing feature-based model.

1 Introduction

Children acquire language at a remarkable rate despite many layers of complexity in their learning environment. Previous computational models of human vocabulary learning have been primarily aimed at the mapping problem or the problem of "referential indeterminacy" (Quine, 1969), namely, determining which word maps onto which object within a noisy context (Siskind, 1996; Trueswell et al., 2013; Stevens et al., 2017; Smith et al., 2014; Fazly et al., 2010; Frank et al., 2009). These models explicitly make the simplifying but counter-factual assumption that each word can map to only one meaning in order to address how it is that learners determine which meaning a word refers to from among multiple potential referents in a scene. The models further assume that each possible meaning competes with every other possible meaning. For example, in a scene depicting "A cat drinking milk", the meaning of the word *cat* competes with the meaning of *milk*, *bowl* and every other potential meaning evoked in the scene. This perspective emphasizes the richness of visual scenes, but it overlooks the complexity associated with word meanings which very commonly refer to multiple distinct senses or meanings (Piantadosi et al., 2012). For example, a *bowl* can refer to a "dish used for feeding" in the cat scene, but to a "toilet bowl" within a different context. That is, the meaning of a word cannot be a winner-takes-all affair in which meanings compete with one another across contexts, because people learn to assign multiple meanings to many words in their vocabularies.

Multiple meanings of one word can typically not be subsumed under a general definition or rule. This is clearly true in the case of **homonyms**, which have multiple, unrelated meanings (e.g., baseball *bat* vs. flying *bat*). It is also true of many **polysemes**, which evoke conventional senses that are related to one another yet distinct. Natural language polysemy often involves extensions along multiple dimensions that are not completely predictable on the basis of a general definition or rule. For example, while baseball *caps* and bottle *caps* both cover something tightly, English speakers must learn that corks and lids, which also cover things tightly, are not called *caps*, while mushroom caps are, even though the latter do not cover anything tightly (for discussion of rule-based polysemy see e.g., (Srinivasan and Rabagliati, 2015; Srinivasan et al., 2017)). Notably, polysemes are much more frequent than homonyms, insofar as 40% of frequent English words are polysemous (Durkin and Manning, 1989), while closer to 4% of words are homonyms (Dautriche, 2015).

Even though homonyms are relatively rare, children as young as 3 years old have been found to know a number of them (Backscheider and Gelman, 1995). At least for these words, preschoolers have managed to overcome their reluctance to as-

Proceedings of the Society for Computation in Linguistics (SCiL) 2019, pages 216-225.
New York City, New York, January 3-6, 2019

sign a second meaning to a familiar word (Casenhiser, 2005). We also know that children readily generalize the meaning of a word to include new referents that share a single dimension, such as shape (Smith et al., 2002) or function (Gentner, 1978), and Srinivasan et al. (2017) has found that 4-5 year-old children can be taught that a word extends to other referents that share the same material (Srinivasan et al., 2017).

While previous psycholinguistic work has primarily focused on learning words with a single meaning or words that can be generalized along a single dimension (rule-based polysemy), a recent study that we simulate below has investigated words with multiple distinct, conventional meanings (non-rule-based). This work has demonstrated that it is easier to learn conventional polysemy when compared with homonymy, even when the polysemy follows complex, multidimensional extension patterns as in natural language. [1]

We propose a computational model that allows words to be assigned multiple meanings that cannot be generated by a one-dimensional rule, but must instead be learned through exposure (Brocher et al., 2017). We use the results from the behavioral experiment in order to inform and test the proposed model. As reported, the model not only captures the finding that people find it easier to learn polysemous words than ambiguous words, but it also closely approximates human errors. This represents a first step toward addressing the complexity involved in learning more than a single meaning of a given word.

2 Related Work

Only two recent models of human vocabulary learning allow words to evoke multiple senses. The model of Kachergis et al. (2017) implements a bias to prefer a single referent, but allows a second (unrelated) candidate meaning to be represented. Another model, Pursuit, maps each word onto a single candidate meaning per trial, and selects a new candidate meaning (at random) only when the primary meaning is disconfirmed (Stevens et al., 2017). This model retains a stipulation that only a single meaning wins. Importantly, neither of these models is evaluated on their ability to accurately represent multiple meanings. In fact, these and most other models make the simplifying assumption that each sense is represented atomically, without any internal structure or features. This precludes them from even attempting to distinguish polysemy from homonymy, since each meaning is equally (un)related to every other meaning.

It is necessary to allow word meanings to have internal structure if we are to capture relationships among meanings of a single word. The one model of human vocabulary learning that assigns such internal structure is the feature-based associative model of (Fazly et al., 2010), which has been extended in multiple studies to account for patterns of learning complex naturalistic meaning (Nematzadeh et al., 2012, 2014). This model represents a cross-situational learner, acquiring the meaning of each word incrementally by aligning each feature in the context with a probabilistic association to each word. The model learns by ultimately representing each word's meaning as an associated "bag-of-features". We choose this model as a basis for our approach, given its successful application in many word learning tasks and its ability to represent fine-grained properties (features) of meanings.

But critically, we extend the NFS12 model in order to represent the learning of words with multiple distinct meanings that may share overlapping features to varying degrees. The key innovation we add to the bag-of-features model of NFS12 is the following: we assign each distinguishable object a distinct, albeit overlapping, **set of features**. In our version, the model learns words as associations to distinct structured collections of feature-sets rather than learning independent associations of each word to each feature. We replicate the input and tasks of recent experimental multi-meaning word learning work, and compare the performance of the extended model with NFS12 and with the performance of human learners. In the following sections, we describe the original model and our modification of it.

3 Computational Models

3.1 Cross-situational Word Learning Model

We use the implementation of the cross-situational word learner as implemented by Nematzadeh et al. (2012) (**NFS12**) as the best fitting basis for our model. While later versions of the model are also available, these versions encode assumptions regarding hierarchical-categorical learning that are irrelevant to this research and require

[1] Experimental results are under submission.

hand-coded data of the categories in the input. NFS12 learns from <utterance, scene> input pairs that simulate what a language learner hears in the linguistic input: i.e., the utterance, and the features corresponding to the non-linguistic context (the scene). For example, the learner might first encounter the word *cap* accompanied by features that represent the scene of a parent asking a child to put a cap on a summer day, e.g.,

Utterance = "*put your cap on*"
Features = {sun, light, clothing, fabric,
cover, animate,...}

The features for each utterance correspond to all relevant aspects of the understood message and the witnessed scene. This is represented as a bag-of-features in the sense that there are no boundaries to indicate which features represent each object in the visual world. The model learns the probabilistic association between each feature, f, and each word, w, through a bootstrapping process. The model initializes all $P_{t-1}(f|w)$ to a uniform distribution over all words and features. At time t, the model learns the current association of w and f as proportional to their prior learned probability:

$$assoc_t(w, f) = \frac{P_{t-1}(f|w)}{\sum_{w' \in U} P_{t-1}(f|w')} \quad (1)$$

where $P_{t-1}(f|w)$ is the probability of f being part of the meaning of w at the previous learning step. If the association of f with some other word in the utterance is particularly high, the association of f with w will be correspondingly lower. The new evidence is then used to update the probability of all observed features in a smoothed version of:

$$P_t(f|w) = \frac{assoc_t(w, f)}{\sum_{f' \in F} assoc_t(w, f')} \quad (2)$$

where F is the set of all features observed thus far. The associations are thus summed over their occurrences in the input in proportion to the time passed since last occurrence.

$$assoc_t(f, w) = \ln(\sum_{t'=1}^{t} \frac{a_{t'}(w|f)}{(t - t')^d}) \quad (3)$$

The associations are updated with every learning step to account for past experience. The denominator represents the decay of the association over time as memories of the input are assumed to fade in memory. d is proportional to the strength of association such that stronger associations will fade less, even when significant time has passed since a previous encounter of w, i.e., $t - t'$. The learning iterations result in an association score between each feature and each word based on the observed input. The acquisition of word meaning is defined as success on a prediction task over the learned associations as described in Section 4.

3.2 Exemplar-based Learning as Sets of Features

The NFS12 model creates a bank of associations of varying strengths between features and words. It is based on the idea that over many observations of a word, the features that are actually relevant to that word will gain in probability over features that only coincidently co-occurred with the word in some subset of contexts. To date, no version of NFS12 has been evaluated on words with multiple senses. Note that if applied to multiple meanings in its current formulation, all of the features from all of the word's meanings will become associated with the word, without regard to whether certain features tend to occur with one meaning while other features tend to occur with a different meaning. That is, a word with multiple meanings will come to be associated with a merged bag-of-features. For instance, separate occurrences of the word *cap* would be associated with either $\{plastic, cover, bottle\}$ or $\{fabric, head\}$ but the model would predict that a combination of features such as $\{cover, fabric, bottle\}$ would be a reasonable interpretation of *cap*.

We predict that this vague representation will not be sufficient to approach human-like performance in recognizing distinct senses. Based on evidence that people are able to remember particular instances of objects they observe (Allen and Brooks, 1991; Brooks, 1987; Thibaut and Gelaes, 2006; Nosofsky et al., 2018), we modify the input representations to include sets of features for each word in the utterance as follows.

We propose a Structured Multi-Feature **(SMF)** model that extends NSF12, by associating each word with **sets** of features that have been learned on the basis of witnessing potential **referents** (as opposed to features) across scenes.[2] For example, if a scene involved two potential referents (the

[2]Like other models of human word learning, we focus our evaluation on for now the learning of words that correspond to referents in scenes.

sun and a baseball cap), the following feature sets would be candidates for association with the words in the utterance:

Utterance = "*put your cap on*"
Feature sets = {sun, light},
{clothing, fabric, cover}

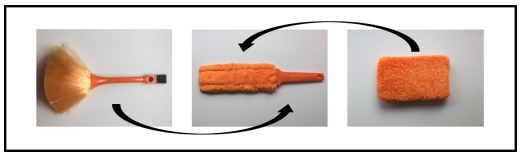
We modify the learning process to estimate the association of a word, w, and a **set** of features, s, following the formulation of the original model.

$$assoc(w, s) = \frac{P_t(s|w)}{\sum_{w' \in U} P_t(s|w')} \quad (4)$$

Thus a set of features, s, essentially represents an hypothesized sense of a referential word. The probability $P_t(s|w)$ is estimated from the previous occurrences of the word, where the probability of each set is proportional to the degree of overlap in features rather than a direct observation of the specific set. The degree of overlap between two sets, $s_f$and s_j is calculated using the Jaccard similarity coefficient, which is the proportion of shared features across the two sets over all features in the two sets.

$$jacc - sim(s_f, s_j) = \frac{|s_f \cap s_j|}{|s_f \cup s_j|} \quad (5)$$

The modification – making use of coherent sets of features rather than independent features – captures a key claim about how people learn referential words. Rather than learning the degree of association between words and individual features, e.g., learning *cap* and fabric, independently of the association between *cap* and clothing, the model assumes that people learn from coherent exemplars. The learner eventually learns a collection of sets of features with various degrees of association strength among the feature sets. The association between fabric and *cap* can only be determined once other features are taken into account as well. In this case, fabric will be more strongly associated with *cap* in the presence of the feature, clothing, and less associated with *cap* if the feature, bottle, is included and clothing is missing. [3]



(a) Polysemy

(b) Homonymy

Figure 1: A sample of the objects used in the novel-word learning experiment. Polysemy (upper panel) - pairs share properties as marked by arrows, with no single core feature shared by all three exemplars. Homonymy (lower panel) - a scrambled selection of three objects with fewer relationships among exemplars.

4 Learning Polysemy vs. Homonymy

We evaluate the NFS12 model and the present SMF model by simulating a novel word learning task in which human participants learned several polysemous or homonymous words as described below. The experiment compared how three populations learn words with multiple meanings: adults, typically developed children, and children with autism spectrum disorder. Since no previous computational models have attempted to capture how humans learn multi-meaning words, we focus on the adult group as a first step since it allows us to minimize assumptions regarding learners' development of cognitive abilities. We follow the experimental design to investigate, for the first time, how words with distinct but related senses (conventional polysemy) are learned, particularly when the range of senses do not follow from any language-wide rule.

4.1 Novel Word Learning Experiment

The experimental work explicitly compared the distinction between homonymy and conventional polysemy. In particular, participants learned 4 novel words, and each novel word was associated with 3 clearly distinct novel objects. Randomly interspersed among the 12 labeled objects were 20

[3] A very recent publication by the authors of NFS12 experiments with the use of sets but remains limited to single-sense word representations and still learns association of word over features rather than sets (Nematzadeh et al., 2017).

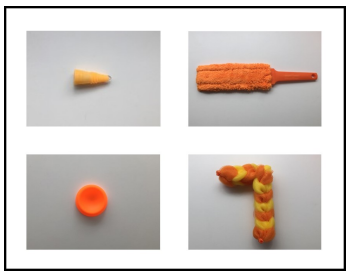


Figure 2: An example of the stimuli presented to participants in the label ID task. The target object was presented along with 3 distractors, which were targets for other novel words.

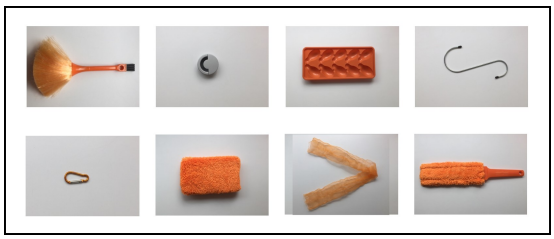


Figure 3: An example of the stimuli presented to participants in the Sense Selection task. All 3 target objects associated with one of the novel words were presented along with 5 filler objects, which had also been witnessed during the exposure but had not been labeled.

unlabeled filler (non-target) objects accompanied by tones. Novel objects were used to avoid interference from familiar words. Half of the participants were randomly assigned to a Polysemy condition in which the objects were related to one another with the 3 objects sharing distinct features with one another. The other half was assigned to a Homonymy condition, in which the 3 objects assigned to each word did now share any distinguishing features that could distinguish them from the filler objects in terms of a stronger feature-relation. (See Figure 1 for an example). The "polysemous" meanings of words were confirmed to be more similar than the "homonymous" meanings, as intended, using a separate norming study with a new group of participants.

After brief exposure, participants completed the following two tasks designed to determine whether polysemous words were easier to learn than homonymous words.

1. Label ID task - Participants were asked to select, from 4 available options, the one object that corresponded to a given label. The 3 foil objects had been labeled by each of the 3 other labels (see Figure 2). Results showed significantly higher accuracy for the polysemy condition over the homonymy condition.

2. Sense Selection task - Participants were presented with the label of one of the 4 words, and shown 8 objects (see Figure 3). Three of the objects corresponded to the 3 senses of the word and the 5 additional objects were fillers that had been witnessed during exposure. Accuracy was lower on this task, showing only slight polysemy advantage due to task difficulty.

The Label ID task allows a comparison of the two conditions, polysemy and homonymy, but participants may have used the memory of one or two words to perform well by a process of elimination by recognizing an object as related to a different label. The Sense Selection task allows for a more thorough error analysis; importantly, the particular objects selected by humans was made available to the computational analysis. We perform an error analysis on the selection rate of each filler object. The results of this task provide a crucial test of the bag-of-features meanings learned by the NFS12 model.

4.2 Experimental Simulation

We trained the model on input that reflected the exposure in the novel word study. In particular, two annotators hand-coded each object with 4 to 5 features to compose a joint a list of 40 features that jointly described all 12 labeled and 20 unlabeled (filler) objects. The features included properties related to shape, size, color, texture, material, symmetry, etc. We trained the NSF12 and SMF models independently: the features were used as a bag-of-features for the NFS12 model, and as structured sets of features for SMF. Recall that although each input item consisted of a single word associated with observable features, the models differ in the way they learn. NFS12 learns the association of a word with each feature, while SMF learns the association of a word to a subset of features.

At the end of training, we tested the models by simulating each of the two tasks described above. We first estimate the association of each word to each of the items, using the cosine distance between the learned associations and the feature rep-

	Polysemy	Homonymy
NFS12	0.88	0.37
SMF	**0.92**	**0.51**

Table 1: Pearson correlation between results from participants on the task with NSF12 and proposed SMF models.

resentation of the word. For the NFS12 model, we calculated the cosine similarity between all the associated features. For the SMF model, we calculated the maximum cosine similarity score over all the sets of features associated with the word and the feature representation of the item (i.e., we considered the sense of the word most similar to the object in question).

The likelihood of choosing an object as a target is measured by the proportional similarity of each object compared with the other objects presented in the task. For each stimuli set of 4 items used in the Label ID task (see Figure 2) and 8 items used for the Sense Selection task (see Figure 3) , we calculate

$$P(object|w) = \frac{cos(o, w)}{\sum_{o' \in O} cos(o', w)} \quad (6)$$

where, w is the word presented as visual stimuli at test. o ranges over all the objects presented at test (4 or 8 items), and O is the full set of objects for this test set.

5 Results

The experimental settings kept the alignment of sets of objects constant across participants while randomizing the word labels and the order of object presentation. For example, the same set of 4 objects in Figure 2 was used to test all 4 words. We replicated the combinations of objects to test each label in order to compare the computational models to people's choices. We used the default parameter settings included in the configuration files for NFS12.

5.1 Label ID Task

We first evaluate each model on its ability to replicate the polysemy advantage observed in human data. We obtain the item selection probability using Equation 6 for the target items only. Following the results from human experiments, we average the item probability over all targets to get the results from each model (see Figure 4).

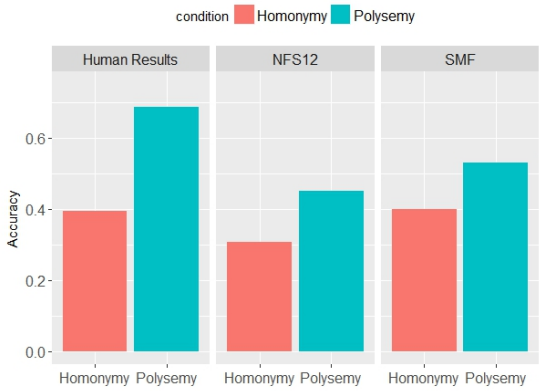


Figure 4: The likelihood of choosing an object corresponding to a target sense for Homonymy vs. Polysemy conditions in the Label ID task for (1) the human results, (2) the model of NFS12, (3) our extension of that model, SMF.

As can be seen in the middle panel of Figure 4, NFS12 replicates the human polysemy advantage in identifying the target meanings in the polysemy condition more accurately than targets in the homonymy condition. However, the accuracy of NFS12 in choosing the target object for the given label is considerably lower compared with the human results in the left most panel. The low accuracy of NFS12 compared with human performance suggests that NFS12 is prone to selecting non-target objects which do not resemble a specific sense of the learned word. Recall that because NSF12 does not maintain each exemplar's associated set of features, the model creates a superset of weighted features without respecting the co-variance among features that are associated with a particular sense. Thus, NSF12 assigns high probability to objects which share features from distinct target meanings, whereas humans are much less likely to do this.

To illustrate, Figure 1 provides example stimuli representing three senses of a polysemous vs. homonymous word. The middle object in the upper panel of Figure 1 shares some features with the left-most object (e.g., handle and overall shape), and other features with the right-most object (e.g., color, texture, and rectangular shape). However, the homonymous senses of the word share fewer features with one another (lower panel of Figure 1). Since almost no features overlap between pairs of homonymous objects, the number of features included in the superset for this word is

higher than for the polysemous word.

As a result, NFS12 under-performs in accuracy for all targets presented in the upper panel of Figure 1 in both the homonymy and polysemy conditions. In homonymy (lower panel of Figure 1), the bag-of-features consists of more features compared with the polysemous condition. Probabilistically, this bag-of-features will generate a higher number of subsets that happen to coincide with the features associated with fillers, which results in lower the accuracy of NFS12 for homonymy. For instance, NFS12 accuracy is significantly lower for the translucent yellow item in the middle of the panel because it simply aggregates the highly frequent features, learning a strong association between the word label and the feature *orange*. On the other hand, SMF preserves the co-occurrence statistics of features, preventing the orange feature from being incremented in isolation from the other features of that object.

The SMF model also captures the polysemy over homonymy advantage with higher accuracy than NFS12. Overall then, accuracy more closely matches human performance more closely when compared with NFS12 (see right panel on Figure 4). To quantify the correlation of each model with the particular selections made by human participants, we calculate the Pearson correlation over all objects (targets and fillers), using the results from Equation 6. (We use the Pearson correlation as the results from both human and models have normal distributions with kurtosis values close to 3).

The correlations with human errors for both models are given in Table 1. SMF offers significant improvement over NFS12 in the homonymy condition, and mirrors human errors in the polysemy condition slightly better as well. The weaker absolute correlation in the homonymy condition of the SMF model when compared with polysemy (.51 vs. 92) stems from the model over-performing on some items while under-performing on others, when compared with humans. We hypothesize that people differ from the model in the weights they give particular features, e.g., color vs. size. For example, SMF has higher accuracy than humans in selecting the leftmost item in the homonymy condition in Figure 1, possibly by forming a bias towards large-size items, while people may not attend to size to the same degree.

The models increase probability with every

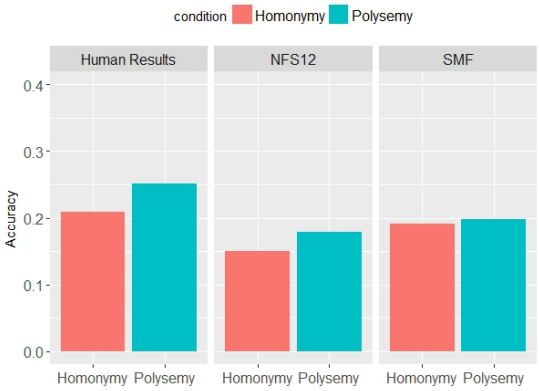


Figure 5: Sense Selection Task: The likelihood of choosing an object corresponding to a target sense for Homonymy vs. Polysemy conditions in (1) the human results, (2) the model of NFS12, (3) our extension of the model, SMF.

	Polysemy	Homonymy
NFS12	0.82	0.67
SMF	**0.91**	**0.90**

Table 2: Pearson correlation between errors produced by human participants with NSF12 and SMF models on the Sense Identification task.

overlapping feature regardless of what the feature denotes (shape, color, size, etc.). It is well known that children learn to attend to shape in learning referential novel nouns by two years of age (Smith et al., 2002). Moreover, people learn to attend to certain dimensions of meaning more closely given certain categories, e.g., using color to distinguish fruits and vegetables but not dogs and cats (Sloutsky et al., 2016). The SMF model overcomes this difficulty to some degree by having distinct memories of individual items. In order to capture these sorts of biases toward certain features for certain types of words, future models need to learn such biases over time, as we discuss further in Section 6.

5.2 The Sense Selection task

Following the results in subsection 5.1, we aim to further our analysis of the learning pattern of each model using a second task which challenges participants to recognize all senses of a word simultaneously, and includes more distractors (filler objects). The accuracy of choosing all three target items is presented in Figure 5.

Human's polysemy advantage was less pro-

nounced in the Sense Selection task compared with the Label ID task. As shown in Figure 5, both models also show less difference between polysemy and homonymy than they did on the Label ID task. While the polysemy advantage is higher in NFS12, SMF actually shows closer performance to the human data, due to more comparable levels of accuracy.

We again evaluate the probability of choosing each of the objects over both targets and fillers. That is, we compare the probability of each model selecting each object with human performance. In particular, we calculate the Pearson correlation between each of the two models and human results; see Table 2. The correlations of SMF with human results are much better than NFS12 in both the Polysemy and Homonymy conditions. These results align with our findings in the previous simulation, especially in mirroring NFS12's difficulty in learning unrelated senses (homonymy). The SMF model, on the other hand, approaches a 0.9 correlation in both conditions. Thus, although the SMF model has a lower overall probability of choosing targets compared to people, it closely mirrors human error patterns. These results support the role of distinct memories of exemplars, while taking into account the overlap among sets of features during selection. Note that the high correlations can be attributed to similarity in the relative ranking across items for the human results and SMF. At the same time, SMF still underestimate the overall probability of predicting certain items, which results in a lower accuracy compared with the human results.

6 Discussion and Future Directions

We have presented a computational analysis of the acquisition of word meaning for words with multiple senses. Despite the growing interest in computational models for analyzing human word learning, this aspect has remained under-studied due to the complexity of the problem. Our analysis is the first, to our knowledge, to directly model differences in the acquisition of multi-sense words with varying degree of overlap across senses. The computational design enables a closer analysis of the strengths and weaknesses involved in the human learning of multi-sense words, though the analyses of human errors.

The model of Nematzadeh et al. (2012) learned the association between independent features and words. It was chosen as the benchmark for our analysis because it represents the rare model which goes beyond atomic meanings by offering feature-based representations. Results demonstrate, however, that its bag-of-features representation is not sufficient to account for human-like learning of multi-meaning words, particularly in the case of homonymy, where combining the features of unrelated senses results in a particularly noisy representation. Our modified version which is a Structured Multi-Feature model, changes both the input representation and how the model learns to associate words with meanings. In particular, SMF preserves the co-occurrence statistics of the features associated with particular objects (exemplars), as motivated by evidence in human memory research (Allen and Brooks, 1991; Brooks, 1987; Thibaut and Gelaes, 2006; Nosofsky et al., 2018).

This study offers only the first step toward a computational model that fully captures the way that human learn realistic words, which commonly evoke a range of senses that importantly include function and metaphorical extensions that are not part of the interpretation of our novel stimuli. We recognize that our hand-coding of features makes both NFS12 and SMF impractical, but insofar as words meaningfully differ on a number of distinct dimensions, the reliance on features–however they are to be determined–is reasonable. Given the quite short exposure phase in the experimental work, the current analysis has not explored the role of memory or attention mechanisms included in the original model of NFS12 (Nematzadeh et al., 2014).

We believe that correlations with human performance could potentially be improved with suprisal or novelty affecting the weights of features. We also know that people pay more attention to some features over others in a way that depends on linguistic cues, the domain involved, and their prior knowledge. For example, people attend to colors to distinguish fruits, while color is less important when identifying dogs vs. cats.

The addition of structured sets of features offers an improvement over a general bag-of-features approach and has demonstrated strong correlations with human performance. Learning words with multi-meanings is a common occurrence in natural languages so it behooves models that aim to capture this basic fact.

Future extensions of SMF should incorporate a mechanism to simulate attention, including primacy and recency effects, in order to investigate how people weight different features or dimensions of meaning in various contexts. Although, NFS12 included a mechanism in the model to encode higher attention to novel words, this only captures item-based novelty, i.e., how frequently an item is observed, which does not play a significant role within the context of our experiment.[4] The multi-meaning words, however, introduce the challenge of attending to new meanings of familiar words over a short span of time. To more fully understand the relevant mechanisms and their roles in word learning, we plan to simulate the tasks discussed here using real-world polysemes with much richer sets of features. The conclusions of this study will be further used to guide extensions of the experimental designs in order to consider the role of attention in human word learning as well.

References

Scott W Allen and Lee R Brooks. 1991. Specializing the operation of an explicit rule. *Journal of experimental psychology: General*, 120(1):3.

Andrea G Backscheider and Susan A Gelman. 1995. Children's understanding of homonyms. *Journal of Child language*, 22(1):107–127.

Andreas Brocher, Jean-Pierre Koenig, Gail Mauner, and Stephani Foraker. 2017. About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, pages 1–24.

Lee R Brooks. 1987. Decentralized control of categorization: The role of prior processing episodes.

Devin M Casenhiser. 2005. Children's resistance to homonymy: An experimental study of pseudo-homonyms. *Journal of Child Language*, 32(2):319–343.

Isabelle Dautriche. 2015. *Weaving an ambiguous lexicon*. Ph.D. thesis, Sorbonne Paris Cité.

Kevin Durkin and Jocelyn Manning. 1989. Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research*, 18(6):577–612.

Afsaneh Fazly, Afra Alishahi, and Suzanne Stevenson. 2010. A probabilistic computational model of cross-situational word learning. *Cognitive Science*, 34(6):1017–1063.

[4] We use the original settings and keep this constant with future studies.

Michael C Frank, Noah D Goodman, and Joshua B Tenenbaum. 2009. Using speakers' referential intentions to model early cross-situational word learning. *Psychological science*, 20(5):578–585.

Dedre Gentner. 1978. A study of early word meaning using artificial objects: What looks like a jiggy but acts like a zimbo? *Reading in developmental psychology*.

George Kachergis, Chen Yu, and Richard M Shiffrin. 2017. A bootstrapping model of frequency and context effects in word learning. *Cognitive science*, 41(3):590–622.

Aida Nematzadeh, Barend Beekhuizen, Shanshan Huang, and Suzanne Stevenson. 2017. Calculating probabilities simplifies word learning. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*.

Aida Nematzadeh, Afsaneh Fazly, and Suzanne Stevenson. 2012. A computational model of memory, attention, and word learning. In *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics*, pages 80–89. Association for Computational Linguistics.

Aida Nematzadeh, Afsaneh Fazly, and Suzanne Stevenson. 2014. A cognitive model of semantic network learning. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Robert M Nosofsky, Craig A Sanders, and Mark A McDaniel. 2018. Tests of an exemplar-memory model of classification learning in a high-dimensional natural-science category domain. *Journal of Experimental Psychology: General*, 147(3):328.

Steven T Piantadosi, Harry Tily, and Edward Gibson. 2012. The communicative function of ambiguity in language. *Cognition*, 122(3):280–291.

Willard V Quine. 1969. Word and object. *Cambridge, Mass*.

Jeffrey Mark Siskind. 1996. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61(1-2):39–91.

Vladimir M Sloutsky et al. 2016. Selective attention, diffused attention, and the development of categorization. *Cognitive psychology*, 91:24–62.

Linda B Smith, Susan S Jones, Barbara Landau, Lisa Gershkoff-Stowe, and Larissa Samuelson. 2002. Object name learning provides on-the-job training for attention. *Psychological Science*, 13(1):13–19.

Linda B Smith, Sumarga H Suanda, and Chen Yu. 2014. The unrealized promise of infant statistical word–referent learning. *Trends in cognitive sciences*, 18(5):251–258.

Mahesh Srinivasan, Catherine Berner, and Hugh Rabagliati. 2017. Childrens use of lexical flexibility to structure new noun categories. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society*.

Mahesh Srinivasan and Hugh Rabagliati. 2015. How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, 157:124–152.

Jon Scott Stevens, Lila R Gleitman, John C Trueswell, and Charles Yang. 2017. The pursuit of word meanings. *Cognitive science*, 41(S4):638–676.

Jean-Pierre Thibaut and Sabine Gelaes. 2006. Exemplar effects in the context of a categorization rule: Featural and holistic influences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6):1403.

John C Trueswell, Tamara Nicol Medina, Alon Hafri, and Lila R Gleitman. 2013. Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive psychology*, 66(1):126–156.