Discovering Conceptual Structure Through Explicit and Implicit Cues in Child-Directed Speech

Anonymous CogSci submission

Abstract

[Write the abstract here].

Keywords: Conceptual learning, child-directed speech, language and cognition

Introduction

A hallmark of conceptual knowledge is its hierarchical organization. For example, a husky can be categorized as a dog, but it can also be categorized as a mammal, an animal, or a living being. Hierarchical organization is fundamental to human cognition as it allows, among other things, the generalization of knowledge through inference. For example, upon learning that all living beings are made out of cells, one can conclude that dogs are made of cells, too.

How do children acquire conceptual hierarchy? Early accounts considered conceptual hierarchy to be the consequence of the emergence of a domain-general logic of class-inclusion [explain a bit more] (Inhelder and Piaget 1964). However, researchers have noted that children can acquire hierarchy in a specific domain before mastering the domain-general logic of classes (Chi et al., 1989; Carey, 1985; Inagaki and Hanato,2002; Keil, 1981).

There are signs that children as young as 3 years old show hierarchical knowledge in various domains (e.g., animals, clothes, and food). Such signs include using superordinate words like "food" and "animal" according to parental report (Fenson et al., 1994), using different words to label the same object at different levels of conceptual hierarchy (Waxman and Hatch, 1992; Clark, 1997), and being able to extend the meaning of novel words to superordinate categories even controlling for perceptual similarity (Liu, Golinkoff, & Sak, 2001).

Some researchers proposed that children can learn conceptual hierarchy from cues in parental speech. For example, analyses of parent-child interactions have shown that parents rarely introduce words at the superordinate-level without also providing the basic level term (Blewitt, 1983; Shipley et al., 1983; Callanan, 1985). For example, parents rarely point to an object and say "this is an animal!". Instead, they usually anchor the super-ordinate word "animal" at the basic level by saying something along the lines of "This is a duck; a duck is a kind of animal." Such an anchoring strategy provides children with a categorization of the same object at different levels, which may help children understand the underlying hierarchical organization.

In a different line of research prompted by recent advances in data science (Landauer and Dumais, 1997; Mikolov et al., 2013), researchers have found that the statistical distribution of basic-level terms in parental speech can lead to a coherent structures at the super-ordinate level (Huebner and Willits, 2018; Fourtassi et al., 2019). To illustrate, one can learn that "horse" and "dog" are part of a higher-level category just by observing that these words co-occur in similar contexts. This cue can be a powerful source of conceptual hierarchy because it is based on pure co-occurrence and does not require the presence of a lexicalized label for the higher-level category; the latter emerges in a bottom-up fashion as a cluster of related words at the lower-level.

These cues could in principle be helpful for children. In the case of the "is-a-kind-of" anchoring, there is evidence that preschool children ably use this cue to interpret the meaning of a novel word at the super-ordinate level (Callanan, 1989). In the case of pure co-occurrence, extensive research in the last couple of decades has shown that children are capable of tracking distributional statistics of various linguistic units (Saffran et al., 1996). Besides, children appear to rely on the way words co-occur in speech to make conceptual generalizations (Fisher et al. 2011; Malten et al. 2015).

The current study

The cues reviewed above can be thought of as ends in a continuum that varies from explicit to implicit. The "is-a-kind-of" cue is the most explicit cue since both the terms (i.e., the basic and super-ordinate labels) and their hierarchical relationship are explicitly stated. The pure co-occurrence cue is the most implicit cue since, on the one hand, the super-ordinate term is not required and, on the other hand, the hierarchical relationship (that is, the fact that co-occurring basic level terms are part of a higher-level category) can only be induced.

While previous studies have focused on these extremes, other cues are available that have an intermediate status on this continuum. Here, we also study the way parents hint to the hierarchical relationship between two concepts pragmatically, i.e., without using an explicit inclusion expression. For example, instead of saying "a cow is a kind of animal" parents can say the following (e.g., in the context of a play session): "Do you want a cow or do you want another animal?" (see Table 1 for more examples). We also study a cue based on affordance. For example, food items can be cued by the verb "eat" and clothing items by the verb "wear".

Further, previous studies vary in terms of both the datasets and methods used, which has made their comparison difficult. Implicit cues have generally been studied using large-scale data and have been evaluated based on their ability to provide an accurate similarity space for words. In contrast, explicit cues have been studied mainly in the context of small-scale experiments and have been tested mainly through counting the frequency of a given linguistic expression (e.g., "X is a kind of Y").

In this work, we make a systematic comparison of explicit and implicit cues using similar methods. Such comparison is crucial as it allows us, for instance, to quantify the relative role that each cue could play in development. More precisely, we take a classification approach: We operationalize different cues as features that can be used to compute similarity. We then evaluate this continuous similarity measure by using it for a classification task, deciding whether different basic-level categories are part of the same superordinate category. Thus we can assign a classification accuracy to each cue type.

This paper is organized as follows. . .

Analyses

Data

We constructed a large-scale corpus by aggregating over all English-language transcripts from CHILDES (MacWhinney, 2000; Sanchez et al., 2018). These transcripts involved the caregivers' speech addressed to children up to three years of age. We had a total of XXX transcripts, across XXXX unique children, yielding a total of XXX utterances.

We decided to study the six following super-ordinate categories: "animal", "furniture", "clothes", "food", "toys" and "vehicles". For each of these categories, we used the corresponding basic-level terms available in the Child Developmental Inventory (CDI) containing the list of words that children can produce up to 3 years. These categories were chosen because they were the optimal set of superordinate categories that had been studied previously and had CDI data available. Most previous experimental work used only a subset of these categories for a given study (e.g., Callanan, 1985; Bornstein and Arterberry, 2010; Wright et al., 2015).

Cues to Conceptual Hierarchy and their Feature Vectors

As indicated above, we explored four cues to conceptual hierarchy: "is-a-kind-of", pragmatic, affordance-based, and pure co-occurrence. We represented each cue as a set of features and we tested how these features allow us to classify basic-level terms into super-ordinate categories. To this end, we started by representing each basic-level word in the CDI lexicon as a feature vector (thus, each basic-level word had a different feature vector for each cue). In the case where the cue relied on an explicit category marker (i.e., the first three cues), the feature vectors were based on the super-ordinate categories introduced above. Otherwise (e.g., the fourth cue), the feature vector was an embedding in a high dimensional space derived based on the words' pattern of co-occurrence only. In the following, we explain how we computed the feature vectors for each cue (see also Figure 1).

Is-a-kind-of This cue tests the extent to which parents use explicit expressions of class inclusion (Callanan, 1985). For each word at the basic label, X, we construct a feature vector of length six, where every cell corresponds to a super-ordinate category, Y, and the entry in each cell corresponds to the frequency with which X appears with Y is in one of the follow-

ing expressions: "X is a/an Y" and "X is a kind of Y" (we kept the same expressions used in previous studies).

Animals	Do you want a cow or do you want another animal?
Furniture	Furniture means sofa and chair and...
Clothes	This is another clothes. See, it's just like this shirt.
Food	She asks Lily what her favorite food is. If Lily says chocolate I am in trouble.
Toys	You close the book and get another toy because I think we are tired of this.
Vehicles	The only vehicle you cut out so far is the train.

Table 1: Examples of utterances from CHILDES where parents hint at a hierachical relations between a basic- and superordinate- level terms.

Pragmatic Parents can express conceptual hierarchy between X and Y without necessarily using an "is-a-kind-of" expression. In many cases, parents can hint at this hierarchy, in a pragmatic fashion, using a diversity of linguistic expressions (Table 1). To capture this diversity, we relax grammatical constraints between X to Y, and we keep only the requirement that X and Y should co-occur.

More concretely, we represent each basic-level term, X, with a feature vector where each entry represents the frequency with which X co-occurs with the corresponding super-ordinate term Y. This co-occurrence is determined using a fixed window of k utterances. Values of $k > 1$ allow us to capture the case where a relationship between X and Y is established across more than one utterance. For example:

– Mother : What kind of animal is this?
– Child : ??
– Mother : It's a giraffe!

Affordance-based The super-ordinate label is not the only category marker that can cue conceptual hierarchy for a basic level term, especially when this category can be characterized by an affordance. For example, "food" can be characterized as the category of things we eat and "clothes" as things we wear. Thus, children can learn that some concepts (e.g., "apple" and "bread") are parts of a higher-level category ("things we eat") by observing how these concepts co-vary with a cue of their common affordance (i.e., the verb "eat").

We computed the feature vectors for this cue as follows. In a first step, we tried to find a single verb that could be used as an affordance marker for an entire category. We used "eat" for food, "wear" for clothes, "play" for toys, and "ride" for vehicles. The category "furniture" has no such obvious function verb. We decided to use the verb "use" because if there were a verb that could fit every member of the category of furniture, it would be that (even though it can also fit things that are not members of the category). For the animal category, we could find no verb that could categorize the instances.
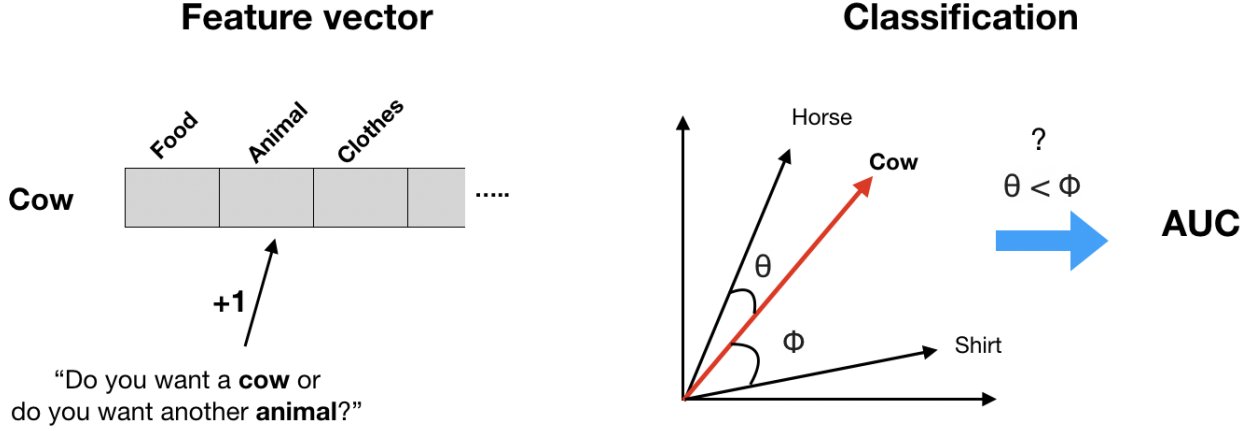
Feature vector

Classification

Figure 1: A schematic description of the task. For each basic-level word (here, 'cow') a feature vector is derived from child-directed speech based on how the cue is defined. Here, the vector cells correspond to the superordinate categories. The entry in a given cell (e.g., animal) is incremented when the word 'cow' co-occurs with the corresponding category label. The cue is evaluated based on its ability to classify pairs of words into 'same' or 'different' superordinate categories. Here, the pair 'cow'-'horse' belongs to the same category. The corresponding vectors should be closer to each other than the vectors of a pair that belongs to different categories (e.g., 'cow'-'shirt'). This evaluation is quantified by a standard measure in signal detection theory called the Area Under the ROC Curve (AUC).

We detected the concept-affordance relationship, syntactically, based on their occurrence in a verb-complement structure.[1] For example, in the utterance "the bird eats the berries", the word "berries" was categorized as "eat"-able. For each basic-level term, we computed a feature vector where entries correspond to the frequency with which this term occurs in a verb-complement relationship with the verb/affordance at hand.

Pure Co-occurrence Unlike the first three cues, the pure co-occurrence cue is not based on an explicit category marker at the super-ordinate level. It is based, instead, on the way basic-level terms are distributed together in speech (The distributional hypothesis; Harris 1957). Following previous research (Fourtassi et al., 2019), we quantified this cue using a word embedding model (Word2Vec, Mikolov, 2013). We used this model to represent basic-level words as vectors in a high-dimensional space, representing the distribution of these words in a latent semantic structure.

Task and Evaluation

Above, we characterized all cues in a vectorial framework. This framework allows us to directly compare the cues in terms of how they quantify the similarity between words (defined as the cosine of the angle formed by their vectors). Based on this similarity, we test the ability of each cue to predict which pairs of words belong to the same superordinate category (e.g., "apple" and "bread") and which pairs of words belong to different categories (e.g., "apple", "horse")

(Figure 1).

The task was as follows. First, we listed all pairs of basic-level words in the CDI dataset and their cosine similarity (according to each cue). Then, we evaluated the ability of the similarity measures to accurately predict whether the pairs belonged to "same" or "different" categories, across various discrimination thresholds. We quantified performance in the task using the standard Area Under the ROC curve (hereafter AUC, Fawcett, 2006). The AUC score can be interpreted as the probability that, given two pairs of words, of which one is from the same category, the pairs are correctly classified based on the similarity. For each cue, we derived both a global AUC score across all categories and a category-specific AUC score where we evaluated only the subset of pairs of words that contained at least an instance of a target category.

Results and Discussion

The "is-a-kind-of" cue is rare The instance of this cue were so rare that we could not even build feature vectors for basic-level words. In total, we found only four instances, all of them were characterizing the "animal" category. This finding contrasts with previous studies that found this cue in parental speech. This contrast can be explained by the fact that these previous studies were done in the context of a controlled experiment and parents were aware of the task (e.g., teaching words at the super-ordinate level), whereas here we tested a large-scale corpus containing a diversity of situations. Thus, it is possible that, in these previous studies, parents used a teaching strategy that they thought could optimize the short-term outcome (as determined by the experimenter),
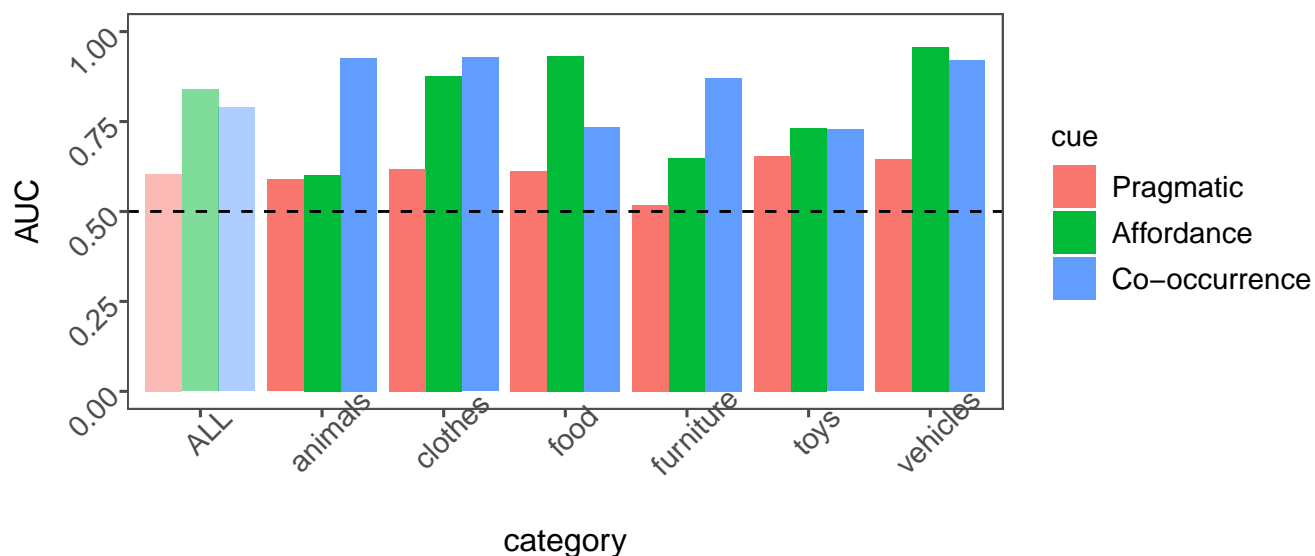
---

[1]There are more complex structures that could, in principle, be used by parents. We kept what appears to be the simplest one.

Figure 2: The Aurea Under the ROC Curve (AUC) scores for each cue across all categories ('ALL') and for each category. A value of 0.5 represents pure chance, and a value of 1 represents perfect performance. The AUC score can be interpreted as the probability that, given two pairs of basic-level words, of which one is from the same superordinate category, the pairs are correctly classified using their cue-based similarity.

rather than a strategy that reflects their spontaneous interaction with children in daily life

The pragmatic cue is noisy Figure 2 shows the global AUC score across categories as well as the AUC scores specific to each category. The accuracy of the pragmatic cue was generally low. The reason this cue performed so poorly is primarily due to the fact that we relaxed explicit grammatical constraints. While this operationalization allowed us to capture all possible ways the hierarchical relation between two concepts can be expressed linguistically, it also made the representation susceptible to errors, mainly by increasing the rate of false alarms: A basic level term (e.g., "Juice") can also co-occur with a super-ordinate label of which it is not an instance (e.g., "Don't poor the juice on your clothes!"). [Note: Increasing the size k of the sliding window (i.e., the number of adjacent utterances within which the basic- and superordinate-level terms should co-occur) only did not improve the performance (We need to show a graph here].

The affordance-based cue is more accurate but not universal The accuracy of this cue was relatively high. Nevertheless, it was only noticeable in the cases where we had an obvious verb to cue the affordance of the superordinate category, i.e., "food", "clothes", "vehicles", and "toys". The accuracy was low in the case of the "furniture" category since the verb "use" is not exclusive to this category and can also be used with instances of the other categories. This fact increased the rate of false alarms. The accuracy for the "animal" category was almost at chance as it was not characterized with affordance, though its instances occurred with verbs from other categories (e.g., "ride a horse", "play with the dog", and "eat the chicken").

The pure co-occurrence cue is the most reliable This cue is the most implicit but also the most powerful: The AUC score is generally high, including for the "animals" and "furniture" categories, which were not accurately captured with any of the previous cues. This finding means that for at least some categories, one can learn their common high-level categorization less through explicit linguistic markers, and more through the patterns of their usage. This strategy seems even more plausible for high-level categories that do not have an explicit label, or for which the label could not be available to young learners such as "animate" vs. "inanimate". Further should work should be done to study such cases.

The cues are stable across development The results we showed concern cues derived from parental speech up to 3 years old, as this is the age when signs of conceptual hierarchy start to emerge in the developmental literature. But we were also interested in how information in these cues may change as children grow older. Results of this analysis, presented in Figure 3, show that both all cues remain stable across development, at least up to 6 years old.

The cues provide non-redundant information Logistic regressions predicting the binary classification of pairs of basic-level words as belonging to same or different superordinate categories. The predictors are the pairs' similarity measures derived from each cue. We fit a different regression for each superordinate category.

We explored the extent to which explicit and implicit cues provided complementary vs. redundant information. To this end, we fit logistic regressions predicting the binary classification of pairs of basic-level words as belonging to same or different superordinate categories. The predictors are the

pairs' similarity measures derived from each cue. They were centered and scaled. The results of the regressions, summarized in table 2, indicate that, overall, each cue remains highly significant when controlling for the other cues, suggesting that the cues provide non-redundant information.
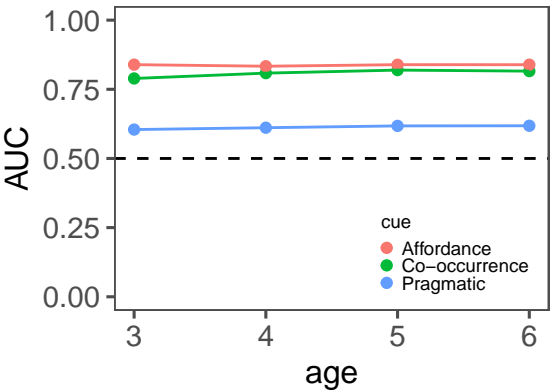

Figure 3: The Area Under the ROC Curve (AUC) scores for each cue (across all categories) using speech heard by children up to 6 years of age. A value of 0.5 represents pure chance, and a value of 1 represents perfect performance.

General Discussion

Table 2: Logistic regressions predicting the binary classification of pairs of basic-level words as belonging to same or different superordiante categories. The predictors are the pairs' similarity measures derived from each cue. We fit a different regression for each superodinate category.

	Animals	Furniture	Toys	Food	Clothing	Vehicles
(Intercept)	−2.741***	−3.195***	−3.244***	−2.616***	−3.101***	−4.663***
	(0.085)	(0.138)	(0.155)	(0.112)	(0.183)	(0.348)
Co-occurrence	2.285***	2.040***	1.178***	0.905***	1.644***	1.249***
	(0.074)	(0.127)	(0.136)	(0.060)	(0.171)	(0.193)
Affordance	0.022	0.547***	0.620***	2.112***	1.535***	2.211***
	(0.057)	(0.094)	(0.113)	(0.092)	(0.153)	(0.245)
Pragmatic	0.179***	−0.104	0.722***	0.325***	0.359*	0.159
	(0.050)	(0.080)	(0.120)	(0.059)	(0.146)	(0.138)

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$