UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
When close isn't enough: Semantic similarity does not facilitate cross-situational word-learning

Permalink
https://escholarship.org/uc/item/0fr8353p

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
LaTourrette, Alexander S
Yang, Charles
Trueswell, John

Publication Date
2022

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

When close isn't enough:
Semantic similarity does not facilitate cross-situational word-learning

Alexander LaTourrette (alatour@sas.upenn.edu)
Department of Psychology, 425 S. University Ave.
Philadelphia, PA 19104 USA

Charles Yang (charles.yang@ling.upenn.edu)
Departments of Linguistics, Computer Science, and Psychology, 3401-C Walnut Street
Philadelphia, PA 19104 USA

John Trueswell (trueswel@psych.upenn.edu)
Department of Psychology, 425 S. University Ave.
Philadelphia, PA 19104 USA

Abstract

Infants' earliest words are learned by observation of the referent world, but substantial research suggests such learning is highly error-prone. However, recent work suggests that even learners' incorrect guesses may fall within the correct meaning's semantic neighborhood—enabling learners to converge on the correct meaning across exposures. Here, we evaluate the semantic similarity of adults' hypothesized word meanings in a cross-situational word-learning task. We find evidence for a weak semantic neighborhood effect: incorrect guesses are judged as similar to correct meanings (Study 1). However, this effect is not associated with successful word-learning. While learners tend to provide similar, internally consistent guesses across exposures, their accurate guesses are not similar to their previous guesses (Study 2). Moreover, incorrect guesses similar to the target do not increase accuracy on the subsequent exposure (Study 3). These results suggest early word-learning is driven by cues available in-the-moment, not by gradual exploration of semantic space.

Keywords: word learning; reference; cross-situational; semantic similarity

Introduction

Infants learn their first words by mapping the sounds they hear to physically co-present referents (e.g., Tsui et al., 2019). However, infants' early word learning environments rarely afford just one potential mapping between a novel word and a referent. Instead, previous research estimated that infants have an average of around 8 types of objects in view at any given moment when interacting with caregivers (Clerkin et al., 2017). Thus, infants are likely to face a daunting challenge in mapping a novel word to its target referent.

To illustrate this challenge, previous work has used the Human Simulation Paradigm (HSP), in which adults guess word meanings from muted videos of caregiver-child interactions. These findings reveal that even adults struggle to accurately guess a caregiver's intended meaning from only the visual context of an utterance (Cartmill et al., 2013; Gillette et al., 1999; Medina et al., 2011). However, while adults are quite inaccurate overall, averaging approximately 25% accuracy for nouns and under 10% for verbs, a small subset of these caregiver interactions allowed a majority of learners to make accurate mappings. Indeed, exposure to even one of these high-quality "referential gem" videos was sufficient to drive success in a cross-situational word learning task (Medina et al., 2011). When learners viewed a single "gem" learning instance at the beginning of a series of 5 vignettes corresponding to that same word, they were more successful in guessing the meaning correctly not just on the first "gem" trial, but also on subsequent exposures, indicating they successfully maintained their initial hypothesis. In contrast, when adults viewed a series of vignettes without a "gem" exposure, they were less accurate and showed no improvement across trials. This suggests early word learning may succeed via a series of momentary insights, facilitated by these relatively rare moments of referential clarity.

From Junk to Gems?

Thus, as children encounter new words repeated across different contexts, they will often be incorrect in guessing their meanings. It is critical, then, what children do with these incorrect guesses: whether they are merely treated as red herrings, or if children can learn from these guesses—using them as stepping-stones to converge on the target meaning.

Indeed, recent work suggests that such a stepping-stone strategy might be viable. Johnson, Schalla, & Suanda (2021) found that in an HSP task examining concrete nouns, even learners' incorrect guesses tended to fall into the target noun's semantic neighborhood (e.g., a learner might guess "bread" when the target meaning is "sandwich"). To demonstrate this, a separate group of adults was shown a guess (e.g., "bread") and asked whether its corresponding target word ("sandwich") or another target word ("ball") was more similar in meaning. Strikingly, participants selected the guess's corresponding target meaning as more similar on 75% of trials. A similar effect was observed when participants were shown the target word (e.g., "sandwich") and asked to choose the most similar word from its corresponding guess ("bread") or a guess from an unrelated trial ("donut"). Thus, even incorrect guesses fell close to the target word's semantic neighborhood.

It remains an open question, however, whether learners use such a stepping-stone strategy to succeed in cross-situational

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. ©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

word learning. The strategy is quite intuitive: it would seem counter-productive if a child who associated bread with the word "sandwich" later rejected this guess and simply started from scratch, instead of guessing a new, bread-related meaning. Indeed, a similar kind of similarity-based fine-tuning of word meanings is evident in studies of polysemy (Floyd & Goldberg, 2021) and in older children's lexical development, as they begin to distinguish between words with similar meanings (Ameel et al., 2008; Saji et al., 2011).

However, the best evidence that word learners engage in similarity-based inferences in cross-situational word learning comes from Zhang et al. (2020), who presented adults with an HSP task featuring six consecutive vignettes referring to the same verb. As learners progressed through the vignettes, their answers became more semantically similar to the target, even if they never arrived at the correct answer (e.g., guessing "twist" instead of "turn"). While this suggests learners did home in on the target semantic neighborhood across trials, it is not clear whether learners who made semantically close incorrect guesses were then more successful in arriving at the target meaning. Additional exposures likely enabled learners to make better, more semantically appropriate guesses, resulting in an increase in semantic similarity to the target. However, it is unclear whether such semantically close guesses ultimately facilitated successful word learning.

Current Work

In three studies, we directly examine the viability of semantic similarity-based cross-situational word learning, in which incorrect guesses serve as stepping-stones to the target meaning. In Study 1, we assess the semantic similarity of incorrect guesses to target meanings in a non-cross-situational HSP study. As in Johnson et al. (2021), the HSP task simply asked adults to guess the word uttered in a single context, not to learn words across contexts. Notably, these contexts were all video-recorded utterances in homes, and the videos' target words included nouns and verbs. Whereas Johnson et al. (2021) found a semantic similarity effect for nouns used in laboratory play contexts, here we seek to extend this effect to more realistic and varied settings and to test whether similar effects emerge for nouns and verbs.

In Study 2, we turn to an HSP cross-situational word-learning task, in which participants learned words from a series of HSP exposures. Here, we ask whether learners' guesses do, in fact, tend to remain in the same semantic neighborhood across exposures. Furthermore, we test whether this internal consistency between guesses is beneficial for learners, resulting in more accurate guessing.

Finally, in Study 3, we directly test whether semantic similarity to the target word on one trial of the cross-situational word-learning task is associated with accuracy on the next trial. That is, when a learner's guess is close to the target meaning, is their next guess more likely to be correct?

By examining the role of semantic similarity on a trial-by-trial basis, we provide a clearer picture of the mechanisms that drive cross-situational word learning. If semantic similarity plays a role in cross-situational word learning, then learners' incorrect guesses should be similar to the target meaning (Study 1), as well as internally consistent across exposures (Study 2). Moreover, both this internal consistency (Study 2) and semantic similarity to the target (Study 3) should be associated with successful cross-situational word learning outcomes.

Study 1

To begin, we evaluated whether the incorrect guesses from a previously conducted HSP task (Medina et al., 2011) were semantically similar to their corresponding target meanings. In this HSP task, adults were not asked to learn words, simply to guess what word had been uttered in a particular context. For each target word, they viewed a 40 s, muted video of a caregiver-child interaction, with a beep indicating when the caregiver uttered the target word. All videos were filmed in the children's homes, with both nouns and verbs serving as target words. If even the adults' incorrect guesses about the uttered words tend to be in the target word's semantic neighborhood, then a new group of participants should reliably identify these guesses as semantically similar to their corresponding target words, across both lexical categories.

Method

Design. Participants evaluated the semantic similarity of the previous HSP learners' incorrect guess-target word pairings. The target word's lexical category was manipulated within subjects, with each participant evaluating guesses for both nouns and verbs. Following Johnson et al. (2021), we also assessed the guess-target similarity bidirectionally, assigning participants to either the Guess-Based or Target-Based condition (see Figure 1). In the Guess-Based condition, participants were shown the learner's guess (e.g., "yummy") and asked which of two target words was more similar to the guess: its corresponding target word ("bite"), or a target word from another trial (e.g., "book"). In contrast, participants in the Target-Based condition were shown the target word (e.g., "bite") and asked which of two guesses was more similar to it: the corresponding guess ("yummy"), or a guess from a different trial (e.g., "toy"). This resulted in a 2 (Lexical Category: Noun vs. Verb) x 2 (Directionality: Guess-Based vs. Target-Based) mixed design.

Participants. Three hundred English-speaking adults were recruited from two sources: 181 from a university participant pool who participated for course credit and 119 from Prolific who were paid $1.50 to participate in an 8-minute experiment. Data were collected through PCIbex (Zehr & Schwarz, 2018). An additional 31 participants were excluded for failing to accurately answer at least 3 of the 4 attention-check questions featuring high-similarity word pairs (e.g., identifying "sheep" as more similar to "goat" than "beak").

Materials. The previously conducted HSP task yielded 3184 incorrect guess-target word pairings, produced by 37 participants. Each guess was provided after the HSP learner viewed a single, muted video of a caregiver-child interaction, without feedback on their guess. All videos were constructed such that the target word occurred 30 s into the 40 s video and was marked by a beep. Target words were the 48 most common nouns and verbs (24 of each) in caregiver speech. While this HSP task used third-person recordings, which differ from a child's view in many ways, previous work has found little numerical difference in HSP accuracy with first- or third-person videos, indicating both can facilitate word learning (Yurovsky et al., 2013).

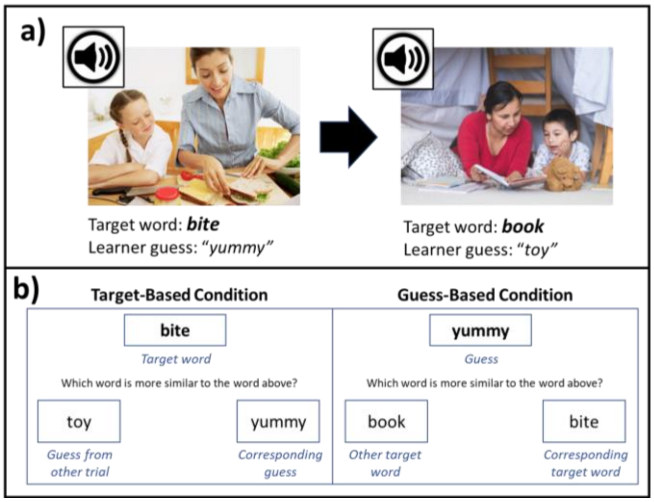

Figure 1. a) Design of original non-cross-situational HSP task: learners guessed a different target word on each trial, indicated only by a beep. For confidentiality, images do not depict actual stimuli. **b)** Study 1 Design. Sample trial for the learner guess shown in (a) (i.e., "yummy" guessed for "bite"), presented in Target-Based or Guess-Based condition. Blue text not shown in task.

Procedure. Participants were told they would see a series of words and be asked to "choose another word that they are similar to." Each participant evaluated 101 incorrect guess-target word pairings (e.g., "yummy"-"bite"), resulting in each pairing being evaluated approximately 5 times in each condition. In the Target-Based condition, participants were shown the target word, then asked to select the more similar word from a corresponding guess, or a guess from a different vignette (with a different target word). In the Guess-Based condition, participants were shown the guess, then selected between the corresponding target and another target word.

Guess-target pairings were shown in random order to each participant. Intermixed with these pairings, participants also completed 4 attention-check trials. These trials had the same structure as the others but featured highly semantically similar word pairs (i.e., sheep-goat, dive-swim, cough-hiccup, fence-gate) identified in previous work (Vigliocco et al., 2004). Participants who failed to correctly pair at least 3 of these 4 word pairs were excluded.

Preregistration. All analyses were preregistered on OSF (https://osf.io/2a4g3).

Results

Similarity Judgments. To test whether participants rated the incorrect guesses as similar to their targets, we constructed a binomial mixed effects model predicting the selection of the target-guess pairing, containing deviation-coded fixed effects of Lexical Category and Directionality and random effects of participant, participant-by-lexical category, target word, and target-word-by-directionality. All analyses reported here followed a maximal random effects structure, simplifying if needed for convergence, and fit with lme4 (Bates et al., 2014).

This analysis revealed a non-significant effect of Lexical Category, $\beta = .044$, SE = .077, $p = .57$, with target Verbs (M = .56, SD = .08) and Nouns (M = .55, SD = .08) similarly likely to be matched with their corresponding guesses. The effect of Directionality also failed to reach significance, $\beta = .038$, SE = .040, $p = .33$. However, a significant Lexical Category x Directionality interaction did emerge, $\beta = .39$, SE = .078, $p < .0001$, with target Nouns more successfully paired with guesses in the Guess-Based condition, while target Verbs were more successfully paired in the Target-Based condition. This effect was not predicted, nor especially large—all cell means fell between .53 and .58—but it suggests that the similarity of target-guess pairs was not always robustly evident to participants.

Crucially, however, when combining across conditions, participants showed a significantly above-chance (greater than 50%) tendency to pair learners' incorrect guesses with the corresponding target Noun ($\beta = .20$, SE = .070, $p = .0037$) or target Verb ($\beta = .25$, SE = .035, $p < .001$). Thus, participants' incorrect guesses were judged to be semantically similar to the target words.

Correlations with HSP Accuracy. Finally, it is possible that especially informative contexts (i.e., the "gems" which facilitate correct guesses) also lead learners to make incorrect guesses that are similar to the target meaning. To test this possibility, we conducted an exploratory analysis examining whether the accuracy of guesses on a vignette predicted the similarity of its incorrect guesses to the target. We constructed a binomial mixed effects model predicting the selection of the target-guess pairing and including a fixed effect of HSP Accuracy and random effects of participant, participant-by-accuracy, and target word.

We observed a significant effect of HSP Accuracy, $\beta = .40$, SE = .17, $p = .019$, with higher accuracy predicting more similar incorrect guesses. This suggests that rather than incorrect guesses being uniformly similar to the target meaning, this tendency is significantly stronger in more referentially transparent contexts.

Discussion

Overall, learners' incorrect guesses did tend to fall within the target meaning's semantic neighborhood. This suggests similarity-based word learning is, at least in principle, possible: after all, incorrect guesses must be similar to the target meanings if those incorrect guesses are to be helpful. Moreover, these results provide encouraging new evidence that this effect is present with both nouns and verbs, even when learners were not told the lexical category of the target word. This effect also appears to be stronger for vignettes on which other learners were likely to accurately guess the target word, suggesting a new way in which such "referential gems" might be especially useful to learners.

On the other hand, the overall semantic similarity effect observed here (M = .55, SD = .05) is also substantially smaller than that observed in Johnson et al. (2021) (M = .75, SD = .06) despite relying on the same similarity measure. This suggests that semantic similarity may be a weaker cue for learners in more varied environments. Perhaps the greater variety of target words in the current study, which went beyond words for concrete objects to frequent but less concrete nouns like "time" and verbs like "have," resulted in less obvious semantic pairings. Alternatively, home contexts may offer more semantically variable referents than laboratory play spaces, where caregivers and children might structure their interactions more: e.g., by choosing semantically similar toys to play with (a horse and a cow, rather than a horse and a yo-yo). While these differences clearly warrant further investigation, our results nevertheless suggest that the semantic similarity of incorrect guesses could, in principle, offer a viable word-learning strategy in children's everyday environments.

Study 2

Study 1 found that one-exposure guesses about a word's meaning from a given context are semantically similar to the target word meaning, even when the guess was incorrect. In Study 2, to test whether learners utilize this semantic similarity to learn words across situations, we re-analyzed the results of a previously conducted cross-situational word-learning study (Medina et al., 2011). This study also gave learners an HSP task, but in this case, learners heard a novel word, rather than a beep, in each vignette and viewed 5 vignettes for each of 12 target words. If learners make new guesses using their previous guess's semantic neighborhood, then each guess should be similar to its preceding one. To test whether learners show this kind of internal consistency, we asked participants to judge the similarity of learners' consecutive guesses. We also go one step further, assessing whether such internal consistency is associated with successful learning. If relying on previous guesses' semantic neighborhoods facilitates word learning, then accurate answers should be similar to the learner's previous guess.

Method

Design. Participants evaluated the semantic similarity of the previous HSP learners' guesses across trials. On each trial, participants were shown an HSP learner's guess on Trial N (e.g., "necklace") and asked whether it was more similar to that same learner's guess on Trial N-1 ("pretty"), or to a different learner's guess on Trial N-1 ("put").

Participants. We recruited 135 English-speaking adults from Prolific, paying each $1.50 to participate in an 8-minute experiment. An additional 8 participants were excluded for failing to accurately answer at least 3 of the 4 attention-check questions, as in Study 1.

Materials. As in Study 1, each guess was provided after the learner saw an HSP vignette. Here, however, learners heard a novel word at 30 s, as the parent uttered the target word. The same novel word was used across 5 vignettes corresponding to each of the 12 target words. Target words were presented interleaved. All target words were concrete nouns.

This previously conducted HSP task yielded 2640 trial-to-trial word pairings, produced by 64 participants. Each participant contributed at most 4 guess pairs from the 5 exposures per target word (comparing Trials 1-2, 2-3, 3-4, and 4-5). Notably, this included both correct and incorrect guesses. However, trials where learners guessed the same meaning as on the previous trial were dropped, as learners were presumably not generating new hypotheses.

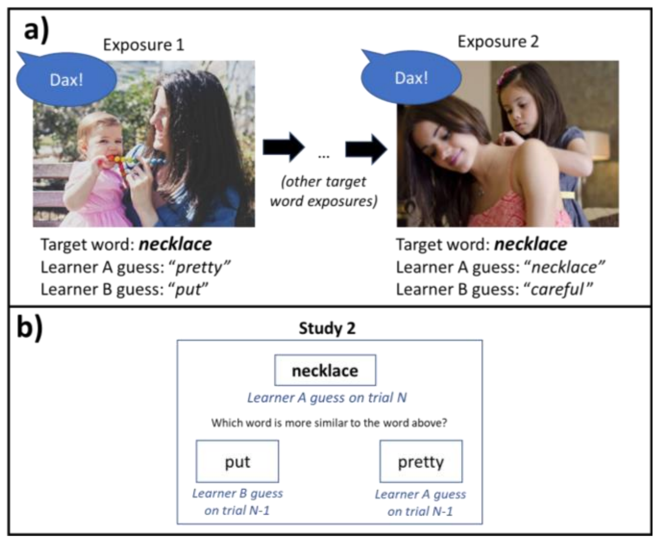

Figure 2. a) Design of original cross-situational HSP task: learners saw 5 vignettes per target noun (2 shown above). Vignettes were separated by 11 vignettes for other nouns. For confidentiality, images do not depict actual stimuli. **b)** Study 2 Design. Sample trial for learner A guesses above. Subjects chose if learner A's guess on Exposure 2 ("necklace," which happened to be correct), was more similar to learner A's previous guess ("pretty" on Exposure 1) or another learner's previous guess ("put" on Exposure 1).

Procedure. Procedure was largely identical to Study 1. Each participant evaluated 102 guess pairings from consecutive trials, resulting in each pairing being evaluated approximately 5 times. Participants were always shown the guess from Trial N and then asked to choose which of two words was more similar to it: the same learner's guess from

Trial N-1, or a different learner's guess from Trial N-1.

Preregistration. All analyses were preregistered on OSF (https://osf.io/mqvp7).

Results

Similarity Judgments. To analyze participants' similarity judgments, we constructed a binomial mixed effects model predicting the selection of the learner's previous guess. The intercept of this model was significantly greater than zero, β = .093, SE = .026, $p < .001$, indicating that participants were more likely than chance to identify a learner's current and previous guess as semantically similar (M = .52, SD = .04). This indicates learners' guessing was internally consistent: learners' guesses were more similar to their previous guess than to another learner's previous guess. However, the small size of the effect, with consecutive guesses identified as similar on only 52% of trials, indicates participants likely diverged from previous semantic neighborhoods with some frequency.

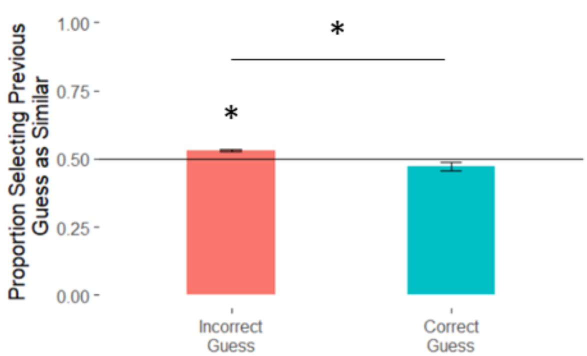

Figure 3. Study 2 Results. Subjects judged the learner's current guess to be similar to their previous guess more often when the current guess was incorrect than correct, $p < .001$. Only incorrect guesses were judged to be similar to their previous guesses, $p < .001$. Error bars represent +/- 1 SEM.

Association with Accuracy. Of course, internal consistency is not necessarily advantageous to learners. To test whether remaining in the same semantic neighborhood proved helpful to learners, we used a binomial mixed effects model to test whether the accuracy of learners' guesses predicted their semantic similarity to the learner's previous guess. This model yielded a significant effect of guess accuracy on semantic similarity, β = -.27, SE = .06, $p < .0001$; however, correct guesses were significantly *less* likely than incorrect guesses to be judged similar to the previous guess (see Figure 3). This result runs counter to the predictions of a similarity-based learning account. Indeed, only incorrect guesses were judged as similar to the learner's previous guess (M = .52, SD = .05, β = .12, SE = .03, $p = .0002$); correct guesses were not (M = .47, SD = .18, β = .0023, SE = .16, $p = .99$). Thus, while learners tended to be internally consistent, this tendency only emerged when their guesses were incorrect. Correct answers were unrelated to the previous semantic neighborhood.

Discussion

Learners in a cross-situational word-learning task showed reliable, albeit limited, internal consistency across exposures to a word, with new guesses remaining within the previous guess's semantic neighborhood. However, this internal consistency showed no signs of benefiting word-learning. In particular, correct guesses showed no semantic neighborhood effect at all, indicating such similarity-based inference was not associated with successful learning.

Of course, the directionality of this effect is unclear. Learners focusing too much on a previous guess's semantic neighborhood may have neglected cues in the present context that pointed to a correct, more semantically dissimilar meaning. Alternatively, learners may have resorted to similarity-based inferences more when the present context yielded no clear referent. In either case, there is little evidence that learners' tendency to rely on a previous guess's semantic neighborhood facilitated word learning.

Study 3

To provide a more direct test of the role of semantic similarity in cross-situational word-learning, we evaluated the most straightforward prediction of such an account: that incorrect guesses that are semantically similar to the target word should result in more accurate guesses on the following exposure. To test this, we used the same cross-situational HSP data as in Study 2, testing whether the semantic similarity of incorrect guesses to the target predicted accuracy on the next exposure.

Method

Design. Participants evaluated the semantic similarity of the HSP learners' guesses to the corresponding target words, as in Study 1. On each trial, participants were shown an HSP learner's guess and its target word, presented in either a Guess-Based or Target-Based condition, identical to Study 1 (see Figure 1). These similarity judgments were collapsed across conditions to provide a composite semantic similarity measure. We then asked whether guesses that were semantically similar to the target resulted in more accurate guesses on the next trial.

Participants. We recruited 60 English-speaking adults from Prolific, paying each $1.50 to participate in an 8-minute experiment. Another 9 subjects were excluded for failing at least 3 of the 4 attention-check questions, as in Study 1.

Materials. Learners in the cross-situational HSP task generated 3,186 incorrect answers, consisting of 2,389 unique guess-target word pairs (some pairs were given repeatedly or by multiple participants). Of this set, 1,755 pairs had already been evaluated by participants in Study 1 (e.g., "yummy"-"bite," see Figure 1). In these cases, we used the data collected in Study 1 to provide estimates of the guess-target semantic similarity. For the remaining 634 pairs, we collected judgments from new participants, run in the same paradigm as Study 1. By combining the two datasets, we ensured we averaged over 10 similarity judgments for each guess-target word pairing.

Procedure. Procedure was identical to Study 1. Each participant evaluated 102 guess-target pairings, in either a Guess-Based or Target-Based condition (see Figure 1).

Preregistration. All analyses were preregistered on OSF (https://osf.io/kagfn).

Results

We calculated each guess's semantic similarity to the target word as the proportion of participants who selected the guess and target word as being similar, across Guess-Based and Test-Based conditions. We then asked whether a guess's similarity predicted the accuracy of the original learner's subsequent guess in a binomial mixed effects model with random effects of original HSP learner, target word, and random slopes of similarity-by-learner and similarity-by-word. Critically, the previous guess's semantic similarity to the target did not significantly influence accuracy, β = .35, SE = .55, p = .52. That is, learners' closer-to-target guesses did not lead to more accurate responses on the next trial.

Although semantically close guesses may have failed to lead learners to the target meanings, perhaps learners were nonetheless gradually converging on these meanings across exposures. To test this possibility, we conducted a secondary analysis examining guesses' semantic similarity to the target across all 5 exposures. A linear mixed effects model predicting semantic similarity found no significant effect of the exposure number, β = .0040, SE = .0058, p = .50. Thus, learners' incorrect guesses did not become any closer to the target word over the course of learning.

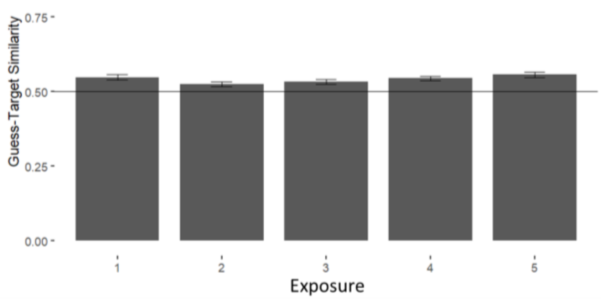

Figure 4. Study 3 Results. Learners' guesses did not significantly increase in similarity to the target across 5 exposures to the word, p = .5. Error bars represent +/- 1 SEM.

Discussion

These results provide no indication that learners successfully used semantically close incorrect guesses to home in on target meanings across exposures. Learners were not more accurate after close guesses than distant ones, and their guesses showed no gradual increase in similarity to the target. Of course, these results do not rule out the possibility that learners constrained new guesses to previous guesses' semantic neighborhoods. However, these results do suggest that such inferences did not systematically improve adults' word learning. Perhaps learners do employ this strategy—but it does not appear particularly successful.

General Discussion

As children hear new words repeated in different contexts, most of their hypothesized meanings for these words will inevitably be incorrect. Here, we evaluated the viability of a "stepping-stone" learning strategy by which learners could use these incorrect guesses to inform their next guess, eventually converging on the target meaning.

Results from Experiment 1 suggest such a strategy could, in principle, succeed: learner's incorrect guesses were judged to be semantically similar to the target meaning, although this effect was not as large as in previous work (cf. Johnson et al., 2021). In addition, Study 2 revealed that cross-situational word-learners did show some degree of internal consistency, with consecutive guesses also judged to be semantically similar. However, learners' *correct* guesses in this task were *not* found to be similar to the preceding guess, suggesting that learning was not facilitated by these similarity-based inferences. Indeed, a focus on previous guesses may have led learners astray, preventing them from identifying helpful cues in the current context. Finally, Study 3 evaluated the critical prediction that semantically close-to-target incorrect answers would facilitate subsequent learning—and found no evidence to support this prediction. Thus, while learners sometimes make guesses that are close to the correct meaning, they do not successfully capitalize on these moments to gradually converge on that correct meaning.

Of course, there may be multiple reasons for such a failure. Some similar meanings may co-occur too reliably to distinguish using only physical context (e.g., "computer" and "monitor"). In other cases, the guess may simply be synonymous with the target word (perhaps especially for verbs: e.g., "twist" and "turn"; Zhang et al., 2020). Finally, some types of similarity may make for better stepping-stones. For instance, taxonomically similar meanings (e.g., "horse" and "dog") might provide better stepping-stones than thematically similar meanings (e.g., "horse" and "saddle"). Future work should examine such similarity-based inferences in children's word-learning. After all, young children may be even less likely to engage in such a sophisticated, and largely unhelpful, word-learning strategy.

In sum, these results suggest learners' previous incorrect guesses do not play a large role in successful cross-situational word-learning. While incorrect guesses do tend to fall in the correct semantic neighborhood, this does not ultimately lead learners to the target meaning. Nor does a reliance on previous guesses' semantic neighborhoods serve learners well: correct guesses, unlike incorrect guesses, were not semantically similar to the learner's previous, incorrect guess. Thus, when learners correctly identify the meaning of a word, they may do so largely by relying on the cues available in that particular context, not by making a new guess based on incorrect guesses made beforehand.

References

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, *58*(2), 262–285. https://doi.org/10.1016/j.jml.2007.01.006

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*(7), 1–23.

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, *110*(28), 11278–11283. https://doi.org/10.1073/pnas.1309518110

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160055. https://doi.org/10.1098/rstb.2016.0055

Floyd, S., & Goldberg, A. E. (2021). Children make use of relationships across meanings in word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(1), 29–44. https://doi.org/10.1037/xlm0000821

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Johnson, S. C., Schalla, T., & Suanda, S. H. (2021, July 26-29). In the Right Neighborhood: Semantic Coherence in Parents' Ambiguous Naming Events. Poster presented at the 43rd Annual Meeting of the Cognitive Science Society.

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*(22), 9014–9019. https://doi.org/10.1073/pnas.1105040108

Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition*, *118*(1), 45–61. https://doi.org/10.1016/j.cognition.2010.09.007

Tsui, A. S. M., Byers-Heinlein, K., & Fennell, C. T. (2019). Associative word learning in infancy: A meta-analysis of the switch task. *Developmental Psychology*, *55*(5), 934–950. https://doi.org/10.1037/dev0000699

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, *16*(6), 959–966. https://doi.org/10.1111/desc.12036

Zehr, J., & Schwarz, F. (2018). *Penncontroller for internet based experiments (ibex).* https://doi.org/10.17605/OSF.IO/MD832

Zhang, Y., Amatuni, A., Cain, E., & Yu, C. (2020). Seeking Meaning: Examining a Cross-situational Solution to Learn Action Verbs Using Human Simulation Paradigm. *Proceedings of the Cognitive Science Society*, 7.

Acknowledgments

We thank the members of UPenn's Language Learning Lab for insightful discussion of these ideas, with special thanks to Lila Gleitman, Victor Gomes, and Daniel Swingley. The research reported here was supported by the Eunice Kennedy Shriver National Institute of Child Health and Development of the National Institutes of Health under award number F32HD103448-02. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.