Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The acquisition of subordinate nouns as pragmatic inference

June Choe *, Anna Papafragou

Department of Linguistics, University of Pennsylvania, United States

ARTICLE INFO

Keywords:
Subordinate level meaning
Word learning
Semantic contrast
Pragmatics
Informativeness

ABSTRACT

Word learning is characterized by a bias for mapping meanings at the “basic” level (‘dog’), as opposed to a subordinate level (‘poodle’; Markman, 1986, 1990; Clark, 1987; Waxman et al., 1991, 1997). The fact that learners nevertheless acquire subordinate nouns has been attributed to properties of the referential world across multiple labelling events (e.g., Xu & Tanenbaum, 2007b; Spencer et al., 2011). Here we propose that the acquisition of subordinate-level meanings requires pragmatic reasoning that allows learners to take informative relevant alternatives into consideration. In support of this hypothesis, in a series of experiments we find that adult learners exploit information about semantic alternatives to generalize word meanings beyond the basic level. In Experiment 1, the introduction of a labelled alternative at the subordinate level eliminated the basic-level bias. In Experiment 2, this effect was found to be specific to labelled but not unlabeled alternatives. In Experiment 3, the availability of alternatives affected conjectures about subordinate-level word meanings even when these alternatives were presented well after the initial moment of ostensive labeling. Lastly, Experiment 4 replicated the semantic contrast effect using exclusively novel language input, highlighting the general communicative nature of these inferences. We conclude that the acquisition of subordinate nouns relies on pragmatic inferences about the informativity of labels as intentional linguistic-pragmatic acts, as opposed to simple word-to-world co-occurrences.

Introduction

Learning subordinate nouns

Word learning is a challenging task, in part because words do not just map onto objects and events that exist in the world but rather invoke specific meanings that the speaker intends to convey in the moment. A major aspect of word learning involves identifying the level of *specificity* encoded by word meanings. This is especially relevant for word meanings that enter into a subset-superset relationship, such as ‘dog’ vs. ‘poodle’: these meanings differ in specificity (‘poodle’ is more specific than ‘dog’) but are difficult to disambiguate using evidence from the referential world alone (Quine, 1960). Research suggests that young learners show a bias for mapping novel nouns to the so-called “basic”-level meaning (‘dog’), as opposed to the narrower, subordinate-level meaning (‘poodle’; Markman, 1984, 1990). The tendency to map novel words to basic-level meanings can be so strong that it even hinders the classification of a subordinate-level category when it is labelled, compared to when no labels are presented at all (Waxman, 1990). Despite these challenges, learners must eventually overcome the strong basic-level bias to acquire labels for subordinate-level categories.

How do learners manage to acquire subordinate nouns from their input over the course of language development? Early work highlighted the role of structured linguistic support that “anchors” the basic-level category as the relevant semantic domain, after which the subordinate-level distinction is explicitly introduced (e.g., “This is a dog. It is a *terrier*”; Shipley & Kuhn, 1983; Blewitt, 1983; Callanan, 1985). Waxman and colleagues (Waxman et al., 1991, 1997, among others) have also shown that this ability for basic-level categories to serve as an anchor for inferences about novel exemplars has direct implications for the acquisition of subordinate-level categories in experimental settings. In Waxman et al. (1997), preschool-age children were presented with an exemplar from a familiar basic-level category (‘dog’) paired with a novel label (“nooc”), accompanied by a short description. When the exemplar was described with an enduring and generalizable attribute which explicitly contrasted with that of the broader category (e.g., “Noocs help us pull sled, not help us find birds”), children overwhelmingly arrived at a subordinate-level meaning for the novel label. This inference was absent when the attribute was not contrastive or was merely incidental (e.g., “They just took a bubble bath”), suggesting that children can use

* Corresponding author at: Department of Linguistics, University of Pennsylvania, 3401-C Walnut St., Suite 300C, Philadelphia, PA 19104, United States.
E-mail address: yjchoe@sas.upenn.edu (J. Choe).

https://doi.org/10.1016/j.jml.2023.104432
Received 19 September 2022; Received in revised form 16 May 2023; Accepted 17 May 2023

semantic contrast to generalize to the subordinate-level when they are encouraged to focus on the relevant, defining distinctions.

Other studies have shown that children can integrate such cues with additional evidence from conventions in linguistic form, such that they are even more eager to assign a subordinate-level meaning to compound nouns than simple nouns when presented in the same anchoring context, exploiting the pattern that compound nouns are often used to express subordination (Gelman et al., 1989; Clark, 1992). More generally, the morphological characteristics of basic- and subordinate-level terms reveal how a language encodes contrast between subcategories within a conceptual domain (Cruse, 1986; Lyons, 1977), and this gives rise to crosslinguistic differences in how children produce and understand linguistic cues to subordination. For example, Clark and Berman (1984, 1987) find that both English- and Hebrew-speaking children as young as two years of age understand compound nouns as consisting of a head that selects the (basic level) category and a modifier that encodes a restrictive and contrastive information (cf. *Dalmatian-dog* in English). However, Hebrew-speaking children produce novel compounds at a later age due to the availability of other linguistic devices to express subordinate-level contrast such as affixation, which is also the preferred form for adult speakers of the language (Berman & Clark, 1989).

Even though carefully designed linguistic support can provide a helpful cue for subordination, others have argued that it is not necessary to rely on such richness of input in child-directed speech as the primary mechanism for the acquisition of subordinate nouns. Instead, it has been proposed that attending to the physical circumstances of a referent's presentation in the perceptual world (e.g., Spencer et al., 2011; Jenkins et al., 2015, 2021) or to the distribution of label-referent pairings in the input (e.g., Xu & Tenenbaum, 2007a, 2007b; Lewis & Frank, 2016, 2018) is sufficient for the learner to shift away from the basic-level hypothesis for a word's meaning, especially when the evidence accumulates across situations. For example, one class of accounts proposes that the acquisition of subordinate nouns can be captured largely via bottom-up perceptual and memory mechanisms, without needing to invoke rational processes (Spencer et al., 2011; Jenkins et al., 2015, 2021). Under this view, objects are construed at the basic or subordinate level via the style in which they are presented (i.e., circumstances of observation). For example, if three dalmatians are presented to the learner simultaneously with each labelled "fep," the close proximity of exemplars invites discrimination and enhances memory for fine-grained features that are shared (e.g., black spots); these happen to be the features that define a dalmatian, resulting in a subordinate-level interpretation of "fep" as meaning 'dalmatian.' However, if the same three dalmatians are presented sequentially across space and time, the most detailed features decay and learners are only able to maintain a coarse-grained category defined by the more "global" features (e.g., has a tail and four legs); these happen to be the features that define the basic-level category 'dog.' In this sense, it is the referential world which presents objects at the subordinate- and basic-level, and the task of the child is to discover the dalmatian-ness and dog-ness of exemplars from the world via a largely bottom-up strategy.

A second class of accounts building on Bayesian models of word learning (e.g., Frank et al., 2009) capitalize on the sheer number of examples for acquiring subordinate-level meanings through a mechanism of rational inference over sampling processes. This is formalized in the so-called *size principle* (Tenenbaum, 1999), which assigns the highest probability to the most restrictive hypothesis that is consistent with the data, which becomes increasingly more likely with more data. For a learner, a bigger category (like the basic-level category 'dog') is more likely to be used for any referential expression *a priori* due to its larger size in the conceptual space, so it is assigned a higher *prior* than a narrower category (like the subordinate-level category 'dalmatian') which is consistent with a smaller subset. But for the same reason, the *likelihood* of repeatedly observing exemplars consistent with a narrower subordinate-level meaning (e.g., 'dalmatian') is much greater when assuming the subordinate-level meaning rather than the basic-level meaning. Thus, when learners encounter multiple exemplars consistent with a subordinate-level meaning, they capitalize on the "suspicious coincidence" of that arrangement to infer that the word most likely means the subordinate-level category as opposed to the basic-level category (Xu & Tenenbaum, 2007b; Lewis & Frank, 2018). In this way, the model formalizes a kind of rational inference over a conceptual space of possible word meanings, given the data.

Despite their differences, both the bottom-up and the suspicious coincidence kinds of accounts treat exemplars with a target (subordinate) label as the fundamental unit of evidence for the learner: what matters for the acquisition of subordinate nouns is the number of such labelled exemplars, their distribution, presentation style and so on. However, this premise has recently been challenged by reports that the basic-level bias can be strongly modulated by the presence of other exemplars in the task. In a series of cross-situational word learning experiments, Wang and Trueswell (2019, 2022) found that adults and three-to-five-year-old children overwhelmingly generalized the meaning of a novel label to the basic-level exemplars ('dogs') even when the label exclusively co-occurred with exemplars consistent with a narrower subordinate-level meaning (e.g., 'dalmatians'). Instead, the crucial determinant for subordinate-level generalizations was the simultaneous learning of a second label when that label was paired with other members from the same basic-level category (e.g., non-dalmatian dogs). Critically, this effect disappeared when the second label was paired with members from a different basic-level category (e.g., birds), suggesting that learners generate task-specific inferences about which category levels are being highlighted in an ostensive labelling event, independently of the information about the distribution of the target label. In other words, there appears to be a strong contribution of other exemplars even in low-context, cross-situational word learning paradigms. Therefore, understanding the acquisition of subordinate nouns requires explanations beyond the number of exemplars, the spatiotemporal dynamics of the learning event, and the like.

The acquisition of subordinate nouns as a pragmatic puzzle

This paper begins with the assumption that learners make pragmatically driven inferences about the hypothesis space over which possible word meanings are proposed and evaluated (for a review of the evidence, see Grigoroglou & Papafragou, 2021). Within this context, we view the acquisition of subordinate nouns as posing a specific kind of pragmatic puzzle. Unlike accounts that frame the acquisition of subordinate-level nouns as a question of how various cues apparent in the physical world interact and converge on a specific concept, we ask under what discourse contexts learners *expect to hear* a label with a narrower meaning. We propose that the crucial task for the learner is to discover the intended level of informativity that is assumed in the labelling event (Grice, 1975; Clark, 2017). Under this framing, the meaningful difference between basic- and subordinate-level categories is not in the inherent size of the area that they carve up in the conceptual space. Rather, the distribution of basic- and subordinate-level labels is primarily governed by speaker intent, which makes it first and foremost a linguistic-pragmatic act.

When the situation presents many choices for labelling a referent, a label for a narrower category (e.g., the subordinate-level category) is typically the more informative one (Murphy & Brownell, 1985; Rosch, 1978). Yet basic-level meanings are by definition preferred across many contexts of use, perhaps because – other things being equal – they are informative enough to satisfy the needs of a generic addressee (Brown & Dell, 1987; Lockridge & Brennan, 2002; Grigoroglou & Papafragou, 2019a, 2019b). For such word meaning conjectures that enter into a subset-superset relationship, a subordinate meaning is only recognizable when the learner needs to consider a more informative (and crucially, *relevant*) level of describing a referent (versus the basic-level alternative). Thus, we predict that the acquisition of subordinate nouns should benefit from contexts that highlight the relevance of being more informative than usual. This is where contrast can help: the choice of a contrastive label can invoke a more restricted alternative set and indirectly point to otherwise less accessible subordinate-level categories and away from more accessible basic-level categories. Even though not typically studied in the context of pragmatic theory, both the classic finding that prior mention of the basic-level term anchors and promotes a subordinate alternative (Gelman et al., 1989; Waxman et al., 1991, 1997) and the more recent finding that the presence of semantic alternatives in cross-situational learning modulates the strength of the basic-level bias (Wang & Trueswell, 2019, 2022) can be understood as addressing this common task of identifying the relevant level of informativeness within a semantic domain with respect to which the utterance is to be interpreted.

The general idea that contrast facilitates conjectures at the subordinate-level is, of course, not new, given the known role of contrast in language acquisition (Markman, 1984, 1990; Clark, 1987, 1988, 1990). For example, Clark's (1987, p.2) Principle of Contrast states that "any difference in form in a language marks a difference in meaning," and has been proposed as the driving force allowing learners to discover new mappings between concept and form, for learning word meanings and beyond. The role of alternatives in facilitating inferences that are normally difficult to access has also been extensively explored in the scalar implicature literature. For example, we find a striking parallel to the study of young children's difficulty with generating the pragmatic interpretation of "some" as meaning 'some but not all' (Noveck, 2001; Papafragou & Musolino, 2003; among others). This difficulty has sometimes been argued to reflect children's limitations in processing (Guasti et al., 2005; Tieu et al., 2015), similar to how the task of learning subordinate-level categories has been treated by some in the literature on word learning (e.g., Ross & Murphy, 1996; Sloutsky et al., 2007; Sloutsky, 2010). For present purposes, we note that, when the stronger alternative "all" was introduced in a prior context, children were more successful in generating the pragmatic 'not-all' interpretation of "some" – but only when that alternative was relevant for the purpose of the conversation (Skordos & Papafragou, 2016; cf. Barner et al., 2011). Furthermore, children arrived at the 'not-all' interpretation of "some" even when previously exposed to "none" (instead of "all"; Skordos & Papafragou, 2016), further suggesting that establishing the relevance of the appropriate scale, not necessarily the accessibility of a particular stronger alternative, is key to computing scalar implicatures. In other words, despite the fact that "none" itself was not intended as an actual alternative of "some" in the scene, it nevertheless constrained children's search for the intended meaning of "some" by making quantification expressions salient alternatives within the discourse context. Relatedly, older children have been shown to even integrate evidence about the epistemic state of a speaker to make such inferences, by reasoning about the informational strength of a statement given a speaker's possibly limited knowledge state (Kampa & Papafragou, 2020; Papafragou, Friedberg, & Cohen, 2018).

In the present study, we explicitly set out to investigate the acquisition of novel subordinate nouns as a pragmatic task. Our goal is to go beyond previous cross-situational word learning research on the acquisition of subordinate nouns, where the role of pragmatic inference was not explored systematically. To do so, we devise new tasks to investigate the role, limits, and potency of semantic contrast in adult learners' basic- vs. subordinate-level generalization of novel words from single learning trials. We posit that a communicative act gives evidence for the meaning of a word, because it affords the learner some insights about the informativity encoded in the word (depending on whether a speaker's intent is to be more specific as reflected in their choice of a subordinate-level noun, as opposed to the basic-level lexical alternative).

In four word learning experiments, we raise and test three interlocking predictions of the above pragmatic position. First, we predict that the presence of a semantic alternative at the subordinate-level should facilitate subordinate-level conjectures for a target label (assuming that the presence of the alternative makes it clear that the more informative, subordinate-level categories are relevant to the task). Accordingly, Experiment 1 asks whether the rate of basic-level generalizations for an ostensive target label (e.g., "mipen" paired with a red apple) decreases if the target is immediately followed by a labelled semantic alternative at the subordinate level (e.g., "kalmick" paired with a green apple). Second, we hypothesize that this effect of contrast should be primarily linguistic, as opposed to merely conceptual (see also Clark, 1987, 1988, 1990). Accordingly, Experiment 2 asks whether the effect of contrast is stronger when the alternative is labelled rather than simply present but unlabelled. Third, we predict that these inferences about contrast and alternative sets need not be limited to the introduction of labelled referents in ostensive contexts; a word meaning conjecture can be revised and updated outside of the "official" point of referent introduction, if the learner discovers the speaker's intent to be more specific in the use of the word. Thus, Experiment 3 explores whether the use of a word beyond ostensive labeling statements can be construed as additional evidence for (the specificity of) its meaning. Lastly, Experiment 4 replicates the semantic contrast effect under input from an unfamiliar language, which more closely mimics the natural circumstances of early word learning. In concluding, we consider alternative interpretations of our findings and discuss the implications of our data for word learning more broadly.

Experiment 1

Participants

Fifty-three English-speaking adults participated in Experiment 1. Sample size was informed by prior experiments measuring the basic-level bias using the grid paradigm (around 50 in Spencer et al., 2011, Lewis & Frank, 2018, and Wang & Trueswell, 2022). Participants were recruited from the undergraduate subject pool at the University of Pennsylvania (n = 26) and on Prolific (n = 27), a platform for online subject recruitment.[1] The experiment was hosted and conducted online via PCIbex (Zehr and Schwarz, 2018).

Materials and procedure

To study learners' generalization of word meanings, we adopted the test-grid paradigm (also called the Immediate Generalization Paradigm; Caplan, 2022), a standard experimental design in the literature on the acquisition of subordinate nouns (Xu & Tenenbaum, 2007b; Spencer et al., 2011; Jenkins et al., 2015; Lewis & Frank, 2016, 2018; Wang & Trueswell, 2019; among others). In this task, learners are given an unambiguous referent-label mapping as an opportunity to learn a target label, after which they are presented with an array of images (presented simultaneously, often in a grid layout) and asked to select all images that match the label. The makeup of the set of images used at test is such that the previously observed referent is simultaneously a member of multiple hierarchical categories (often at the subordinate, basic, and superordinate levels). Thus, the choice of selection at test can reveal which specific meaning the learner generalized the meaning of the novel label to.

In Experiment 1, we used the test-grid paradigm to probe word meaning generalizations from a single exposure to a label. There were ten trials: eight critical trials and two catch trials which tested for attention and color vision deficiency. The eight critical trials involved novel words from eight superordinate-level domains (e.g., fruits, animals, vehicles, etc.), and all words were taught once using a single image of the referent.

Each trial was divided into the *learning phase* and the *testing phase*. Images from eight distinct "semantic domains" (i.e., superordinate-level categories) corresponding to each critical trial were prepared. The domains were balanced in the number of natural (e.g., fruits) and artifact

[1] Later comparisons showed that data from the two groups were identical.

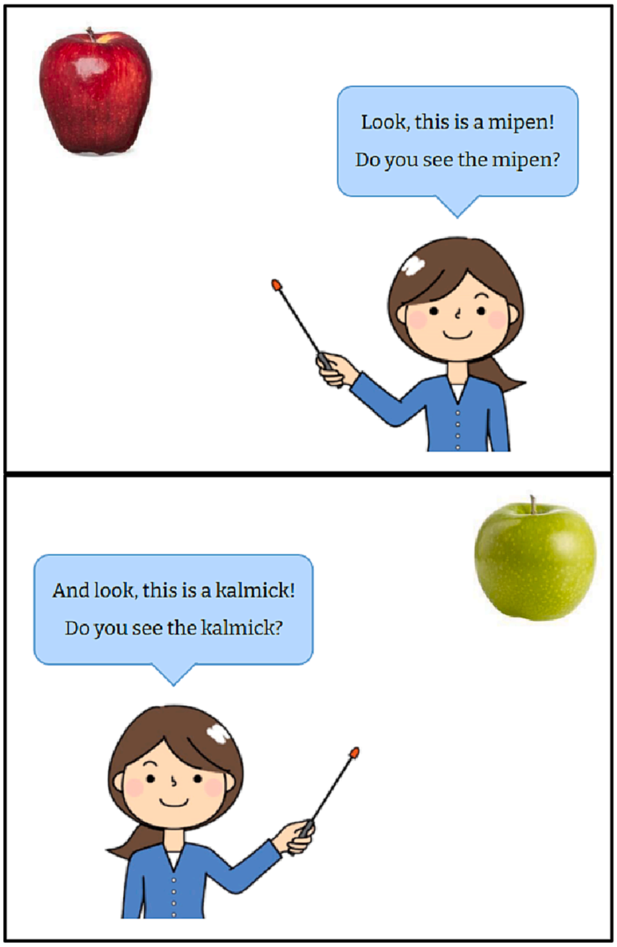

Fig. 1. Presentation of the target (red apple, "mipen", top panel) and the alternative (green apple, "kalmick", bottom panel) in the learning phase of the *Contrast* condition in Experiment 1. In the *No Contrast* condition, only the target was shown and labelled, as in the top panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(e.g., cars) kinds. For each domain, a total of fourteen images of single exemplars were created. Twelve of these images (all set against a naturally occurring background) were used in the testing phase: two were from the target subordinate category (e.g., red apples), two from an alternative subordinate category (e.g., green apples), three other members from the same basic category (e.g., other apples), and five members from other basic-level categories within the same semantic domain (e.g., non-apple fruits). The remaining two images – one additional member from each of the two subordinate-level categories – showed the exemplar without a background and were used in the learning phase.

At the beginning of the experiment, a cartoon character, Sally, appeared on the center of the screen and introduced herself as a foreign language speaker. Before proceeding to the trials, Sally told participants that they would be learning words from her native language and that they should pay attention as she would ask questions about these words later. All communication from Sally was delivered in written form, appearing in a speech bubble.

The learning phase began with Sally on the screen by herself for one second, after which objects appeared at Sally's sides one-by-one. When the learning phase introduced two objects, one appeared to the left and then the other appeared to the right of Sally. When the learning phase only introduced one object, it appeared once to the left of Sally. Only one object was visible at any given time and each object stayed on screen for seven seconds while Sally labelled the object.

We manipulated the presence of a semantic alternative in two conditions in a within-participant design. In the *No Contrast* condition, only the target was shown and labelled (Fig. 1a). In the *Contrast* condition, the target was followed by a semantic alternative at the subordinate-level with a different label (Fig. 1a-b). In effect, participants learned one novel word (the target) in *No Contrast* trials and two novel words (the target and alternative) in *Contrast* trials. After all object(s) had been presented, the learning phase concluded with Sally returning back to the upright position for one second before the trial moved on to the testing phase.

In the testing phase of each trial, eighteen images were presented in a 3-by-6 grid (Fig. 2). Each image was placed inside a 150-pixel square cell with 15 pixels of row and column gaps. The grid included two matches to the target subordinate category, two matches to the alternative subordinate category, three other matches to the basic category, three matches to the superordinate category, and eight non-matches (i.e., members of other superordinate categories). The images were laid out in the grid in randomized order between design groups. Fig. 2 shows an image grid from the test phase for the fruit domain, where participants

Fig. 2. Images shown in the testing phase for a trial involving the fruit domain. Selections representing the basic-level generalization (all apples) are marked with a blue outer border and selections representing the target subordinate-level generalization (red apples) are marked with an additional inner yellow border. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

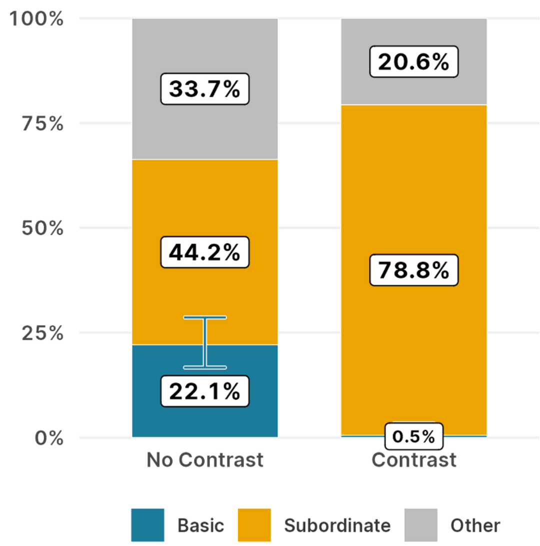

Fig. 3. Responses at test for Experiment 1.

are asked to find matches to the label paired with a red apple ("mipen") in the learning phase.

Sally instructed the participants to select all matches for one of the novel labels (i.e., the target label) introduced in the learning phase. For example, if the target label was "mipen" paired with a red apple in the learning phase, Sally asked participants: "Do you see any other mipens below? Click on all mipens you see!". Participants interacted with the image grid by clicking one image at a time. The testing phase was not timed, and participants could freely select and unselect any number of images as many times as they wished, as long as at least one image was selected before proceeding. The final set of selections as well as the target and timestamp of individual clicks made in the testing phase were recorded for analysis.

Coding

For each trial-level response, the final set of selections at test was coded into one of three categories: Subordinate, Basic, and Other. A response was coded as Subordinate ("mipen" means red apple) if it reflected a narrow generalization to only the subordinate-level category (e.g., both the red apples and no other images selected). Responses were coded as Basic ("mipen" means apple) if they included both subordinate target exemplars as well as all other members from the basic-level category (e.g., all apples, including both the red apples and the five other kinds of apples). Consistent with prior work, we applied a strict coding criterion such that partial selections of members from the subordinate- and basic-level categories were excluded from the count for Subordinate and Basic responses.[2] Instead, these were coded as Other, a catchall category consisting of responses where participants selected at least one of the subordinate-level exemplars but otherwise did not exhibit the intent to generalize to a particular noun category (such as selecting all red objects or all apples except the green apple alternative; for details on Other responses for all experiments, see Appendix A). The coding scheme also included Superordinate responses (e.g., all fruits), but none were observed in any of our experiments. Lastly, selections which included exemplars from other semantic domains (e.g., a planet) or failed to include any of the target subordinate-level exemplars (e.g., included no red apples) were judged to reflect inattentiveness and were excluded from analysis.

Table 1
Mixed-effects logistic regression model fitted to Basic responses in Experiment 1.

	β (SE)	t	p
(Intercept)	−4.04 (0.8)	5.4	<0.0001
Contrast	−2.26 (0.5)	−4.3	<0.0001

Results

Data from three participants who answered incorrectly on the two catch trials were discarded. Additionally, 12 trial-level responses were filtered out according to the exclusion criteria for coding. In total, 388 responses from the remaining 50 participants were entered into the analysis. The distribution of these responses used for analysis is shown in Fig. 3. For the purposes of our analysis, we use the strict coding scheme where only exhaustive selections count for Basic and Subordinate responses – i.e., participants select all seven basic-level exemplars (a Basic response) or both the subordinate-level exemplars (a Subordinate response). This follows the practice of prior studies using the same test grid paradigm and gives a more conservative estimate of the effect. For completeness, Appendix A1 shows the distribution of responses with a fine-grained breakdown of the Other category.

Following previous experimental work on the acquisition of subordinate nouns (e.g., Lewis & Frank, 2016, 2018; Wang & Trueswell, 2019, 2022), we first conducted an analysis of Basic responses. We fitted a mixed-effects logistic regression model to the rate of Basic responses using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages in R (R Core Team 2021). The model summary is shown in Table 1.[3] We found a significant difference in the proportion of Basic responses between the *Contrast* condition and the *No Contrast* condition (β = -2.26, SE = 0.5, p < 0.0001; Table 1), with Basic responses decreasing from 22% to 0.5% when a semantic alternative with a different label was present. A subsequent analysis confirmed that this difference was also reflected in the same model fitted to the proportion of Subordinate responses, where Condition had a similar effect in the opposite direction (β = 0.84, SE = 0.13, p < 0.0001), increasing Subordinate responses from 44% in the *No Contrast* condition to 79% in the *Contrast* condition.

Discussion

We proposed that a crucial task for the acquisition of subordinate nouns is to discover the intended level of informativity or specificity assumed of the labelling event. Consistent with our prediction that a semantic alternative at the subordinate-level makes the less-accessible subordinate-level distinction relevant to the conversation, we find that the presence of the alternative facilitates subordinate-level generalizations, suggesting that learners expected a greater degree of specificity encoded in the target label. In other words, reasoning about the role of the semantic alternative in the choice of the target label makes accessible the vertical contrast between the subordinate- and basic-level meanings for the target label and highlights its relevance in the learning context. Crucially, we achieve this effect from a single exposure to the target noun without explicitly grounding (or "anchoring") the space of possible meanings within the basic-level semantic domain (e.g., Waxman et al., 1991, 1997) nor exposing participants to multiple labelled referents for the target word (e.g., Wang & Trueswell, 2019; 2022). This effect cannot be accounted for by either bottom-up or "suspicious-coincidence" models of subordinate word learning, since the semantic alternative does not contribute to either the perceptual (e.g., Spencer et al., 2011; Jenkins et al., 2015) nor the distributional (e.g., Xu & Tenenbaum, 2007b; Lewis & Frank, 2016) profile of the target label

[2] Throughout our experiments, even if we were to include these more lax coding categories into the analyses, our results would not change.

[3] Model formula: Basic ~ Contrast + (1 | Participant) + (1 | Item). Contrast was sum coded with *Contrast* at 1 and *No Contrast* at −1.

itself.

Additionally, we observed two somewhat unexpected patterns. First, the rate of Basic responses was low, both on the absolute scale and also relative to the rate of Subordinate responses. Empirically, the magnitude of the basic-level bias has been reported to vary widely (see discussion in Jenkins et al., 2015), putting 22% on the low end but still comparable to that of other studies.[4] In Experiment 1, the pattern may be driven by the high threshold for a selection to be coded as a Basic response in our design: both of the target subordinate-level exemplars (e.g., dalmatians) plus all five other members from the basic-level category (e.g., other dogs) had to be selected for such a response as opposed to two or three basic-level exemplars in prior work (Xu & Tenenbaum 2007b; Spencer et al., 2011; Lewis & Frank, 2018; Wang & Trueswell, 2022). In other words, there is simply greater cognitive effort required to search and click on a larger set of images. Second, and relatedly, the rate of Other responses was relatively high at 20%-30% of all responses, with an internal make-up that differed substantially between conditions (see Appendix A1). Of the 67 Other responses in the *No Contrast* condition, two-thirds were incomplete subsets of the Basic response (i.e., both subordinate targets and some but not all of the other basic-level exemplars), partially explaining the overall low rate of Basic responses.[5] Interestingly, these responses were entirely absent among the 39 Other responses in the *Contrast* condition. Here, over half were the so-called "mutually exclusive" responses (Gelman et al., 1989; see Appendix A1), which included all basic-level members except those from the alternative subordinate-level category presented during the learning phase of *Contrast* trials (e.g., all apples except for the two green apples).[6] We revisit these points in later experiments.

Experiment 2

Experiment 1 offered strong evidence for the role of lexical contrast in promoting subordinate-level conjectures. However, the precise role of contrast is currently open to two interpretations. One possibility is that this effect is limited to (or more strongly connects to) linguistic contrast. This is consistent with our pragmatic account: unlabeled exemplars should not invoke alternative word meanings for the target label, since that they do not constitute evidence for intent underlying the specific choice of the target label (cf. also Clark, 1987). A second possibility is that the mere presence of an alternative, subordinate-level referent suffices for making subordinate-level meanings accessible, for example by offering a perceptual contrast that highlights subordinate-level distinctions. This possibility is reminiscent of the position that the arrangement of objects in the world (along with other bottom-up cues) has consequences for the generalization of word meaning (Spencer et al., 2011). Under such accounts, the contrast between exemplars from different subordinate-level categories may be sufficient to trigger low-level mechanisms that highlight the kinds of features that define subordinate-level categories. Experiment 2 was similar to Experiment 1 but sought to adjudicate between these possibilities by disentangling the effect of labelling from the mere presence of the alternative referent.

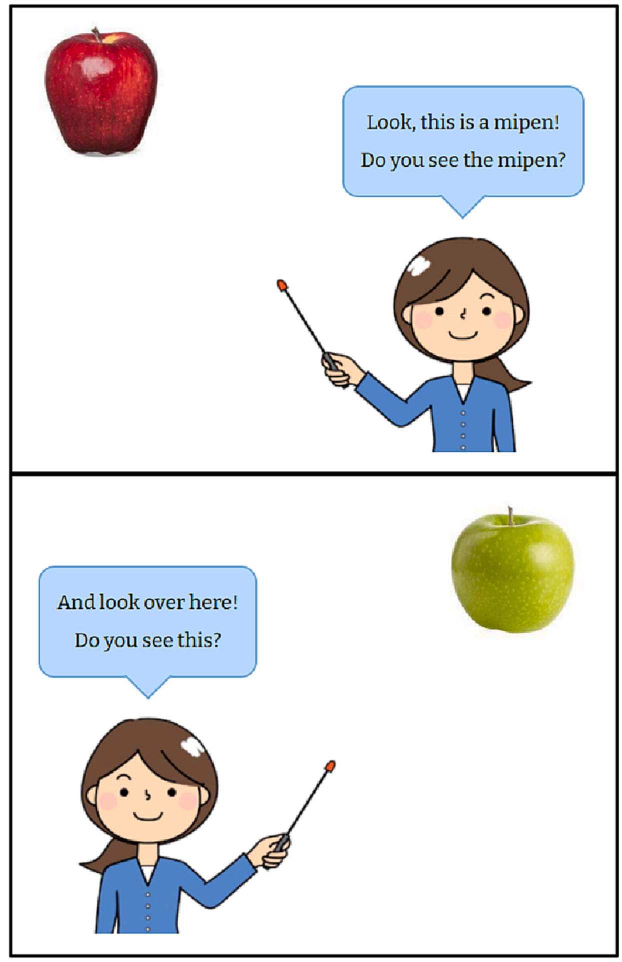

Fig. 4. Presentation of the target (top panel) and the alternative (bottom panel) in the *Unlabelled Alternative* condition of Experiment 2. The *Labelled Alternative* condition is the same as Fig. 1.

Participants

Ninety English-speaking adults who did not previously participate in Experiment 1 were recruited on Prolific. Sample size was approximately doubled from Experiment 1 to test the statistical significance of the two main effects (but not the interaction effect) in the more complex 2-by-2 design.

Materials and procedure

Participants were exposed to two conditions in a within-subject design (Fig. 4). The *Labelled Alternative* condition replicated the *Contrast* condition of Experiment 1. In the *Unlabelled Alternative* condition, the alternative was present but not labelled: Sally simply drew attention to the object by saying, "(And) look over here! Do you see this?" To guard against possible presentation effects, the order in which the target referent appeared relative to the alternative referent was

[4] For example, Jenkins et al. (2015) reports the magnitude of the Basic-level bias to be 26% in their "single exemplar" trials, which is equivalent to the *No Contrast* condition of Experiment 1. For more comparison, Spencer et al. (2011) reports the magnitude of the bias ranging from 30% to 50% while the estimate from Xu & Tenenbaum (2007b) is higher at 76% in one of the experiments.

[5] Under more liberal coding schemes which account for partial selections of the basic-level set (e.g., Lewis & Frank, 2018), the proportion of Basic response in the *No Contrast* condition of Experiment 1 increases to approximately 40%, closer to the empirical average for the magnitude of the basic-level bias measured using this paradigm. See Appendix A1.

[6] While both "incomplete Basic" and "mutually exclusive" responses reflect a failure to generalize to the basic-level category, we make this distinction because the source of the error differs. The "incomplete Basic" responses are likely to be driven by the failure to identify the appropriate basic-level category to generalize to. In the case of "mutually exclusive" responses, however, the basic-level category is correctly identified but participants nevertheless pursue a narrower interpretation which excludes the semantic alternative. Notably, none of the "incomplete Basic" selections in the *No Contrast* condition pattern like the "mutually exclusive" selections in the *Contrast* condition (e.g., no participants selected all apples except for the two green apples after seeing a single red apple labelled in the learning phase).

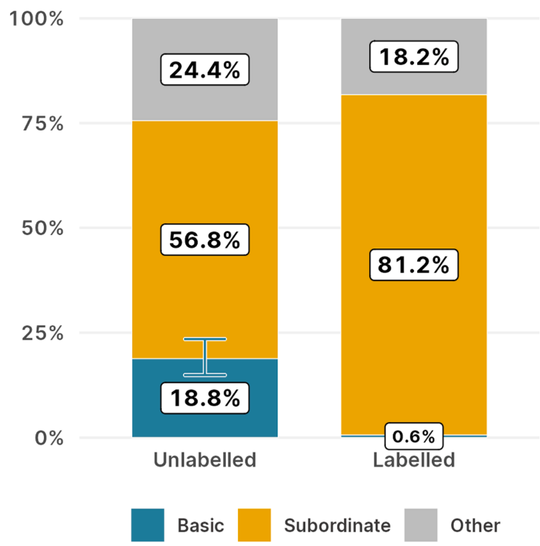

Fig. 5. Responses at test for Experiment 2.

Table 2
Mixed-effects logistic regression model fitted to Basic responses in Experiment 2.

	β (SE)	t	p
(Intercept)	−7.6 (1.7)	−4.5	<0.0001
Label	−2.7 (0.5)	−5.3	<0.0001
Order	2.5 (0.9)	2.9	0.0033

counterbalanced, such that the target was introduced before the alternative in half of the trials (*Target First*) and the target was introduced after the alternative in the other half of the trials (*Target Second*), resulting in a 2-by-2 design.

Coding

The coding scheme followed that of Experiment 1.

Results

After applying the same filtering criteria as in Experiment 1, 4 participants were excluded and 19 trial-level responses from the rest were removed from the analysis. A total of 669 responses from 86 participants entered the analysis. Results are shown in Fig. 5. (See also Appendix A2 for the distribution of responses with a fine-grained breakdown of the Other category.).

We first fitted a mixed-effects logistic regression model to the rate of Basic responses. The model summary is shown in Table 2.[7] We found a significant main effect of Label ($\beta = -2.7$, $SE = 0.5$, $p < 0.0001$) such that there was an overall lower rate of Basic responses when the alternative was labelled compared to when it was simply present and unlabelled. There was also a main effect of the Order nuisance variable ($\beta = 2.5$, $SE = 0.9$, $p < 0.01$), indicating that the basic-level interpretation of the target label was more likely when the target was shown first in the learning phase, before the alternative was introduced (24.1% vs. 14.7%). Similar to Experiment 1, this difference in Basic responses between the conditions was straightforwardly reflected in the rate of Subordinate responses, which increased from 57% to 81%. In the same model fitted to Subordinate responses,[8] the effect of Label was again significant ($\beta = 2.3$, $SE = 0.42$, $p < 0.0001$).[9]

Discussion

Experiment 2 investigated the nature of the semantic contrast effect that constrains the generalization of word meanings. We found that the presence of a linguistically marked alternative at the subordinate-level category was more likely to facilitate subordinate-level generalizations of the target label compared to the presence of a mere conceptual alternative (*Unlabelled Alternative*). This finding is consistent with the proposal that *linguistic* – and not mere conceptual - contrast drives informativity calculations (see also Clark, 1987).

Two additional observations are in order. First, we found an effect of presentation order. This effect is consistent with prevailing hypothesis-testing models of word learning, whereby a learner could initially posit a basic-level interpretation of the target label, which may or may not be revised upon encountering the semantic alternative (e.g., Trueswell et al., 2013; Stevens et al., 2017), especially if the first label was understood as introducing a basic-level category with respect to which the following (alterative) referent was to be interpreted (Waxman et al., 1991, 1997). Second, generalizations to the basic-level category were overall infrequent in Experiment 2. Looking at Appendix A2, just as in Experiment 1, we again find that "incomplete Basic" responses in which some but not all basic-level exemplars were selected from the grid (classified under Other in Fig. 5) were just as frequent as the more strictly coded Basic responses. Experiment 3 directly addressed this issue with a change in task design.

Experiment 3

If, as we have proposed, the acquisition of subordinate nouns crucially involves assuming that the speaker was informative in using a word, then *any* context in which the word is used by a knowledgeable speaker can also provide important clues to the specificity of meaning it encodes. Beyond the presence of a semantic alternative within the learning phase, as in Experiments 1 and 2, other evidence that highlights the relevance of narrower semantic categories might promote subordinate conjectures. Thus, in Experiment 3 we asked whether the presentation of images and questions *at test* itself may implicitly set up expectations about the specificity of a novel word's meaning.

Within the test-grid paradigm, the idea that expectations about informativity may matter after learners posit their initial hypothesis for word meaning may appear counterintuitive. This is because the paradigm implicitly assumes that testing a word meaning conjecture is benign and insulated from the rest of the conversational exchange. Indeed, the possibility that information beyond the brief moment of exposure to a label has consequences for word learning is unexpected under many accounts of the acquisition of subordinate nouns (Spencer et al., 2011; Lewis & Frank, 2018; among others) and even many classes of word learning models more broadly (e.g., Yu & Smith, 2007; Fazly et al., 2010): the use or observation of a word in strictly non-labelling

[7] Model formula: Basic ~ Label + Order + (1 + Order | Participant) + (1 | Item). Both predictors were sum coded. The interaction effect could not be estimated with the model's logit link function because no Basic response was observed in the condition crossing *Labelled Alternative* and *Target Second* conditions.

[8] The model formula is the same as the model for Basic responses, but with the correlation term removed in the subject random effects. The Bonferroni correction for multiple comparisons has been applied to the reported p-value.

[9] Because the semantic alternative was always present across both conditions, we were able to inspect the distribution of "mutually exclusive" responses under different training conditions, as first observed in the *Contrast* condition of Experiment 1. These responses constituted<10% of Other responses in the *Unlabelled Alternative* condition, but over 50% of Other responses in the *Labelled Alternative* condition (see Appendix A2). While the mechanism driving these "mutually exclusive" responses is unclear, they pattern specifically with the labelling of the semantic alternative, not just its presence, suggesting that these "mutually exclusive" interpretations are specific to the processing of linguistic, as opposed to conceptual, contrast.

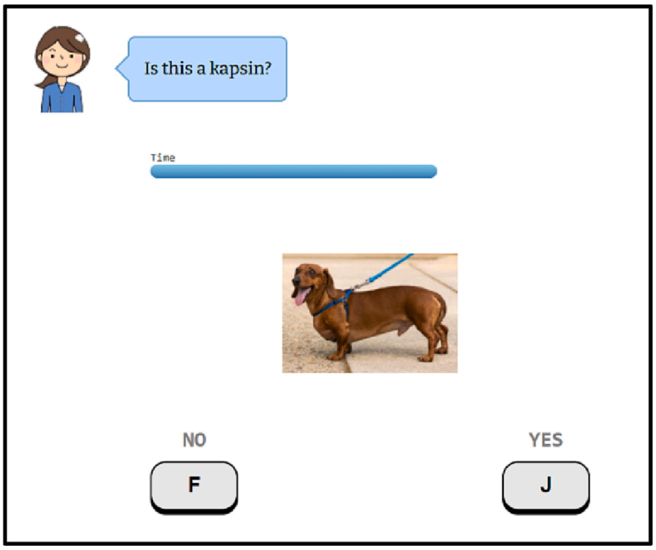

Fig. 6. Sequential design for Experiment 3 test phase.

contexts is not considered to be additional evidence for word meaning. However, our account, along with findings from Experiments 1 and 2, challenges the assumption that learning is put on hold when learners are asked to reason about the meaning of a word. Instead, we assume that effects of contrast on subordinate-level conjectures can appear even at the test phase as the learner considers what the speaker might have in mind within a set of possible meanings.

To probe this idea, we adopted the basic design of Experiment 1 (*Contrast* and *No Contrast* conditions) but made several changes to the test phase. First, we deconstructed the test-grid into a sequence of forced-choice trials (one for each image), so as to get a clearer picture of how learners reason through the choice of generalizing a word's meaning as they navigate the generalization set with each new exemplar. Second, to examine how learners make inferences about contrast *at test* to guide expectations about what the speaker might have in mind, we manipulated the order of exemplars in the test sequence. We expected this order to shape the course of word meaning generalization online, under the assumption that the basic- and subordinate-level exemplars from the generalization set differ in highlighting the relevance of the less-accessible vertical contrast between basic- and subordinate-level meanings (see also Skordos & Papafragou, 2016, on the role of alternative order in scalar inference). For example, inviting learners to select the subordinate-level exemplars first and withholding the basic-level exemplars until later may shift their hypothesis to a narrower category, as the speaker's emphasis on members from the same subordinate-level category is odd assuming the speaker's intent to highlight the basic-level meaning of the word. On the other hand, inviting learners to consider the inclusion of other exemplars from the basic-level category first may not have such an effect, as the basic-level is the preferred (generic-addressee) level of informativity. Thus, being tested with basic- versus subordinate-level exemplars first may give rise to different inferences about the level of specificity encoded in the word.

Participants

One hundred adults of the same demographic background who did not previously participate in Experiments 1 and 2 were recruited on Prolific. The analysis for Experiment 3 tested the statistical significance of the two main effects in a 2-by-2 design, so the sample size followed that of Experiment 2.

Materials and procedure

Experiment 3 replicated the within-subject learning conditions of Experiment 1 (*Contrast* and *No Contrast*) but made three changes to the design of the test phase. First, the simultaneous presentation of the grid of images at test was replaced with a sequential presentation of each image. In this new sequential design, participants were instructed to indicate whether each image matched the target label from the learning phase (e.g., "Is this a kapsin?") using keypresses, under a five-second timeout (Fig. 6).

Second, we crossed the presence of a semantic alternative (*Contrast* vs. *No Contrast*) in the learning phase with an additional Test Sequence manipulation (*Basic First* vs. *Subordinate First*) over the order in which the basic- and subordinate-level exemplars were presented in the test phase. In the *Basic First* condition, learners had to determine whether the label generalized to the three other basic-level exemplars first, before seeing the two matches to the target subordinate-level category. In the *Subordinate First* condition, the blocks of basic- and subordinate-level exemplars were presented in reverse order. The sequence of exemplars presented in the test set for the two conditions are shown in Fig. 7. Recall that the coding of both Basic and Subordinate responses requires the learners to accept at least the target subordinate-level exemplars at the test phase; we now manipulated the point at which responding to these target trials happens within the testing sequence – before (*Subordinate First*) or after (*Basic First*) being presented with other exemplars from the basic-level category. Given data from Experiments 1 and 2, we expected participants to quickly and reliably respond "Yes" to all subordinate-level exemplars and "No" to all other non-match exemplars, with responses to the basic-level exemplars being the crucial determiner for whether learners generalized to the subordinate- vs. the basic-level meaning.

Lastly, the number of exemplars in the test phase was reduced from 18 to 12 to alleviate the sequential test phase's demands on working memory. This reduced set maintained the relative proportion of exemplar types, including two matches to the target subordinate-level category and three other basic-level matches (in the *Contrast* condition, one of these was a member of the alternative subordinate-level category).

Coding

The coding scheme followed that of the Experiment 1.

Results

After applying the same filtering criteria from the previous two experiments, 776 responses from 97 participants entered the analysis. The distribution of these responses is shown in Fig. 8. (See also Appendix A3

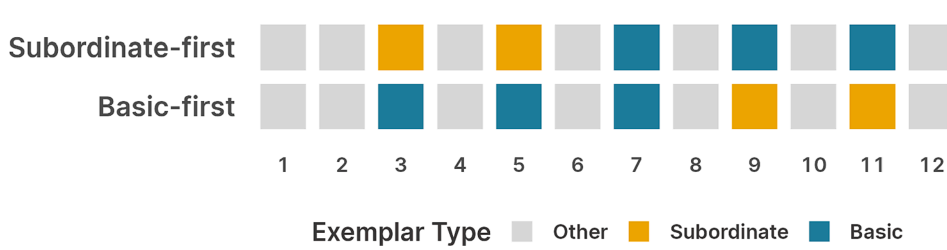

Fig. 7. Sequential presentation of exemplars during the test phase of Experiment 3.

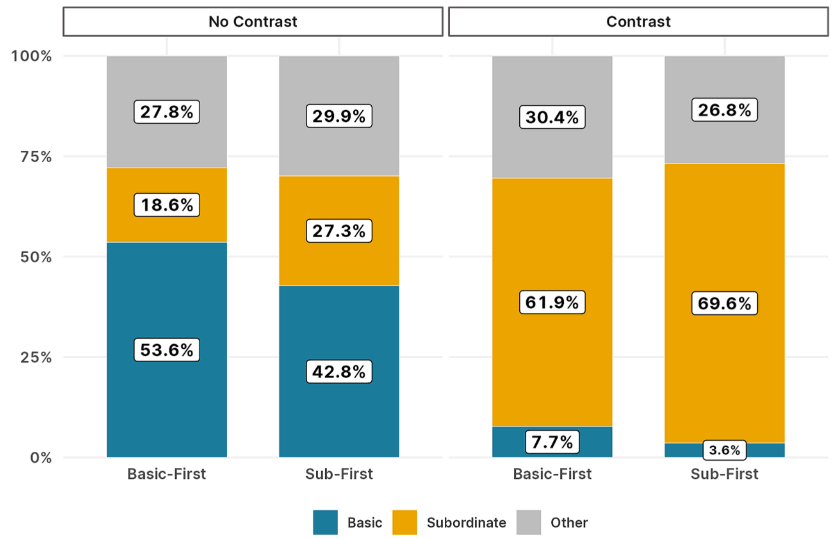

Fig. 8. Responses at test for Experiment 3.

Table 3
Mixed-effects logistic regression model fitted to Basic responses in Experiment 3.

	β (SE)	t	p
(Intercept)	−2.2 (0.5)	−4.1	<0.0001
Contrast	−2.1 (0.2)	−10.6	<0.0001
Test Sequence	−0.45 (0.2)	−3.0	0.003
Contrast * Test Sequence	−0.12 (0.1)	−0.8	0.44

Table 4
Mixed-effects logistic regression model fitted to Subordinate responses in Experiment 3.

	β (SE)	t	p
(Intercept)	−0.3 (0.1)	−2.5	0.013
Contrast	1.1 (0.1)	11.5	<0.0001
Test Sequence	0.25 (0.1)	2.8	0.005
Contrast * Test Sequence	−0.04 (0.1)	−0.5	0.64

for the distribution of responses with a fine-grained breakdown of the Other category.).

We fitted a mixed-effects logistic regression model to the rate of Basic responses, with the Contrast (*Contrast* vs. *No Contrast*) and Test Sequence (*Basic First* vs. *Subordinate First*) conditions and their interaction as fixed effects, and random intercepts by subject and item. The model summary is shown in Table 3.[10]

We found a significant main effect of Contrast ($\beta = -2.1$ $SE = 0.2$, $p < 0.0001$) replicating the findings from Experiment 1; there was an overall lower rate of Basic responses in the presence of a labelled semantic alternative. We also found a significant effect of Test Sequence ($\beta = -0.45$, $SE = 0.2$, $p = 0.003$), such that the rate of Basic responses was lower in the *Subordinate First* condition than in the *Basic First* condition. The interaction term did not reach significance ($\beta = -0.16$, $SE = 0.1$, $p = 0.44$). The same pattern was reflected in the model fitted to the rate of Subordinate responses, as shown in Table 4.

Discussion

The present findings suggest that learners considered information from the test phase as additional evidence for generalizing word meanings beyond ostensive labelling in the learning phase. As predicted by our pragmatic account, when learners were prevented from the opportunity to generalize to the basic-level category immediately after learning the word (by initially only being exposed to exemplars from the subordinate-level category in the *Subordinate First* condition), they were led to believe that the speaker had in mind a narrower, subordinate-level meaning of the target label. In comparison, no such expectation arose in the *Basic First* condition where participants were able to commit early to their initial basic-level conjecture. Together, these data show that the basic-level bias is fragile and may be abandoned or inhibited under sufficient evidence that the label in fact encodes a different (e.g., narrower) level of specificity than initially thought. By adopting sequential testing, where learners must continually collect evidence and update their inferences about the informativity level encoded in the use of the word, we were able to probe such effects within the structure of the test phase.

The present data are surprising for many subordinate-word learning models, where acting on a hypothesis about a word's meaning is not counted as additional information for that word's meaning. For example, this prediction is explicitly formalized in the notion of strong versus weak sampling in the suspicious coincidence account, which states that a learner acting on their existing knowledge (e.g., extending a learned label to other, unlabeled exemplars from the same subordinate-level category) does not contribute to the generalization of word meanings because there is no "suspicious coincidence" to be uncovered from that self-driven generative process (Tenenbaum & Griffith, 2001; Xu & Tenenbaum, 2007a; Lewis & Frank, 2016). However, on the current pragmatic account, any use of the word can be taken as an indication of what the speaker might have in mind.

Finally, we note that the rate of Basic responses in the *No Contrast* condition of Experiment 3 (40 ~ 50%) was much higher than that observed in Experiments 1 (22%). This is consistent with our initial suspicion that the suppression of the basic-level bias in those prior experiments was the consequence of the large number of test trials that necessitated a strict coding threshold for consistent Basic responses. Sequential testing with "Yes" and "No" responses in Experiment 3 imposed fewer cognitive demands compared to the more open-ended task of searching for and clicking on many images in the grid. In turn, we find that the data from the *No Contrast* condition pattern closer to the empirical average for the magnitude of the basic-level bias effect. Inspection of Appendix A3 reveals that Incomplete Basic responses were still present among Other responses in both the *No Contrast* ($M_{\text{Basic-first}} =$ 22.30%, $M_{\text{Sub-first}} =$ 25.80% of all responses) and *Contrast* ($M_{\text{Basic-first}} =$

[10] Model formula: Basic ~ Contrast + * TestSequence + (1 | Participant) + (1 | Item).

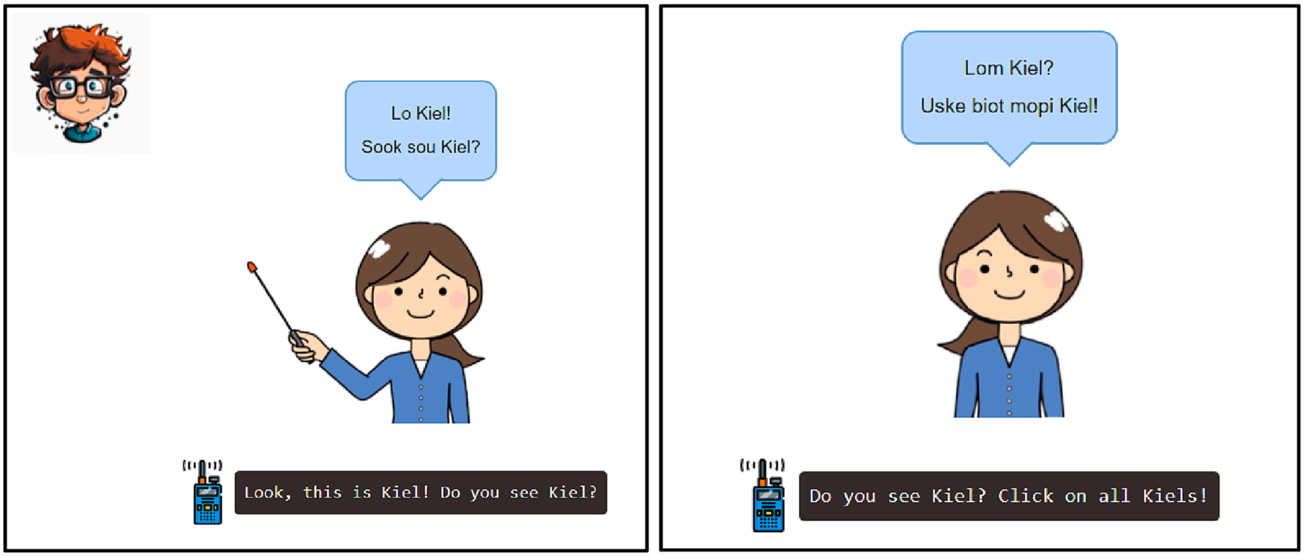

Fig. 9. Presentation of Sallu's friend Kiel (left) and the instructions to find more of them (right) in the first trial of the exposure session in the *Foreign* condition of Experiment 4. English subtitles generated by a translator device appeared at the bottom of the screen in the first exposure trial and were removed from the second exposure trial onwards.

11.90%, $M_{Sub\text{-}first} = 12.90\%$) conditions, but to a lesser degree compared to our earlier task; even though their frequency was comparable to that of Basic responses in Experiments 1 and 2, it was halved in Experiment 3.

Experiment 4

A remaining question about the present paradigm in both our own work and in previous literature on subordinate nouns is whether the degree to which English is used to teach the novel words plays a role in the generalization of word meanings. In a departure from our earlier experiments and prior work, Experiment 4 made a change to the task to better approximate the natural circumstances of early word learning. Specifically, Experiment 4 replicated Experiment 1 in two versions: an *English* version where the novel words were embedded into English carrier sentences (an exact replication of Experiment 1) and a *Foreign* version, where all input was delivered in an artificial language. Thus Experiment 4 was a 2-by-2 design crossing the presence of a labelled contrast (*Contrast* vs. *No Contrast*) and the type of exposure language (*English* vs. *Foreign*). If the role of contrast truly characterizes mechanisms of subordinate noun acquisition, then the results of Experiment 4 should broadly replicate the patterns of our earlier study.

Participants

Participants who did not previously participate in our experiments were recruited from the undergraduate subject pool at the University of Pennsylvania (n = 105). The number of participants was roughly balanced between the *English* (n = 50) and *Foreign* (n = 55) variants of the experiment; each sought to replicate the effect of the *Contrast* condition in Experiment 1 with similar sample sizes.

Materials and procedure

All materials and procedure for the *English* version were identical to Experiment 1. In the *Foreign* version, the prior context for the task was communicated in English by a narrator. Participants were introduced to the same cartoon character (now named Sallu) and were told that she was a native speaker of a foreign language called "Uffish" (in actuality, a string of nonce words). As in the *English* condition, participants were told that Sallu would like to play a game to teach them some Uffish words. An exposure session was designed to familiarize participants to Uffish. This session consisted of two trials. In the first trial, Sallu introduced and named her friend Kiel (learning phase), and then asked whether there were any Kiels in an image grid where there were, in fact, two other images of Kiel alongside pictures other faces (testing phase; see Fig. 9). Throughout the experiment all communication from Sallu within a trial appeared in two speech bubbles: one was used for labelling ("Lo X! Sook sou X?" meaning 'This is a X. Do you see the X?') and another for asking participants to find matches to the target label ("Lom X? Uske biot mopi X!" meaning 'Do you see the X? Click on all the X!'). These two speech bubbles were the only instances of Uffish within each trial. To help participants interpret the instructions, for the first ('Kiel') exposure trial only, participants were provided with English subtitles generated by a translator device (see Fig. 9).[11]

In the second exposure trial, the narrator took away the translator and told participants that Sallu would now teach them the word for her favorite food.[12] Sallu re-appeared to introduce and label a pizza ("mouli") and asked participants to select more "mouli" from the image grid, where there were five additional pizzas of different kinds alongside other unrelated objects (e.g., beach, ice cream, statue, etc.). There was no number morphology in Uffish, and by virtue of using the exact same carrier phrase for proper names (first trial) and nouns (second trial), participants were led to believe that Uffish lacked definite articles (whereas in the *English* version, labelling was always via definite reference). This design allowed the *Foreign* version to depart more strongly from English.[13] Participants who did not select both Kiel- and all five "mouli"-instances in the exposure phase were removed from analysis. After the exposure session, the narrator told participants that they would

[11] The narrator's instructions were: "Sallu would like to play a little game with you and teach you some Uffish words. Sallu says "Hi!"." (speech bubble of a nonce word appears on top of Sallu). "You'll get a translator to help you understand Sallu," (subtitles appear below Sallu) "starting the with names of her friends. Pay attention to the words because you'll be asked questions about them later!".

[12] The narrator's instructions were: "Great job! Let's play the game again, now without the translator" (translator icon disappears) "This time, Sallu will teach you the Uffish word for her favorite food!".

[13] Of course, the proper name for Kiel (and finding other snapshots of this individual) is not comparable to naming a kind (e.g., pizza). This was an intentional choice to not interfere with the main manipulation of the experiment.

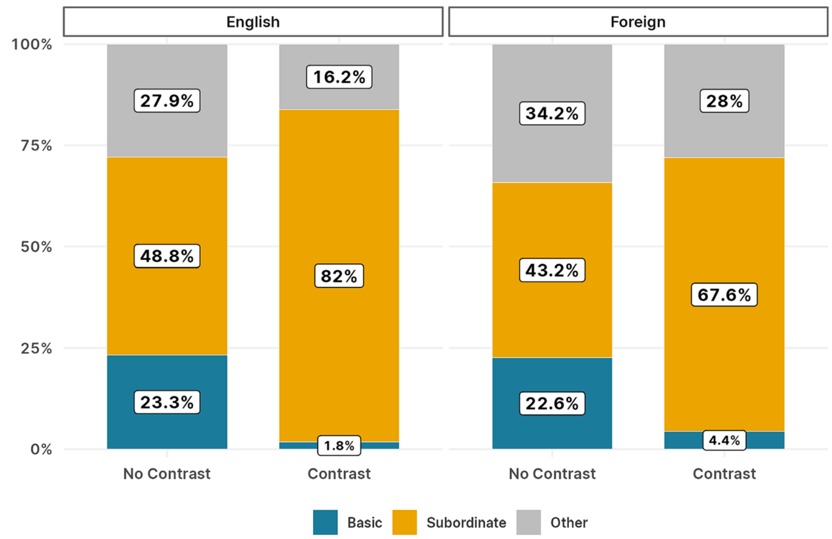

Fig. 10. Responses at test for Experiment 4.

continue playing the same game with some more Uffish words, and the same ten trials (8 critical and 2 filler) proceeded in the same manner as in the *English* version, except with speech bubbles in Uffish without the translator and subtitles. The novel nouns used in prior versions of this experiment were therefore now interpreted as part of a foreign language (e.g., "Lo mipen! Sook sou mipen?", i.e., 'This is a mipen. Do you see the mipen?', and "Lom mipen? Uske biot mopi mipen!", i.e., 'Do you see the mipen? Click on all the mipens!').

Coding

The coding scheme followed that of the Experiment 1.

Results

After applying the same filtering criteria as in prior experiments, six participants from the *English* version and four participants from the *Foreign* version were removed. Additionally, 33 trial-level responses were dropped. In total, 711 responses from 93 participants were entered into the analysis. Results are shown in Fig. 10. (See also Appendix A4 for the distribution of responses with a fine-grained breakdown of the Other category.).

We fitted a mixed-effects logistic regression model to the rate of Basic responses, with Contrast (*Contrast* vs. *No Contrast*), Version (*English* vs. *Foreign*), and Trial Block (*First half* vs. *Second half*) and all two-way interactions as fixed effects. The model also included random intercepts and slope of Contrast by subject and random intercepts by item.[14] The model summary is shown in Table 5. We found a significant main effect of Contrast ($\beta = -3.8$ $SE = 0.8$, $p < 0.0001$), successfully replicating the findings from Experiment 1. No other predictors reached significance.

The same model fitted to the rate of Subordinate responses is shown in Table 6. We found a significant main effect of Contrast ($\beta = 0.8$, $SE = 0.1$, $p < 0.0001$) in the opposite direction, such that there were more Subordinate responses in the *Contrast* condition than in the *No Contrast* condition. We also found a significant main effect of Version ($\beta = 0.4$, $SE = 0.2$, $p = 0.035$), such that the *Foreign* condition had an overall lower rate of Subordinate responses than the *English* condition. No other predictors reached significance.

Table 5
Mixed-effects logistic regression model fitted to Basic responses in Experiment 4.

	β (SE)	t	p
(Intercept)	−5.2 (0.9)	−6.0	<0.0001
Contrast	−3.8 (0.8)	−4.5	<0.0001
Version	−0.2 (0.5)	−0.3	0.736
Trial Block	−0.3 (0.2)	−1.4	0.170
Contrast * Version	−0.2 (0.5)	−0.4	0.673
Version * Trial Block	−0.4 (0.2)	−1.6	0.102
Trial Block * Contrast	−0.1 (0.1)	−0.9	0.369

Discussion

Experiment 4 successfully replicated the effect of semantic contrast on the rate of Basic and Subordinate responses using a different linguistic format with minimal English exposure. Importantly, the design of the *Foreign* version brings the immediate generalization task closer to the natural circumstances of early word learning than iterations of the task in prior studies, where English has traditionally dominated participants' interactions with the interlocutor (whose motivations for introducing novel words in an English frame were often left unclear given the possible English lexical alternatives). These findings thus confirm the generalizability of the semantic contrast effect to unfamiliar language input and its nature as stemming from general principles of communication which includes, but are not limited to, expectations about linguistic form.

As a final note, while the rate of Basic responses was not significantly different between the two language versions, Subordinate responses were more frequent in the *English* version compared to the *Foreign* version, which instead showed more Other responses (see Appendix A4 for details). In other words, minimizing exposure to English increased the overall difficulty of the (already hard) task of generalizing to the narrower, subordinate-level categories. One possibility is that the use of English makes the familiar subordinate-level semantic categories more accessible for mapping, which is consistent with pragmatic accounts where learners exploit familiar linguistic devices for subordination to infer and encode semantic contrast. However, this finding is unexpected under bottom-up accounts, where the salience of perceptual features determines the ease of mapping words to subordinate-level categories.

[14] Model formula: Basic ~ (Contrast + Version + Order)^2 + (1 + Contrast | Participant) + (1 | Item). All predictors were sum coded with *Contrast* at 1 and *No Contrast* at −1, *English* at 1 and *Foreign* at −1, and *Block1* at −1 and *Block2* at 1.

Table 6
Mixed-effects logistic regression model fitted to Subordinate responses in Experiment 4.

	β (SE)	t	p
(Intercept)	0.6 (0.3)	2.0	0.049
Contrast	0.8 (0.1)	7.1	<0.0001
Version	0.4 (0.2)	2.1	0.035
Trial Block	0.1 (0.1)	1.4	0.170
Contrast * Version	0.1 (0.1)	1.4	0.673
Version * Trial Block	−0.04 (0.1)	−0.4	0.102
Trial Block * Contrast	−0.04 (0.1)	−0.4	0.369

General discussion

In this study, we proposed that the challenge of acquiring subordinate nouns involves to a large extent a pragmatic puzzle. Assuming that learners make inferences about the interlocutor's intent behind the choice of a word (Grice, 1975), we reasoned that generalizing to the narrower, subordinate-level meaning would greatly benefit from linguistic cues to the specificity of a word in the face of the basic-level bias. Such cues could involve contrast to other lexical alternatives that can support inferences about the target label. We tested a set of predictions flowing from the hypothesis that the introduction of a semantic alternative at the subordinate level would make subordinate-level conjectures accessible to a learner who would otherwise produce a basic-level generalization for a novel label.

These predictions were confirmed. We found that the presence of a labelled alternative at the subordinate level eliminated the basic-level bias during word learning with adult learners (Experiment 1); furthermore, the mere presence of the alternative referent without a label was less likely to suppress the basic-level bias compared to cases where the alternative was labelled (Experiment 2). Additionally, the role of alternatives in shaping expectations of specificity extended well beyond the moment that a label was ostensibly introduced (Experiment 3). As we had anticipated, these results show that learners choose between subordinate- and basic-level meanings for a label by inferring the level of informativity encoded in the use of the label (and not just from any type of contrast). Lastly, this effect of semantic contrast was not an artifact of task adaptation nor driven by language familiarity; pragmatic inferences about the level of specificity encoded in a word could be calculated even from novel language input (Experiment 4). These results underscore the fact that labels for objects invoke concepts and meanings that the speaker intends to convey, and that this information is often delivered to the listener by highlighting (in verbal and non-verbal ways) how the label contrasts with a specific set of other relevant lexical options that the speaker could have used but did not.

Alternative accounts

The robust finding that the presence of a semantic alternative at the subordinate level facilitates subordinate-level generalizations is unexpected under accounts of subordinate word learning where information outside of labelling and referent introduction is not considered in forming hypotheses about word meaning. This is because the semantic alternative does not contribute to either the perceptual (e.g., Spencer et al., 2011; Jenkins et al., 2015, 2021) or the distributional (e.g., Xu & Tenenbaum, 2007b; Lewis & Frank, 2016, 2018) profile of the target label itself, which has been assumed to be the fundamental unit of evidence for word learning in much prior work. Instead, not only do the availability and salience of alternatives in the communicative episode guide word meaning generalizations in our data, but we also observe a strong effect of a labelled alternative even after just a single instance of the target label, without needing to expose the learner to multiple exemplars cross-situationally.

Furthermore, the idea that aspects of communication outside of ostensive labelling can affect the interpretation of a word upends the traditional divide between learning and testing, where the latter is not thought to contribute additional evidence for word meaning because it is driven by the learner acting on their own hypothesis (hence the long-standing assumption that the test grid, as popularized by Xu & Tenenbaum 2007b, simply probes the hypothesized word meaning). In fact, some have explicitly proposed built-in mechanisms that prevent such learning, including the distinction between weak and strong sampling under the suspicious coincidence account which restricts evidence for word meaning to label-referent pairings provided by a knowledgeable speaker (Tenenbaum & Griffith, 2001; Xu & Tenenbaum, 2007b; Lewis & Frank, 2016). To a degree, this skepticism is warranted, as allowing such a mechanism to take hold may result in a positive feedback loop of confirming hypotheses just via opportunities to act on them. Certainly, this kind of unconstrained learning is undesirable, though there is something to be said about how opportunities to test hypotheses about word meaning can sometimes reveal insights to the word's meaning itself. If learners can integrate information about speaker intent in hypothesizing word meanings, then there may be non-labelling or even non-verbal contexts inviting the learner to respond to or reason about the word that in fact highlight an alternative word meaning that is also consistent with the data (namely, observed exemplars) but was simply more difficult to access at the moment of labelling. We argued that this may be the case for subordinate- vs. basic-level meanings, and indeed, Experiment 3 showed that learners can be led to consider the narrower, subordinate-level meaning when the opportunities for the learner to act on their current conjecture for a word's meaning encourage a narrower interpretation of the word. This model of subordinate-word learning seems reasonable under the assumption that natural circumstances that invite the learner to act on a hypothesized word meaning are often intertwined with some task or goal relevant to a conversation, where the behavior of the interlocutor can also reveal insights into word meaning (Wang & Mintz, 2018).

In sum, information from contrast and alternatives can come from any point in a communicative episode, and the specific question about word meaning that the learner is being encouraged to consider may provide additional evidence to word meaning beyond ostensive labelling. This is a challenge for bottom-up accounts of learning subordinate nouns (e.g., Spencer et al., 2011) as well as associationist word learning models more broadly (e.g., Yu & Smith, 2007; Fazly et al., 2010), but the data may be captured by a class of so-called hypotheses-testing models where learners reason over semantic categories rather than exemplars (Gleitman & Trueswell, 2020) and posit a single conjecture at a time rather than a range of possible meanings (e.g., Stevens, Trueswell, Yang, & Gleitman, 2017; Trueswell, Medina, Hafri, & Gleitman, 2013; Medina et al., 2011). Other rational models of word learning such as those based in Bayesian inference also share our emphasis on speaker intent (e.g., Frank & Goodman, 2012, 2014) and may thus account for our findings in terms of the likelihood of the data under possible word meanings, although the specific mechanism for the acquisition of subordinate nouns must consider information beyond the sampling statistics of a label as the primary cue to the level of specificity encoded in a word (i.e., beyond the "suspicious" nature of label-referent pairings). Indeed, the suspicious coincidence effect stemming from the observation of multiple positive evidence for the subordinate-level meaning has been argued to be fragile and indirect (Caplan, 2022), which also explains its disappearance under richer contexts providing information about semantic contrast that relieve the learner from reasoning about the specificity of word meaning from such unreliable cues in the referential world (Wang & Trueswell, 2019, 2022).

Pragmatics and the acquisition of subordinate nouns

Inferences about speaker intent are pervasive, and arguably inevitable in word learning, since labels do not simply describe the world (Grigoroglou & Papafragou, 2021). We have argued that this is especially salient for the challenge of acquiring subordinate nouns, where overcoming the basic-level bias requires reasoning about the level of specificity encoded in the word or the level of informativity intended by the speaker in the use of the word. Our work refines the often-cited role of contrast in facilitating word meaning generalizations beyond the basic-level (e.g., Clark, 1987, 1988, 1990; Waxman et al., 1991, 1997; Wang & Trueswell, 2019, 2022) by proposing that subordinate-level generalizations are facilitated by not just any contrast to the preferred basic-level meaning. Rather, the successful inference depends on the learner's ability to recognize which particular scale is being highlighted in the task – the lateral contrast between mutually exclusive categories at the same level, or the harder-to-access vertical contrast between basic- and subordinate-level categories.

This perspective is reminiscent of approaches to the development of the pragmatic interpretation of "some" as meaning 'some but not all' from the literature on scalar inference; such pragmatic interpretations in children have been shown to benefit from prior exposure to relevant lexical alternatives such as "all" or even "none" (Skordos & Papafragou, 2016; cf. also Barner et al., 2011). Likewise, we find that information about semantic contrast and the relevance of specific alternatives, despite not revealing properties about the target label itself, nevertheless constrained the interpretation of the target label to a subordinate-level category. The assumption that children are Gricean has also been extensively explored in other parts of the word learning literature. For instance, children as young as two years old show sensitivity to discourse context and speaker intent in positing hypotheses about word meaning (Diesendruck et al., 2004; Diesendruck & Markson, 2005; Diesendruck, 2005). Our findings affirm the strong contribution of such socio-pragmatic constraints in the case of acquiring subordinate nouns, above and beyond other attested factors such as bottom-up perceptual processes and statistical learning over surface-level properties of label-referent pairings.

In natural input, we expect these inferences built on the contrast between subordinate- and basic-level meanings to be especially crucial for early word learning when the referential world rarely offers direct negative evidence to rule out the basic-level interpretation in favor of a narrower, subordinate-level meaning (see discussion in Jenkins et al., 2015). These sources of information may come from the anchoring of the basic-level category as the relevant semantic domain (Waxman et al., 1991, 1997) and the presence of other labelled exemplars in cross situational learning (Wang & Trueswell, 2019, 2022), but may also come from the presence of labelled alternatives at the subordinate-level in single learning trials (Experiments 1 and 2) and even from cues to the relevance of certain alternatives made available outside of ostensive labelling (Experiment 3). Thus, we posit that the unifying principle underlying these various effects of contrast on the acquisition of subordinate nouns is about guiding learners to discover the relevant set of specific alternatives. Furthermore, although we have primarily focused on cases where the learner is mapping existing concepts to new labels, such as in the task of associating the concept 'dog' with the sound /dɒg/ (Gleitman & Trueswell, 2020), this primarily pragmatic learning mechanism could also be useful in explaining how unfamiliar or *newly formed* concepts are mapped to novel labels (Waxman & Markow, 1995; Lupyan et al., 2007; Lupyan & Thompson-Schill, 2012; LaTourrette & Waxman, 2019; Caplan, 2022). The replication of the semantic contrast effect from even novel language input (Experiment 4) offers a promising lead towards this generalized account of learning specific word meanings via inferences about the relevant alternatives involved in the speaker's choice of label. More broadly, it may be that the conceptual and semantic (i.e., linguistically encoded) categories co-develop as children master the conventions of their native language (Berman & Clark, 1989; Clark, 1992, 2017), which would require learners to also integrate evidence from uses of the word, speaker intent, communicative goals, and the like in forming and updating word meaning conjectures.

Conclusions

In this study, we proposed that the generalization of word meanings is – among other things – driven by assumptions about the relevant level of specificity intended in the use of the word, and that contrast can offer strong cues for such inferences. In a series of word learning experiments, we showed that semantic alternatives facilitate mappings to subordinate-level meanings, and especially so when the alternative is labelled, suggesting that learners can use linguistically marked contrast to reason about the level of informativity for a word's meaning that is expected from an ostensive labelling event. Furthermore, we showed that learners can integrate evidence from the accessibility of alternatives elsewhere in the communicative episode beyond the moment of labelling to generate such inferences about contrast to constrain word meaning generalizations. This sensitivity to the informativeness of an utterance offers a possible mechanism for the acquisition of subordinate nouns despite the apparent sparsity of evidence for subordinate-level meanings offered by the referential world.

CRediT authorship contribution statement

June Choe: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization. **Anna Papafragou:** Supervision, Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data, analysis codes, stimuli, and experiment scripts for the present study are available through the Open Science Framework (https://osf.io/x8cuf/, "The acquisition of subordinate nouns as pragmatic inference" project).

Acknowledgments

The authors are grateful to the audiences at HSP 2022, ELM 2022, CogSci 2022, and BUCLD 2022, as well as members of the Language & Cognition lab at the University of Pennsylvania for valuable feedback on this project. We also thank Florian Schwarz for guidance on the technical implementation of the experiments.

Appendix

Note: Appendix Figures show the distribution of responses at test for each experiment with a detailed breakdown of the Other category: "Inc. Basic" and "Inc. Subordinate" indicating incomplete responses (partial selections of subordinate and basic-level members) and "Mutually Exclusive" responses (selections of all basic-level members except the alternative when it was present in the learning phase) are now marked separately (see Fig. A1, Fig. A2, Fig. A3, Fig. A4).

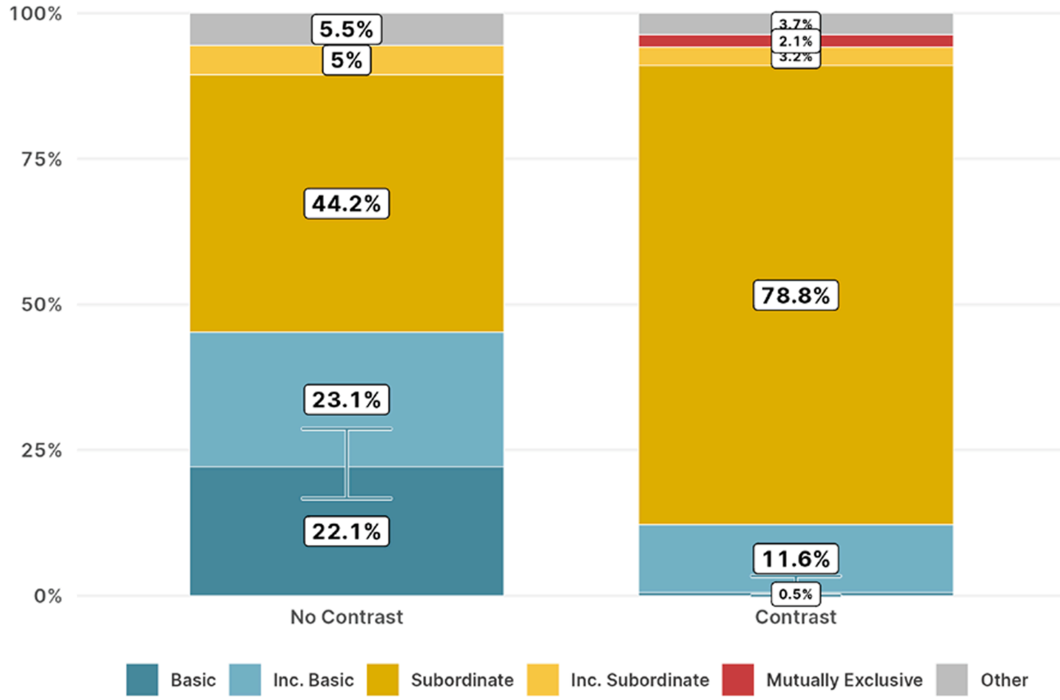

Fig. A1. Detailed responses at test for Experiment 1.

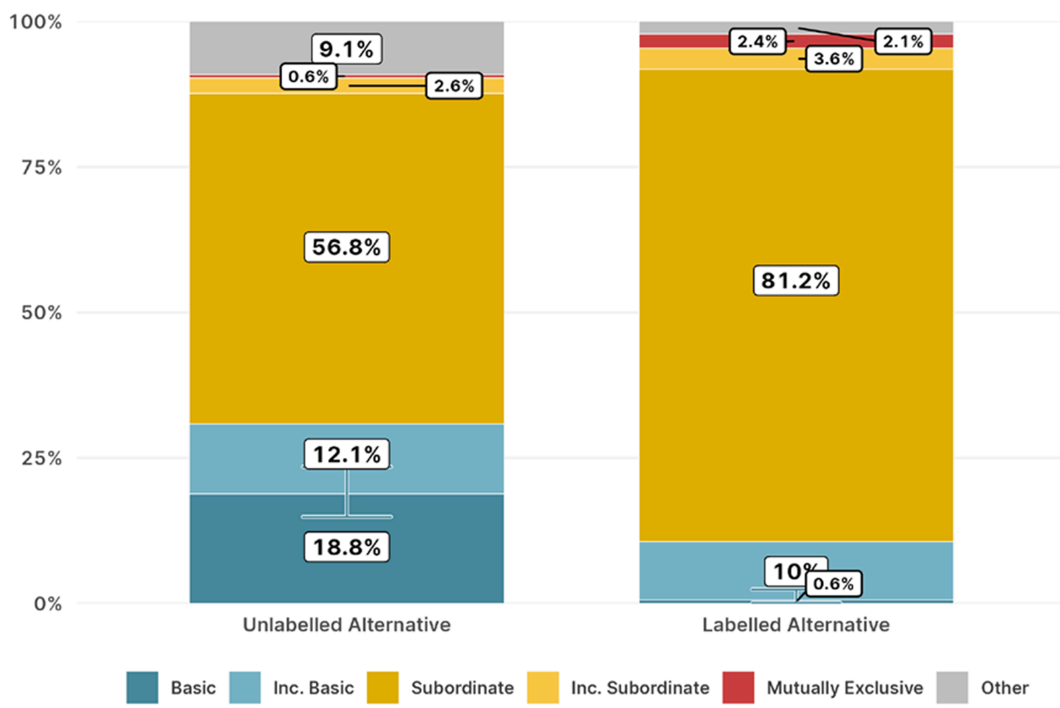

Fig. A2. Detailed responses at test for Experiment 2.

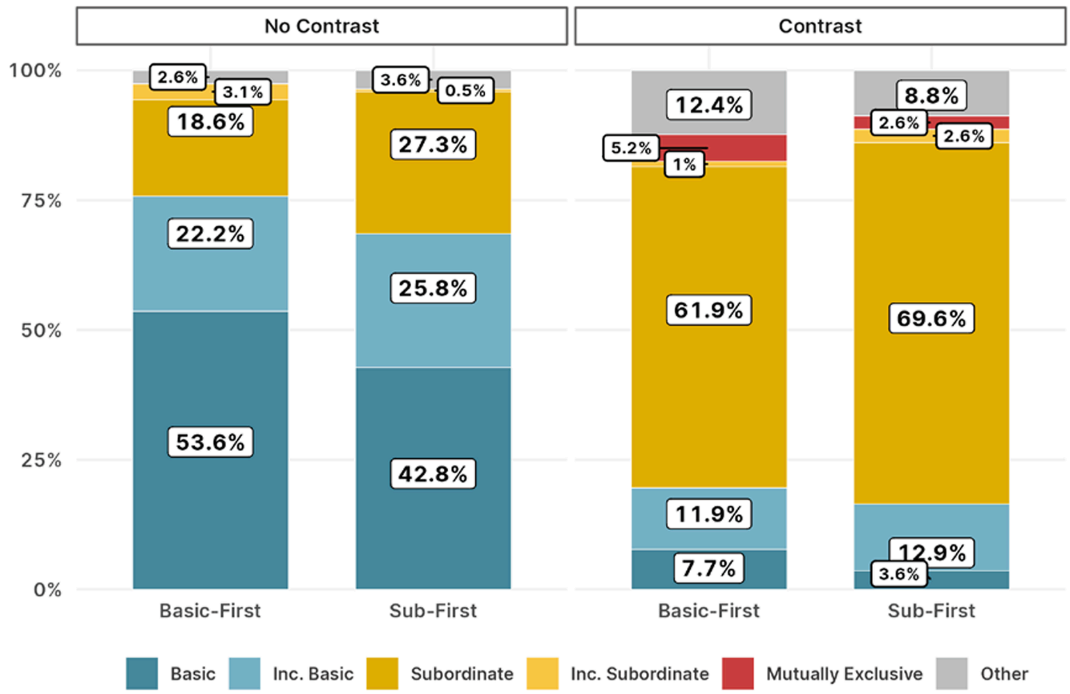

Fig. A3. Detailed responses at test for Experiment 3.

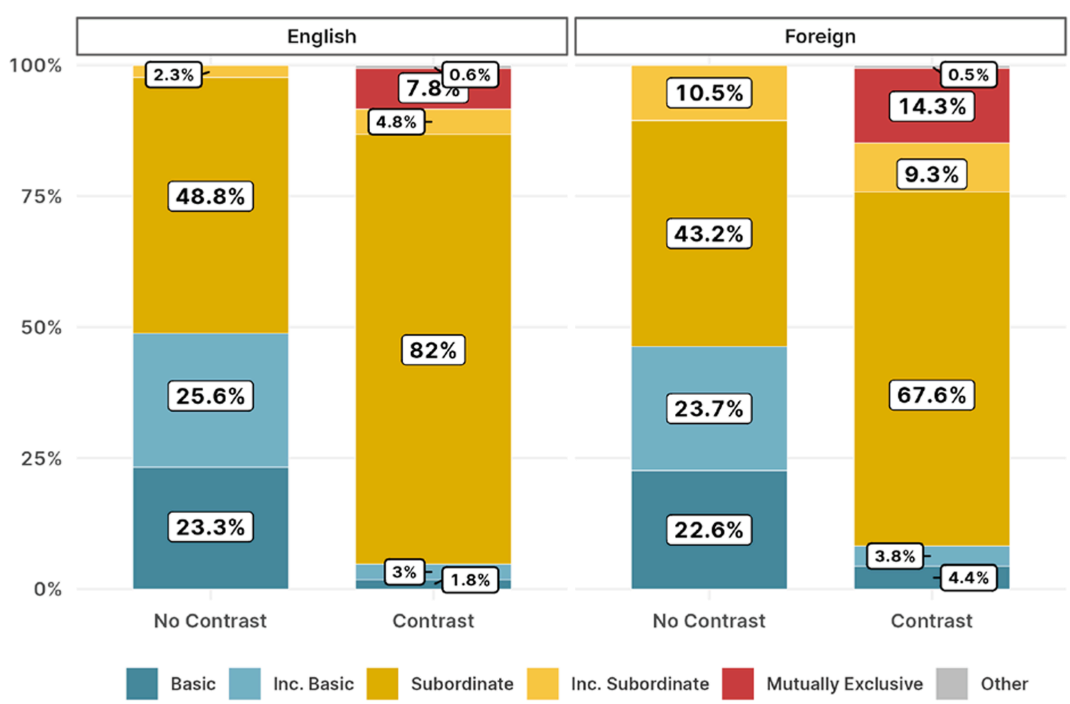

Fig. A4. Detailed responses at test for Experiment 4.

References

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition, 118*, 84–93.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Berman, R. A., & Clark, E. V. (1989). Learning to use compounds for contrast: Data from Hebrew. *First Language, 9*(27), 247–270.

Blewitt, P. (1983). Dog versus collie: Vocabulary in speech to young children. *Developmental Psychology, 19*, 602–609.

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology, 19*(4), 441–472.

Caplan, S. (2022). *The Immediacy of Linguistic Computation* [Doctoral dissertation, University of Pennsylvania].

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development, 56*(2), 508–523.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language, 60*, 547–590.

Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language, 14*, 547–567.

Clark, E. V. (1987). The Principle of Contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Lawrence Earlbaum Assoc: Hillsdale, NJ.

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language, 15*(2), 317–335.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*, 417–431.
Clark, E. V. (1992). Conventionality and contrast: Pragmatic principles with lexical consequences. In A. Lehrer, & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 171–188). Lawrence Erlbaum Associates Inc.
Clark, E. V. (2017). Semantic categories in acquisition. In H. Cohen, & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (pp. 397–421). Elsevier.
Cruse, D. A. (1986). *Lexical Semantics*. New York: Cambridge University Press.
Diesendruck, G. (2005). The principle of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology, 41*(3), 451–463.
Diesendruck, G., & Markson, L. (2005). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*(5), 630–641.
Diesendruck, G., Markson, L., Akhtar, N., & Reudor, A. (2004). Two-year-olds' sensitivity to speakers' intent: An alternative account of Samuelson and Smith. *Developmental Science, 7*(1), 33–41.
Fazly, A., Alishahi, A., & Stevenson, S. (2010). A probabilistic computational model of cross-situational word learning. *Cognitive Science, 34*, 1017–1063.
Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science, 336*(6084), 998.
Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology, 75*, 80–96.
Frank, M. C., Goodman, N. D., & Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science, 20* (5), 578–585.
Gelman, S. A., Wilcox, S. A., & Clark, E. V. (1989). Conceptual and lexical hierarchies in young children. *Cognitive Development, 4*, 309–326.
Gleitman, L. R., & Trueswell, J. C. (2020). Easy Words: Reference Resolution in a Malevolent Referent World. *Topics in Cognitive Science, 12*(1), 22–47.
Grigoroglou, M., & Papafragou, A. (2019a). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology, 55*, 951–966.
Grigoroglou, M., & Papafragou, A. (2019b). Children's (and adults') production adjustments to generic and particular listener needs. *Cognitive Science, 43*, e12790.
Grigoroglou, M., & Papafragou, A. (2021). Contributions of pragmatics to word learning and interpretation. In A. Papafragou, J. Trueswell, & L. Gleitman (Eds.), *The Oxford Handbook of the Mental Lexicon*. Oxford: Oxford University Press.
Grice, P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics* (Volume 3). New York: Academic Press.
Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes, 20*(667), 696.
Jenkins, G. W., Samuelson, L. K., Smith, J. R., & Spencer, J. P. (2015). Non-Bayesian noun generalization in 3-to 5-year-old children: Probing the role of prior knowledge in the suspicious coincidence effect. *Cognitive Science, 39*(2), 268–306.
Jenkins, G. W., Samuelson, L. K., Penny, W., & Spencer, J. P. (2021). Learning words in space and time: Contrasting models of the suspicious coincidence effect. *Cognition, 210*.
Kampa, A., & Papafragou, A. (2020). Four-year-olds incorporate speaker knowledge into pragmatic inferences. *Developmental Science, 23*(3), Article e12920.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1–26.
LaTourrette, A., & Waxman, S. R. (2019). A little labelling goes a long way: Semi-supervised learning in infancy. *Developmental Science, 22*(1).
Lewis, M. L., & Frank, M. C. (2016). Understanding the effect of social context on learning: A replication of Xu and Tenenbaum (2007b). *Journal of Experimental Psychology: General, 145*(9), 72–80.
Lewis, M. L., & Frank, M. C. (2018). Still suspicious: The suspicious-coincidence effect revisited. *Psychological Science, 29*(12), 2039–2047.
Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review, 9*(3), 550–557.
Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*, 1077–1083.
Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology, 141*, 170–186.
Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge: Cambridge University Press.
Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX).
Medina, T.N., Snedeker, J., Trueswell, J.C., & Gleitman, L.R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108*, 9014-9019.
Markman, E. M. (1984). The acquisition and hierarchical organization of categories by children. In C. Sophian (Ed.), *Origins of cognitive skills: The 18th Annual Carnegie Symposium on Cognition*.
Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*(1), 57–77.
Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 70–84.
Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition, 78*, 165–188.
Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition, 86*, 253–282.
Papafragou, A., Friedberg, C., & Cohen, M. (2018). The role of speaker knowledge in children's pragmatic inferences. *Child Development, 89*, 1642–1656.
R Core Team. (2021). R: A Language and Environment for Statistical Computing. https://www.R-project.org/.
Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 28–49). Hillsdale, NJ: Erlbaum.
Ross, B. H., & Murphy, G. L. (1996). Category-based predictions: Influence of uncertainty and feature associations. *Journal of Experimental Psychology, 22*, 736–753.
Shipley, E. F., & Kuhn, I. (1983). A constraint on comparisons: Equally detailed alternatives. *Journal of Experimental Child Psychology, 35*, 195–222.
Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicature: Alternatives and relevance. *Cognition, 153*, 6–18.
Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179–185.
Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science, 34*, 1244–1286.
Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning words in space and time probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science, 22*(8), 1049–1057.
Stevens, J. S., Trueswell, J. C., Yang, C., & Gleitman, L. R. (2017). The pursuit of word meanings. *Cognition, 41*(4), 638–676.
Tenenbaum, J. (1999). Bayesian modeling of human concept learning. In M. S. Kearns, S. A. Solla, & D. A. Cohn (Eds.), *Advances in Neural Information Processing Systems 11*. Cambridge, MA: MIT Press.
Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2015). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 1–30.
Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*(1), 126–156.
Wang, F. H., & Mintz, T. H. (2018). The role of reference in cross-situational word learning. *Cognition, 170*, 64–75.
Wang, F. H., & Trueswell, J. C. (2019). Spotting dalmatians: Children's ability to discover subordinate-level word meanings cross-situationally. *Cognitive Psychology, 114*, Article 101226.
Wang, F. H., & Trueswell, J. C. (2022). Being suspicious of suspicious coincidences: The case of learning subordinate word meanings. *Cognition.*
Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development, 5*, 123–150.
Waxman, S. R., Lynch, E. B., Casey, K. L., & Baer, L. (1997). Setters and Samoyeds: The emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology, 33*(6), 1074–1090.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13- month-old infants. *Cognitive Psychology, 29*(3), 257–302.
Waxman, S. R., Shipley, S., & Shepperson, B. (1991). Establishing new subcategories: The role of category labels and existing knowledge. *Child Development, 62*, 127–138.
Xu, F., & Tenenbaum, J. B. (2007a). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288–297.
Xu, F., & Tenenbaum, J. B. (2007b). Word learning as Bayesian inference. *Psychological Review, 114*(2), 245–272.
Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420.