



## OPEN ACCESS

## EDITED BY

Roberto Filippi,  
University College London, United Kingdom

## REVIEWED BY

Catherine Sandhofer,  
University of California, Los Angeles,  
United States  
Stanka A. Fitneva,  
Queen's University, Canada

## \*CORRESPONDENCE

Tanja C. Roembke  
✉ tanja.roembke@psych.rwth-aachen.de

RECEIVED 27 February 2023

ACCEPTED 09 June 2023

PUBLISHED 04 July 2023

## CITATION

Roembke TC, Simonetti ME, Koch I and  
Philipp AM (2023) What have we learned from  
15 years of research on cross-situational word  
learning? A focused review.  
*Front. Psychol.* 14:1175272.  
doi: 10.3389/fpsyg.2023.1175272

## COPYRIGHT

© 2023 Roembke, Simonetti, Koch and Philipp.  
This is an open-access article distributed under  
the terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# What have we learned from 15 years of research on cross-situational word learning? A focused review

Tanja C. Roembke\*, Matilde E. Simonetti, Iring Koch and  
Andrea M. Philipp

Chair of Cognitive and Experimental Psychology, Institute of Psychology, RWTH Aachen University,  
Aachen, Germany

In 2007 and 2008, Yu and Smith published their seminal studies on cross-situational word learning (CSWL) in adults and infants, showing that word-object mappings can be acquired from distributed statistics despite in-the-moment uncertainty. Since then, the CSWL paradigm has been used extensively to better understand (statistical) word learning in different language learners and under different learning conditions. The goal of this review is to provide an entry-level overview of findings and themes that have emerged in 15 years of research on CSWL across three topic areas (mechanisms of CSWL, CSWL across different learner and task characteristics) and to highlight the questions that remain to be answered.

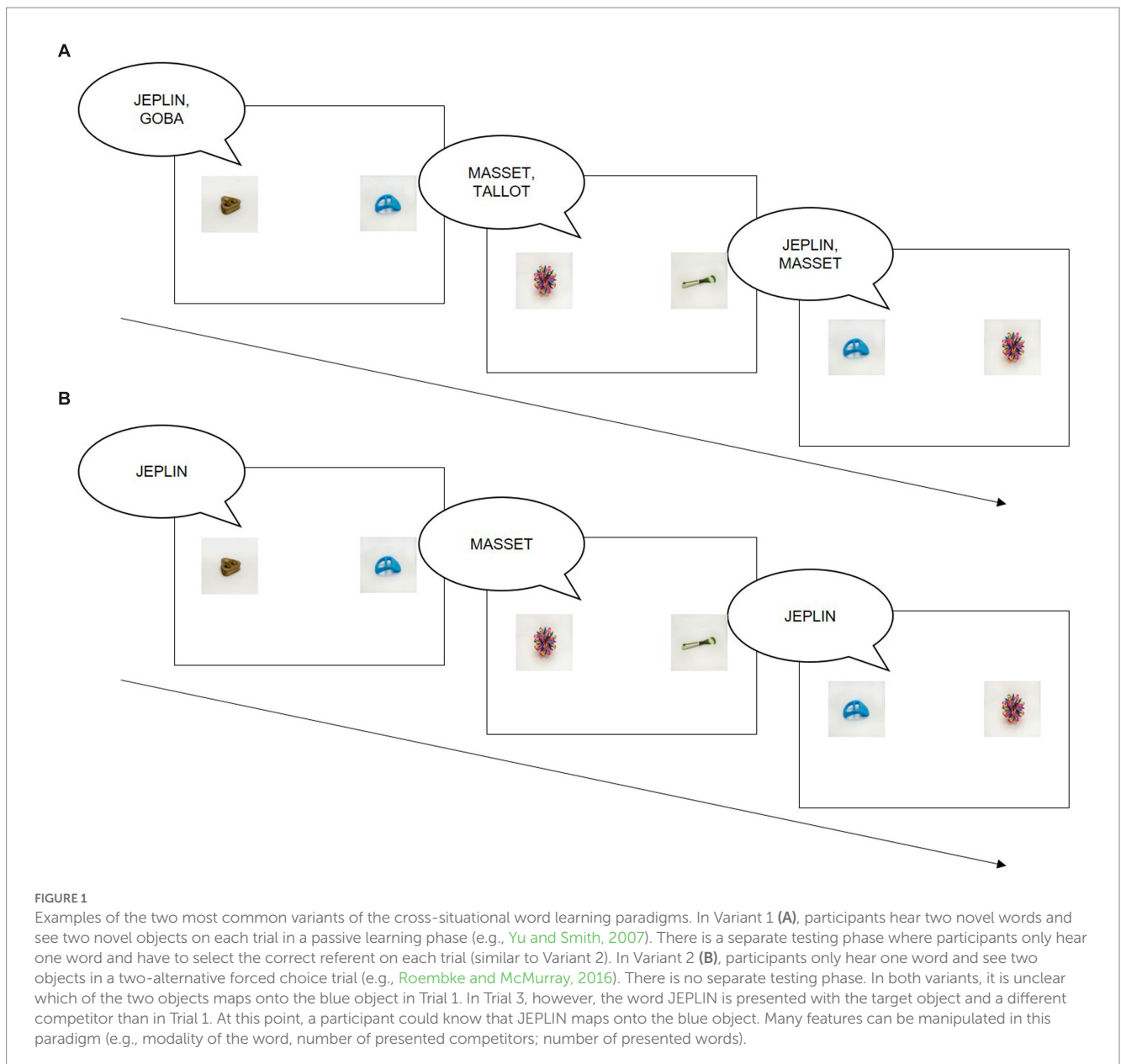
## KEYWORDS

cross-situational word learning, cross-situational statistical learning, word learning, language acquisition, review, referential ambiguity

## Introduction

As described by Quine (1960), any linguistic learning situation by itself may be ambiguous because a novel word form can map on many different present referents or even absent, abstract entities: When you hear the novel word GAVAGAI, you do not know if it refers to the bunny you see, its ears or the grass it is hopping on. However, as pointed out by Yu and Smith (2007; but see also Siskind, 1996; Akhtar and Montague, 1999), the so-called problem of referential ambiguity only exists if a single situation is considered by itself; across many time points, enough distributed statistics may be available to allow for the learning of the correct word-meaning-mappings under the assumption that a word and its meaning are more likely to co-occur than a word and other potential referents. That is, as you hear the word GAVAGAI over and over again, you may be able to extract what it means by learning what other objects or context it co-occurs with. Yu and Smith (2007) tested this hypothesis in adults and in 12- and 14-month-old infants (Smith and Yu, 2008), and found that all age groups were able to acquire the word-object-mappings based on co-occurrence statistics only. This type of word learning has since been termed cross-situational word learning (CSWL).<sup>1</sup>

<sup>1</sup> Cross-situational word learning can also be referred to as cross-situational learning, cross-situational statistical learning, statistical word learning or observational word learning. We will use the term cross-situational word learning throughout this manuscript.



In a typical CSWL experiment, participants hear one or more novel words and see several novel objects. While any trial by itself is ambiguous, the correct word-object-mappings can be learned by tracking word-object-co-occurrence over time. While a word and its target object are always present on the same trial, foil objects differ across trials. Participants are typically instructed that their task is to learn which object each word maps onto, but are not told that co-occurrence indicates a correct mapping (Yu and Smith, 2007). In some variations of the paradigm, participants complete a separate (passive) learning and (active) testing phase (e.g., Yu and Smith, 2007; Escudero et al., 2016c; Poepsel and Weiss, 2016; Figure 1A); in others, they always have to select an object on each trial (e.g., Trueswell et al., 2013; Dautriche and Chemla, 2014; Roembke and McMurray, 2016; Figure 1B). Final accuracy is typically assessed right after learning, though there is a small number of studies that also tested retention at a later time point (Vlach and Sandhofer, 2014; Vlach and DeBrock, 2019; Walker et al., 2020; McGregor et al., 2022). In each variation of

the CSWL paradigm, importantly, participants never receive feedback as to whether they selected the correct referent or not.

At the time of writing this review, Yu's and Smith's papers were cited 695 (2007) and 939<sup>2</sup> (2008) times, respectively, reflecting the wide interest in CSWL. The standard CSWL paradigm has been combined with other methods, such as eye-tracking (e.g., Fitneva and Christiansen, 2011, 2017; Yu and Smith, 2011; Yu et al., 2012; Trueswell et al., 2013; Roembke and McMurray, 2016), event-related potentials (e.g., Angwin et al., 2022; Mangardich and Sabbagh, 2022) or fMRI (Berens et al., 2018), and adapted to allow for more detailed trial-by-trial analyses of behavior (e.g., Suanda and Namy, 2012; Trueswell et al., 2013; Dautriche and Chemla, 2014; Roembke and

<sup>2</sup> These numbers were returned by a Google Scholar ([www.google.com/scholar](http://www.google.com/scholar)) search on May 12, 2023.

McMurray, 2016, 2021), where accuracy on a current trial is predicted by characteristics of preceding ones. CSWL has also inspired a number of computational models of word learning (e.g., Frank et al., 2009; Fazly et al., 2010; Vogt, 2012; Yu and Smith, 2012; Yurovsky and Frank, 2015; Blythe et al., 2016; Kachergis et al., 2017; Stevens et al., 2017; Bhat et al., 2022).

Now, approximately 15 years after the publication of Yu's and Smith's seminal papers, it is our goal to take stock of the literature, to identify the themes that have emerged in the research on CSWL and to highlight the questions that remain to be answered. We will also highlight differences between results from CSWL studies and unambiguous word learning studies (i.e., paradigms without referential ambiguity, such as the explicit pairing of a word with its meaning) when appropriate. The goal of this focused review is not to provide a comprehensive overview of the research that has been conducted on CSWL—this would be beyond the scope of this paper—but rather to provide a broad entry-level road map to past, present and future research on CSWL. As such, we hope that this overview will be helpful for both researchers that have already conducted research on CSWL but also those that are new to the field. We will review research within three sections: (1) Mechanisms of CSWL, (2) CSWL across different learner characteristics, and (3) CSWL across different task characteristics.

## Mechanisms of cross-situational word learning

In Yu and Smith's (2007) original study, CSWL was considered a type of statistical learning with an underlying associative, domain-general mechanism: Across trials, associations between co-occurring words and referents are gradually strengthened, whereas associations between non-co-occurring items remain weak (Figure 2A; Yu and Smith, 2007, 2012; Kachergis et al., 2012). One core assumption of this gradual associative mechanism is that people can maintain multiple hypotheses (e.g., between a word and its eventual target object as well as competitors) simultaneously. During learning, different mappings may then compete with each other, both within the same trial as well as across trials (Fazly et al., 2010; Yurovsky et al., 2013; Benitez et al., 2016). Another assumption of the gradual associative mechanism is that at least some of the learning is implicit with CSWL being a form of statistical learning (Frost et al., 2019; Weiss et al., 2020). In Yu and Smith's (2007) original study, for example, awareness was not necessary to acquire the word-object-mappings.

As an alternative, the propose-but-verify (hypothesis-testing) account was put forward (Figure 2B; Medina et al., 2011; Trueswell et al., 2013; Woodard et al., 2016). In this account, people are thought to only maintain one hypothesis about each word's meaning, which is verified and, if needed, revised. As a result, in this account, CSWL is more adequately described as a type of fast-mapping procedure than as a gradual associative one (Trueswell et al., 2013). In propose-but-verify, learning is generally thought to be more dependent on participants' awareness (i.e., to be explicit).

Several studies have tested the conflicting predictions that people can maintain multiple hypotheses per word (gradual associative account) or not (propose-but-verify/hypothesis-testing account). Evidence for propose-but-verify originally came from trial-by-trial

autocorrelation analyses showing that participants were at chance on a current trial if they had not picked the correct referent previously (Trueswell et al., 2013): If only one hypothesis is maintained, learning must be at chance if the proposed target is not available for selection on a current trial and a new target hypothesis has to be formed. However, since then, it has been shown across several different paradigm variants that multiple hypotheses can be maintained in parallel (e.g., Dautriche and Chemla, 2014; Yurovsky et al., 2014; Roembke and McMurray, 2016). For example, Yurovsky et al. (2014) first trained participants on a set of word-object-mappings in a CSWL task. Subsequently, they exposed participants to a set of new word-object-mappings as well as ones from the first training that had not been learned. Participants were found to be better at acquiring the words that they had received training on before, suggesting that they had retained some partial knowledge from the previous training besides having been at chance performance for these words then. Similarly, Roembke and McMurray (2016) observed that participants were more likely to look at object competitors that had been more frequently paired with the word than a baseline object, even as they clicked on the correct target object. This looking behavior is consistent with the maintenance and parallel in-the-moment activation of multiple mappings per word. Given that multiple hypotheses are tracked per word, one question that remains to be answered is how incorrect meanings may be unlearned to facilitate the activation of the correct referent (McMurray et al., 2012).

More recently, instead of seeing the two accounts as opposing, it was suggested that they may represent two distinct learning systems that work in parallel during CSWL. Here, the core of learning is associative and gradual, but the formation of explicit hypotheses and attention allocation can impact how associations are formed (McMurray et al., 2012; Yurovsky and Frank, 2015; Roembke and McMurray, 2016). Consistent with a mixed account, for example, Kachergis et al. (2014) found that adults were able to acquire word-object-mappings via CSWL even in the absence of an explicit effort to learn; at the same time, acquisition was also superior under explicit study instructions. The relative reliance on different learning mechanisms (e.g., gradual accumulation of statistics versus propose-but-verify/hypothesis-testing) may vary with a number of factors, including the degree of ambiguity, the number of unfamiliar words, familiarity of visual referents (e.g., whether they can be easily described or not), context (e.g., cover task or not) and task (e.g., task instructions, time pressure to respond) but also people's beliefs and confidence during learning (e.g., Wang and Mintz, 2018; Wang, 2020; Dautriche et al., 2021).

In addition, CSWL mechanisms may also differ across development: In younger children, there is evidence that only one hypothesis is maintained per word (Woodard et al., 2016; Aravind et al., 2018). At the same time, it has also been suggested that implicit, associative learning is actually *more* common in young children than adults (Ramscar et al., 2013). Evidence for a qualitative change in CSWL mechanism across development also comes from a study by Fitneva and Christiansen (2017). They manipulated to what extent three age groups (4-year-olds, 10-year-olds and young adults) were exposed to more initially incorrect (mismatched) word-object-mappings or not. Surprisingly, it was found that while 4-year-olds' CSWL benefitted from initial accuracy, young adults' learning was best if they were exposed to an incorrect mapping at first (no learning difference due to initial accuracy/inaccuracy was observed for

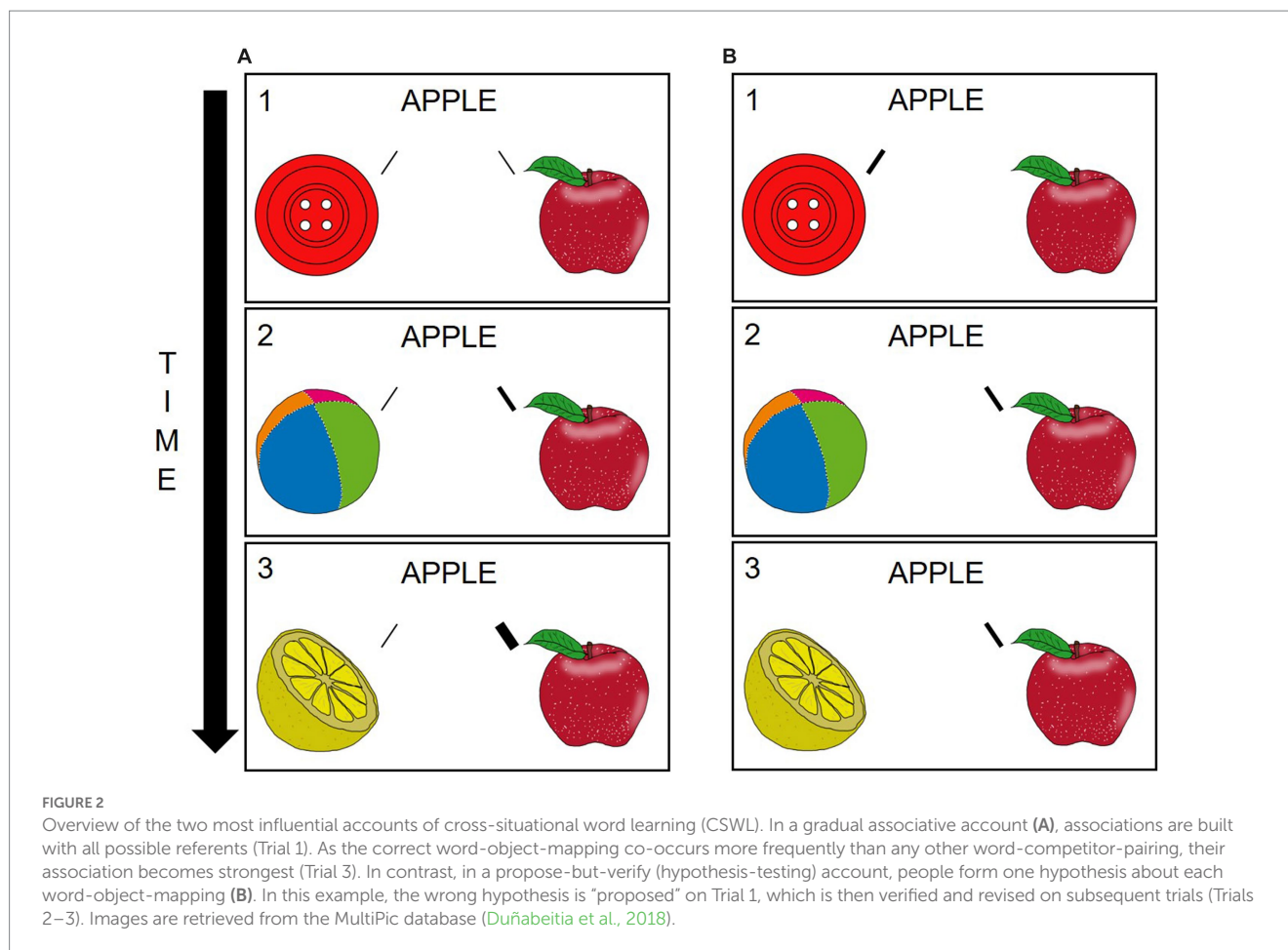


FIGURE 2

Overview of the two most influential accounts of cross-situational word learning (CSWL). In a gradual associative account (A), associations are built with all possible referents (Trial 1). As the correct word-object-mapping co-occurs more frequently than any other word-competitor-pairing, their association becomes strongest (Trial 3). In contrast, in a propose-but-verify (hypothesis-testing) account, people form one hypothesis about each word-object-mapping (B). In this example, the wrong hypothesis is "proposed" on Trial 1, which is then verified and revised on subsequent trials (Trials 2–3). Images are retrieved from the MultiPic database (Duñabeitia et al., 2018).

10-year-olds). Moreover, the learning benefit due to accurate (4-year-olds) and inaccurate (young adults) initial mappings was not specific to the manipulated items but applied to the whole set of to-be-learned words. While the exact mechanism behind the initial inaccuracy/accuracy benefit is unclear, it suggests an important role of memory and attention during CSWL (Fitneva and Christiansen, 2017).

Consistent with this interpretation, recent work by Vlach and DeBrock (2017, 2019) suggests that memory may be the best predictor of 2- to 6-year-olds' CSWL performance (more so than their age or vocabulary size). At the same time, in adults, neither working memory nor phonological short-term memory predicts overall CSWL (Walker et al., 2020). Yet, our understanding of CSWL across development is currently limited by not knowing when statistical word learning is most common: Do children learn words cross-situationally when they are infants and more limited in their ability to actively shape their environment (Kachergis et al., 2012; Smith et al., 2014)? Or is it the most common learning mechanism of new words in older children and adults who acquire most of their vocabulary by reading (Nagy et al., 1985, 1987)? Answering these questions may give us insights into the type of mechanisms that are most likely to be engaged during CSWL at different ages.

As mentioned previously, the experimental research on CSWL has also inspired a number of computational models, which have helped test the plausibility of CSWL in general (e.g., Blythe et al., 2010; Vogt, 2012) and different CSWL mechanisms more specifically (e.g., Frank et al., 2009; Fazly et al., 2010; Yu and Smith, 2012; Yurovsky and Frank,

2015; Kachergis et al., 2017; Stevens et al., 2017; Bhat et al., 2022). In fact, Bhat et al. (2022) recently identified 19 different models that range in their theoretical founding (gradual associative, propose-but-verify/hypothesis-testing or mixed), the input they take (e.g., symbolic stimuli) as well as their computational formalism (e.g., connectionist or Bayesian; see Bhat et al., 2022 for a detailed review of the different models), with the associative models by Kachergis et al. (2012, 2013, 2017) being highlighted as very successful at explaining data from several studies. However, many of these models (including the ones by Kachergis et al., 2012, 2013, 2017) fail to consider children's developing cognitive abilities (Vlach and DeBrock, 2019). One recent exception is the WOLVES, a dynamics field theory model by Bhat et al. (2022), which was able to model infants', toddlers' and adults' (looking) behavior across development.

To summarize, the two original competing accounts—gradual associative/statistical and propose-but-verify/hypothesis-testing—have been very helpful in framing research on CSWL, as they provided specific testable hypotheses. The resulting data suggest that CSWL is most accurately described by a mixed account that can incorporate findings that are more in line with gradual or statistical learning as well as ones in line with propose-but-verify (hypothesis-testing). Moving forward, taking into account the developmental time line of this learning represents an important new step. Moreover, most research on CSWL has been concerned with the mechanisms of how initial word-object-mappings are established but less so with how newly acquired meanings are retained (as reflected in the small



number of studies that test retention of word meanings, e.g., Walker et al., 2020; McGregor et al., 2022) and integrated in the broader lexicon.

## Cross-situational word learning across different learner characteristics

The plausibility of CSWL as a form of word learning is limited if it can only be shown in (neurotypical) adults. As such, Smith and Yu (2008) provided a litmus test by adapting the original adult paradigm that used accuracy as the dependent variable to an eye-tracking paradigm to be used with infants. As highlighted previously, they showed that 12- and 14-month-old infants were able to learn word-objects-mappings from cross-situational statistics after a short exposure. Since then, CSWL has been observed in children of different age groups (e.g., Smith and Yu, 2008; Scott and Fisher, 2012; Suanda et al., 2014; Bunce and Scott, 2017; Vlach and DeBrock, 2017; Roembke et al., 2018; Benitez et al., 2020; Crespo and Kaushanskaya, 2021; Mangardich and Sabbagh, 2022), children with developmental language disorder (DLD; Ahufinger et al., 2021; McGregor et al., 2022), children with autism (Venker, 2019; Hartley et al., 2020) and late talking children (Cheung et al., 2022) as well as older adults (Bulgarelli et al., 2021), adults with hippocampal amnesia (Warren et al., 2020) and aphasia (Peñaloza et al., 2017). CSWL was also effective when learning words in a second language (Hu, 2017; Tuninetti et al., 2020). For an exception in which no evidence for learning was found when testing CSWL in an isolated community in Papua New Guinea, the unfamiliarity with laboratory-based experiments of these participants is a likely explanation (Mulak et al., 2021). Overall, it is clear that language learners with different cognitive profiles can acquire word-object-mappings via distributed statistics as in CSWL, even as learning is typically lower in a CSWL than in a unambiguous word learning task where word and meaning are explicitly paired (Mulak et al., 2019). Nevertheless, there are (at least) two methodological limitations to many of these studies: (1) they often only assess learning right after exposure, leaving unclear how long-lasting the acquired representations are; and (2) because the goal of these studies was to show if CSWL was possible at all, the to-be-learned statistical relationships are very simple (i.e., low number of words to be learned; low referential ambiguity within each trial); thus, it is an open question whether CSWL would also be possible if more complex statistical relationships had to be tracked.

While the previous studies on different learners found performance above chance at the group level, there is often considerable variability at the individual level, sometimes to the extent that it is not clear whether successful CSWL is universal enough on a person to person basis to contribute meaningfully to an all-encompassing explanation of word learning. In young children, for example, it has been questioned whether CSWL is robust and whether its resulting representations are long-lasting (Vlach and Johnson, 2013; Aravind et al., 2018; Vlach and DeBrock, 2019). Similarly, adult participants with hippocampal amnesia (Warren et al., 2020) and aphasia (Peñaloza et al., 2017) were able to acquire words cross-situationally; however, their learning was at a slower rate than age-matched control groups. In addition, while learning was above chance at a group level/in

some participants<sup>3</sup>, this was not true for a substantial subset of the samples. For example, seven out of 16 participants with aphasia performed at chance in a simple CSWL task (Peñaloza et al., 2017)—the result that a majority did learn is impressive, but still calls into question why CSWL was not accessible to the non-learners.

Investigations into the impact of individual differences on CSWL are limited in number. Consistent with the idea that attention is an important determinant of whether statistical co-occurrences are encoded, differences in selective sustained attention can explain some variability in CSWL, with strong learners having fewer fixations with longer durations than weaker learners (Yu and Smith, 2011; Yu et al., 2012; Smith and Yu, 2013), which may reduce the number of incorrect associations that are encoded (Bhat et al., 2022). In addition, it has been hypothesized that a person's language skills drive CSWL ability (rather than the other way around; Vlach and DeBrock, 2017). Data from Scott and Fisher (2012) were interpreted to be consistent with this hypothesis, as they found that toddlers with larger vocabularies also performed better on more difficult CSWL tasks than children with smaller vocabularies. Investigating individual differences in CSWL ability is not straightforward; to our knowledge, there is currently no data on whether a typical CSWL task is a reliable measure of an individual's statistical word learning ability (versus group-level ability; see the general statistical learning literature on a discussion of this; e.g., Siegelman et al., 2017). In a typical CSWL experiment, participants are naïve to the existence of co-occurrence statistics; this would no longer be the case if they repeatedly participated in CSWL tasks.

Furthermore, CSWL does not only differ between individuals within a specific group but also between groups. Such differences in CSWL performance across groups are multi-faceted: For example, children with DLD (McGregor et al., 2022) were found to score lower than age-matched controls when learning words cross-situationally. Interestingly, for children with DLD, a performance gap emerged early on during the CSWL task, while later learning occurred at a similar rate as in the controls (McGregor et al., 2022). This suggests that initial encoding during CSWL may be the bottleneck, whereas the actual mapping of words onto meanings based on co-occurrence is not affected in children with DLD. At the same time, late-talking children learned words at the same rate as controls during CSWL training, but were worse at retention (Cheung et al., 2022). Finally, children with autism performed similarly to vocabulary-matched controls in a CSWL task, but were slower to pick the correct referents (Hartley et al., 2020). It is possible that some of these CSWL learning differences across groups contribute to observed language delays (e.g., smaller overall vocabularies), though a causal relationship is currently not well-established. Further research should also explore whether CSWL ability is an appropriate intervention target to mitigate a language delay (c.f., Alt et al., 2014).

Group differences are also observed between monolinguals and bilinguals: Consistent with results from unambiguous word learning paradigms, bilinguals outperformed monolinguals in some CSWL studies (Escudero et al., 2016c; Poepsel and Weiss, 2016; Crespo

<sup>3</sup> Due to the small number of amnesic participants, no group analyses were conducted by Warren et al. (2020).

et al., 2023) but not all (Benitez et al., 2016; Crespo and Kaushanskaya, 2021), suggesting that differences in language learning history may impact how easily words are acquired statistically. However, the circumstances under which bilinguals outperform monolinguals are currently not well understood, with some suggesting that a bilingual advantage may only exist when more complex word-object mappings are acquired (Poepsel and Weiss, 2016) or when words include multiple sources of phonological variability (Crespo et al., 2023). Given these mixed results, there is a need for more research on how bilingualism may affect statistical word learning. In this context, it is also important to consider how more complex mappings between words and objects (e.g., one word maps onto several meanings), which tend to be more common for bilinguals, can be acquired via CSWL (Poepsel and Weiss, 2016).

In summary, language learners with different cognitive profiles can acquire words cross-situationally. However, at this point, these studies often represent an existence proof (i.e., are the mappings learned at all?), and are thus limited to relatively simple learning situations. In addition, most of the studies are conducted with English-speaking participants (but see exceptions in subsequent section), leaving unclear to what extent statistical word learning may be common in other languages and cultures (but see Mulak et al. (2021) for an exception). While there is some suggestive evidence that a deficit in CSWL ability may contribute to language delay in some learners, the underlying mechanisms are not well-understood. As the relationship between individual differences and CSWL has been mostly investigated in young children, it is currently unclear to what extent it is predicted by individual differences in vocabulary or memory in older children and adults. Thus, there currently is no good understanding of what individual differences predict CSWL performance.

## Cross-situational word learning across different task characteristics

Another theme that has emerged in the research on CSWL is what characteristics within a learning context impact how easily a word is acquired. We will first review research on the impact of trial-by-trial characteristics on CSWL followed by a closer look at the impact of stimulus characteristics.

### Trial-by-trial characteristics

Different trial-by-trial characteristics like the amount of referential ambiguity (number of objects on the screen), number of to-be-learned words, how often each word is repeated, and the complexity of the word-object-mappings impact CSWL: Observed learning rate will be highest if there are only two visual referents on each screen and a small number of to-be learned one-to-one-mappings (i.e., each word maps onto one referent only) that are repeated often (e.g., Yu and Smith, 2007; Poepsel and Weiss, 2016; Roembke and McMurray, 2016). Some studies have combined the learning of cross-situational statistics with other cues, such as morphological (Finley, 2022) or social ones (Frank et al., 2013; MacDonald et al., 2017); these can facilitate how easy it is to learn or what type of information is encoded.

In a typical CSWL experiment, trials are randomized completely or pseudo-randomized to avoid direct repetitions of the same word within a larger block (e.g., Roembke and McMurray, 2016; Hartley et al., 2020; Escudero et al., 2022; Yip, 2022). Both children and adults are more likely

to be correct on a current trial if they completed a trial with the same word more recently, suggesting recency facilitating retrieval of previous learning episodes (Roembke and McMurray, 2016; Roembke et al., 2018). Research has found that children perform similarly for when CSWL trials are presented in massed order (i.e., no or few interleaved trials between repetitions) and interleaved order (i.e., many interleaved trials between repetitions) (Smith and Yu, 2013; Vlach and Johnson, 2013). At the same time, however, adults substantially benefited from massing (Benitez et al., 2020). In infants, massing may hurt CSWL, as it leads to the visual habituation to the referents which in turn then hurts their encoding (Smith and Yu, 2013). Adults' cross-situational learning is best if they can decide which objects are presented to them next, thus optimizing the order in which they receive information (Kachergis et al., 2013). Temporal order has also been found to matter on a smaller time-scale: within-each trial. Apfelbaum and McMurray (2017) observed that whether words and objects were presented synchronously or not during CSWL impacted how mappings were linked, as participants were more likely to form spurious incorrect associations with competitor objects when auditory and visual stimuli were present at the same time.

Another common feature of CSWL experiments is that each word is presented an equal number of times. However, it has been argued that such a uniform frequency distribution is not representative of the real world (Blythe et al., 2010; Vogt, 2012; Hendrickson and Perfors, 2019), where some words are much more likely to be encountered than others (Zipfian distribution). Thus, a recent study by Hendrickson and Perfors (2019) compared CSWL when both words and meaning were either distributed uniformly or not. Surprisingly, it was found that participants learned better in the non-uniform, Zipfian distribution than the uniform one, likely because they were able to use high-frequency words to disambiguate the meaning of low-frequency ones (Hendrickson and Perfors, 2019).

### Stimulus characteristics

In a typical CSWL study, words are newly generated non-words that have minimal overlap, follow legal phonotactics (e.g., FEP, DAX or GOBA for speakers of English) and are presented in spoken form by a native speaker (e.g., Fitneva and Christiansen, 2011; Vlach and Sandhofer, 2014; Roembke and McMurray, 2016). In addition, the vast majority of CSWL studies was conducted with English-speaking participants, thus implementing and investigating word characteristics that are common in alphabetic, Western languages (e.g., CVCV structure of to-be-learned words; c.f. Yip, 2022).

Work with minimal word pairs (e.g., BON/TON or DEET/DIT; Escudero et al., 2016b) showed that English-speaking adults and even infants encode fine phonological consonant and vowel detail when learning words via CSWL (Escudero et al., 2016a,b; Mulak et al., 2019). Yip (2022) also found that native speakers of Cantonese Chinese encoded phonological detail during CSWL but that tonal information was more critical to learning than other types of phonological information. It is not clear to what extent phonological familiarity and variability of words can facilitate CSWL as a downstream benefit of easier word form encoding (as has been observed in unambiguous word learning studies; e.g., Rost and McMurray, 2010): While it was found to be easier to map words than non-linguistic beeps onto objects for adults (but not for children; Roembke et al., 2018), phonotactic legality of words only had a small impact on CSWL, if at all (Dal Ben et al., 2022). Similarly, no learning benefit was observed when words were presented via multiple speakers versus a single one (Crespo and Kaushanskaya, 2021). Nevertheless,

words are more easily acquired cross-situationally if they are presented visually rather than auditorily, which is consistent with results from other unambiguous word learning paradigms (Escudero et al., 2022).

There is only very little work on how variability in the visual referents may impact cross-situational word learning. One exception is a study by Wang (2020) which reported that participants were more likely to use implicit learning mechanisms if objects could not be described easily. In addition, Monaghan et al. (2015) found that CSWL was robust even when objects moved on the screen (also see Walker et al., 2020; Rebuschat et al., 2021). It is currently not known whether shapes that are more closely related to already existing semantic representations are easier/harder to encode and/or how variability in the visual exemplars may impact how word-object-mappings are acquired statistically.

CSWL is mostly tested with non-words and novel objects as referents; thus, the non-words are created to mimic concrete basic-level nouns. However, there are also studies that explore how different types of words are acquired via CSWL. Recent investigations suggest that learning subordinate-level word meanings (e.g., the word DALMATIAN) may be more difficult during CSWL than the learning of basic-level ones (e.g., DOG; Wang and Trueswell, 2019, 2022). Scott and Fisher (2012) showed that toddlers could learn novel verbs with the help of distributed statistics. Similarly, Monaghan et al. (2015) found that nouns and verbs embedded in short “sentences” can be learned simultaneously cross-situationally in the absence of syntactic cues. In a new variant of the CSWL paradigm (Rebuschat et al., 2021), participants are exposed to multi-word utterances and complex scenes. Thus, participants’ task is to not only learn word-object-mappings but to extract both grammar and vocabulary from cross-situational statistics without receiving feedback. Despite the increased complexity, participants were able to learn the sentence-to-scene correspondences (Rebuschat et al., 2021). These findings have since been replicated and extended by Walker et al. (2020) who also tested retention after a 24h delay, observing performance improvements for some word types after consolidation. While these studies still do not contain the complexity of natural language acquisition, they represent an intriguing next step in showing that CSWL is not an artificial phenomenon limited to a small set of word types under low referential ambiguity (Walker et al., 2020; Rebuschat et al., 2021).

## Summary

To summarize, many different factors impact how easy it is to acquire a word during CSWL. The variables investigated range from the methodologically important (e.g., number of objects on screen/referential ambiguity which impacts chance level) to the potentially more theoretically interesting ones (e.g., variability in word exemplars). At the same time, it is surprising to see that some factors that have been robustly found to influence word learning in other paradigms (e.g., massing/interleaving for children, phonotactic familiarity) are potentially less important in CSWL. Our understanding of CSWL will further improve as future studies move beyond investigating the acquisition of isolated nouns of English-sounding words (Monaghan et al., 2015; Walker et al., 2020; Rebuschat et al., 2021; Yip, 2022).

## Conclusion

The basic phenomenon of CSWL is well-established across a high number of populations, stimuli and learning conditions. It is exciting

to see how far the field has come since Yu’s and Smith’s original seminal publications in 2007 and 2008, testing different mechanisms of CSWL and adding complexity to the CSWL paradigm at different levels of the learning process. However, there also remain questions that deserve more attention in the next 15 years of research (see Textbox 1 for a non-exhaustive list).

### TEXTBOX 1: Overview of some open questions on cross-situational word learning (CSWL).

1. When (i.e., at what age) is CSWL most common? How does CSWL differ across different age groups?
2. How does CSWL relate to other forms of word learning and word learning outcomes?
3. Is a deficit in CSWL implicated in language delay? Is CSWL a possible target for word learning interventions?
4. What individual differences predict CSWL performance?
5. To what extent do differences in learning experience, such as bilingualism, impact CSWL?
6. How does CSWL occur for more complex statistical relationships (e.g., multiple mappings per words)?
7. How are incorrect mappings unlearned as part of CSWL?
8. How does CSWL occur in languages other than English? Is it more common in some populations/languages than others?
9. How well are words retained that are acquired via CSWL? How quickly are they integrated into the broader lexicon?
10. How does CSWL relate to other types of implicit/statistical learning?
11. What is the role of awareness in CSWL?

## Author contributions

TR drafted the first version of the manuscript. MES, IK, and AMP commented on and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The research reported in this article was supported by the Deutsche Forschungsgemeinschaft (Grant No. 505754094) awarded to TCR and IK.

## Acknowledgments

We would like to thank two reviewers for their feedback on earlier versions of this manuscript. In addition, we would like to thank Alina Mummenhoff for helping with researching and annotating recent publications on cross-situational word learning.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.



## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ahufinger, N., Guerra, E., Ferinu, L., Andreu, L., and Sanz-Torrent, M. (2021). Cross-situational statistical learning in children with developmental language disorder. *Lang. Cogn. Neurosci.* 36, 1180–1200. doi: 10.1080/23273798.2021.1922723
- Akhtar, N., and Montague, L. (1999). Early lexical acquisition: the role of cross-situational learning. *First Lang.* 19, 347–358. doi: 10.1177/014272379901905703
- Alt, M., Meyers, C., Oglivie, T., Nicholas, K., and Arizmendi, G. (2014). Cross-situational statistically based word learning intervention for late-talking toddlers. *J. Commun. Disord.* 52, 207–220. doi: 10.1016/j.jcomdis.2014.07.002
- Angwin, A. J., Armstrong, S. R., Fisher, C., and Escudero, P. (2022). Acquisition of novel word meaning via cross situational word learning: an event-related potential study. *Brain Lang.* 229:105111. doi: 10.1016/j.bandl.2022.105111
- Apfelbaum, K. S., and McMurray, B. (2017). Learning during processing: word learning doesn't wait for word recognition to finish. *Cogn. Sci.* 41, 706–747. doi: 10.1111/cogs.12401
- Aravind, A., de Villiers, J., Pace, A., Valentine, H., Golinkoff, R. M., Hirsh-Pasek, K., et al. (2018). Fast mapping word meanings across trials: young children forget all but their first guess. *Cognition* 177, 177–188. doi: 10.1016/j.cognition.2018.04.008
- Benitez, V. L., Yurovsky, D., and Smith, L. B. (2016). Competition between multiple words for a referent in cross-situational word learning. *J. Mem. Lang.* 90, 31–48. doi: 10.1016/j.jml.2016.03.004
- Benitez, V. L., Zettersten, M., and Wojcik, E. (2020). The temporal structure of naming events differentially affects children's and adults' cross-situational word learning. *J. Exp. Child Psychol.* 200:104961. doi: 10.1016/j.jecp.2020.104961
- Berens, S. C., Horst, J. S., and Bird, C. M. (2018). Cross-situational learning is supported by propose-but-verify hypothesis testing. *Curr. Biol.* 28, 1132–1136.e5. doi: 10.1016/j.cub.2018.02.042
- Bhat, A. A., Spencer, J. P., and Samuelson, L. K. (2022). Word-object learning via visual exploration in space (WOLVES): a neural process model of cross-situational word learning. *Psychol. Rev.* 129, 640–695. doi: 10.1037/rev0000313
- Blythe, R. A., Smith, K., and Smith, A. D. M. (2010). Learning times for large lexicons through cross-situational learning. *Cogn. Sci.* 34, 620–642. doi: 10.1111/j.1551-6709.2009.01089.x
- Blythe, R. A., Smith, A. D. M., and Smith, K. (2016). Word learning under infinite uncertainty. *Cognition* 151, 18–27. doi: 10.1016/j.cognition.2016.02.017
- Bulgarelli, F., Weiss, D. J., and Dennis, N. A. (2021). Cross-situational statistical learning in younger and older adults. *Aging Neuropsychol. Cognit.* 28, 346–366. doi: 10.1080/13825585.2020.1759502
- Bunce, J. P., and Scott, R. M. (2017). Finding meaning in a noisy world: exploring the effects of referential ambiguity and competition on 2.5-year-olds' cross-situational word learning. *J. Child Lang.* 44, 650–676. doi: 10.1017/S0305000916000180
- Cheung, R. W., Hartley, C., and Monaghan, P. (2022). Multiple mechanisms of word learning in late-talking children: a longitudinal study. *J. Speech Lang. Hear. Res.* 65, 2978–2995. doi: 10.1044/2022\_JSLHR-21-00610
- Crespo, K., and Kaushanskaya, M. (2021). Is 10 better than 1? The effect of speaker variability on children's cross-situational word learning. *Lang. Learn. Dev.* 17, 397–410. doi: 10.1080/15475441.2021.1906680
- Crespo, K., Vlach, H. A., and Kaushanskaya, M. (2023). The effects of bilingualism on children's cross-situational word learning under different variability conditions. *J. Exp. Child Psychol.* 229:105621. doi: 10.1016/j.jecp.2022.105621
- Dal Ben, R., Souza, D. D. H., and Hay, J. F. (2022). Combining statistics: the role of phonotactics on cross-situational word learning. *Psicologia: Reflexao e Critica* 35:30. doi: 10.1186/s41155-022-00234-y
- Dautriche, I., and Chelma, E. (2014). Cross-situational word learning in the right situations. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 892–903. doi: 10.1037/a0035657
- Dautriche, I., Rabagliati, H., and Smith, K. (2021). Subjective confidence influences word learning in a cross-situational statistical learning task. *J. Mem. Lang.* 121:104277. doi: 10.1016/j.jml.2021.104277
- Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., et al. (2018). MultiPic: a standardized set of 750 drawings with norms for six European languages. *Q. J. Exp. Psychol.* 71, 808–816. doi: 10.1080/17470218.2017.1310261
- Escudero, P., Mulak, K. E., Fu, C. S. L., and Singh, L. (2016c). More limitations to monolingualism: bilinguals outperform monolinguals in implicit word learning. *Front. Psychol.* 7:1218. doi: 10.3389/fpsyg.2016.01218
- Escudero, P., Mulak, K. E., and Vlach, H. A. (2016a). Infants encode phonetic detail during cross-situational word learning. *Front. Psychol.* 7, 1–11. doi: 10.3389/fpsyg.2016.01419
- Escudero, P., Mulak, K. E., and Vlach, H. A. (2016b). Cross-situational learning of minimal word pairs. *Cogn. Sci.* 40, 455–465. doi: 10.1111/cogs.12243
- Escudero, P., Smit, E. A., and Angwin, A. J. (2022). Investigating orthographic versus auditory cross-situational word learning with online and laboratory-based testing. *Lang. Learn.* 73, 543–577. doi: 10.1111/lang.12550
- Fazly, A., Alishahi, A., and Stevenson, S. (2010). A probabilistic computational model of cross-situational word learning. *Cogn. Sci.* 34, 1017–1063. doi: 10.1111/j.1551-6709.2010.01104.x
- Finley, S. (2022). Morphological cues as an aid to word learning: a cross-situational word learning study. *J. Cogn. Psychol.* 35, 1–21. doi: 10.1080/20445911.2022.2113087
- Fitneva, S. A., and Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cogn. Sci.* 35, 367–380. doi: 10.1111/j.1551-6709.2010.01156.x
- Fitneva, S. A., and Christiansen, M. H. (2017). Developmental changes in cross-situational word learning: the inverse effect of initial accuracy. *Cogn. Sci.* 41, 141–161. doi: 10.1111/cogs.12322
- Frank, M. C., Goodman, N. D., and Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychol. Sci.* 20, 578–585. doi: 10.1111/j.1467-9280.2009.02335.x
- Frank, M. C., Tenenbaum, J. B., and Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Lang. Learn. Dev.* 9, 1–24. doi: 10.1080/15475441.2012.707101
- Frost, R., Armstrong, B. C., and Christiansen, M. H. (2019). Statistical learning research: a critical review and possible new directions. *Psychol. Bull.* 145, 1128–1153. doi: 10.1037/bul0000210
- Hartley, C., Bird, L. A., and Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition* 200:104265. doi: 10.1016/j.cognition.2020.104265
- Hendrickson, A. T., and Perfors, A. (2019). Cross-situational learning in a Zipfian environment. *Cognition* 189, 11–22. doi: 10.1016/j.cognition.2019.03.005
- Hu, C. F. (2017). Resolving referential ambiguity across ambiguous situations in young foreign language learners. *Appl. Psycholinguist.* 38, 633–656. doi: 10.1017/S0142716416000357
- Kachergis, G., Yu, C., and Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychon. Bull. Rev.* 19, 317–324. doi: 10.3758/s13423-011-0194-6
- Kachergis, G., Yu, C., and Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Top. Cogn. Sci.* 5, 200–213. doi: 10.1111/tops.12008
- Kachergis, G., Yu, C., and Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Front. Psychol.* 5:588. doi: 10.3389/fpsyg.2014.00588
- Kachergis, G., Yu, C., and Shiffrin, R. M. (2017). A bootstrapping model of frequency and context effects in word learning. *Cogn. Sci.* 41, 590–622. doi: 10.1111/cogs.12353
- MacDonald, K., Yurovsky, D., and Frank, M. C. (2017). Social cues modulate the representations underlying cross-situational learning. *Cogn. Psychol.* 94, 67–84. doi: 10.1016/j.cogpsych.2017.02.003
- Mangardich, H., and Sabbagh, M. A. (2022). Event-related potential studies of cross-situational word learning in four-year-old children. *J. Exp. Child Psychol.* 222:105468. doi: 10.1016/j.jecp.2022.105468
- McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbi-Kelm, T., et al. (2022). What children with developmental language disorder teach us about cross-situational word learning. *Cogn. Sci.* 46:e13094. doi: 10.1111/cogs.13094
- McMurray, B., Horst, J. S., and Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychol. Rev.* 119, 831–877. doi: 10.1037/a0029872
- Medina, T. N., Snedeker, J., Trueswell, J. C., and Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9014–9019. doi: 10.1073/pnas.1105040108
- Monaghan, P., Mattock, K., Davies, R. A. I., and Smith, A. C. (2015). Gavagai as gavagai does: learning nouns and verbs from cross-situational statistics. *Cogn. Sci.* 39, 1099–1112. doi: 10.1111/cogs.12186



- Mulak, K. E., Sarvasy, H. S., Tuninetti, A., and Escudero, P. (2021). Word learning in the field: adapting a laboratory-based task for testing in remote Papua New Guinea. *PLoS One* 16:e0257393. doi: 10.1371/journal.pone.0257393
- Mulak, K. E., Vlach, H. A., and Escudero, P. (2019). Cross-situational learning of phonologically overlapping words across degrees of ambiguity. *Cogn. Sci.* 43, e12731–e12719. doi: 10.1111/cogs.12731
- Nagy, W. E., Anderson, R. C., and Herman, P. A. (1987). Learning word meanings from context during normal reading. *Am. Educ. Res. J.* 24, 237–270. doi: 10.2307/1162893
- Nagy, W. E., Herman, P. A., and Anderson, R. C. (1985). Learning words from context. *Read. Res. Q.* 20, 233–253. doi: 10.2307/747758
- Peñalosa, C., Mirman, D., Cardona, P., Juncadella, M., Martin, N., Laine, M., et al. (2017). Cross-situational word learning in aphasia. *Cortex* 93, 12–27. doi: 10.1016/j.cortex.2017.04.020
- Poepsel, T. J., and Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition* 152, 9–19. doi: 10.1016/j.cognition.2016.03.001
- Quine, W. V. O. (1960). *Word and Object: An Inquiry Into the Linguistic Mechanisms of Objective Reference* Cambridge, MA, USA: The MIT Press.
- Ramscar, M., Dye, M., and Klein, J. (2013). Children value informativity over logic in word learning. *Psychol. Sci.* 24, 1017–1023. doi: 10.1177/0956797612460691
- Rebuschat, P., Monaghan, P., and Schoetensack, C. (2021). Learning vocabulary and grammar from cross-situational statistics. *Cognition* 206:104475. doi: 10.1016/j.cognition.2020.104475
- Roembke, T. C., and McMurray, B. (2016). Observational word learning: beyond propose-but-verify and associative bean counting. *J. Mem. Lang.* 87, 105–127. doi: 10.1016/j.jml.2015.09.005
- Roembke, T. C., and McMurray, B. (2021). Multiple components of statistical word learning are resource dependent: evidence from a dual-task learning paradigm. *Mem. Cogn.* 49, 984–997. doi: 10.3758/s13421-021-01141-w
- Roembke, T. C., Wiggs, K. K., and McMurray, B. (2018). Symbolic flexibility during unsupervised word learning in children and adults. *J. Exp. Child Psychol.* 175, 17–36. doi: 10.1016/j.jecp.2018.05.016
- Rost, G. C., and McMurray, B. (2010). Finding the signal by adding noise: the role of noncontrastive phonetic variability in early word learning. *Infancy* 15, 608–635. doi: 10.1111/j.1532-7078.2010.00033.x
- Scott, R. M., and Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition* 122, 163–180. doi: 10.1016/j.cognition.2011.10.010
- Siegelman, N., Bogaerts, L., and Frost, R. (2017). Measuring individual differences in statistical learning: current pitfalls and possible solutions. *Behav. Methods* 49, 418–432. doi: 10.3758/s13428-016-0719-z
- Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* 61, 39–91. doi: 10.1016/S0010-0277(96)00728-7
- Smith, L. B., Suanda, S. H., and Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends Cogn. Sci.* 18, 251–258. doi: 10.1016/j.tics.2014.02.007
- Smith, L. B., and Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106, 1558–1568. doi: 10.1016/j.cognition.2007.06.010
- Smith, L. B., and Yu, C. (2013). Visual attention is not enough: individual differences in statistical word-referent learning in infants. *Lang. Learn. Dev.* 9, 25–49. doi: 10.1080/15475441.2012.707104
- Stevens, J. S., Gleitman, L. R., Trueswell, J. C., and Yang, C. (2017). The pursuit of word meanings. *Cogn. Sci.* 41, 638–676. doi: 10.1111/cogs.12416
- Suanda, S. H., Mugwanya, N., and Namy, L. L. (2014). Cross-situational statistical word learning in young children. *J. Exp. Child Psychol.* 126, 395–411. doi: 10.1016/j.jecp.2014.06.003
- Suanda, S. H., and Namy, L. L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cogn. Sci.* 36, 545–559. doi: 10.1111/j.1551-6709.2011.01218.x
- Trueswell, J. C., Medina, T. N., Hafri, A., and Gleitman, L. R. (2013). Propose but verify: fast mapping meets cross-situational word learning. *Cogn. Psychol.* 66, 126–156. doi: 10.1016/j.cogpsych.2012.10.001
- Tuninetti, A., Mulak, K. E., and Escudero, P. (2020). Cross-situational word learning in two foreign languages: effects of native language and perceptual difficulty. *Front. Commun.* 5:602471. doi: 10.3389/fcomm.2020.602471
- Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition* 183, 181–191. doi: 10.1016/j.cognition.2018.10.025
- Vlach, H. A., and DeBrock, C. A. (2017). Remember Dax? Relations between children's cross-situational word learning, memory, and language abilities. *J. Mem. Lang.* 93, 217–230. doi: 10.1016/j.jml.2016.10.001
- Vlach, H. A., and DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 45, 700–711. doi: 10.1037/xlm0000611
- Vlach, H. A., and Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition* 127, 375–382. doi: 10.1016/j.cognition.2013.02.015
- Vlach, H. A., and Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cogn. Sci.* 38, 757–774. doi: 10.1111/cogs.12092
- Vogt, P. (2012). Exploring the robustness of cross-situational learning under Zipfian distributions. *Cogn. Sci.* 36, 726–739. doi: 10.1111/j.1551-6709.2011.1226.x
- Walker, N., Monaghan, P., Schoetensack, C., and Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: an individual differences approach. *Lang. Learn.* 70, 221–254. doi: 10.1111/lang.12395
- Wang, F. H. (2020). Explicit and implicit memory representations in cross-situational word learning. *Cognition* 205:104444. doi: 10.1016/j.cognition.2020.104444
- Wang, F. H., and Mintz, T. H. (2018). The role of reference in cross-situational word learning. *Cognition* 170, 64–75. doi: 10.1016/j.cognition.2017.09.006
- Wang, F. H., and Trueswell, J. C. (2019). Spotting Dalmatians: Children's ability to discover subordinate-level word meanings cross-situationally. *Cogn. Psychol.* 114:101226. doi: 10.1016/j.cogpsych.2019.101226
- Wang, F. H., and Trueswell, J. (2022). Being suspicious of suspicious coincidences: the case of learning subordinate word meanings. *Cognition* 224:105028. doi: 10.1016/j.cognition.2022.105028
- Warren, D. E., Roembke, T. C., Covington, N. V., McMurray, B., and Duff, M. C. (2020). Cross-situational statistical learning of new words despite bilateral hippocampal damage and severe amnesia. *Front. Hum. Neurosci.* 13:448. doi: 10.3389/fnhum.2019.00448
- Weiss, D. J., Schwob, N., and Lebkuecher, A. L. (2020). Bilingualism and statistical learning: lessons from studies using artificial languages. *Biling. Lang. Cogn.* 23, 92–97. doi: 10.1017/S1366728919000579
- Woodard, K., Gleitman, L. R., and Trueswell, J. C. (2016). Two- and three-year-olds track a single meaning during word learning: evidence for propose-but-Verify. *Lang. Learn. Dev.* 12, 252–261. doi: 10.1080/15475441.2016.1140581
- Yip, M. C. W. (2022). Cross-situational word learning of Cantonese Chinese. *Psychon. Bull. Rev.* doi: 10.3758/s13423-022-02217-7
- Yu, C., and Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychol. Sci.* 18, 414–420. doi: 10.1111/j.1467-9280.2007.01915.x
- Yu, C., and Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Dev. Sci.* 14, 165–180. doi: 10.1111/j.1467-7687.2010.00958.x
- Yu, C., and Smith, L. B. (2012). Modeling cross-situational word-referent learning: prior questions. *Psychol. Rev.* 119, 21–39. doi: 10.1037/a0026182
- Yu, C., Zhong, Y., and Fricker, D. (2012). Selective attention in cross-situational statistical learning: evidence from eye-tracking. *Front. Psychol.* 3:148. doi: 10.3389/fpsyg.2012.00148
- Yurovsky, D., and Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition* 145, 53–62. doi: 10.1016/j.cognition.2015.07.013
- Yurovsky, D., Fricker, D. C., Yu, C., and Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychon. Bull. Rev.* 21, 1–22. doi: 10.3758/s13423-013-0443-y
- Yurovsky, D., Yu, C., and Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cogn. Sci.* 37, 891–921. doi: 10.1111/cogs.12035