Conceptual Hierarchy in Child-Directed Speech: Implicit Cues are More Reliable

Kyra WILSON[1], Michael C. FRANK[1], & Abdellah FOURTASSI[2]

[1] Department of Psychology, Stanford University

[2] Department of Computer Science, Aix-Marseille University

Author Note

'All data and analytic code are available at https://github.com/afourtassi/concepts'

'None of the authors have any financial interest or a conflict of interest regarding this work and this submission.'

Correspondence concerning this article should be addressed to Abdellah FOURTASSI, Postal address. E-mail: abellah.fourtassi@univ-amu.fr

Abstract

"In order for children to understand and reason about the world in an adult-like fashion, they need to learn that conceptual categories are organized in a hierarchical fashion (e.g., a dog is also an animal). The caregiver's linguistic input can play an important role in this learning, and previous studies have documented several cues in parental talk that can help children learn the conceptual hierarchy. However, these previous studies used different datasets and methods which made difficult the systematic comparison of these cues and the study of their relative contribution. Here, we use a large-scale corpus of child-directed speech and a classification-based evaluation method which allowed us to investigate, within the same framework, various cues that vary radically in terms of how explicit the information they offer is. We found the most explicit cues to be too sparse or too noisy to support robust learning. In contrast, the implicit cues offered, overall, a reliable source of information. Our work confirms the utility of caregiver talk for conveying conceptual information. It provides a stepping stone towards a cognitive model that would use this information in a principled way, leading to testable predictions about children's conceptual development."

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

Conceptual Hierarchy in Child-Directed Speech: Implicit Cues are More Reliable

Introduction

A hallmark of the lexicon is its hierarchical organization. For example, the same entity can have differnt labels with a nested structure, e.g., a husky, a dog, a mammal, and an animal.[1] Such hierarchical organization is fundamental to human cognition as it allows, among other things, the generalization of knowledge through inference. For example, upon children's learning that all animals are made of cells, they can conclude that dogs are also made of cells because the category "dog" is *included* in the category "animal" (e.g., Markman, 1989; Inhelder & Piaget, 2013; Sloutsky, 2010) (add Murphy 2004). The current study asks how this hierarchical kniwledge develop? More speicifically, we ask whether children can learn conceptual hierarchy from the language they hear around them.

Previous reseach has provided evidence that children as young as 3 years old show sings of hierarchical knowledge in their early lexicon. For example, observationl studies show that children use superordinate words like "food" and "animal" (e.g., Fenson et al., 1994) (add Mike book), and they use different labels to refer to the *same* object, shifting the coconceptual perspective from one level of abstracness to another (e.g., using the labels "dog" and "animal" to refer to a dog) depednding on the context (e.g., Clark, 1997). In more controlled in-lab experiments, preschool children are able, depending on the situation, to interpret the meaning of a novel word (e.g., "dax") either at the basic or at the superordiante level (Callanan 1989; Markman & Hutchinson, 1984). Critically, 3-year olds do not simply extend the meanings of words based only on how objects look, that is, young children do not simply use a novel word (e.g., "dax") to name both a car and a truck (but not, say, a banana) because the shape of a car is more similar to that of truck than to that of a banana. Rather, they do take into account taxonomic relations even controlling for

[1] This is not the only possible organization though, different domains within the lexicon may call upon different structures .

perceptual similarity (Liu, Golinkoff, & Sak, 2001).

How do children begin acquiring the hierarchical strucure of their lexicon? Early accounts considered this learing to be the consequence of the emergence of a *domain-general* logic of class inclusion – that one category can be part of a larger one – which develops relatively late, through middle childhood (Inhelder & Piaget, 2013) (Winer, 1980). However, and though the logic of classes may represent a mature, adult-like form of lexical organization, enabling reasoning on both famliar and non-familair domaines, the above-mentioned evidence suggests that children begin showing ability for hierarchal reasoning in *specific domains* they are familair with or interested in (e.g., the domaine of food or dinosaurs) much earlier. This fact suggests that the input they receive in specific domaines (whether through perceptual or linguistic means) plays an important role in shaping this organization (e.g., Chi, Hutchinson, & Robin, 1989; Carey, 1987; Inagaki & Hatano, 2002).

Research on the role of children's input in the development of a hierarchical lexicon can be summarized into a perceptual-based account and a language-based account (though these accounts are more complementary than mutually exclusive). In the perceptual-based account, children are understood to rely on their first-hand observation of the world, allowing them — using primarily their ability to pick up on perceptual simialrity — to from concepts (e.g., Madole & Oakes, 1999; McClelland & Rogers, 2003; Quinn & Eimas, 2000; Sloutsky, 2015, Smith and Heise (1992), 2010) (Saxe et al., 2019). Second, using their inductive reasoning skills, they synthesize these concepts into a nested tree structure via perfoming some sort hiereachical clustering whereby concepts that look similar (e.g., a dog and a cat) are placed nearby in the tree strucutre (Kemp et al., 2004, 2007, Tenebaum et al., 2006). Finally, childern are understood learn how to attribute mutliple labels to the same entity (e.g., dog, mamal, animal), whereby more abstract/broad labels map on to larger regions in the nested tree strucure (e.g., Xu and Tenenbaum, 2007, but see Spencer, Perone, Smith, & Samuelson, 2011 and Lewis and Frank, 2018).

The current study is, however, best situated within the framework of the language-based account, whereby children are understood to rely, not only on their own perceptual data, but *also* on knowledge transmission from other people, primarily via the means of language (Gelman, 2009; Harris, 2012). In particular, superordiante concepts (more than basic-level ones) do not necessarily share similar — sensory accessible — features, and their learning thus require additional, culture-specific input. For example, the categories "animal" and "plants" are organized for Groote Eylandt Australian aborigines rather into three categories: "biological", "food", and "totemic" (Waddy, 1982).

Regarding how children may learn conceotual hierachy from language, studies have shown that when parents introduce words at the superordinate level, they typically also provide the basic level term (Callanan, 1985; Shipley, Kuhn, & Madden, 1983). For example, parents rarely point to an object and say "this is an animal!" Instead, they usually *anchor* the superordinate 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 them understand the hierarchical organization.

Always within the language-based account, more recent research, prompted by advances in machine learning (e.g., Mikolov, Sutskever, Chen, Corrado, & Dean, 2013), found that the statistical distribution of basic-level terms in parental speech can lead to coherent structures at the superordinate level (Fourtassi, Scheinfeld, & Frank, 2019, Frermann and Lapata (2016); Huebner & Willits, 2018). For example, though "fish" and "bird" do not look very similar, people talk about them in similar linguistic contexts, typically leading to similar distributions in speech. Having a shared pattern of co-occurrence (i.e., a shared distribution) can cue taxonomic relationships (e.g., Sloutsky, Yim, Yao, & Dennis, 2017). Such distributional information can be a powerful source of conceptual knowledge because it does not require the presence of a lexicalized label for the higher-level category. On this kind of account, categories emerge in a bottom-up fashion as a cluster of related words.

These linguistic cues 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 labels (i.e., the basic and superordinate labels) and their hierarchical relationship are explicitly stated. The "distributional cue" is the most implicit cue since, on the one hand, the superordinate 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 the above-reviewed studies have focused on these extremes, other cues are available that have an intermediate status on this continuum. We introduce and test two such cues. The first, which we name the "pragmatic cue," is when parents hint at the hierarchical relationship between two concepts pragmatically, that is, without using an explicit inclusion expression. For example, instead of saying "a cow is a kind of animal" parents can say the following: "Do you want a cow or do you want another animal?" (see Table 1 for more examples drawn from real child-caregiver conversational data). Finally, we study an "Affordance-based cue" whereby action affordances provide another – more implicit – linguistic cue for category membership. For example, food items could be identified as members of a single superordinate category by virtue of their affordance of "eating" and clothing items by their their affordance of "wearing." (see Figure 1).

This study

Previous studies examining individual linguistic cues to conceptual hierachy in the lexicon have varied in terms of both the datasets and methods they used, which has made comparison difficult, hindering cumultative scientific progress on on this question. On the one hand, 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. On the otger hand, explicit cues have been studied mainly in the context of small-scale experiments and have been tested mainly by counting the frequency of a given linguistic expression (e.g., “X is a kind of Y”).

In this study, we make a systematic comparison of explicit and implicit cues using similar methods. In particular, we quantify the information provided by each of these cues in Figure 1 using spontaneous – rather than scripted – child-directed speech, using language data from naturalistic child-caregiver interactions (CHILDES 2018.1, MacWhinney, 2000; Sanchez et al., 2019). We test information containd in this input data across the entire normative lexcon as characterizd by the MacArthur-Bates Communicative Inventory (CDI) (Fenson et al., 1994). We focus on the hierecahical relations that the basic-level terms in this early lexicon bears on six superordiante labels: Animals, Furniture, Clothes, Food, Toys, and Vehicle. Our research question is to investigate the relative usefulness of each lingusitic cue in providing hierachical conceptual information that children can utlize for learning.

The paper is organized as follows. We begin by introducing our child-directed data and the derived linguistic cues for concetual hierachy. We then explain the task and evaluation methods that we use to study and compare these cues. Next, we present the results quantifying the role of each of the four cues and their relative informativeness by 3 three-years old (i.e., the age where early sings of conceptual hierachy appears, according to the above-reviewd developmental litteraure) as well as across development up to 6 years old. Finally, we discuss our findings in the light of the literature on early linguistic and conceptual development.

Analyses

Data

We constructed a large-scale corpus by aggregating all English-language transcripts from CHILDES (MacWhinney, 2000; Sanchez et al., 2019). We selected all utterances in which the speaker was not tagged as “Child” or “Target_Child” and which were addressed to children up to three years of age. The final corpus contained 1.9 million utterances and 7.9 million words from a collection of 4,939 transcripts across 660 children with an average age of 26.8 months.

We decided to study the six following superordinate categories: “animal”, “furniture”, “clothes”, “food”, “toys” and “vehicles”. For each of these categories, we used the corresponding basic-level terms available in the English-language MacArthur-Bates Communicative Inventory (CDI) (Fenson et al., 1994), a parent-report instrument that provides a partial listing and categorization of words produced by children 18–30 months. This set of categories was chosen because it combines all sets of superordinate categories that had been studied previously and for which CDI data were available. Indeed, most previous experimental work (which we partly reviewed above) used only a subset of these categories for a given study. Further, a Data-driven approach using CDI words and input from CHILDES (Fourtassi et al., 2019) led a set of categories made of “animal”, “food,” “clothes,” and a broad category of “artifacts” which included instances of toys, vehicles, and furniture. Here, we kept the last three categories differentiated as in previous experimental work.

Cues to Conceptual Hierarchy and their Feature Vectors

As indicated above, we explored four cues to conceptual hierarchy: “is-a-kind-of”, pragmatic, verb affordance-based, and distributional. We represented each cue as a set of features and we tested how these features allow us to classify basic-level terms into superordinate categories. To this end, we started by using each cue to derive a feature vector for each basic-level word in the CDI lexicon. 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 superordinate 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’ shared pattern of co-occurrence only. In the following, we explain how we computed the feature vectors for each cue (see also Figure 2).

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 6, where every cell corresponds to a superordinate category, Y, and the entry in each cell corresponds to the frequency with which X appears with Y is in one of the following expressions: "X is a/an Y" or "X is a kind/type/sort of Y" (we kept the same expressions used in previous studies).

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 using a wide diversity of linguistic expressions (Table 1). Detecting these expressions at scale is a challenge given their complexity, so as a first attempt 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 superordinate 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?

– Mother: It's a giraffe!

Affordance-based. The superordinate 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 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.

In addition to these verbs, we also identified synonymous verbs for each category that could also signify an affordance for the category. However, these verbs were either not found in the corpus (words like “devour” and “utilize” are not used in child-directed speech) or they occurred in too many contexts to be useful for categorization (“have” can be used synonymously with “eat”, but it has additional meanings that make this not a useable cue). Because of this, only one verb was used for computing the feature vectors for each of the categories.

We detected the concept-affordance relationship, syntactically, based on their occurrence in a verb-complement structure.[2] 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.

Distributional cue. Unlike the first three cues, the pure distributional cue is not based on an explicit category marker at the superordinate level. It is based, instead, on the way basic-level terms are distributed together in speech (Harris, 1957). Following previous research (Fourtassi et al., 2019), we quantified this cue using the word embedding model

[2] There are more complex structures that could, in principle, be used by parents. We used the simplest as a first approximation, though the performance of this cue could likely be enhanced by considering a wider variety of constructions.

Word2Vec (Mikolov et al., 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 (from the normative vocabulary) belong to the same superordinate category (e.g., “apple” and “bread”) and which pairs of words belong to different categories (e.g., “apple”, “horse”).

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 the full range of possible discrimination thresholds. We quantified performance in the task using the standard Area Under the ROC curve (hereafter AUC). 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 (see Figure 2).[3]

[3] A similar task and evaluation method have been used in previous work to evaluate cues to phonological categories in early development (Fourtassi, Dunbar, & Dupoux, 2014; Fourtassi & Dupoux, 2013).

Results

Individual Cue Results

Instances of our most explicit cue type, the *"is-a-kind-of" 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 characterizing the "animal" category. Thus, we did not have meaningful results to report for this cue. As for the other cues, Figure 3 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. For this cue, we only report the results with $k = 1$, which captures relations between basic and superordinate words within a single utterance. Increasing the value of k lead to worse, noisier performance. Regarding the *affordance-based cue*, the accuracy was relatively high for some categories, i.e., "food", "clothes", "vehicles", and "toys" and low for others, i.e., "furniture" and "animal." Finally, the *distributional cue* leads to the best overall results across most superordinate categories.

Developmental change?

We were also interested in how information in these cues may change as children grow older. For this analysis, we followed the same approach as above but included progressively more data in the corpus by adding utterances addressed to older children (Table 2). Results of this analysis, presented in Figure 4, show that the performance of all cues remained stable across development, at least up to 6 years old.

Cross-cue results

Here 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 were the pairs' similarity measures derived from each cue. The results of the regressions, summarized in Table 3, indicate that, overall, each cue remains highly significant when controlling for the other cues. Thus, although the distributional cue were highest performing when alone, each cue type provided *non-redundant* information;

Discussion

A crucial question in the study of cognitive development is understanding how children acquire the complex hierarchical relationships that characterize mature human conceptual knowledge. A particularly challenging task is learning how basic-level categories are grouped into abstract superordinate categories. The difficulty of this task stems from the fact that sensorimotor experience does not always provide direct evidence for learning. Here we build on and extend previous work investigating the role of *language* in learning about abstract concepts (Gelman, 2009; Harris, 2012). The novelty of our work is that we used a unified computational framework that has allowed us to directly compare the relative importance of different linguistic cues present in child-directed speech with respect to their ability to help categorize six common superordinate categories.

The most explicit cue — where caregivers state the relationship between a basic-level term and its superordinate category label (e.g., "a dog is a kind of animal") — did not scale up well to the naturalistic dataset we used. This finding contrasts with previous work that found this cue in parental speech (Callanan, 1985; Shipley et al., 1983). This contrast can be explained by the fact that these previous studies were done in the context of rather controlled settings and parents were aware of the task (e.g., teaching words at the superordinate level), whereas here we tested a large-scale corpus containing a diversity of situations.

Caregivers can hint at conceptual hierarchy without necessarily stating it explicitly, however (Table 1). To capture this pragmatic cue, we quantified the co-occurrence between basic-level and superordinate terms within utterances. While this simple operationalization was meant to capture all possible ways the hierarchical relationship 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 superordinate label of which it is not an instance (e.g., “Don’t pour the juice on your clothes!”). The rate of such false alarms was quite high, which explains to overall low — though not random — scores of this cue. From a developmental point of view, this finding highlights the limitation (at scale) of a simple co-occurrence-based strategy: A deeper understanding of the utterance is necessary if children are to learn conceptual hierarchy from pragmatic cues while avoiding false alarms.

Another linguistic cue was based on verb affordance (e.g., basic-level terms for food can all occur as the grammatical object of the verb “eat”). The accuracy of the affordance-based cue was relatively high for the categories which had an obvious verb to cue its affordance, i.e., “food” (eat), “clothes” (wear), “vehicles” (ride), and “toys” (play). The accuracy was low in the case of “furniture” since the verb “use” is not exclusive to this category and can also be used with instances of other categories, leading to false alarms. The accuracy for “animal” was also low as it was not characterized by any particular affordance verb. [4]

The distributional cue was the most implicit one since it did not require a label for the abstract category. The score for this cue was generally high, including for “animal” and “furniture”, two categories that were not accurately captured with any of the previous cues. This finding suggests that children can learn that certain basic-level terms share common abstract properties by realizing that they have a similar distribution in speech, i.e., that they are used in similar contexts. 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

[4] At the same time, the performance of the cue on this category was not totally random as animal instances tend to co-occur consistently with some verbs from other categories (e.g., "ride a horse", "play with the dog", and "eat the chicken").

the young learners (e.g., “animate” vs. “inanimate”).

An important finding was that the linguistic cues were largely non-redundant across the categories, suggesting that children can combine several cues to refine their knowledge about superordinate concepts. In addition, the cues did not fare similarly across categories, suggesting that children can rely more on different cues to learn different categories, e.g., they may use the affordance-based cue more to learn about “food”, “clothes” or “vehicle” and the distributional cue more to learn about “animals” and “furniture.”

For all the cues discussed above, our goal was to test their ability to provide a reliable source of information in child-directed speech. However, our instantiation of the cues abstracted away from the children’s cognitive and information processing limitations, providing only an “ideal observer” point of view (Marr, 1982). That said, previous experimental work has provided evidence for the cognitive plausibility of the learning mechanisms that underly each of these cues.

For instance, preschool children ably use the “is-a-kind-of” cue to interpret the meaning of a novel word at the superordinate level (Callanan, 1989) (though the current study shows this cue not to be pervasive in the natural input). Both the pragmatic cue and affordance-based cue rely on the ability to track co-occurrence between pairs of words. Extensive research in the last couple of decades has shown that even infants are capable of tracking distributional statistics of various linguistic units (Saffran, Aslin, & Newport, 1996). In particular, Bannard & Matthews (2008) have shown toddlers encode together in memory collocational words such as “sit” and “chair” and Wojcik & Saffran (2015) have shown that they encode relationships between novel words co-occurring in a sentence. Previous work also provided evidence showing that patterns of co-occurrence do influence children’s conceptual associations and inference (Fisher, Matlen, & Godwin, 2011).

The distributional cue requires not only sensitivity to co-occurrence but also sensitivity to their *shared* patterns of co-occurrence. For example, learners should be sensitive to the fact that “raven” and “salmon” co-occur with similar words and phrases such as “lay egg”, “live,” and “reproduce”, although the pair of words “raven” and “salmon” may not themselves co-occur with each other. There is evidence that even infants are sensitive to shared patterns of co-occurrence (e.g., Lany & Saffran, 2011), although the mechanism by which this sensitivity shapes conceptual knowledge emerges slowly and continues developing through middle childhood (see review by Unger & Fisher, 2021).

This work takes a first step towards integrating different cues to conceptual hierarchy from caregiver language. Future work should go further by also considering the role of sensorimotor data (Andrews, Frank, & Vigliocco, 2014; Bruni, Tran, & Baroni, 2014). Ultimately, the goal would be to integrate both sources into a *cognitive* model that uses these cues in a principle way. Such a model should make precise developmental predictions about how the cues interact and how their role in learning changes across development.

Another direction for future work is to compare the findings we obtained with data in English to other languages. Understanding variation in children’s experiences allows us to determine which aspects of development are universal and which are culture-specific (Rogoff, Dahl, & Callanan, 2018). The case of learning abstract, superordinate categories is an excellent case where such comparison could be very enlightening as we know such categories have both similarities and differences across cultures (e.g. Waddy, 1982). Thus, it is crucial that we investigate whether and how variation in caregivers’ language induces variations in children’s categorical structure of the world.

References

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*, *6*(3), 359–370.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241–248.

Bruni, E., Tran, N.-K., & Baroni, M. (2014). Multimodal distributional semantics. *Journal of Artificial Intelligence Research*, *49*, 1–47.

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, 508–523.

Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschoolers' hypotheses about word meanings. *Developmental Psychology*, *25*(2).

Carey, S. (1987). *Conceptual change in childhood.* MIT Press.

Chi, M. T., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly (1982-)*, 27–62.

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, *64*(1), 1–37.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i–185.

Fisher, A. V., Matlen, B. J., & Godwin, K. E. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition*, *118*(3).

Fourtassi, A., Dunbar, E., & Dupoux, E. (2014). Self-consistency as an inductive bias in early language acquisition. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 36).

Fourtassi, A., & Dupoux, E. (2013). A corpus-based evaluation method for distributional semantic models. In *51st annual meeting of the association for computational linguistics proceedings of the student research workshop* (pp. 165–171).

Fourtassi, A., Scheinfeld, I., & Frank, M. C. (2019). The development of abstract concepts in children's early lexical networks. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 129–133).

Frermann, L., & Lapata, M. (2016). Incremental bayesian category learning from natural language. *Cognitive Science*, *40*(6), 1333–1381.

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115–140.

Harris, P. L. (2012). *Trusting what you're told.* Harvard University Press.

Harris, Z. S. (1957). Co-occurrence and transformation in linguistic structure. *Language*, *33*(3).

Huebner, P. A., & Willits, J. A. (2018). Structured semantic knowledge can emerge automatically from predicting word sequences in child-directed speech. *Frontiers in Psychology*, *9*, 133.

Inagaki, K., & Hatano, Giyoo. (2002). *Young children's naive thinking about the biological world.* New York : Psychology Press.

Inhelder, B., & Piaget, J. (2013). *The early growth of logic in the child: Classification and seriation.* Routledge.

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science*, *14*(5), 1207–1219.

Liu, J., Golinkoff, R. M., & Sak, K. (2001). One cow does not an animal make: Young children can extend novel words at the superordinate level. *Child Development*, *72*(6), 1674–1694.

MacWhinney, B. (2000). *The childes project: Tools for analyzing talk. Transcription format and programs* (Vol. 1). Psychology Press.

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review*, *19*(2), 263–296.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Henry Holt; Co., Inc.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310–322.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).

Quinn, P. C., & Eimas, P. D. (2000). The emergence of category representations during infancy: Are separate perceptual and conceptual processes required? *Journal of Cognition and Development*, *1*(1), 55–61.

Rogoff, B., Dahl, A., & Callanan, M. (2018). The importance of understanding children's lived experience. *Developmental Review*, *50*, 5–15.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). Childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*(4), 1928–1941.

Shipley, E. F., Kuhn, I. F., & Madden, E. C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language*, *10*(3).

Sloutsky, V. (2015). Conceptual development. *Handbook of Child Psychology and Developmental Science*, 1–50.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*(7), 1244–1286.

Sloutsky, V. M., Yim, H., Yao, X., & Dennis, S. (2017). An associative account of the development of word learning. *Cognitive Psychology*, *97*, 1–30.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In *Advances in psychology* (Vol. 93, pp. 233–272).

Unger, L., & Fisher, A. V. (2021). The emergence of richly organized semantic knowledge from simple statistics: A synthetic review. *Developmental Review*, *60*, 100949.

Waddy, J. A. (1982). Biological classification from a groote eylandt aborigine's point of view. *Journal of Ethnobiology*, *2*(1), 63–77.

Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition*, *138*, 10–20.

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*, *6*(3), 359–370.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241–248.

Bruni, E., Tran, N.-K., & Baroni, M. (2014). Multimodal distributional semantics. *Journal of Artificial Intelligence Research*, *49*, 1–47.

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, 508–523.

Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschoolers' hypotheses about word meanings. *Developmental Psychology*, *25*(2).

Carey, S. (1987). *Conceptual change in childhood.* MIT Press.

Chi, M. T., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly (1982-)*, 27–62.

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, *64*(1), 1–37.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i–185.

Fisher, A. V., Matlen, B. J., & Godwin, K. E. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition*, *118*(3).

Fourtassi, A., Dunbar, E., & Dupoux, E. (2014). Self-consistency as an inductive bias in early language acquisition. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 36).

Fourtassi, A., & Dupoux, E. (2013). A corpus-based evaluation method for distributional semantic models. In *51st annual meeting of the association for computational linguistics proceedings of the student research workshop* (pp. 165–171).

Fourtassi, A., Scheinfeld, I., & Frank, M. C. (2019). The development of abstract concepts in children's early lexical networks. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 129–133).

Frermann, L., & Lapata, M. (2016). Incremental bayesian category learning from natural language. *Cognitive Science*, *40*(6), 1333–1381.

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115–140.

Harris, P. L. (2012). *Trusting what you're told*. Harvard University Press.

Harris, Z. S. (1957). Co-occurrence and transformation in linguistic structure. *Language*, *33*(3).

Huebner, P. A., & Willits, J. A. (2018). Structured semantic knowledge can emerge automatically from predicting word sequences in child-directed speech. *Frontiers in Psychology*, *9*, 133.

Inagaki, K., & Hatano, Giyoo. (2002). *Young children's naive thinking about the biological world*. New York : Psychology Press.

Inhelder, B., & Piaget, J. (2013). *The early growth of logic in the child: Classification and seriation*. Routledge.

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science*, *14*(5), 1207–1219.

Liu, J., Golinkoff, R. M., & Sak, K. (2001). One cow does not an animal make: Young children can extend novel words at the superordinate level. *Child Development*, *72*(6), 1674–1694.

MacWhinney, B. (2000). *The childes project: Tools for analyzing talk. Transcription format and programs* (Vol. 1). Psychology Press.

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review*, *19*(2), 263–296.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Henry Holt; Co., Inc.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310–322.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).

Quinn, P. C., & Eimas, P. D. (2000). The emergence of category representations during infancy: Are separate perceptual and conceptual processes required? *Journal of Cognition and Development*, *1*(1), 55–61.

Rogoff, B., Dahl, A., & Callanan, M. (2018). The importance of understanding children's lived experience. *Developmental Review*, *50*, 5–15.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). Childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*(4), 1928–1941.

Shipley, E. F., Kuhn, I. F., & Madden, E. C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language*, *10*(3).

Sloutsky, V. (2015). Conceptual development. *Handbook of Child Psychology and Developmental Science*, 1–50.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*(7), 1244–1286.

Sloutsky, V. M., Yim, H., Yao, X., & Dennis, S. (2017). An associative account of the development of word learning. *Cognitive Psychology*, *97*, 1–30.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In *Advances in psychology* (Vol. 93, pp. 233–272).

Unger, L., & Fisher, A. V. (2021). The emergence of richly organized semantic knowledge from simple statistics: A synthetic review. *Developmental Review*, *60*, 100949.

Waddy, J. A. (1982). Biological classification from a groote eylandt aborigine's point of view. *Journal of Ethnobiology*, *2*(1), 63–77.

Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition*, *138*, 10–20.

Table 1

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

Category	Utterance	Interlocutors	Corpus
Animals	Do you want a cow or do you want a different animal?	Mother to Max, 30 months	EllisWeismer
Furniture	Furniture means sofas and chairs and...	Mother to Naima, 23 months	Providence
Clothes	This is another clothes. See it's just like this shirt.	Investigator to Shem, 30 months	Clark
Food	She asked Lily what her favorite food was. If Lily says chocolate I am in trouble.	Mother about Lily, 24 months	Providence
Toys	You close the book and we'll get a different toy cause I think you're tired of this.	Mother to child, 13 months	Ambrose
Vehicles	The only vehicle you cut out so far is the train.	Mother to Warren, 30 months	Manchester

Table 2

Information about the corpora used in the analysis of developmental change.

Age	Age (months)	Children	Transcripts	Utterances	Words
<4	30.1	843	6,750	2.657 M	10.827 M
<5	33.5	971	7,889	3.093 M	12.807 M
<6	34.8	1,046	8,654	3.221 M	13.325 M

Table 3

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. The predictors were centered and scaled for comparability.

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

Note: *p<0.05; **p<0.01; ***p<0.001

Figure 1. The cues to conceptual hierarchy in the linguistic input can be understood as falling on a continuum from most explicit to most implicit.

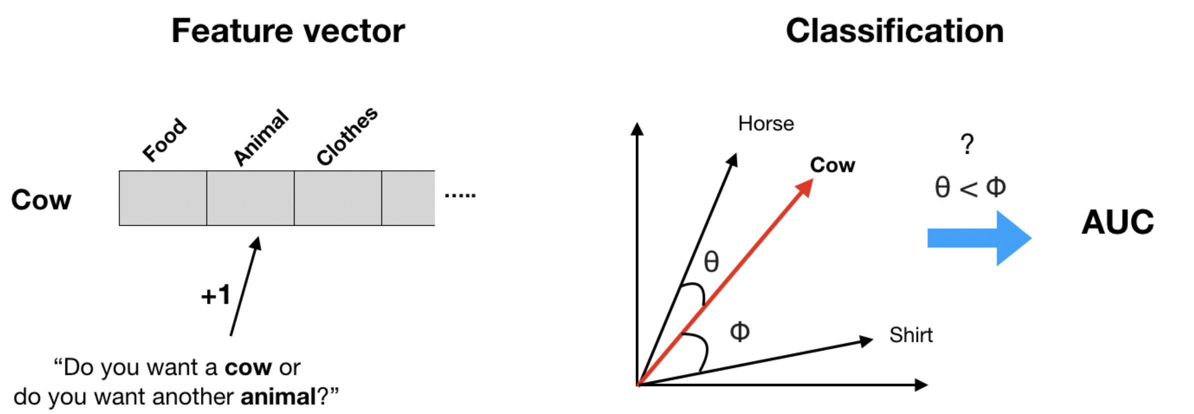

Figure 2. A schematic description of the task. For each basic-level word (here, 'cow') a feature vector is derived from child-directed speech based on how the cue is defined. Here, the vector cells correspond to the superordinate categories. The entry in a given cell (e.g., animal) is incremented when 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).

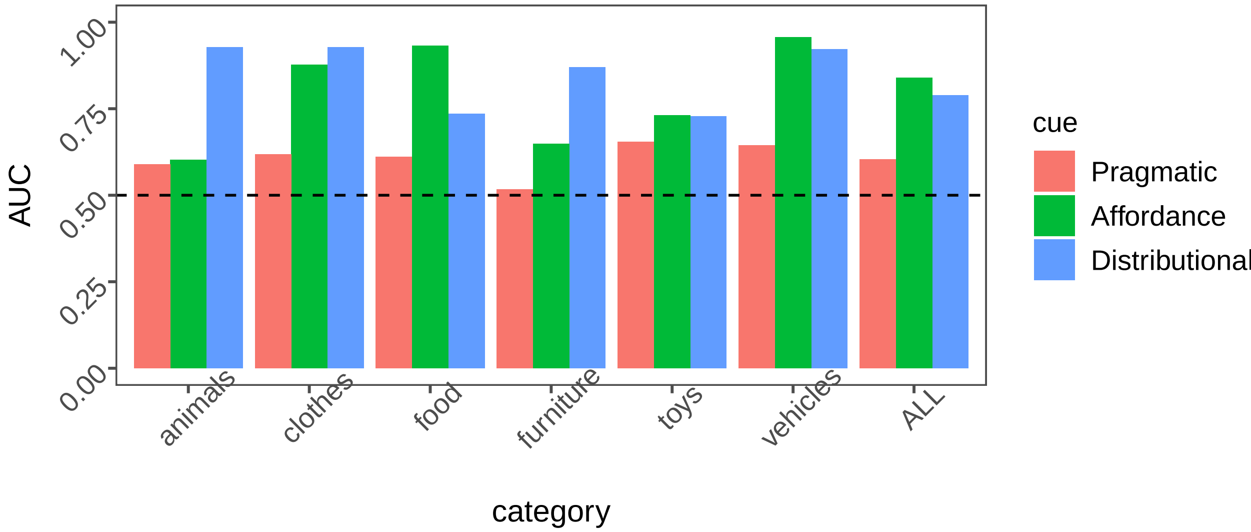

Figure 3. The Area Under the ROC Curve (AUC) scores of the cues for each category and across all categories ('ALL'). A value of 0.5 represents pure chance, and a value of 1 represents perfect performance. The AUC score can be interpreted as the probability that, given two pairs of basic-level words, of which one is from the same superordinate category, the pairs are correctly classified using their cue-based similarity.

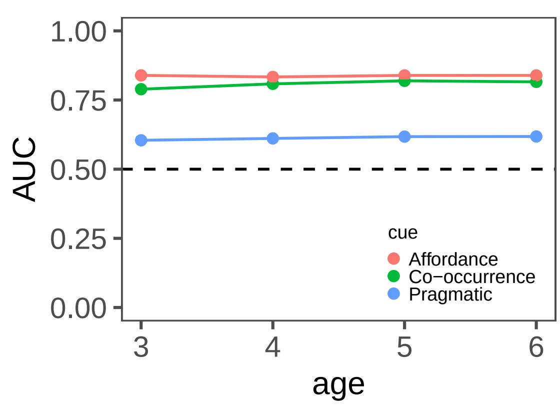

Figure 4. The Area Under the ROC Curve (AUC) scores for each cue (across all categories) using speech heard by children up to a particular age. A value of 0.5 represents pure chance, and a value of 1 represents perfect performance.