Extraneous visual noise facilitates word learning

Katherine E. Twomey (k.twomey@lancaster.ac.uk), Lizhi Ma (malizhi110@hotmail.com)
Gert Westermann (g.westermann@lancaster.ac.uk)

Department of Psychology, Lancaster University
Lancaster, LA1 4YF UK

Abstract

Variability is important to learning; however, whether it supports or hinders language acquisition is unclear. 3D object studies suggest that children learn words better when target objects vary, however storybook studies indicate that contextual variability impairs learning. We tested a dynamic systems account in which background variability should boost learning by speeding the emergence of new behaviors. Two groups of two-year-old children saw arrays of one novel and two known objects on a screen, and heard a novel or known label. Stimuli were identical across conditions, with the exception that in the constant condition objects appeared on a white background, and in the variable condition backgrounds were colored. Only children in the variable condition showed evidence of word learning, suggesting that extraneous variability supports learning by decontextualizing representations, and indicating that adding low-level entropy to the developmental system can trigger a change in behavior.

Keywords: word learning; language acquisition; variability; memory decontextualization; dynamic systems

Children’s early word learning has long fascinated researchers. When a child hears the new word *spaceship*, linking it with a new toy flying machine rather than their toy dog – or indeed the flying machine’s wings, its color, the way it moves, and so on – seems to pose little problem. Given that the space of potential referents is theoretically infinite (Quine, 1960), this ability to quickly map a novel word to a novel object is impressive; indeed, robust referent selection has been observed in children as young as 18 months (Carey & Bartlett, 1978; Halberda, 2006; Houston-Price, Plunkett & Harris, 2005; Markman & Wachtel, 1988).

However, there is mounting evidence that a single episode of referent selection is not sufficient for full word learning; rather, children learn word-object associations incrementally, forming in-the-moment mappings between labels and objects and strengthening memories of these mappings across repeated encounters via cross situational learning (Horst & Samuelson, 2008; Smith & Yu, 2008; Yurovsky, Fricker, Yu, & Smith, 2014). Clearly, then, memory and language are linked from very early in development (Taylor, Liu, & Herbert, 2016): learning a new word depends critically on children’s ability to form and retain word-object associations. Consequently, the field has recently focused on the multiple factors that affect children’s ability to retain word-object mappings, demonstrating that referent selection and word learning are flexible, even fragile processes which depend heavily on the temporal and visual availability of information in the learning environment, for example repetition, competition, and timing (e.g., Arias-Trejo & Plunkett, 2010; Horst, Scott, & Pollard, 2010; Mather & Plunkett, 2009).

Developmental research has demonstrated that variability of to-be-learned items is a key influencing factor in early learning. For example, visual variability encountered across stimuli facilitates categorization in 6- to 7-month-old infants (Quinn & Bhatt, 2010), and phonological variability in affect or speaker has been shown to support early word recognition (Rost & McMurray, 2009). Recent work has revealed a similar effect of variability on word learning: when shown a novel 3D object category with exemplars that varied in color, 30-month-old children learned category labels, but did not when exemplars were identical, or varied in shape and color simultaneously (Twomey, Ranson, & Horst, 2014). Thus, while some target variability supports word learning, too much variability appears to disrupt it.

In addition to target variability there is good theoretical reason to expect extraneous, non-target variability – entropy – to support word learning. Evidence from adult problem-solving studies suggests that introducing entropy to a task facilitates learning. For example, adults solving a series of gear system problems presented on a computer screen learned a short-cut solution faster when the task contained entropy in the form of variability in spatial location of the stimuli than when stimuli were presented in a consistent spatial location (Stephen, Dixon, & Isenhower, 2009). On dynamic systems theories of cognition and development, cognitive structure emerges from the dynamic interactions of multiple, coupled components including the learner’s body, learning history and in-the-moment characteristics of the task (Thelen & Smith, 1996). Cognitive structure is instantiated as a stable state (“attractor”) in the behavior of this complex system. Dynamic systems of this type exhibit “phase shifts” from one attractor to another, resulting in qualitative and quantitative changes in the system’s behavior. Because phase shifts result in behavioral change, from the dynamic systems perspective, they index learning. As Stephen et al. (2009) demonstrate, extraneous entropy during learning destabilizes attractor states, speeding the onset of a phase shift. During development, then, non-target variability should speed up learning by helping new cognitive structure emerge via a shift from one behavioral state to another.

Despite this strong theoretical prediction, evidence for an effect of non-target variability in early learning is mixed. The categorization literature suggests that non-target variability helps learning. For example, Goldenberg & Johnson (2015) presented 16- to 20-month-old infants with a looking time task. Children saw novel category exemplars
on backgrounds which (a) repeated, (b) varied randomly, or (c) varied within interleaved blocks. Only infants who saw backgrounds which varied in interleaved blocks correctly generalized category labels at test. In contrast, the word learning literature suggests that lack of contextual variability supports word learning: when learning words from a storybook, repeating the context in which 3-year-old children encounter novel words by reading from the same book repeatedly boosts word retention relative to teaching children the same novel words from multiple different books (Horst, Parsons & Bryan, 2011; Williams & Horst, 2014). More broadly, however, the prediction from dynamic systems theory that additional entropy should boost word learning in a single task has yet to be explicitly tested. Critically, if background variability helps children learn words, this would provide evidence for continuity in the low-level mechanisms driving learning, from toddlerhood to adulthood. The current study addressed this gap by presenting children with a word learning task in which objects appeared either on a white background or on multiple colored backgrounds. We selected two-year-old children in line with previous research which demonstrates this age group’s success in similar looking-based referent selection tasks (Bion, Borovsky & Fernald, 2013). On the dynamic systems account, children in the variable color condition should show stronger retention of label-object associations than children in the constant color condition.

Method

Participants

Thirty typically developing, monolingual English-learning two-year-old children (14 girls, M = 22.77 months, SD = 1.87 months; range = 20.0 – 26.0 months) with a mean productive vocabulary of 176.04 words (SD = 117.50 words, range = 4 – 413 words) and no family history of colorblindness participated. Half of the children were randomly assigned to the constant color condition, and half to the variable color condition. Children’s ages and productive vocabularies were the same in either condition (p > .30). Data from six children were excluded due to fussiness (1), parental interference (3), bilingualism (1), and an eye tracker sample rate of under 25% (1). Parents were reimbursed for travel expenses and children received a small gift for participating.

Stimuli

Each child saw a warm-up, referent selection and test phase. Critically, stimuli for each phase were identical across conditions with the exception that during warm-up and referent selection in the variable color condition objects appeared on colored backgrounds, and in the constant color condition backgrounds were always white. Children also saw engagement and attention-getting stimuli. Overall, warm-up, referent selection and retention stimuli were videos containing 2D photographic images of known and/or novel objects (depicted in Fig. 1).

<table>
<thead>
<tr>
<th>Set</th>
<th>Known 1</th>
<th>Known 2</th>
<th>Novel</th>
<th>Novel label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>apple</td>
<td>spoon</td>
<td>zorch</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>banana</td>
<td>cup</td>
<td>tife</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>clock</td>
<td>block</td>
<td>blick</td>
<td></td>
</tr>
</tbody>
</table>

Fig 1. Object depicted in the current study.

Known objects were an apple, a ball, a banana, a car, a cup and a fork, and were selected because their labels are familiar to children of this age group (Fenson et al., 1993). Novel objects were a purple, green and black foam rocket (labeled zorch), a spherical yellow object with multiple flexible legs capped with pink and green balls (labeled tife), and a blue kazoo with raised orange spots (labeled blick), selected from an online database of objects unfamiliar to children of this age (NOUN Database; Horst & Hout, 2015). Each trial consisted of a single video of three objects. Videos were created in Microsoft Powerpoint 2010, and converted to .avi format using Microsoft Windows Live Movie Maker 2011. Each video was accompanied by embedded audio consisting of the same female speaker saying Can you find the [label]? Look at the [label]! Where’s the [label]? , as well as sound effects to keep children engaged in the task. Known labels were the appropriate English labels for those objects, and novel labels were blick (kazoo), tife (legs/balls) and zorch (rocket), selected as plausible but unfamiliar English object names. Auditory stimuli commenced 5 s after the start of each trial. First label onsets occurred from 0.78 – 0.90 s after the beginning of the auditory stimulus and offsets from 1.27 to 1.58 s; second label onsets from 2.20 – 2.48 s and offsets from 2.65 – 3.25 s; and third label onsets from 3.54 s – 4.21 s and offsets from 4.22 s to 5.19 s.

Engagement. Engagement stimuli consisted of a 7 s video of a female experimenter on a white background, smiling and saying Hello! Let’s play a game! Can you find what I’m looking for? in child-directed speech.

Warm-up. Warm-up stimuli were 16 s videos, each depicting a set of three of the known objects, designed to familiarize children with the task. In the first 0.5 s, a small colored rectangle appeared in the middle of a black screen and spun in an anticlockwise circle, expanding until it filled the whole screen, at which point it became the background on which the objects would appear. In the constant color condition, the background on each of the three warm-up trials was white. In the variable color condition, the background was blue, green, pink, purple or red. In the next 2 s the three objects appeared in the top left-hand corner of the screen and bounced diagonally downwards accompanied by a boing sound, coming to a rest in the center of the screen and remaining there for 9.5 s, during which time the
target object was labeled three times (e.g., Can you find the apple? Look at the apple! Where’s the apple!). During the next 3 s the target object rotated accompanied by a twinkling sound, followed by ostensive auditory feedback (e.g., There’s the apple!). In the final 1 s the objects bounced diagonally towards the bottom right hand corner and offscreen, accompanied by the sound of children cheering.

Referent selection. Referent selection trials were 13 s long, and identical to warm-up trials with the exception that children saw one novel and two known stimuli, and there was no ostensive feedback phase. Background colors were either white (constant color) or pseudorandom (variable color), as in the warm-up trials. Object location was pseudorandomized.

Retention. Retention trials were 9.5 s long and proceeded in an identical manner to referent selection trials except that the background was always gray and appeared immediately (i.e., there was no 0.5 s period where the background appeared) and all three objects were novel. Each object was labeled on two trials.

Procedure and Design

Before the experiment began the experimenter showed caregivers pictures of the known and novel objects to ensure they were appropriately known and novel to the child. All children were familiar with the known objects and unfamiliar with the novel objects. Caregivers were asked to complete a UK adaptation (Hamilton, Plunkett, & Schafer, 2000) of the MacArthur-Bates Communicative Development Inventory (Fenson et al., 1994), a vocabulary inventory commonly used to score toddlers’ receptive and productive vocabulary. Caregivers completed the vocabulary inventory either before or after the experiment, depending on the child’s level of engagement.

The eyetracking session took place in a quiet, dimly-lit room. Children sat on their caregiver’s lap 50-70 cm in front of a 21.5” 1920 x 1080 computer screen. Beneath the screen a Tobii X120 eyetracker recorded the child’s gaze location at 17 ms intervals, and a video camera above the screen recorded the caregiver and child throughout the procedure. Caregivers were instructed not to interact with their child or look at the screen during the task to avoid biasing their child’s behavior, and were asked to sit at a 90° angle from their child to ensure the eyetracker tracked the child’s eyes only.

The eyetracker was first calibrated using a five-point infant calibration procedure available in Tobii Studio. Immediately following calibration, children saw the engagement stimulus once.

Warm-up. The three warm-up trials immediately followed the engagement stimulus. The warm-up phase in each condition was identical with the exception that in the constant variable color condition, backgrounds were multiple, uniform colors, while in the constant color condition, backgrounds were white. Which objects appeared, which served as targets, and left-right positioning of objects were pseudorandomized across children such that no object appeared on more than two successive trials.

Referent selection. Fifteen referent selection trials immediately followed the warm-up phase. An example referent selection phase for the variable color condition is depicted in Fig. 1. Again, the corresponding warm-up phase in the constant color condition was identical with the exception that backgrounds were white. Referent selection trials were presented in three blocks of five trials for each set. Sets were kept constant across trials to maximize children’s retention of novel labels (Axelsson & Horst, 2014); thus, one child might see a block of five repetitions of the apple + fork + zorch set, followed by the banana + cup + tife set, and finally the car + ball + blick set, with block order Latin square counterbalanced across children. In each referent selection block children were asked to look at a known object on two trials and a novel object on three trials. Known/novel trial order and background color (variable color condition only) was pseudorandomized such that no more than two of the same trial type appeared in succession.

During referent selection an attention getting stimulus appeared six times pseudorandomly such that it was always succeeded by at least one referent selection trial, and consisted of a 3 s video of the speaker saying What’s next?. Finally, after the referent selection phase, children saw a 5 s “Well done” video of the speaker saying Well done! All finished! See you soon!

Break. Following referent selection, children took a five-minute break. During this time they either remained on their caregiver’s lap and watched an age-appropriate animation or moved to a seating area in the same room and colored pictures from a book.

Warm-up. After the break children saw a further warm-up trial, presented on a gray background.

Test. Three memory recativtaion and three retention trials immediately followed the warm-up trial, each depicting the three novel objects seen during referent selection. Trial order and object location were pseudorandomized.

Coding and data cleaning. Left, middle and right AOIs were square and centered on each object’s stationary position after they had bounced into the screen. Unreliable/offscreen and non-AOI looks were discarded,
resulting in a final dataset of 115,762 referent selection and 61,247 test gaze samples. Individual gaze samples were numerically coded (1 = target look, 0 = non-target look), creating a raw looking time measure, which was further collapsed into 100ms time bins for statistical tractability. All subsequent analyses use this target looking measure, and are standardized from the offset of the first label plus 233ms (Swingley, Pinto, & Fernald, 1999) to 6733ms post-labeling.

**Results**

Because the focus of the current paper is the effect of extraneous variability on children’s word learning, and due to space constraints, we present here the results from the test phase. Analyses of looking during referent selection are reported separately and discussed in detail in Twomey, Ma & Westermann (under review); overall, however, we found chance level looking and no difference between conditions. At test, each novel object served as a target on one memory reactivation trial and one retention trial. Fig. 3 depicts looking times during the memory reactivation trials and shows little difference in target looking in the two conditions. This conclusion was supported by a linear mixed effects model with main effects of time bin (treated as continuous) and condition and their interaction, with by-participant random slopes and intercepts for condition and by-item random intercepts to rule out item effects (Barr, Levy, Scheepers & Tily, 2013). As in referent selection, there was a small but robust increase in looking with time (beta = 0.0019, SE = 0.00063, t = 2.99, χ²(1) = 10.49, p = .0012). However, condition had no independent effect on looking times, and did not interact with time bin (main effect of condition: beta = 0.043, SE = 0.00098, t = 0.66, χ²(1) = 0.12, p = .73; time bin x condition interaction: beta = -0.0080, SE = 0.00098, t = -0.81, χ²(1) = 0.67, p = .41).

Data from the three retention trials show a markedly different pattern, however. As Fig. 4 illustrates, children in the variable color condition looked at the target at above-chance levels immediately following labeling and again at around 4000ms, suggesting that encountering variable colored backgrounds during referent selection facilitated their retention of the novel label-object mappings. A mixed effects model with the same fixed effects structure as above and by-participant and by-item random intercepts and slopes for condition revealed that target looking decreased over time (time bin: beta = -0.029, SE = 0.00078, t = -3.65, χ²(1) = 32.55, p < .001). This effect was constant for children in either condition (time bin x condition: beta = -0.0010, SE = 0.0011, t = -0.87, χ²(1) = 0.75, p = .39). Critically, however, proportion target looking was greater for children in the variable color condition than in the constant color condition (beta = -0.26, SE = 0.090, t = 2.85, χ²(1) = 5.41, p = .020).

**Discussion**

The current study explored whether extraneous variability would boost young children’s word learning. We trained two groups of two-year-old children with novel label-object associations via multiple referent selection trials. Stimuli presented to both groups were identical except that half the children saw arrays of novel objects displayed on a white background (constant color condition), and half saw objects on multiple colored backgrounds (variable color condition). Analyses of test trials revealed a clear effect of background variability: while children did not appear to correctly identify previously-seen novel objects during the memory reactivation trials, on retention trials children who had seen variable backgrounds during referent selection looked for longer at target objects than did children who had seen constant colored backgrounds, and did so at levels greater would be expected by chance. Thus, infants who had seen objects on variable backgrounds learned and retained the novel object-label mappings, but infants who had seen the objects on a constant background did not. These results offer converging evidence that following reactivation of memory traces, background variability facilitates learning, raising several interesting issues (see also Twomey et al., under review).

The importance of memory reactivation Children in the variable color condition looked at target objects at chance levels on the three memory reactivation trials, but at levels...
greater than expected by chance on the subsequent three retention trials. Typically in word learning studies children see only a single retention trial for each object. Our results suggest that null findings in these studies could be due to a lack of recall ability rather than a lack of learning. These data indicate that including memory reactivation trials in future studies could help shed light on whether children are failing to learn, or failing to recall. Establishing the locus of memory reactivation in the word learning field is therefore critical for a thorough understanding of the delicate memory processes underlying early language acquisition.

Decontextualization in early learning The fact that only children in the variable color condition retained novel label-object associations may seem unexpected in light of recent work in word learning indicating that consistency in context supports, not impairments, word learning (e.g., Axellsson & Horst, 2014). Given these results, why should what seems to be a more challenging task (i.e., variable color versus constant color backgrounds) lead to better learning? In fact, our results are in line with a wealth of adult literature demonstrating that background variability supports recall (e.g., Godden & Baddeley, 1975). More recent work has explored the effect of context on adults’ category learning; for example, Finch, Carvalho and Goldstone (2016) showed that variable backgrounds led to better retention of previously seen exemplars of a bird category.

These results are attributed to a decontextualization mechanism. When memories are formed after a single encounter, both context and target are encoded. On subsequent encounters, if the context stays the same, it remains part of the representation. These context-dependent memories are harder to recall when the context changes. Godden and Baddeley (1975) describe a classic example of this effect, showing that divers who had learned word lists either on dry land or underwater were better at recalling words learned underwater when tested underwater, and better at recalling words learned on dry land when tested on dry land. When an item is encountered in multiple different environments, however, the representation becomes decontextualized: the context becomes less important to the representation. If an item with a decontextualized representation is encountered in a new environment, then, it is easier to recall than if the representation were context-dependent.

The same mechanisms that explain these adult data can account for children’s word learning in the current study. During referent selection, children in the constant color condition learned context-dependent representations, while children in the variable color condition learned decontextualized representations. At test children encountered objects on a gray screen – and critically, neither group had seen objects presented on a gray screen until this point. Thus, recall was possible for children in the variable color condition, who were able to generalize their decontextualized memory traces to the new test context. This raises the question of why contextual consistency in existing studies supports word learning – the opposite of the current findings. It is possible that different types of context have qualitatively different effects. Here, in line with Stephen et al. (2009), “context” was low-level, extraneous variability. In contrast, the contexts in the existing literature were rich and salient: in the storybook studies, books were constructed from photograph-like images, resulting in a complex visual scene that varied from page to page. In addition, the sentence contexts in which novel words appeared also varied (Horst et al., 2011). Similarly, in the referent selection work, “context” consisted of the competitor objects presented alongside the targets, which were considerably more complex than a simple block of color (Axellsson & Horst, 2014). Thus, it may be that in rich learning environments, restricting complexity supports learning (Radesky & Christakis, 2016), while in simpler learning environments, increasing complexity by adding background noise helps learning.

This decontextualization account provides a mechanism by which added variability can support learning, as predicted by the dynamic systems account. Importantly, decontextualization is one among many potential mechanisms by which learning under the dynamic systems account may be shaped. As noted above, this theory predicts that background entropy should facilitate learning by speeding up the emergence of new stable behavioral states (Stephen et al., 2009). However, the dynamic systems account also suggests that other types of variability should support learning, raising the intriguing possibility for future work that entropy introduced in a different modality, for example sound or spatial location, could also support word learning. Work is underway to test these predictions. Overall, however, on either the specific decontextualization account or the broader dynamic systems approach, the current work extends a well-established phenomenon in adult cognition to children with a new task, pointing to a view of development as a continuous process driven by domain-general mechanisms.

Acknowledgements

Data are available on request from the authors. The support of the Economic and Social Research Council is gratefully acknowledged (LuCiD: ES/L008955/1; Future Research Leaders grant to KT: ES/N01703X/1). GW was further supported by a British Academy/Leverhulme Trust Senior Research Fellowship (SF150163). We thank the parents and children who made this work possible.

References


Twomey, Ma & Westermann (under review). All the right noises: Background variability helps early word learning. *Infancy.*
