IN PRESS, Cognitive Science

Part of the special issue in honor of Linda Smith's Rumelhart Prize

Moving Word Learning to a Novel Space:

A Dynamic Systems View of Referent Selection and Retention

Larissa K. Samuelson*+

Sarah C. Kucker~+

John P. Spencer *+

*School of Psychology, University of East Anglia

~ The Callier Center for Communication Disorders, School of Behavioral and Brain Sciences,

The University of Texas at Dallas

+DeLTA Center

Corresponding Author: Larissa K. Samuelson
Room 0.09b
Lawrence Stenhouse Building
University of East Anglia
Norwich NR3 7TJ
email: l.samuelson@uea.ac.uk

keywords: word learning, dynamic systems, Dynamic Field Theory, fast mapping, attention, spatial memory, referent selection

Abstract

Theories of cognitive development must address both the issue of how children bring their knowledge to bear on behavior in-the-moment, and how knowledge changes over time. We argue that seeking answers to these questions requires an appreciation of the dynamic nature of the developing system in its full, reciprocal complexity. We illustrate this dynamic complexity with results from two lines of research on early word learning. The first demonstrates how the child's active engagement with objects and people supports referent selection via memories for what objects were previously seen in a cued location. The second set of results highlights changes in the role of novelty and attentional processes in referent selection and retention as children's knowledge of words and objects grows. Together this work suggests understanding systems for perception, action, attention, and memory and their complex interaction is critical to understand word learning. We review recent literature that highlights the complex interactions between these processes in cognitive development and point to critical issues for future work.

1. Introduction

Two research questions are central to the field of cognitive development: how do children bring their knowledge to bear on behavior in-the-moment, and how does this knowledge change over time? Research seeking answers to these questions must appreciate the dynamic nature of the developing system in its full, reciprocal complexity. That is, we must appreciate that the child's ability to bring knowledge to bear is based on bidirectional interactions with the context, including people and things around the child (Smith & Thelen, 2003). Likewise, knowledge change is the emergent product of the accumulation of many small moments of perceiving, attending, remembering, and behaving embedded in that context (Samuelson & Smith, 2000). These ideas are well illustrated in the field of early word learning. In particular, recent work examining how children select the referent of a novel word in-the-moment demonstrates how a history of perceiving, acting, and remembering in a context guides initial mappings, and how accumulating vocabulary knowledge refines the behavior of this system over development.

Here we review data from two lines of work examining how children's knowledge is brought to bear in determining the referent of a novel name and how that process changes over development. The first line of work showcases how the child's active engagement with objects and people supports referent selection via memories of what objects have been seen where. This work points to a new understanding of how children can use the shared space of early word learning interactions to guide the mapping of novel names to novel objects. In so doing this work links research on memory with research on intention-understanding to suggest that space can serve as the substrate for children's ability to understand aspects of communicative intent. The second line of work highlights changes in the role of novelty and attentional processes in referent selection and retention. This work suggest that the contribution of attraction to novelty and accumulating vocabulary knowledge in early word learning are not fixed but instead mutually influential and dynamically influenced by the specifics of the mapping context. We argue that these lines of work complement a growing literature demonstrating the importance of the developmental cascade by which children learn to learn words. Together this work argues for an understanding of language that is about more than just the linguistic system; we must also understand systems for perception, action, attention, and memory (Samuelson & Smith, 2000). We review recent work in the field that highlights both the complex interactions between general cognitive processes and word learning and points to key issues and questions for future work.

2. Remembering what was where can bring objects and names together

In a recent paper (Samuelson, Smith, Perry, & Spencer, 2011), we demonstrated how children's actions on objects—looking, reaching, and exploring them—created memories that could be used to link novel words to novel objects in a referent selection task. In particular, acting on objects involves looking and reaching to an object's location in space which in turn, creates a memory for where the object was seen. We have found that children can use these memories of what object had been seen and explored in a cued location to bind a name and object that do not occur together in time (Samuelson et al., 2011). Our studies are based on a seminal study by Baldwin (1993) examining young children's ability to read the referential intent of a speaker. The task proceeds as follows (see also Fig. A in the supplementary materials and the associated text in the supplementary materials): a novel object is presented to a 20-month-old child for exploration and manipulation on one side of a table. This object is then removed and a second novel object is presented on the other side of the table. The child is allowed to reach for, grasp, and explore this second object. This is repeated and then both objects are placed in separate opaque buckets on their respective sides of the table. The experimenter looks into one bucket and says “Modi!” The object from the *other* bucket is then taken out and placed on its side of the table. It is removed after the child examines it and the object from the bucket that the experimenter had looked into is placed on the table. After the child examines this object, it is also removed. Both objects are then placed on a tray at the center of the table. The tray is pushed toward the child, and the experimenter asks, “Can you get me the modi?” Children retrieve the object that was in the bucket the experimenter was looking in when she said the novel word 67% of the time. Baldwin suggested this result indicates that children understand the pragmatic use of eye gaze as an intentional cue (Baldwin, 1993).

2.1 A dynamic neural field model of word-object binding

We have used Baldwin’s design to explore the processes that support children’s smart performance in tasks such as these. In particular, we have argued this result reflects children’s use of spatial memory to bind words to objects (Samuelson et al., 2011). We have implemented this proposal in a Dynamic Neural Field (DNF; Schöner, Spencer, & The DFT Research Group, 2015) model that provides a process account of how children can use the shared space of social interactions to link novel names to referents, even when the two are not presented simultaneously in space and time. Dynamic Field Theory is an embodied, dynamic systems approach to cognitive-level processes based on an understanding of brain function at the level of neural population dynamics (Erlhagen, Bastian, Jancke, Riehle, & Schoner, 1999; Jancke et al., 1999).

In particular, this approach uses fields of metrically-organized neural sites that interact according to a local excitation/lateral inhibition function (Durstewitz, Seamans, & Sejnowski, 2000; Spencer, Austin, & Schutte, 2012). Neural fields, like local neural populations in the brain (Amari, 1977; Cohen & Newsome, 2009; Fuster, 2003), move into and out of attractor states, reliable patterns of activation that the neural population maintains in the context of inputs. Note that DFT is an embodied approach in two senses. First, this theoretical perspective was explicitly developed to solve the cognitive grounding problem, that is, to explain how populations of neurons at the cognitive level can be tightly coupled to the sensori-motor surfaces. Second, to demonstrate that this theoretical language is embodied, our colleagues have built robotic implementations of cognitive architectures like the one presented here (see, e.g., Faubel & Schöner, 2008).

Panel A of Fig. 1 illustrates a central concept in DFT – a 'peak' within a dynamic field. In this example, a neural population in a simulated visual cortex creates a stable peak of activation representing an estimate of an object's location. The x-axis represents the location in a spatial frame (e.g., a retinal frame or the workspace of a task) and the circled peak of activation represents the detection of a stimulus at that location in space. Note that the blue line in A shows the activation level in the field, while the red line shows neural sites above the activation threshold. These local decisions—peaks—share activation with other neural populations—other peaks—creating a macro-scale attractor state. In DFT, thinking is the movement into and out of these states. Behaving is the connection of these states to sensory and motor systems.

--------------------------------Insert Fig. 1 about here ------------------------------------

The six fields pictured in Fig. 1 comprise our full model of Word-Object Learning (WOL). This is the same model presented in Samuelson et al. (2011) with the input fields shown (see Samuelson et al., 2011 for model equations, parameters and a step-by-step account of processing). The model captures processes at the second-to-second and developmental timescales and provides a process-based account of how children bring memories of what they have seen where to link names and objects. The model is composed of four 1D fields— scene attention (A), label input (D), shape-based attention (G), and color attention (H)—and four 2D fields—space-shape (B), space-color (C), label-shape (E), and label-color (F). The scene attention field encodes the spatial position of the item in the scene, the shape and color attention fields encode the feature values of those items (e.g., circular and red), and the label input field encodes the word (as an abstract label). The 2D fields, by contrast, integrate or 'bind' features using special binding dimensions—spatial position and labels. In particular, the space-feature fields (e.g., space-color) represent the presence of colors and shapes at particular locations in the workspace. By contrast, the word-feature fields (e.g., word-color) represent the mapping of a particular shape or color with a word. Fields are coupled such that activation passes along four shared dimensions: space (A↔B↔C), words (D↔E↔F), shape (B↔E), and color (C↔F). This can be seen as a light blue vertical "ridge" of activation on the right of panels E and F that shows activation from the attended stimulus (circled peak in panel A) being passed into the two space-feature fields. Coupling across the spatial dimension creates a "bound" object representation—a pattern of peaks representing a specific color and shape at a rightward location. Similar coupling across the label dimension (thin peak in panel D; vertical ridge in panel E, panel F), binds a word with associated features. Hebbian learning in each field enables trial-to-trial learning of which objects are where and what features go with each word, building a vocabulary of position-object and word-object mappings.

To examine the processes that support children's use of memories of what was seen where to map names to objects, we simulated each step in Baldwin's task in the model (see supplementary materials, Fig. B for step-by-step processing in the model). Behaviorally, the model captures the moment-by-moment interactions of the child in the task environment. That is, when the objects are presented during familiarization, children look at them in particular locations in space. They reach for them in space. They manipulate them in space. And then they attend as each object is removed from its side of the table. These actions create associations between each of the novel objects and unique locations. Then, when the experimenter looks into a bucket placed at one of those unique locations and says the name, the child's memory of the object previously seen and acted on at that location is recalled and bound to the novel name. Thus, the child is able to link the novel name to the correct object via space.

2.2 Children use space to bind words to objects

With 17- 20-month old children, we have tested several predictions of this theoretical proposal and quantitatively simulated children's behavior with the DNF model (see Samuelson et al. 2011). In our first experiment (see Fig. A in the supplementary materials), we replicated Baldwin's task in a control condition (No Switch) and disrupted space as a cue in an experimental condition by changing the location of the objects on the second familiarization trial (Switch). The 24 children tested performed identically to those in Baldwin's study in the replication condition, choosing the object from the named bucket on 73% of test trials (compared to 67% in Baldwin's study, see Fig. C in the supplementary materials). In contrast, children in the Switch condition performed at chance levels. Note that if binding the object and label depended only on understanding the intentions of the experimenter at the time of labeling, it should not matter where the objects had been beforehand. That it did demonstrates the importance of space in binding labels to objects.

In our second experiment we removed the hidden object component of the task. Following the familiarization presentation, the experimenter pointed to the empty place on the table where one of the objects had been and said, “Modi!” Children linked the object that corresponded to the named location at the same rate as those in the No Switch condition of Experiment 1 (see supplementary Materials, Fig. B).

In our third experiment, we pitted space against temporal congruence. We used four familiarization trials for each object to create strong spatial memories, but then put a single object on the opposite side from where it was presented during familiarization, pointed directly to it and said the name. Children chose the temporally linked object significantly *less* than predicted by chance. In other words, they selected the object that had previously been in the labeled location even though it was not there during the labeling event itself. In a control condition during which the object and label were presented together at that location without prior familiarization, children bound the name and object at high levels (see Fig. B in supplementary materials).

Finally, as a critical test that space plays a central role in this phenomenon, we examined whether another salient cue, color, would yield the same result. Instead of presenting the same object at a unique spatial location each time, the objects were always presented centrally on a uniquely colored tray (see supplementary Materials, Fig. A). Thus, each object was associated with a distinctive cue, but the cues were not separated in space and did not afford differences in children’s actions in space. During the labeling event, no objects were present (as in the prior experiments) and one of the two colored trays was placed at the center of the table. The experimenter pointed at the tray and said “Modi!” The test was the same as in the other experiments. Children performed at chance levels in this task. Critically, when tested children were asked to put the objects on the correct trays they performed well above what would be expected by chance (70% correct). Thus, the color cue was distinctive and memorable, but it was not used to bind the objects and labels. We contend that this is because, unlike with spatial cues, the color cues did not allow the child to orient toward the objects differentially in space.

2.3 Computational models, parents, and robots use space to bind names to objects

Over 100 runs of 12 simulations with different random starting points (corresponding to testing 12 individual children), the DNF model captured children’s performance in this task very well, using the same parameters for all tasks (see Fig. B in supplementary materials). Together, the data and simulations show that children can use consistency in spatial location to bind a novel name to a novel object in an ambiguous naming situation. These studies suggest that the child’s attention and actions in space can be used as deictic references to bind objects in the physical environment to cognitive representations of names. In a final study, we asked parents to teach their children names and coded how consistent the parents were in keeping the two objects on two different sides of the table. Children whose parents kept objects in consistent spatial locations demonstrated better learning of the novel names when later tested by an experimenter (see Samuelson et al., 2011).

An important question left unanswered by this series of studies is the spatial frame children are using in these tasks. For instance, are children encoding and remembering words and objects in a body-centered frame of reference (to my left), or are they remembering objects in a table-centered frame of reference (to the left of the table)? A recent study by Morse and colleagues (2015) suggests that children are encoding locations in this task relative to the body. In particular, children were familiarized with objects while seated. Then, during the naming event, children were asked to stand. If children are using a body frame of reference, this should disrupt their ability to link the name with their previous experience of the objects. Results were consistent with this prediction.

In addition to these data, Morse et al. proposed a theoretical model and implemented this model on a robotic platform (an iCub). The robotic model generally mimicked children's performance across conditions. This makes an important contribution in that it demonstrates word learning can be embodied, consistent with earlier work in robotics using a DNF model similar to the model reported here (Faubel & Schöner, 2008; for discussion see Samuelson & Faubel, 2015). An important question is how the Morse et al. model and DNF models relate. Both provide embodied accounts. A key difference emerges, however, in how words are bound to features. The Morse et al. model only binds features to words via a postural field. By contrast, the DNF model binds words to features directly, and deals with the spatial aspects of sensorimotor interactions with the world via the space-feature fields. This leads to critically different internal representations in the two models. We contend this matters. For instance, the DNF model captures other key aspects of early word learning including a bias to attend to shape in early development (Perone, Spencer, & Samuelson, 2015; Samuelson, Spencer, & Jenkins, 2013) as well as learn words organized in a hierarchy (Jenkins, Samuelson, & Spencer, 2015; Samuelson et al., 2013). Both of these phenomena highlight how words generalize over object features to form categories. It is not clear how the Morse et al. model can capture such effects without an integrated word-feature representation. Having posture as the go-between in every case seems unrealistic, but future work will be needed to probe this issue in greater depth.

More generally, these studies point to a grounded view of how an abstract cognitive processes—interpreting someone else's naming intentions—can unfold (Samuelson et al., 2011). The shared spatial context of naming interactions can serve as the dynamic substrate for what has previously been interpreted as children's ability to infer adults' communicative intent because words occur in time and their referents occur in space. Space serves to organize where children look, and actions like picking up an object to examine or selecting something to hand over in response to a request happen in space. We encode spatial information in memory (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000); thus, cuing locations in space can serve to activate those memories and thereby bring prior knowledge to bear in service of the current task. We have shown how this can help to bind words to objects and allow things in the just previous past to be bound to the present. Over a slightly longer timescale we can see how these representations can evolve continuously in context as current actions cue and bring to bear representations of past events. In this way then, current behavior will be the product of both represented knowledge and the specifics of current context but will also change over development as knowledge becomes stronger. This trajectory of dynamic developmental change is illustrated in another line of work on referent selection that incorporates known objects and examines retention of novel mappings over multiple timescales.

3. The dynamic balance of novelty and knowledge in referent selection and retention

The results reviewed above argue that to understand early word learning we have to understand how language, attention, and memory interact. The next line of work adds to this by showing that these systems interact in complex ways, back-and-forth over multiple timescales from referent selection to retention (Kucker, McMurray, & Samuelson, 2015). Thus, it is not just about multiple systems (attention, memory, language) but multiple systems with reciprocal influences through time.

When confronted with the task of finding a referent for a novel word, children bring to bear both their prior history of learning about words and objects in the world, and biases they have developed from that prior learning. This influences the child's behavior in-the-moment and subsequent learning, as illustrated by the interaction of novelty-driven attention and represented lexical knowledge in referent-selection and retention tasks. Such tasks begin with a series of warm up trials during which children are asked to select each of three familiar objects by name (e.g., "Get the puppy"; see Fig. D in the supplementary materials for an overview of the design). On each experimental trial, children are presented with two of these familiar objects and one novel object. On novel-name referent selection trials, children are asked for an object with a novel name, for instance, "Can you get the toma?" On familiar-name trials, children are asked for a familiar object by name, now in the context of a novel object. Retention of novel word-object mappings is tested after a short delay by presenting children with three novel objects seen previously and asking them to get each, in turn, by name.

Children are typically very good at this task, reliably selecting the novel referent in response to the novel word (Bion, Borovsky, & Fernald, 2013; Horst & Samuelson, 2008; Spiegel & Halberda, 2011; Vlach & Sandhofer, 2012). This behavior is traditionally attributed to word learning constraints. For example, by the mutual exclusivity constraint (Markman, 1990), children understand that objects typically have just one name and thus exclude the two familiar objects with known names as possible referents for the new label. Similarly, according to the

Novel Name-Nameless Category principle (N3C; Golinkoff, Mervis, & Hirsh-Pasek, 1994) children understand that novel things are most likely to be the referent of novel words. Such accounts have provided the basis for understanding how children quickly map novel names to novel referents and, thus, build a lexicon, and children's proficiency at tasks such as the one described above has been taken as evidence of their word learning prowess (see Horst & Samuelson, 2008 for review).

Nevertheless, Horst and Samuelson (2008) found that while 24-month-old infants were very good at both the familiar- and novel-name referent selection trials, they failed to demonstrate retention of the novel name-object mappings five minutes later (see also, Bion et al., 2013; Kucker & Samuelson, 2012). Our recent work has examined the processes that both support referent selection in-the-moment, and the building of strong name-referent mappings that support longer-term retention and word learning. As we review below, this work points to a dynamic interaction between a novelty bias and accumulating lexical knowledge that changes rapidly during early vocabulary development (see, Houston-Price, Caloghiris, & Raviglione, 2010; Kucker et al., 2015; McMurray, Horst, & Samuelson, 2012; Twomey, Horst, & Morse, 2013).

3.1 Novelty-driven referent selection in 18-month-old children

Kucker, McMurray and Samuelson (2016; see also, Kucker 2014) investigated these processes in children very early in vocabulary development by replicating Horst & Samuelson (2008) with 18-month-old children. When asked to select a novel item by name from an array of one novel item and two well-known items, these young children were very good at selecting the novel target item (77% correct). However, when these same children were asked to select a known item from the same array, they failed, selecting the requested item only 30% of the time. Instead, these young children selected the novel item the majority of the time. Given that the children's parents verified the child's knowledge of the known items and names prior to the experiment, and that children selected the known items at very high levels during the warm up trials (94% correct), it is not likely that their failure was due to a lack of knowledge for the known names. Nevertheless, the mechanistic basis for this attraction to novelty is unclear; it could be the product of heightened attention to the most novel object or diminished attention to known objects. Adaptions of McMurray, Horst and Samuelson's (2012) computational model of referent selection and retention show that both possibilities can capture the empirical data (Kucker et al., 2016). However, because that model does not instantiate autonomous processes of visual exploration and attention, it is still unclear how the word learning system, as opposed to the modeler, comes to view novel and familiar stimuli as more or less salient. The more autonomous model of Twomey and colleagues captures related data with older children, but because either the modeler selects the novel object for the robot (Twomey et al., 2013), or the robot looks at and processes all test objects (Twomey, Morse, Cangelosi, & Horst, 2014) it does not inform questions regarding the basis for children's bias.

3.2 Both 18- and 24-month-old children select "supernovel" objects

Data from Kucker et al. (2016) thus present an interesting contrast to that of Horst and Samuelson (2008) and point to rapid changes in dynamic interactions between novelty biases and growing vocabulary knowledge between 18-and 24-months-of-age (Houston-price et al., 2010;

Mather & Plunkett, 2012; Mather, 2013). These data are even more interesting when considered in the context of a study showing that novelty influences the referent selection of 18- and 24-month-old children in similar ways. Kucker, McMurray, and Samuelson (2016) presented children at both ages a referent selection task that only included novel objects. They manipulated the relative novelty of the objects by giving a subset of them to children to examine for two minutes prior to the experimental task. During this pre-familiarization period, the experimenter never named the objects. On subsequent novel-name referent selection trials, two of these familiarized but unnamed objects were presented along with one completely novel object. When asked to "get the toma," both 18- and 24-month-old children selected the unfamiliarized, "supernovel" object at very high levels (69% and 70% respectively; see also Horst, Samuelson, Kucker, & McMurray, 2011). Importantly, both age groups also selected familiar objects by name at equally high levels on trials that included only familiar objects (67%, and 78% for 18- and 24- month-old children respectively). Thus, in a context where items differ only in familiarity, 24-month-old children appear similar to 18-month-old children in their attraction to the most novel object as the referent of a novel name.

3.3 Retention is supported by familiarity

Together these data reveal a complex, dynamic interaction between attraction to novelty and represented vocabulary knowledge as the early vocabulary grows. Smith and Yu have likewise shown that the presence of a relatively more novel item can change the process of forming new name-object mappings in 12- to 14-month-old children using looking-based cross-situational word learning tasks (Smith & Yu, 2013; Yu & Smith, 2011). Likewise, data from

Bion et al. (2013) suggest that this dynamic interaction continues to shift in subsequent months such that 30-month-old children not only reliably select known objects when presented with familiar names, but also retain novel name-referent mappings over delays similar to those tested by Horst and Samuelson (2008; see also, Spiegel & Halberda, 2011). The fact that in Kucker et al's (2016b; see also Horst et al., 2011) study with a "supernovel" object, 24-month-olds can be made to perform similarly to 18-month-olds via a small change in their prior experience with the stimuli is in line with the idea that the underlying developmental processes are continuous.

Further support for this argument comes from data demonstrating that 24-month-old children can be made to retain like 30-month-olds with a similar small change in their prior experience. Kucker and Samuelson (2012) gave 24-month-old children two minutes to explore the novel objects that would later be presented in the standard referent selection and retention paradigm described above. No names were provided during this familiarization period. Referent selection followed the standard procedure of Horst and Samuelson (2008); a novel (but now pre-familiarized) object was presented with two known objects that parents indicated their children could name, and that children retrieved successfully during warm-up trials. Kucker and Samuelson (2012) found that this brief pre-familiarization was enough to boost retention of the novel name-referent mappings formed during referent selection. In particular, 73% of the 24-month-old children tested retained novel words when pre-familiarized with the objects. A second group of children pre-familiarized with the novel words prior to referent selection did not evidence significant retention. Thus it appears a small boost to 24-month-old children's represented knowledge of the objects is sufficient to create a significant change in learning. Critically, when 18-month-old children were pre-familiarized with the novel objects their retention performance did not differ significantly to that from the prior study with no pre-familiarization (33% retention without and 40% with pre-familiarization; Kucker et al., 2016). In addition, pre-familiarization with the novel stimuli did not decrease 18-month-old children's attraction to novelty on known-name trials—they still chose the novel object when familiar items were requested.

The data reviewed above suggest the following developmental picture of referent selection in the context of novel and familiar items. Eighteen-month-old children are attracted to novelty to the extent that they select the most novel item regardless of whether they are given a novel or known name and with arrays of either all novel items or a mix of novel and familiar items. By 24-months-of-age, children are still biased toward novelty enough that they will select the most novel object given a novel name, but they can overcome this bias to select a requested familiar item when presented with two known and one novel object. Furthermore, these older children can retain a novel name-referent mapping formed in this context if they are familiarized with the object prior to the referent selection event. All of this could be taken to suggest a relatively simple picture of the interaction between novelty and knowledge in referent selection and retention: first knowledge must increase enough to overcome the attraction to novelty and with sufficient knowledge (in the form of a prior representation of an object) a robust mapping between a novel word and referent can be formed. This account would then suggest a gradual shift from novelty-driven processes such as N3C to more knowledge-based processes such as mutual exclusivity. A final set of studies suggest, however, that a more dynamic interaction is at play.

3.4 The strength of prior knowledge inversely influences retention of new word-object mappings

Kucker, McMurray, and Samuelson (2016b; see also, Kucker 2014) gave 18- and 24-month-old children a standard referent selection and retention task, but manipulated the strength of children's knowledge of the *known* items. In the well-known condition, the known items were ones the parents reported their children could name prior to the experiment (e.g. shoe, dog). In the weakly-known condition, children had less experience with the items as names of the known foils were taught to the children just before the referent selection task (e.g., whisk, slinky). A within subjects design was used; thus all of the children received label training on the weakly-known items prior to the start of the study and the three trial types (well-known, weakly-known, and novel-name) were randomized. Children's knowledge of the just-taught, weakly-known word-referent mappings was tested by adding three additional warm up trials during which the three newly learned items were placed on a tray and each was requested by name once. These followed the standard warm-up trials during which the well-known items were each requested once. The 18- and 24-month-old children demonstrated good knowledge of both the well- and weakly-known items with performance on warm-up trials ranging from 92-100% across groups and trial types.

Despite this knowledge of the known items, however, 18-month-old children again demonstrated a strong novelty bias during referent selection. They selected the *novel* object on well-known, weakly-known, and novel-name trials at high levels (70%, 81% and 87%, respectively). Interestingly, however, these young children also demonstrated significant retention of novel name-referent mappings when those mappings were formed in the context of weakly-known foils (83% correct selections). Data from 24-month-old children were similarly interesting. They demonstrated a strong novelty bias and selected the novel object on well-known, weakly known, and novel-name trials. And, like 18-month-old children, 24-month-olds demonstrated retention of the novel name-referent mappings formed in the context of weakly-known foils, although at significantly reduced levels compared to the 18-month-old children (52% correct selections). Thus, in contrast to the simple picture of retention being the product of knowledge that is strong enough to overcome attraction to novelty, it appears that it is the dynamic balance between the strength of known lexical representations relative to attraction to novelty that matters (Hollich et al., 2000; Mather, 2013; Zosh, Brinster, & Halberda, 2013).

3.5 The interplay of knowledge and novelty in reference selection and retention

Data from these studies contradict accounts that suggest learning in referent selection and retention tasks is solely based on prior represented lexical knowledge (see, Clark, 1987; Markman, Wasow, & Hansen, 2003; for discussion). Across studies, we have seen multiple cases where children's performance on the known trials was driven by novelty, not knowledge, but it is also clear that mappings formed on the basis of novelty do not always lead to retention. The cases where retention is seen are restricted to those in which older children have a prior memory of the objects or novel-word referent mappings are formed in the context of weakly-known lexical mappings. The first of these cases points to prior memories of the objects driving learning, but it does not fit a mutual exclusivity account because the objects were not named during pre-familiarization. In the second case, it is possible that the weak representations of the just-learned, weakly-known items force a deeper level of engagement with the stimuli—children have to work harder to recall which of the relatively-novel items presented had names and which did not (Craik & Lockhart, 1972; Vlach & Sandhofer, 2014). Notably, however, this possibility does not uniquely support an account based on novelty or mutual exclusivity. It could be that in the process of completing the weakly-known trials, children recall the just-learned names to some extent. However, it is also possible they simply recall that those objects had names, without remembering the mappings to specific words (Axelsson, Churchley, & Horst, 2012; Schafer & Plunkett, 1998). A third possibility is that that the novel targets on weakly-known trials are "supernovel" because they were not pre-familiarized during the word training, which could have enhanced retention. Clearly, it is not simply knowledge or novelty driving referent selection and retention; rather, the dynamic interaction of the two creates developmental changes in both how children bring knowledge to bear and how that knowledge, in the form of robust new novel word-referent mappings, changes over time.

4. Moving to a novel space requires integrating the dynamics of visual attention and word learning

The data reviewed above demonstrate how the child's active engagement with objects and people supports referent selection via memories for which objects had been seen in which locations, and highlight the changing role of novelty and attentional processes as children's knowledge of words and objects grows. More generally, we believe this work argues that systems for perception, action, attention, and memory play critical roles in language development (c.f., Samuelson & Smith, 2000). This is well illustrated by imagining a typical word-learning scenario. A toddler is seated in her highchair eating lunch. In front of her are a number of namable items—a spoon, a plate, a cup, pieces of banana, carrots, and a bowl of applesauce. Mom says, "Can you use the spoon to get some applesauce?" To comply, a number of events must occur. The child needs to process the visual scene, creating a map of the objects in the visual array. This requires binding the correct visual features together such that the circles of banana are not orange and the relative positions of the objects do not change when she shifts her gaze from the cup to the spoon. Likewise, she must parse the auditory stream to pull out the individual word segments, determining which sets of sounds have known referents and which are novel. Finally, to link the right word with the right referent, she needs to coordinate these events so that representations of the novel word "spoon" and the novel referent are co-active allowing associations to form. And later she needs to retrieve and update her representations of the object-word mapping when she sees a different spoon or hears the word in a different context. Clearly then, the success of this word-learning episode will depend on how well the child's attention is allocated, how her perception of the objects and words is coordinated, how well the word-form and object are encoded and how those representations are linked.

This complexity has long been acknowledged in the field but traditionally led to the view that the only way children could build their vocabularies as quickly as they do was via information-laden supportive processes that ensured correct referent selection, accurate auditory parsing, and one-trial learning of new name-object links (Golinkoff et al., 1994; Markman & Wachtel, 1988). The work reviewed above, however, fits with a more recent trend in the literature to examine the multiple processes that support word learning at moment-by-moment timescales and a greater appreciation for the multiple, flexible, general processes that enable children's early word learning prowess. Understanding this complex system requires a theoretical approach that integrates findings across multiple areas of cognitive development and appreciates the changes in processing that occur at both in-the-moment and longer timescales as the system builds itself and learns to learn words. Here we provide a brief overview of this perspective, integrating the work reviewed above with related research in the field, and highlighting critical areas of understanding required to build a more complete picture of the early word learning system.

4.1 Changing the view of the mapping problem in early word learning

One seminal study in this direction took a new perspective on word learning—literally. Using head-mounted cameras and eye-tracking, Yoshida and Smith (2008) challenged the long-held view that word learning is hard because of the infinite number of possible referents the child must consider when a novel word is uttered. These researchers showed that toddlers' short arms and smaller stature means that often there is only one or two objects in view when objects are named (see also, Pereira, Smith, & Yu, 2014). This fits with the fact that even adults viewing complex scenes have a more limited perspective than was previously thought, typically grabbing only 1-3 objects with each shift in attention or gaze (Hollingworth, 2007). Thus, the challenge is not of selecting a referent from among an infinite number. Rather the problem is in coordinating attention to the correct auditory and visual referents in time. This requires understanding how moment-to-moment shifts of attention get mapped to incoming words from the discourse partner.

The field has long acknowledged the importance of coordinating the auditory and visual streams in word learning. Prior work has focused on the role of joint attention (Baldwin, 1991; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1988; Ninio & Bruner, 2008; Tomasello & Todd, 1983), infants' ability to resolve ambiguity in naming contexts (Markman & Wachtel, 1988; Markman, 1990; Mervis & Bertrand, 1994), and parents' support of word learning via presentation of isolated words (Brent & Siskind, 2001; Tomasello & Farrar, 1986), objects

(Harris, Jones, & Grant, 1983; Pereira et al., 2014; Tomasello & Farrar, 1986), or follow-in labeling (Harris et al., 1983; Tomasello & Farrar, 1986). However, parents do not always name what the child is attending to (Pereira et al., 2014; Tomasello & Farrar, 1986; Yu & Smith, 2012a) and children do not always look where parents are attending (Deák et al., 2014; Pereira et al., 2014; Yu & Smith, 2012a), facts that are not surprising given that children and adults can generate upwards of 50,000 eye movements per day (Johnson, Amso, & Slemmer, 2003).

4.2 Understanding how visual and auditory attention interact in early word learning

Thus it is critical that we build on prior work examining joint attention and ambiguity resolution to understand how visual and auditory attention interact and influence each other in both word learning contexts, and more generally. Of course, much work in early cognitive development has already demonstrated how systematic biases in visual attention can impact the word learning context. For instance, it is well known that infants prefer to look at novel things. This bias could enhance object-word mapping in cases where parents notice this bias to novelty and name novel items (e.g., Samuelson & Smith, 1998). Conversely, this bias could detract from object-word mapping if parents name familiar items in the context of novel ones, similar to the finding from Kucker (2014) showing how a novelty bias can detract from novel word learning in a laboratory setting. It is also well known that infants habituate to novelty with continued exposure. This can be clearly seen in recent work demonstrating that what children learn in cross-situational word learning contexts is impacted by the dynamics of their visual exploration during the task (Smith & Yu, 2013; Yu & Smith, 2011). Conversely, the dynamics of looking and habituation are changed by the introduction of words (Baldwin & Markman, 1989; Mather, Schafer, & Houston-Price, 2011; Sloutsky & Fisher, 2012). Thus, it is critical we integrate these findings to understand how early word learning is tied to the dynamics of visual memory processes.

In a similar vein, the work reviewed above on the use of space to bind names to objects, as well as related work (see for example, Richardson & Kirkham, 2004; Vlach & Kalish, 2014) highlights that understanding the role of visual memory can lead to novel insights into how children solve referential ambiguity. Critically, theories of early word learning must understand multiple senses of visual memory: how visual memory modulates attention via habituation to novelty and how visual memory is used to build representations that track which objects are where. This presents a significant challenge. Such theories must bring together processes of visual attention, visual memory, visual binding of what is where, and processes for the formation and updating of word-object links across contexts. Perhaps most centrally, such theories must also be dynamic – able to capture how these processes work together in real time and change over the course of learning (Yu, Zhong, & Fricker, 2012).

4.3 Added complexity: Development over multiple timescales

Critically, there are (at least) two more levels of complexity in this story. First, these systems interact in intricate ways. Smith and Yu (2008) first demonstrated that infants can track the statistics of which words and objects have frequently co-occurred and use this to map words to referents. In these studies of cross-situational statistics, infants (or adults) are presented with multiple objects and hear multiple words on individual trials such that mappings are ambiguous within trials but the correct word-referent mappings can be resolved across multiple trials.

Studies show that 12-month-old infants keep track of these co-occurrences and learn novel mappings in short, 4-minute sessions (Smith & Yu, 2008, see also Fitneva & Christensen, this volume). However, small changes to this basic paradigm can disrupt learning, as was demonstrated when Smith and Yu (2013) manipulated the order of presentation of the 30 trials in their standard cross-situational learning paradigm to create blocks of trials in which a repeating word and object were shown with a more novel word and object. This "novelty trap" disrupted learning—only about a third of infants learned the mappings—even though the overall statistics of the experiment were identical to previous work.

The basis for this disruption in learning is not yet understood. Analysis of infants' accumulated looking statistics showed no difference between strong and weak learners, but weak learners did not show learning of the target words at test. Thus, infants' tracking of what was where interfered with their ability to map co-occurrences. Moreover, examination of infants' looking revealed a novel effect—the word presentation mandatorily cued attention to the most novel object. This is a looking version of our prior finding that that toddlers will map a novel name to the most novel of a set of unnamed objects (Horst et al., 2011; Kucker, 2014; Mather & Plunkett, 2012). These effects show complex interactions between children's representation of objects in context, looking, and word learning.

The second additional level of complexity stems from long-term vocabulary learning. While overall the group of children exposed to the "novelty trap" did not show evidence of learning the target words, a subset of children did—and those children had significantly more words in their vocabularies. This suggests that long-term vocabulary knowledge impacts the interaction between word-object associations and visual attention and memory. Likewise, studies of toddlers' retention of novel name-referent mappings made in the context of known words show that children with larger vocabularies retain but those with smaller vocabularies do not (Bion et al., 2013; Kucker, 2014). Thus, the robustness of the word-object mappings formed will be different depending on the developmental state of the lexicon they are being added to.

On one hand, this is unsurprising—development matters. On the other hand, fully appreciating this means our goal is a moving target—the system is changing itself as we present stimuli and test knowledge in specific tasks. This is perhaps why even as the field has made impressive strides toward documenting the processes involved in early word learning, no theory has yet integrated the component processes together across the relevant timescales (Kachergis, Yu, & Shiffrin, 2013; Yu & Smith, 2012b; Yu et al., 2012). This was recently brought to the foreground in the case of cross-situational word learning. There are two dominant accounts in the literature, one proposing that infants test single hypotheses about mappings and revise when evidence dictates (Trueswell, Medina, Hafri, & Gleitman, 2013), and the other suggesting infants track all possible pairings and let associative learning weight the most frequent (Kachergis, Yu, & Shiffrin, 2012; Yu et al., 2012). However, because these two dominant accounts both operate on the same data (word and object co-occurrences), propose that learning is based on statistical computation, and seek to model the same learning outcomes, these theories cannot be distinguished (Smith, Suanda, & Yu, 2014). Rather, what is needed is a theory of long-term learning processes by which children add and expand their vocabulary representations that also integrates how these processes interact with real-time selection and attention by which children determine the content of those representations (Smith et al., 2014).

This is a daunting prospect, but critical because the data reviewed here, and in many of the other papers in this special issue, demonstrate that the processes and constraints that support word learning do not work in isolation. Fortunately, there are multiple theoretical proposals, computational models, and research programs geared towards examining the mechanistic developmental processes that support word learning and seeking to integrate both current and prior work on how general cognitive processes support word learning. Our work reviewed above adds to these by demonstrating a role for non-linguistic memories of what objects have been seen where in referent mapping and highlighting developmental changes in novelty and knowledge in both referent selection and retention. We believe that theories and models that take seriously the fact that word learning is based on multiple coupled, embedded components that are mutually influential and evolving moment-to-moment, day-to-day, and year-to-year hold the key to understanding how children bring what they know to bear in-the-moment and how knowledge changes over time.

Acknowledgments

This research supported by NIH HD045713 to LS. The content is solely the responsibility of the authors and does not represent the official views of the NIH.

References

Amari, S. I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, *27*(2), 77–87. doi:10.1007/BF00337259

Axelsson, E. L., Churchley, K., & Horst, J. S. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology*, *3*(March), 88. doi:10.3389/fpsyg.2012.00088

Baldwin, D. a. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–890. doi:10.2307/1131140

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832–843. doi:10.1037/0012-1649.29.5.832

Baldwin, D. A., & Markman. (1989). Establishing word-object relations: A first step. *Child Development*, *60*(2), 381–398.

Bion, R. a H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*(1), 39–53. doi:10.1016/j.cognition.2012.08.008

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11376642

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1988). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*.

Clark, E. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language aquisition* (pp. 1–34). Hillsdale, NJ: Erlbaum.

Cohen, M. R., & Newsome, W. T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *The Journal of Neuroscience*, *29*(20), 6635–6648. doi:10.1523/JNEUROSCI.5179-08.2009

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684. doi:10.1016/S0022-5371(72)80001-X

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L., (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science*, *17*(2), 270–281. doi:10.1111/desc.12122

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience*, *3*, 1184–1191. doi:10.1038/81460

Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., & Schoner, G. (1999). The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods*, *94*(1), 53–66. doi:10.1016/S0165-0270(99)00125-9

Faubel, C., & Schöner, G. (2008). Learning to recognize objects on the fly: A neurally based dynamic field approach. *Neural Networks : The Official Journal of the International Neural Network Society*, *21*(4), 562–76. doi:10.1016/j.neunet.2008.03.007

Fuster, J. M. (2003). *Cortex and mind: Unifying cognition*. New York, NY, US: Oxford University Press.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language*, *21*(01), 125–155. doi:10.1017/S0305000900008692

Harris, M., Jones, D., & Grant, J. (1983). The nonverbal context of mothers' speech to infants. *First Language*, *4*(10), 21–30. doi:10.1177/014272378300401003

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., … Rocroi, C. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), i–vi, 1–123. doi:10.1111/1540-5834.00090

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 31–47.

Hoover, M. a, & Richardson, D. C. (2008). When facts go down the rabbit hole: contrasting features and objecthood as indexes to memory. *Cognition*, *108*(2), 533–42. doi:10.1016/j.cognition.2008.02.011

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. doi:10.1080/15250000701795598

Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition*, *118*(2), 234–44. doi:10.1016/j.cognition.2010.10.015

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias, *15*(2), 125–150. doi:10.1111/j.1532-7078.2009.00009.x

Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A., & Schoner, G. (1999). Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *Journal of Neuroscience*, *19*(20), 9016–9028.

Jenkins, G. W., Samuelson, L. K., & Spencer, J. P. (2015). Learning words in space and time: A dynamic neural field account of the suspicious coincidence effect. *Manuscript under Review*.

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 10568–10573. doi:10.1073/pnas.1630655100

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychonomic Bulletin & Review*, *19*(2), 317–24. doi:10.3758/s13423-011-0194-6

Kachergis, G., Yu, C., & Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, *5*(1), 200–13. doi:10.1111/tops.12008

Kucker, S. C. (2014). *The role of vocabulary knowledge and novelty biases in word learning: Exploring referent selection and retention in 18- to 24-month-old children and associative models.* ProQuest Information & Learning, US.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2016a). Modeling novelty: Perceptual biases and associative models. *Manuscript Submitted for Publication*.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2016b). Optimizing ambiguity: The role of knowledge in word learning. *Manuscript in Preparation*.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing Down Fast Mapping: Redefining the Dynamics of Word Learning. *Child Development Perspectives*, *9*(2), 74-78. doi:10.1111/cdep.12110

Kucker, S. C., & Samuelson, L. K. (2012). The first slow step: Differential effects of object and word-form familiarization on retention of fast-mapped words. *Infancy*, *17*(3), 295–323. doi:10.1111/j.1532-7078.2011.00081.x

Markman, E. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77. doi:10.1016/0364-0213(90)90026-S

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. doi:10.1016/0010-0285(88)90017-5

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners, *47*, 241–275. doi:10.1016/S0010-0285(03)00034-3

Mather, E. (2013). Bootstrapping the early lexicon: How do children use old knowledge to create new meanings? *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00096

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science*, *36*(7), 1157–1177. doi:10.1111/j.1551-6709.2012.01239.x

Mather, E., Schafer, G., & Houston-Price, C. (2011). The impact of novel labels on visual processing during infancy. *British Journal of Developmental Psychology*, *29*(4), 783–805. doi:10.1348/2044-835X.002008

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–877. doi:10.1037/a0029872

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the Novel Name — Nameless Category (N3C) Principle Mervis and Bertrand, 1646–1662.

Morse, A. F., Benitez, V. L., Belpaeme, T., Cangelosi, A., & Smith, L. B. (2015). Posture Affects How Robots and Infants Map Words to Objects. *Plos One*, *10*(3), e0116012. doi:10.1371/journal.pone.0116012

Ninio, A., & Bruner, J. (2008). The achievement and antecedents of labelling. *Journal of Child*

Language, *5*(01), 1–15. doi:10.1017/S0305000900001896

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, *21*(1), 178–185. doi:10.3758/s13423-013-0466-4

Perone, S., Spencer, J. P., & Samuelson, L. K. (2015). A Dynamic Neural Field Model of the Shape bias and Its Development. *In Preparation*.

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology. General*, *133*(1), 46–62. doi:10.1037/0096-3445.133.1.46

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, *76*(3), 269–295. doi:10.1016/S0010-0277(00)00084-6

Samuelson, L. K., & Faubel, C. (2015). Grounding word learing in space and time. In G. Schöner, J. Spencer & The DT Research Group, *Dynamic Thinking: a Primer on Dynamic Field Theory* pp 297-325. New York, NY US: Oxford University Press.

Samuelson, L. K., & Smith, L. B. (1998). Memory and Attention Make Smart Word Learning: An Alternative Account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*(1), 94–104. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8624.1998.tb06136.x/abstract

Samuelson, L. K., & Smith, L. B. (2000). Grounding development in cognitive processes. *Child Development*, *71*(1), 98–106. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10836563

Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. *PLoS ONE*, *6*(12), e28095. doi:10.1371/journal.pone.0028095

Samuelson, L. K., Spencer, J. P., & Jenkins, G. W. (2013). A dynamic neural field model of word learning. In *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence* (pp. 1–27). IGI Global. doi:10.4018/978-1-4666-2973-8.ch001

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*(2), 309–320. doi:10.2307/1132166

Schöner, G., Spencer, J. P., & The DFT Research Group (2015). *Dynamic Thinking: A Primer on Dynamic Field Theory*. New York: Oxford University Press.

Sloutsky, V. M., & Fisher, A. V. (2012). Effects of categorical labels on similarity judgments: a critical evaluation of a critical analysis: comment on Noles and Gelman (2012). *Developmental Psychology*, *48*(3), 897–900; discussion 901–6. doi:10.1037/a0027531

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, *18*(5), 251–258. doi:10.1016/j.tics.2014.02.007

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348. doi:10.1016/S1364-6613(03)00156-6

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–68. doi:10.1016/j.cognition.2007.06.010

Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development*, *9*(1), 25–49. doi:10.1080/15475441.2012.707104

Spencer, J. P., Austin, A., & Schutte, A. R. (2012). Contributions of dynamic systems theory to cognitive development. *Cognitive Development*, *27*(4), 401–418. doi:10.1016/j.cogdev.2012.07.006

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*, *109*(1), 132–40. doi:10.1016/j.jecp.2010.10.013

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language, *Child Development* (1), 1454–1463.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197–211. doi:10.1177/014272378300401202

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–56. doi:10.1016/j.cogpsych.2012.10.001

Twomey, K. E., Horst, J. S., & Morse, A. F. (2013). An Embodied Model of Young Children's Categorization and Word Learning. In L. J. Gogate & G. Hollach (Eds.), *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence* (pp. 172–196). Hershey PA: IGI Global.

Twomey, K. E., Morse, A. F., Cangelosi, A., & Horst, J. S. (2014). Competition affects word learning in a developmental robotic system. *Proceedings of the 14th Neural Computation and Psychology Workshop*, *2014*(288382), 1–12.

Vlach, H. a., & Kalish, C. W. (2014). Temporal dynamics of categorization: forgetting as the basis of abstraction and generalization. *Frontiers in Psychology*, *5*(September), 1–9. doi:10.3389/fpsyg.2014.01021

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time : memory processes support children ' s retention of learned words, *3*(February), 1–8. doi:10.3389/fpsyg.2012.00046

Vlach, H. a., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science*, *38*(4), 757–774. doi:10.1111/cogs.12092

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*(4), 553–566. doi:10.1037//0012-1649.30.4.553

Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*(2), 165–180. doi:10.1111/j.1467-7687.2010.00958.x

Yu, C., & Smith, L. B. (2012a). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. doi:10.1016/j.cognition.2012.06.016

Yu, C., & Smith, L. B. (2012b). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review*, *119*(1), 21–39. doi:10.1037/a0026182

Yu, C., Zhong, Y., & Fricker, D. (2012). Selective attention in cross-situational statistical learning: evidence from eye tracking. *Frontiers in Psychology*, *3*(June), 148. doi:10.3389/fpsyg.2012.00148

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science*, *17*(1), 20–28. doi:10.1080/10888691.2013.748420

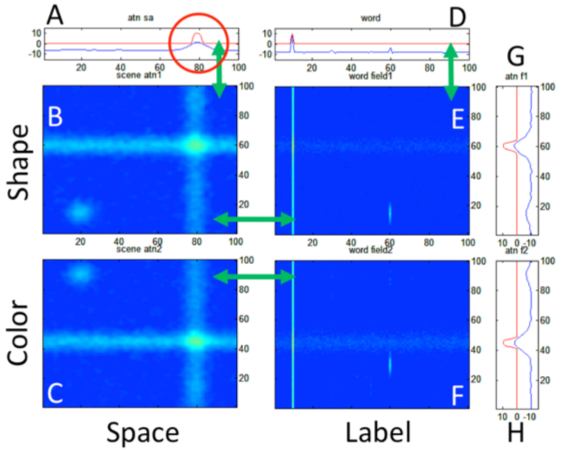

Fig. 1. The Word-Object Learning (WOL) model used by Samuelson, Smith, Perry & Spencer (2011) to capture 17- to 22-month-old children’s use of spatial memories to bind novel names to novel objects. See text for details.

Supplementary Materials

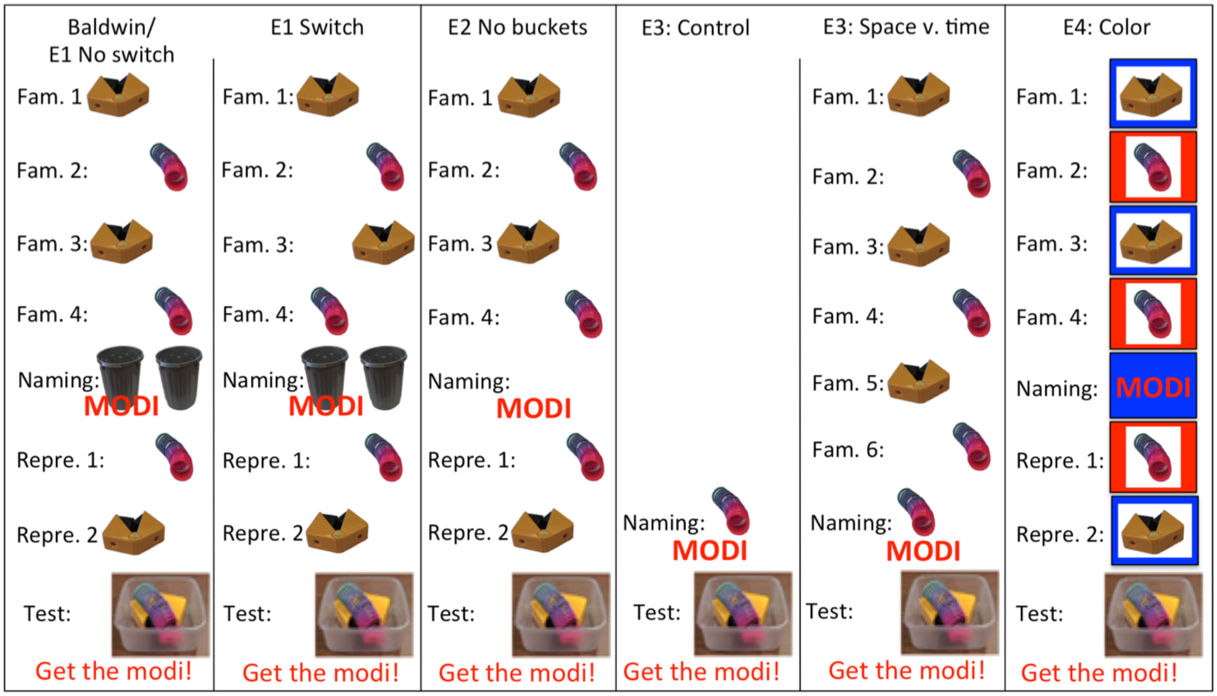

Fig. A: The discussed tasks of Samuelson et al. (2011). In the first experiment Samuelson et al. replicated the original Baldwin (1993) task and implemented a No-Switch experimental condition to test the necessity of spatial consistency for children's performance. In the next experiment the buckets were removed. The experimenter pointed to the empty space on the table where one of the objects had been during familiarization and said the name. A third experiment pitted prior consistency in space against temporal contiguity. During the naming event in the experimental condition the experimenter pointed to and labeled a visibly-present object in an inconsistent spatial position. A control condition confirmed that children this age would bind a name and object presented ostensively. A fourth experiment tested the DNF prediction that children could not use color cues to bind names to objects. See main text for additional details.

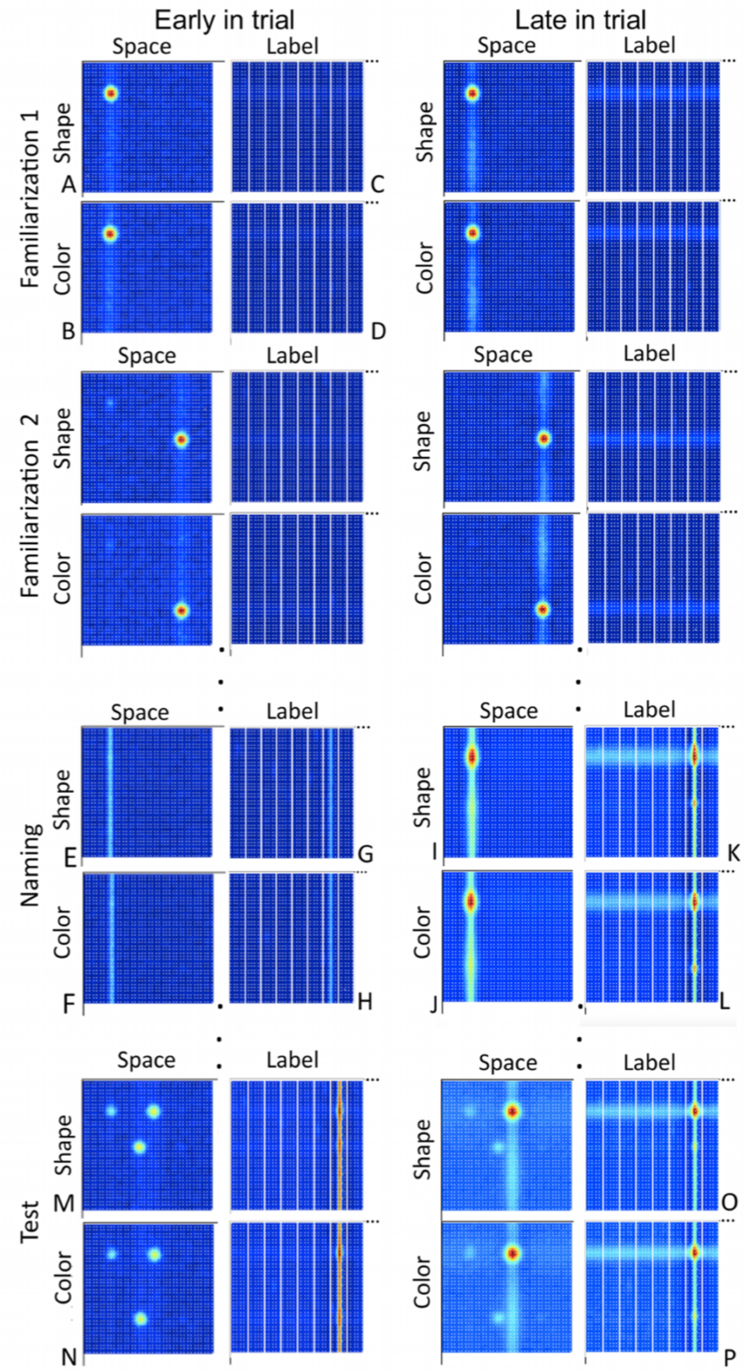

Fig. B: A simulation of the model at key points in time as it captures the events in the experimental task. Note that the 1D attention fields shown in Fig. 1 (main text) have been removed from this depiction.

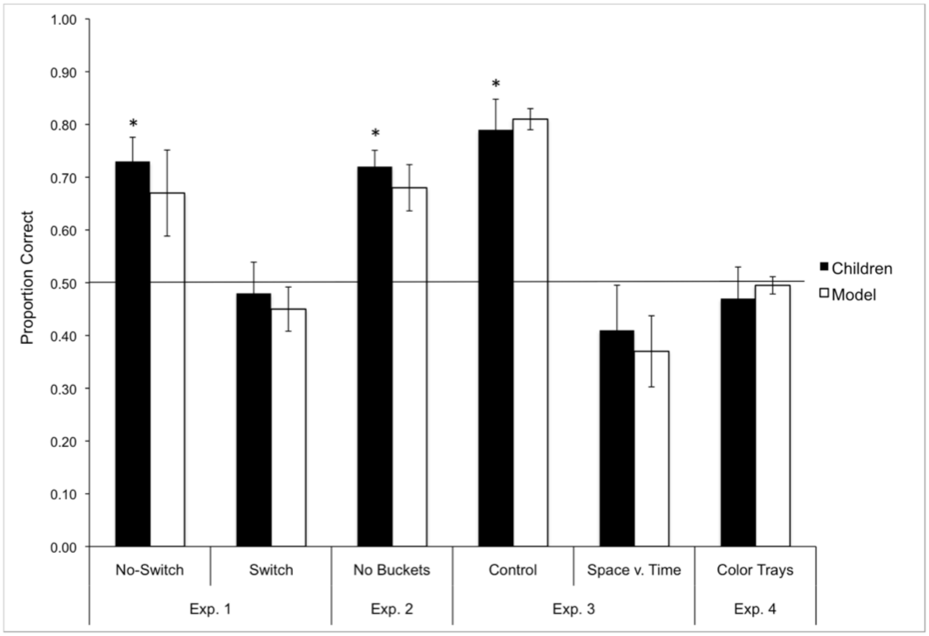

Fig. C: Performance of the 17- 20-month-old children and model in Samuelson et al. 2011. Children's percent correct choices for each experiment (black bars) with standard deviations (range of error bars). Twelve children were run in each condition of each experiment and no child participated in more than one condition. *s indicate performance significantly above chance (.50 in a two item forced-choice task). The mean performance of the Word-Object Learning model (across 12 batches of simulations) for all experiments is also shown (white bars). Error bars show the standard deviation of the model's performance (across 12 batches of simulations) per condition, relative to the target means.

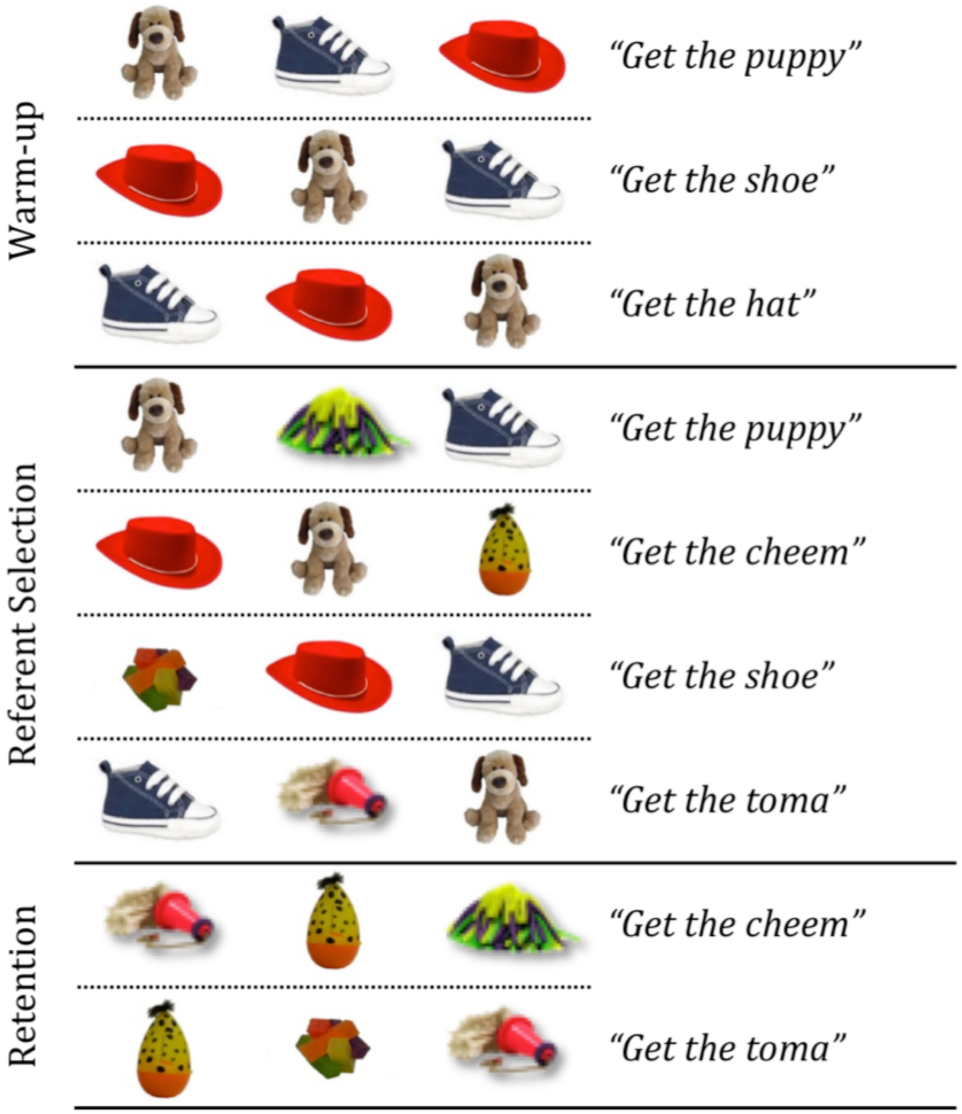

Fig. D: The structure of the referent selection and retention task we use to examine the interaction between novelty-driven attention and lexical knowledge in children's mapping of novel names to referents. Warm-up trials use three objects parents indicate that their children know the name of prior to the task. On referent selection trials children see two objects they know the name of and a novel object. Objects are requested with either a known-name or with a novel name. On retention trials children see three of the novel objects presented during referent selection. Objects are requested with the novel names mapped during referent selection. A 5-minute delay separates the referent selection and retention trials.