

Discovering classes of attitude verbs using subcategorization frame distributions*

Aaron Steven White, Rachel Dudley, Valentine Hacquard & Jeffrey Lidz

University of Maryland, College Park

1. Introduction

In his classic squib on manner-of-speech predicates, Zwicky (1971) poses a cluster of questions foundational to the syntax-semantics literature:

To what extent is it possible to predict certain properties of words (syntactic, semantic, or phonological), given others? [And] insofar as there are such dependencies among properties, what general principles explain them? Put another way: what sorts of word classes are there, and why these and not others?

In response, he enumerates various distributional properties of manner-of-speech predicates—e.g. *shout, scream, yell, holler, etc.*—arguing that these are systematically associated with the semantic property MANNER OF SPEECH.

With similar aims, Fillmore (1970) discusses another canonical example of a systematic relationship between components of verb semantics and syntactic distribution, comparing the syntax of change-of-state verbs with that of surface-contact verbs. For Fillmore, this type of generalization was what linguists had to contribute to the study of meaning. Semantic distinctions beyond those that hold consequence for a word's possible syntactic environments were just as well left to the lexicographer. Embedded in this view is the idea that the syntax is responsive to only some meaning contrasts and that a linguistic theory of the link should speak to exactly which these are and why other conceivable contrasts are excluded (cf. Jackendoff 1972, Grimshaw 1979, Pinker 1989, Levin 1993). That is, though nonlinguistic meanings may be distinguishable to a very fine grain-size, linguistic meanings may not be.

This way of viewing the syntax-semantics interface puts particular stress on the relationships that enable inferences about a word's syntactic distribution *given knowledge of*

*We are grateful to reviewers and audience members at NELS as well as various audiences at the University of Maryland. We would also like to thank Jane Grimshaw and Pranav Anand for useful comments as we were developing the experiments presented here. The authors are jointly supported by an NSF BCS grant (1124338). Aaron Steven White was further supported by an NSF DGE IGERT grant (0801465).

the comcomitant semantic components. We might equally well ask what inferences about a word's semantics can be made *given its syntactic distribution*. This has been recognized as important from an acquisitionist's perspective since at least Brown (1957) on. Arguing for a syntactic bootstrapping approach to word-learning, Gleitman (1990) insists that understanding inferences from syntactic distribution to word meaning is a *sine qua non* of a theory of word-learning. One of her strongest arguments for this view comes from the case of verbs whose "semantic properties are closed to observation." Attitude verbs—e.g. *think, believe, want*, etc.—are a canonical case of such verbs. If one is trying to learn the word *think*, it seems unlikely that they could do so by "titrating discourse situations into ones in which thinking is going on somewhere when you hear /think/ versus those in which no thinking is happening." This problem becomes especially pernicious once one has to make even the most cursory distinctions among mental state predicates, or even attitude predicates in general—e.g. including also verbs of perception, speech act, etc.

This line of reasoning rests on the assumption that syntactic distribution represents an evidence base strong enough to make quite fine-grained inferences about meaning. We know from experimental work by Fisher et al. (1991) and Lederer et al. (1995) that syntactic distribution holds quantifiable information about coarse-grained distinctions, like that between a physical change-of-state verb like *break* and a mental state verb like *believe*; but it is still unclear to what level of granularity syntactic distribution holds information about distinctions between verbs whose semantic properties are closed to observation.

Various syntactic properties have been argued to track semantic distinctions. Bolinger (1968), for instance, argues that both mood selection in Romance, and the ability to be postposed in English, distinguish two semantic classes of attitude verbs: representationals (which describe "mental pictures," and non-representationals, which describe preferences. Hooper (1975) further distinguishes verbs based on assertivity—essentially, whether the predicate can be read parenthetically (cf. Urmson 1952, Reinhart 1983)—and factivity, using various syntactic diagnostics including postposition. One interesting consequence of Hooper's system is that Karttunen's (1971) semifactive verbs—such as *realize* and *discover*—should in principle be distinguishable from true factives—like *love* and *hate*—by their syntax alone. Such claims in the theoretical literature are important for the acquisitionist to take stock of because they provide possible evidence that the learner could use to infer fine-grained semantic distinctions.

In this paper, we address how fine-grained the information about attitude verb semantics present in the syntax is. To do this, we use quantitative measures akin to those Fisher et al. and Lederer et al. used to investigate coarse-grained distinctions. In section 2, we describe two experiments we ran aimed at gathering judgments about a range of attitude verbs' semantic classes and syntactic distribution. In section 3, we present two analyses showing (1) that there is information in the syntax about fine-grained distinctions among attitude verb meanings and (2) what that grain-size is. We conclude in section 4.

2. Experiments

We conducted two experiments to tease out how much information about attitude verb meaning is present in the syntax. The first aimed at discovering what categories naïve

speakers take different attitude verbs to fall into. The second aimed at ascertaining the compatibility of different syntactic constructions with these same verbs.

2.1 Semantic Similarity

We follow Fisher et al. in using a triad methodology to obtain a measure of semantic similarity between verbs. In this methodology, a set of n words of interest are chosen. In our case, we chose 30 attitude verbs from across the attitude verb lexicon (see Figures 1 and 2 for a full list). Every possible 3-combination of these verbs is then constructed, yielding 4060 triads.

60 participants were recruited through Amazon’s crowd-sourcing platform Mechanical Turk. Each participant was asked to “choose the verb that has a meaning least like the others” for 203 triads. For instance, if a participant saw the triad *think*, *believe*, and *say*, they might choose *say*. In this way, we obtained three similarity judgments for each of the 4060 triads.

2.2 Syntactic Compatibility

To obtain a measure of each verb’s compatibility with different syntactic constructions, we used a standard acceptability judgment task on a 7-point Likert scale. The 30 verbs from the semantic similarity experiment were paired with each of 34 different syntactic constructions (see the Appendix for a full list). For each of the resulting 1020 pairings, we created three sentences, yielding a total of 3060 items.

90 participants were recruited through Amazon’s crowd-sourcing platform Mechanical Turk. Sentences were separated into blocks of 102 such that (1) no block contained the same verb-construction pairing twice; (2) every block contained a token of every verb; and (3) every block contained three tokens of each construction each paired with a token of different verb. Each participant was asked to rate the sentences in their assigned block for acceptability—1 being *awful* and 7 being *perfect*.

Prior to analysis we applied a riddit transformation to the acceptability judgments to account for participants’ different uses of the Likert scale. We subsequently applied an item correction to these scores to account for differences between responses to the three items created for each pair. These corrected scores were then averaged, resulting in one score for each verb-construction pair—34 scores per verb.

3. Analysis

The data from both experiments were submitted to two different types of analyses. The first type—hierarchical clustering—is aimed at ascertaining (1) the verb classes latent in the semantic similarity and the syntactic compatibility judgments and (2) the correlation between these classifications. This correlation is interesting because it tells us how much the semantic similarity-based classification and the syntactic compatibility-based classification agree. As such, it gives us a rough measure of how much information about the

semantics is present in the syntax. We find here that the syntax carries a significant amount of information about the semantics.

In the clustering analysis, we make no assumptions about whether certain syntactic features or constructions matter more than others in terms of defining classes. This is both a strength and a weakness of this approach. It is a strength in that it tells us—knowing nothing about what parts of the syntax to pay attention to—how well we can do at grouping verbs. Knowing that we could do well even in this case is good sign, since we can be confident that words' syntactic distributions are indeed correlated with their semantics. However, it would approximate a learner who thinks of syntactic constructions as independent sources of information and who is agnostic about whether any of these sources are more important than any other. That it corresponds to such a learner is also its weakness, since such a learner is not realistic. It seems likely that learners at least know that constructions share components (syntactic features) and that these components are indicative of certain classes over others. Without such knowledge, we are not squeezing as much information out of the syntax as we might otherwise be able to.

To address this issue, we turn to our second type of analysis—maximum entropy modeling—which is aimed at ascertaining just how much information we could squeeze out of the syntax, given an optimal mapping from the syntactic compatibility scores to the semantic similarity scores. That is, if we try to approximate a learner that knows something about the importance of different constructions how much more fine-grained might information in the syntax look? We will approach this question by breaking it in two: (1) *how accurate would a model that was given only syntactic distributions be in predicting semantic classes?*; and (2) *if we tried to get a model to perfectly predict which lexical items were semantically similar given only the syntax, how fine-grained would those predictions be?* We find that the answer to these two questions converge.

3.1 Clustering

We begin with hierarchical clustering, which is a form of unsupervised learning in which a matrix of distances between objects—in our case, verbs—is taken as input and a dendrogram—a tree structure representing different levels of classification—is output. Here, we apply hierarchical clustering, using Ward's method as our linkage criterion, to discover semantic class at different levels of abstraction for both the semantic similarity judgments and the syntactic compatibility judgments we collected.

Semantic clustering

Before running the clustering algorithm on any data, we have to first define what distance between two verbs is in a context. To obtain distances between our verbs from the semantic similarity judgments described in section 2.1, we simply count the number of times a verb was chosen as dissimilar from two others in a triad. For each such judgment, we increment the dissimilarity score between that verb and the two others. This provides us with a symmetric matrix of dissimilarities, which we can think of as representing the distances

Discovering classes of attitude verbs

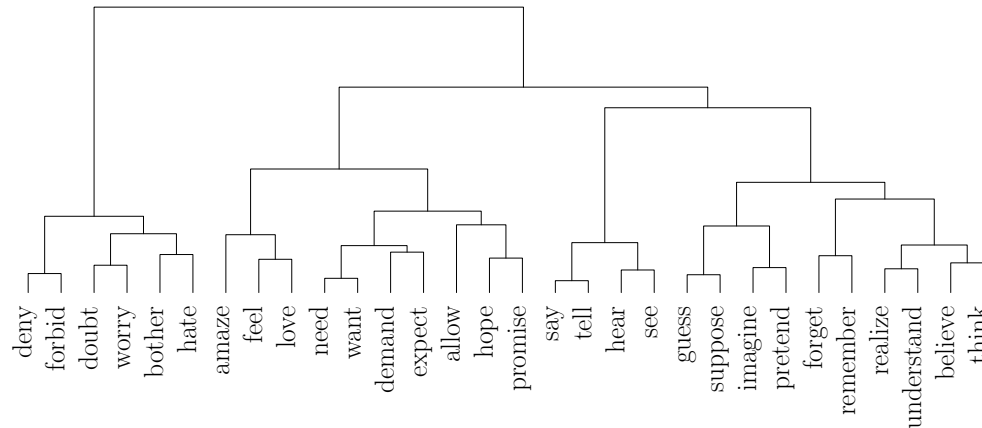


Figure 1: Semantic clusters

between verbs. Figure 1 shows the results of applying hierarchical clustering to the verb dissimilarities.

We find a few major distinctions emerging from these data. Perhaps the most interesting of these is the second-level split between *amaze*, *feel*, *love*, *need*, *want*, *demand*, *expect*, *allow*, *hope*, and *promise*, on the one hand, and *say*, *tell*, *hear*, *see*, *guess*, *suppose*, *imagine*, *pretend*, *forget*, *remember*, *realize*, *understand*, *believe*, and *think* on the other. This split looks much like Bolinger's, discussed in section 1. The first class, we will refer to as the *think*-type class, and second, the *want*-type class.

Within the *think*-type class, we find a further split between speech and perception verbs (*say*, *tell*, *hear*, *see*) and mental state verbs. Within the mental state verbs, we also find a class involving counterfactuality (*guess*, *suppose*, *imagine*, *pretend*) and one involving factive (*forget*, *remember*, *realize*, *understand*) and nonfactive (*think*, *believe*) belief verbs.

Zooming back out to the highest level split, we find a distinction made between verbs with some sort of negative component—*deny*, *forbid*, *doubt*, *worry*, *bother*, *hate*—and those in our *think*- and *want*-type classes. Interestingly, not all verbs in our study that might have conceivably been classified here were: for instance, *forget* seems like plausible candidates for inclusion in this class. We will return to this in section 3.2.

Syntactic clustering

To obtain distances between our verbs from the syntactic compatibility judgments, we can think of each verb as a point with 34 dimensions—one dimension for each of our 34 syntactic constructions. We can then take the Minkowski ($k = 3$) distance¹ between each verb

¹The 3-norm was decided on by first calculating the Minkowski distance for all $k \in [1, 100]$. (The 2-norm is equivalent to Euclidean distance.) For each k the performance according to the Mantel test (correlation

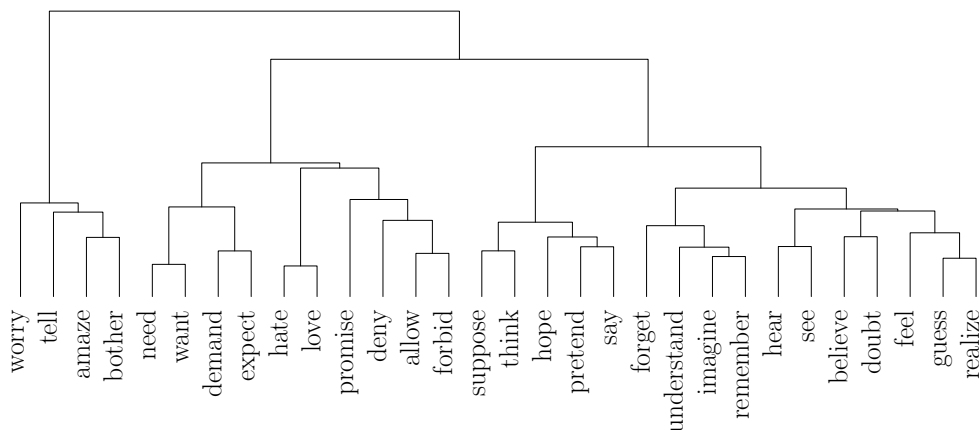


Figure 2: Syntactic clusters

and each other verb, resulting in a dissimilarity (distance) matrix with the same dimensions as the semantic similarity matrix. Figure 2 shows the results of applying hierarchical clustering to these dissimilarities.

At a coarse-grained level, this clustering looks similar to our semantic clustering in that Bolinger’s classes emerge. We see a very similar *want*-type class (*need, want, demand, expect, hate, love, promise, deny, allow, forbid*) and *think*-type classes (*suppose, think, hope, pretend, say, forget, understand, imagine, remember, hear, see, believe, doubt, feel, guess, realize*). The same granularity is not apparent within each of these classes, however. We don’t see a split between counterfactuals and belief verbs; speech and perception verbs (*say, tell*², *hear, see*) do not show up anywhere near each other; and though three of our factives are grouped together (*forget, understand, remember*), the fourth (*realize*) is quite far away.

Despite this lack of agreement at the finer grained levels, we find a significant correlation between our two information sources (Mantel test [$n = 999$] using Pearson’s $r = .301$; $p < .001$). This means that, on the whole, similarities between verbs in the semantic similarity judgments agree with the similarities between verbs derived from the syntactic compatibility judgments. And since the correlation is significant, we can conclude that the syntax does in fact carry information about attitude verb semantics, even absent assumptions about which constructions are important.

with the semantic similarity matrix) was then found. The matrices produced using Euclidean distance (2-norm) and the 3-norm both performed well, with the 3-norm only slightly edging the Euclidean distance out. Importantly, the correlations of the syntactic distance matrices with the semantic similarity matrix between for all k were significant even after Bonferroni corrections.

²In fact, we see that *tell* shows up in a cluster with verbs that take expletive subjects and EXPERIENCER objects. This probably happens because subjects can interpret *it* in these cases referentially and the object as though it had a RECIPIENT and not an EXPERIENCER role.

Discussion

We saw that even if we naïvely—that is, without knowledge of what features are or are not important—discover verb classes from the syntax, there is a significant correlation between the clusterings. This assumption potentially results in information loss, however, since some syntactic constructions may be more important for categorization than others. To take this into account, we will amend these assumptions by allowing different constructions to have different weights, or importances, in terms of telling us about a verb’s semantics. This weighting will allow us to find an upper bound on the information granularity about the semantics contributed by these constructions.

3.2 Maximum Entropy Modeling

Our goal here is to find an upper bound on the granularity of information in the syntax there is about the semantics. That is, we would like to squeeze as much semantic information out of the syntax as possible. To do this, we build two types of models: one that attempts to predict a word’s category given its syntax and another that attempts to replicate human performance on the similarity judgment task using only information about the syntax.

The first will address the question *how accurate would a model that was given only syntactic distributions be in predicting semantic classes like the ones represented in Figure 1?* Doing this at different depths in the tree corresponds to predicting different granularities in the semantics. We can then compare these accuracies to chance to see at what depth in the semantic clustering tree the model fails.

The second will address the question *if we tried to get a model to perfectly predict which lexical items were semantically similar given only the syntax, how fine-grained would those predictions be?* That is, at what point do different lexical items become indistinguishable for a model that knows only about the syntax?

Semantic Category Prediction Models

The tree in Figure 1 was cut at various points to generate semantic class labels. So for instance, a two category split would yield the negative affect³ (*deny, forbid, doubt, worry, bother, hate*) and *want*-type (*need, want, demand, expect, hate, love, promise, deny, allow, forbid*) + *think*-type (*suppose, think, hope, pretend, say, forget, understand, imagine, remember, hear, see, believe, doubt, feel, guess, realize*) classes discussed in section 3.1. A three category split would yield the negative affect, *want*-type, and *think*-type classes, and so on.

We then fit a series of models to predict the categories at each cut depth. For each such depth, we resample the data (n=100) and fit a model, then extract the accuracy of that model in predicting the verb class information. Figure 3 shows the model performance at each depth. These results suggest that the model at cut depth 4, which makes a distinction

³The similarity judgments show a major distinction between verbs with negative affect vs. all the others. But it’s unclear that the syntax would pick up on such a distinction. It may well be that the negative component of these verbs was so salient that it masked other important similarities.

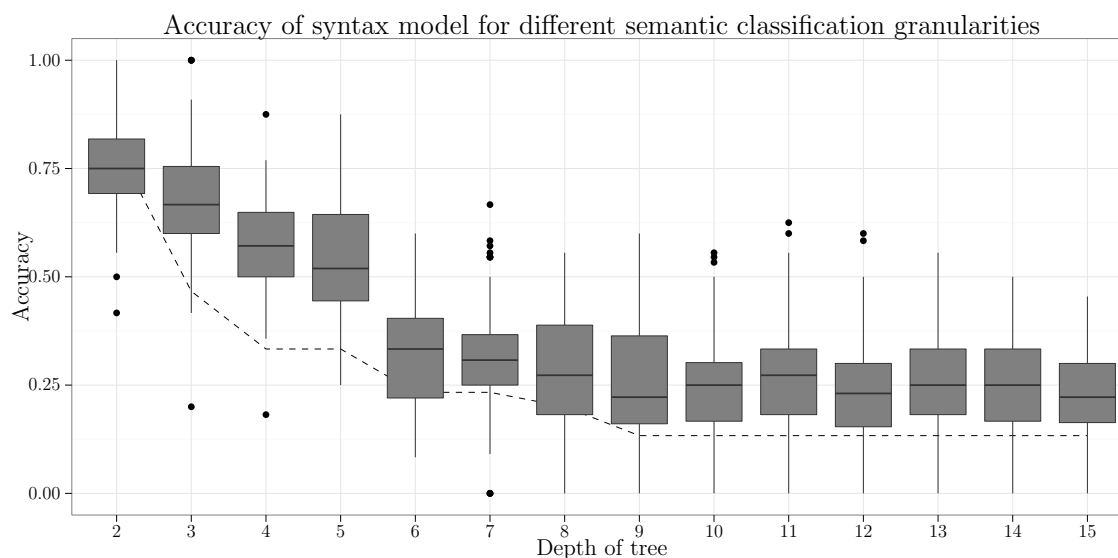


Figure 3: Accuracy the syntax models at various depths of the tree in Figure 1; the dotted line shows chance guessing given the knowledge of the relative class proportions

between a negative affect, *want*-type, *think*-type, and perception/speech class, is the only one that shows resampled accuracy reliably above chance.

This four-way split suggests two things. First, there seems to be slightly more information than is apparent from Figure 2 present in the syntax. We saw earlier that only our major *want*-type/*think*-type split seemed present in the unweighted syntactic clustering. Second, the classes represented in the two-way and three-way splits are not represented in the syntax. This latter result is counterintuitive but is further corroborated with our similarity judgment model in the next section. We discuss the implications in the Discussion section.

Similarity Judgment Model

We fit a multinomial logistic regression⁴ (maximum entropy model) to the raw similarity judgments obtained from participants. The mean syntactic judgments were used for each verb as predictors. So for instance, in the case where *think* and *believe* were judged as similar—i.e. not chosen as the dissimilar in a triad—a data point with *think* as the response and the syntactic compatibility judgments for *believe* and a data point with *believe* as response and the syntactic compatibility judgments for *think* would be entered.

The best model after tuning was then used to predict the probability of each verb given the syntax of each other verb. We can think of the resulting matrix of probabilities as telling us how probable we are to move from one verb to another in a sort of free association. For example, supposing one started with *believe* and was asked to choose another verb with a similar meaning, one might transition from *believe* to *think*; once they were at *think*, they

⁴We used 10-fold cross-validation with 10 repeats to validate this model.

Discovering classes of attitude verbs

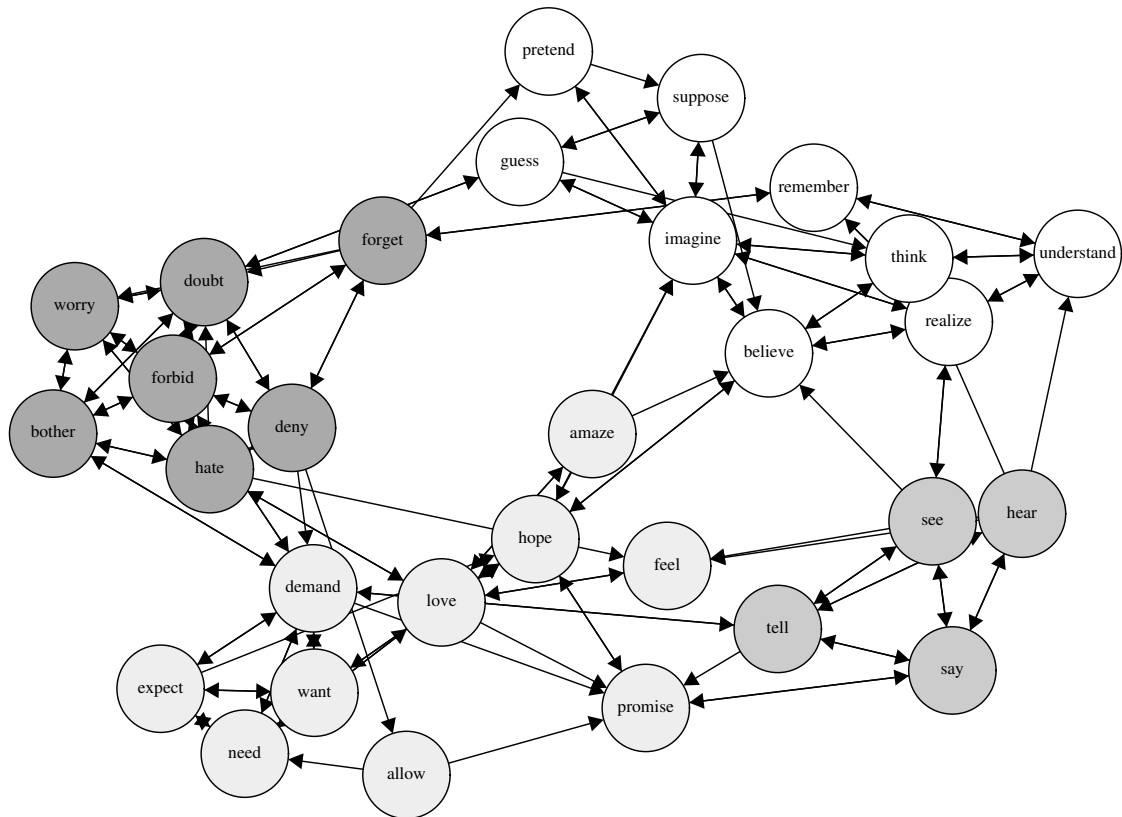


Figure 4: Model’s most likely guesses for similar words; shading represents classes found by Walktrap algorithm

might next transition to *realize*. Figure 4 shows the top 15% of the model’s connections between verbs. This is the percentage that provides a graph with full connectivity—i.e. in which there is a path from any verb to any other verb.

Semantic Class Discovery

To find how fine-grained the encoding of the semantics in the syntactic distribution (given only these constructions) is, we applied Pons and Latapy’s (2005) Walktrap algorithm with 5 steps. This algorithm essentially finds communities—i.e. verb clusters—by having the model do short free associations, starting at many random points, like the one described above. When the free associations often stay within a small group of verbs over many runs, that group is likely a cluster. We find four such communities, all of which correspond quite well to the clusters found in the semantic similarity judgments. This is interesting because it corroborates the accuracy results from our category prediction models.

1. *see, hear, tell, say*
2. *pretend, suppose, guess, imagine, remember, think, realize, understand*
3. *amaze, hope, feel, promise, love, demand, want, need, expect, allow, promise*
4. *worry, doubt, forbid, deny, hate, bother, forget*

First, we see signs of our major split between *think*-type and *want*-type verbs. This is unsurprising since both the syntactic clustering and the semantic clustering discovered these groupings, but it is heartening because it is possible that, in trying to fit our contrasts found in the semantics, we could have lost sensitivity to ones that we found in the syntax originally.

The second major class we see is the speech and perception verb cluster. This was one of the clusters that came out in the semantic clustering, but not in the syntactic clustering. This is interesting because, without weighting our syntactic constructions, this class does not look to be present in the syntactic component. In fact, as noted in footnote 2, *tell* shows up even outside the major *want*- and *think*-type clusters.

Another interesting cluster is the negative affect class we originally found in the semantic similarity judgments. Negative affect for some reason bore very strongly on participants' semantic judgments but did not show up at all in the syntactic judgments; it was spread across the *think*-type and *want*-type classes. Yet here, we see that the semantic feature can be picked out by the syntax with the proper weighting. Perhaps tellingly, this class is where we find some fluidity with respect to the verbs it contains. For instance, whereas *forget* in both the semantic clustering and the syntactic clustering shows up in the *think*-type class, it shows up here with the negative affect verbs. This is probably due to an underlying similarity with negative verbs in the semantic similarity judgments that was washed out by its high similarity to *remember*, which is firmly within the *think*-type class.

3.3 Discussion

We found convergent evidence that weighting syntactic constructions made the granularity finer with respect to our participants' judgments of semantic similarity. However, we also found that the syntax does not carry information about some coarser-grained splits present in the semantics—specifically, the first (two-way) and second (three-way) splits in Figure 1. It seems plausible that the presence of a negative affect class played a role in this. In this case, a good portion of the error in the subsequent models could be due to their inability to predict the negative affect class well.

This raises the question to what extent we should rely on raw semantic similarity judgments from naïve participants. These participants seemed to be sensitive to quite fine-grained semantic features, like counterfactuality and factivity, but they also seemed to be sensitive to features that linguists don't necessarily think bear on the syntax, like negative affect. This could in turn have hindered our model's ability to detect finer-grained distinctions than it did.

This is a worry in one sense but not another. If we move from a model in which we are agnostic about weights on constructions to one where we know the optimal weights, we find that the syntax does in principle carry finer-grained information than it does at first glance—both qualitatively and quantitatively. This suggests that the syntax is mirroring something about naïve similarity judgments. If it were not, we wouldn't expect to find an increase in the ability of the syntax to predict these judgments.

On the other hand, it's possible that we missed something by not taking into account possible weights on semantic features more salient to the semantic system. As discussed

in section 1, it has been argued that the semantic component encodes only a subset of all possible conceptual features, making only this subset available to project into the syntax. In a sense, we would have needed to weight similarity judgments in such a way as to highlight certain features over others. This sort of weight might correspond to higher salience of certain conceptual features to the linguistic system. It is unclear what a principled way of doing this would be, however, so we leave it for future research.

4. General Discussion

In this paper, we have extended experimental findings by Fisher et al. Using a triad methodology along with hierarchical clustering, we showed that naïve speakers are sensitive to fine-grained semantic features of attitude verbs. We also showed that similar classes are encoded in latent in naïve speakers syntactic compatibility judgments. We tested this association quantitatively and found that there was a significant correlation between the two, suggesting shared information.

Acknowledging that some constructions may be more important than others in determining semantic class, we built a model that attempted to simulate the human acceptability judgments given only the syntactic compatibility judgments. Using this model, we investigated what classes were latent in the model's judgments. We found that the model discovered finer-grained classes relative participants' semantic similarity judgments than did the clustering on unweighted syntactic features. This suggests that finer-grained information than is apparent from the initial clustering is present in the syntax.

There are two future directions for this work. First, given that we know there is information in principle available in the syntax about attitude verb semantics, we'd like to know if this is actually true of the input. Lederer et al. showed this was true for coarse-grained distinctions. We would like to know whether it holds for fine-grained ones as well. Second, we'd like to know whether this information is in fact usable in real verb learning—something an adaptation of Gillette et al.'s (1998) human simulation paradigm informed by the judgments gathered in our experiments could address.

References

- Bolinger, Dwight. 1968. Postposed main phrases: an english rule for the romance subjunctive. *Canadian Journal of Linguistics* 14:3–30.
- Brown, Roger W. 1957. Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology* 55:1.
- Fillmore, Charles J. 1970. The grammar of hitting and breaking. In *Readings in english transformational grammar*, ed. R.A. Jacobs and P.S. Rosenbaum, 1–33. Ginn.
- Fisher, Cynthia, Henry Gleitman, and Lila R. Gleitman. 1991. On the semantic content of subcategorization frames. *Cognitive psychology* 23:331–392.
- Gillette, Jane, Henry Gleitman, Lila Gleitman, and Anne Lederer. 1998. Human simulations of vocabulary learning. Technical report, University of Pennsylvania.
- Gleitman, Lila. 1990. The structural sources of verb meanings. *Language acquisition*

1:3–55.

- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic inquiry* 10:279–326.
- Hooper, Joan B. 1975. On assertive predicates. In *Syntax and semantics iv*, ed. John Kimball, 91–124. Academic Press.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. MIT press.
- Karttunen, Lauri. 1971. Some observations on factivity. *Research on Language & Social Interaction* 4:55–69.
- Lederer, Anne, Henry Gleitman, and Lila Gleitman. 1995. Verbs of a feather flock together: Semantic information in the structure of maternal speech. In *Beyond names for things: Young childrens acquisition of verbs.*, 277–297. Lawrence Erlbaum.
- Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*, volume 348. University of Chicago Press.
- Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. The MIT Press.
- Pons, Pascal, and Matthieu Latapy. 2005. Computing communities in large networks using random walks (long version). *arXiv preprint physics/0512106* .
- Reinhart, Tanya. 1983. Point of view in language—the use of parentheticals. In *Essays on deixis*, ed. Gisa Rauh, volume 188, 169–194. Gunter Narr Verlag.
- Urmson, James O. 1952. Parenthetical verbs. *Mind* 61:480–496.
- Zwicky, Arnold M. 1971. In a manner of speaking. *Linguistic Inquiry* 2:223–233.

Appendix

List of syntactic constructions

- | | | |
|---|--|---|
| 1. She _ to put some food in the fish bowl. | 12. Zeke _ them about the bottle. | 23. He _ him they put some food in the fish bowl. |
| 2. He _ there to be a truck at the curb. | 13. He _ they put some food in the fish bowl. | 24. What _ me most of all is that she fit the part. |
| 3. Fletch _ them the bottle. | 14. She _ him eat a sandwich. | 25. Kate was _ to fit the part. |
| 4. It _ her how they put some food in the fish bowl. | 15. He _ to. | 26. She _ about the cup. |
| 5. It _ her to put some food in the fish bowl. | 16. She _ how to put some food in the fish bowl. | 27. He _ it that they put some food in the fish bowl. |
| 6. What _ her most of all is to put some food in the fish bowl. | 17. He was _ that they put some food in the fish bowl. | 28. What he _ most of all was that they put some food in the fish bowl. |
| 7. He _ him that they put some food in the fish bowl. | 18. He _ the cup. | 29. He _ to him that they put some food in the fish bowl. |
| 8. She _ for him to put some food in the fish bowl. | 19. Hilary _ so. | 30. She fit the part, I _. |
| 9. He _. | 20. It _ her that they put some food in the fish bowl. | 31. She _ eating a sandwich. |
| 10. He _ if they put some food in the fish bowl. | 21. He _ him to put some food in the fish bowl. | 32. She fit the part, Fletch _. |
| 11. It _ her how to put some food in the fish bowl. | 22. She _ how they put some food in the fish bowl. | 33. What he _ most of all was to put some food in the fish bowl. |
| | | 34. Gary _ that she fit the part. |