



Social cues modulate the representations underlying cross-situational learning



Kyle MacDonald ^{*}, Daniel Yurovsky, Michael C. Frank

Department of Psychology, Stanford University, United States

ARTICLE INFO

Article history:

Accepted 7 February 2017

Available online 10 March 2017

Keywords:

Statistical learning

Social cues

Word learning

Language acquisition

ABSTRACT

Because children hear language in environments that contain many things to talk about, learning the meaning of even the simplest word requires making inferences under uncertainty. A cross-situational statistical learner can aggregate across naming events to form stable word-referent mappings, but this approach neglects an important source of information that can reduce referential uncertainty: social cues from speakers (e.g., eye gaze). In four large-scale experiments with adults, we tested the effects of varying referential uncertainty in cross-situational word learning using social cues. Social cues shifted learners away from tracking multiple hypotheses and towards storing only a single hypothesis (Experiments 1 and 2). In addition, learners were sensitive to graded changes in the strength of a social cue, and when it became less reliable, they were more likely to store multiple hypotheses (Experiment 3). Finally, learners stored fewer word-referent mappings in the presence of a social cue even when given the opportunity to visually inspect the objects for the same amount of time (Experiment 4). Taken together, our data suggest that the representations underlying cross-situational word learning of concrete object labels are quite flexible: In conditions of greater uncertainty, learners store a broader range of information.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Learning the meaning of a new word should be hard. Consider that even concrete nouns are often used in complex contexts with multiple possible referents, which in turn have many conceptually natural properties that a speaker could talk about. This ambiguity creates the potential for an (in principle) unlimited amount of referential uncertainty in the learning task.¹ Remarkably, word learning proceeds despite this uncertainty, with estimates of adult vocabularies ranging between 50,000 to 100,000 distinct words (Bloom, 2002). How do learners infer and retain such a large variety of word meanings from data with this kind of ambiguity?

Statistical learning theories offer a solution to this problem by aggregating cross-situational statistics across labeling events to identify underlying word meanings (Siskind, 1996; Yu & Smith, 2007). Recent experimental work has shown that both adults and young infants can use word-object co-occurrence statistics to learn words from individually ambiguous

^{*} Corresponding author.

E-mail address: kyle.macdonald@stanford.edu (K. MacDonald).

¹ This problem is a simplified version of Quine's *indeterminacy of reference* (Quine, 1960): That there are many possible meanings for a word ("Gavagai") that include the referent ("Rabbit") in their extension, e.g., "white," "rabbit," "dinner." Quine's broader philosophical point was that different meanings ("rabbit" and "undetached rabbit parts") could actually be extensionally identical and thus impossible to tease apart.

naming events (Smith & Yu, 2008; Vouloumanos, 2008). For example, Smith and Yu (2008) taught 12-month-olds three novel words simply by repeating consistent novel word-object pairings across 10 ambiguous exposure trials. Moreover, computational models suggest that cross-situational learning can scale up to learn adult-sized lexicons, even under conditions of considerable referential uncertainty (Smith, Smith, & Blythe, 2011).

Although all cross-situational learning models agree that the input is the co-occurrence between words and objects and the output is stable word-object mappings, they disagree about how closely learners approximate the input distribution (for review, see Smith, Suanda, & Yu, 2014). One approach has been to model learning as a process of updating connection strengths between multiple word-object links (McMurray, Horst, & Samuelson, 2012), while other approaches have argued that learners store only a single word-object hypothesis (Trueswell, Medina, Hafri, & Gleitman, 2013). In recent experimental and modeling work Yurovsky and Frank (2015) suggest an integrative explanation: learners allocate a fixed amount of attention to a single hypothesis and distribute the rest evenly among the remaining alternatives. As the set of alternatives grows, the amount of attention allocated to each object approaches zero.

In addition to the debate about representation, researchers have disagreed about how to characterize the ambiguity of the input to cross-situational learning mechanisms. One way to quantify the uncertainty in a naming event is to show adults video clips of caregiver-child interactions and measure their accuracy at guessing the meaning of an intended referent (Human Simulation Paradigm: HSP [Gillette, Gleitman, Gleitman, & Lederer, 1999]). Using the HSP, Medina, Snedeker, Trueswell, and Gleitman (2011) found that approximately 90% of learning episodes were ambiguous (<33% accuracy) and only 7% were relatively unambiguous (>50% accuracy). In contrast, Yurovsky, Smith, and Yu (2013) found a higher proportion of clear naming events, with approximately 30% being unambiguous (>90% accuracy). Consistent with this finding, Cartmill et al. (2013) showed that the proportion of unambiguous naming episodes varies across parent-child dyads, with some parents rarely providing highly informative contexts and others' doing so relatively more often.²

Thus, representations in cross-situational word learning can appear distributional or discrete, and the input to statistical learning mechanisms can vary along a continuum from low to high ambiguity. These results raise an interesting question: could learners be sensitive to the ambiguity of the input and use this information to alter the representations they store in memory? In the current line of work, we investigated how the presence of referential cues in the social context might alter the ambiguity of the input to statistical word learning mechanisms.

Social-pragmatic theories of language acquisition emphasize the importance of social cues for word learning (Bloom, 2002; Clark, 2009; Hollich et al., 2000). Experimental work has shown that even children as young as 16 months prefer to map novel words to objects that are the target of a speaker's gaze and not their own (Baldwin, 1993). In an analysis of naturalistic parent-child labeling events, Yu and Smith (2012) found that young learners tended to retain labels that were accompanied by clear referential cues, which served to make a single object dominant in the visual field. And correlational studies have demonstrated strong links between early intention-reading skills (e.g., gaze following) and later vocabulary growth (Brooks & Meltzoff, 2005, 2008; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998). Moreover, studies outside the domain of language acquisition have shown that the presence of social cues: (a) produce better spatial learning of audiovisual events (Wu, Gopnik, Richardson, & Kirkham, 2011), (b) boost recognition of a cued object (Cleveland, Schug, & Striano, 2007), and (c) lead to preferential encoding of an object's featural information (Yoon, Johnson, & Csibra, 2008). Together, the evidence suggests that social cues could alter the representations stored during cross-situational word learning by modulating how people allocate attention to the relevant statistics in the input.

The goal of our current investigation was to ask whether the presence of a valid social cue – a speaker's gaze – could change the representations underlying cross-situational word learning. We used a modified version of Yurovsky and Frank (2015)'s paradigm to provide a direct measure of memory for alternative word-object links during cross-situational learning. In Experiment 1, we manipulated the presence of a referential cue at different levels of attention and memory demands. At all levels of difficulty, learners tracked a strong single hypothesis but were less likely to track multiple word-object links when a social cue was present. In Experiment 2, we replicated the findings from Experiment 1 using a more ecologically valid social cue. In Experiment 3, we moved to a parametric manipulation of referential uncertainty by varying the reliability of the speaker's gaze. Learners were sensitive to graded changes in reliability and retained more word-object links as uncertainty in the input increased. Finally, in Experiment 4, we equated the length of the initial naming events with and without the referential cue. Learners stored less information in the presence of gaze even when they had visually inspected the objects for the same amount of time. In sum, our data suggest that cross-situational word learners are quite flexible, storing representations with different levels of fidelity depending on the amount of ambiguity present during learning.

2. Experiment 1

We set out to test the effect of a referential cue on the representations underlying cross-situational word learning. We used a version of Yurovsky and Frank (2015)'s paradigm where we manipulated the ambiguity of the learning context by

² The differences in the estimates of referential uncertainty in these studies could be driven by the different sampling procedures used to select naming events for the HSP. Yurovsky et al. (2013) sampled utterances for which the parent labeled a co-present object, whereas Medina et al. (2011) randomly sampled any utterances containing concrete nouns. Regardless of these differences, the key point here is that variability in referential uncertainty across naming events exists and thus could alter the representations underlying cross-situational learning.

Download English Version:

<https://daneshyari.com/en/article/5039740>

Download Persian Version:

<https://daneshyari.com/article/5039740>

[Daneshyari.com](https://daneshyari.com)