Learning your language, outside-in and inside-out*

STEPHEN A. HOCKEMA AND LINDA B. SMITH

Abstract

Children learn how to learn language, and they get better as they go along. This article presents an overview of research inspired by a dynamic systems view of language learning that shows it to be a self-organizing process in which children create the units they need from the regularities present in the environment in which they are situated.

1. E-language vs. I-language

The process by which a child learns to communicate with language is remarkable even to the most casual observer. Within a span of about four years, the crying, disorganized newborn transforms into a child who argues and tells stories with inventive fluency. If one were going to build a device that could accomplish this same feat, what kind of device would that be? This article considers an approach to language learning in which children do not begin with a "language-acquisition device" but rather build one as they go. The emerging device becomes increasingly adept at learning the specific language the child needs to learn.

Chomsky (1986) once proposed a distinction between what he termed *I-language* and *E-language*. At the risk of caricature, the distinction was between "internal" and "external" language, with the first representing the linguistic competence of individual speakers and the second representing the observable patterns of utterances across a community of speakers. Following Chomsky, most linguists made I-language the primary focus of study (Chomsky 2000). This has seemed necessary under the assumption that E-language was insufficient to determine I-language. As Chomsky put it:

> ... the child knows *vastly more* than experience has provided. That is true even of simple words. The words are understood in delicate and intricate ways that are far

Linguistics 47–2 (2009), 453–479
DOI 10.1515/LING.2009.016
0024–3949/09/0047–0453

beyond the reach of any dictionary, and are only beginning to be investigated. When we move beyond single words, the conclusion becomes even more dramatic. Language acquisition seems much like the growth of organs generally; it is something that happens to a child, not that the child does. (Chomsky 2000: 6–7 [emphasis added])

Some theorists, not of language but development, however, offer a different view.

If analyzed dynamically, this alloy of speech and action has a very specific function in the history of the child's development; it also demonstrates the logic of its own genesis. *From the very first days of the child's development his activities acquire a meaning of their own in a system of social behavior and, being directed towards a definite purpose, are refracted through the prism of the child's environment. The path from object to child and from child to object passes through another person*. This complex human structure is the product of a developmental process deeply rooted in the links between individual and social history. (Vygotsky 1978: 30 [emphasis added])

In brief, whereas Chomsky's focus is on (internal) *knowledge*, and framing of the problem as *acquisition*, Vygotsky's focus is on language *function* (especially social) and a framing of the problem as *development*.

There are a number of correlated (and alliterative) concepts that tend to "run with" the distinction between external and internal. E-language theories often include other "e-words" concepts, such as *empirical, embedded, extensional, environmental, ecological, explicit, emergent, epigenetic, experiential*, etc. I-language theories in contrast are often associated with the now-traditional viewpoint emphasis on such concepts as: *innate, intensional* and *intentional, intrinsic, implicit, ideal*, and a focus on and the *individual* and the *initial* state.

In this paper, we add the word *embodied* to the list of e-words, as a step toward worldviews dissolving the internal/external distinction (cf. the philosophy of Andy Clark (1997), Edwin Hutchins (1996) and Lev Vygostky (1978). The body is the interface between internal mental life and the external world. This approach is strongly informed and shaped by *dynamic systems theories* of development of cognition and action. The goal is an embodied, *mechanistic* account of how internal language is created from external language (Thelen and Smith 1994).

We build our case primarily concentrating on phenomena related to word learning. The case for this approach concentrates on two ideas raised in the Chomsky quote of the previous section:

(1) That "the child knows vastly more than experience has provided", and
(2) That words really are "understood in delicate and intricate ways"?

The language environment in which the infant is immersed from birth is rife with statistical regularities at multiple levels ideally exploitable by the (domain-general) associative and statistical learning mechanisms that infants bring to the world. Rather than being overly "delicate", words are *learned* via *robust* and yet *simple mechanisms*. Moreover, the *body* and *space* play an important role, binding the temporal/structural regularities in the world and the learning mechanisms in the child together.

2. A grounded dynamic systems approach

Development is a process and it happens in and through real time. The main emphasis in the study of development is understanding and characterizing *change* and the factors that influence it. The dynamic-systems approach is distinguished from other approaches to development in its emphasis on the *multi-causal* nature of these changes, where the causes themselves are temporal, extending and changing through time, and can be both internal and external. The developing organism and its environment, the internal and external, are treated as a fully interacting (coupled) *system*. Changes anywhere in the system affect the dynamics of the whole (Smith and Thelen 2003).

The distributed and highly reciprocal causality means that language cannot be viewed in isolation nor as a segregated special domain. Instead, language is an emergent property of the system as a whole. To quote the philosopher Hubert Dreyfus (arguing against the use of microworlds in AI, but applicable to our views on specialized domains as well):

> My thesis, which owes a lot to Wittgenstein (1953), is that whenever human behavior is analyzed in terms of rules, these rules must always contain a *ceteris paribus* condition, that is, they apply 'everything else being equal'; and what 'everything else' and 'equal' mean in any specific situation can never be fully spelled out without a regress. Moreover, the ceteris paribus condition is not merely an annoyance which shows that the analysis is not yet complete and might be what Husserl called and (sic) 'infinite task'. Rather the ceteris paribus condition points to a background of practices which are the condition of the possibility of all rule-like activity. In explaining our actions we must always sooner or later fall back on our everyday practices and simply say 'this is what we do' or 'that's what it is to be a human being'.... (Dreyfus 1997: 180)

From this point of view, the scientist's task is to characterize what matters, identifying the most important factors through different stages. In the case of language, these factors will necessarily include social and contextual factors, facts about the child's changing abilities at different ages, facts about perception and cognition in general (e.g., short-term memory span, attentional bandwidth, multi-modal perceptual integration, gross and fine motor abilities), and facts about the environment (see, for example, Gathercole et al. 1999 and Samuelson and Smith 2000).

The mechanisms and causal factors that drive development are constantly and continuously changing, and development consists in the accumulation of *moment-to-moment* change — changes with cascading effects. The language and concepts of dynamic systems theory are useful for thinking about changes both at the micro scale, in terms of coupled differential equations and instantaneous parameter values, and at the macro scale, in terms of trajectories, variability and stability. Hence, we can move beyond discrete characterizations of child language (i.e., "either the child knows this about their language — for example, how to make a noun plural — or they do not"). Under a dynamic-systems framework, linguistic knowledge is in the *system* and expressed through interactions between the child and the environment. So whether or not a child can, say, make a noun plural, is a question about a real time event that will always be dependent on multiple factors, including the child's own state ("linguistic" and "nonlinguistic") the current context, and the intrinsic dynamics of the system which in turn depend on the history of the child in the environment. There is no division between competence and performance, no division between knowledge and process.

2.1. *Self-organization: emergence and learning*

The idea of *emergence* is that interesting, beautiful and/or complex higher-level or "global" patterns can be created from systems with constituents whose behavior is governed by simple, lower-level or "local" rules. Thus there is a higher-level structure is nowhere explicitly specified, encoded, or represented in the system, nor is it imposed externally, yet it arises from the *interactions* among many "dumb" components and the environment in which they exist.

Emergence can be seen in how paths form. Given an open, grassy field on a college campus surrounded by buildings, the worn paths that eventually form will reflect the behavior of large numbers of autonomous college students, faculty and staff, none of whom significantly impact the resulting paths individually and none of whom act with the global path structure in mind. Yet, interestingly, in many cases the resulting path structure tends towards minimizing not the path distance between any two buildings, but rather the total path length (Goldstone et al. 2006). The approximated solution (known as a Minimal Steiner Tree [MST]) is mathematically optimal at minimizing the path length traveled between

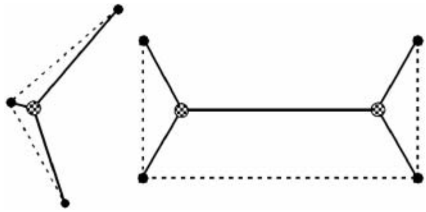

Figure 1. *Examples of Minimal Spanning Trees (dashed lines) and Minimal Steiner Trees (solid lines) connecting arrangements of 3 and 4 black points. The Steiner points, added to the configurations to reduce the total path length connecting the points, are shown as checkered dots (Goldstone et al. 2005).*

any two arbitrarily chosen buildings on average. (Figure 1 shows examples of two such MSTs.) Finding such a path structure is known to be an NP-complete computational problem, yet it will emerge from a collection of agents, "dumb" in the sense that they have no knowledge of the goal, interacting in a common environment. However, the walkers are not the only critical component in this incremental and emergent solution. The environment is also crucial as it is the means by which one agent "leaves its mark" to influence future agents. Each agent's behavior modifies the environment in which it behaves, which then feeds back into future behavior of that and other agents.

Minor changes to the system brought about through the interaction of simple "dumb" mechanisms and a structured environment create higher-level and "smart" structures.

2.1.1. *Language learning as emergence.* In the context of language learning, emergence contrasts sharply with traditional notions of "hypothesis-testing" and the "poverty of the stimulus" for example, see O'Grady 1997: 283–284). Hypothesis testing requires the ability to *formulate* a general hypothesis (either explicitly or implicitly) and then do some form of inference to see if it holds in the environment — that is, actual "input" must somehow be made commensurate to and compared with the "expectations". In contrast, dynamic self-organization only requires gradual, minor, local changes to properties of the constituents (e.g., a slightly trodden path, or connection weights in a connectionist system, or, at a higher level, attentional biases — see Section 2.3 below) caused via their interaction with the environment.

Further, within dynamic systems, interactions do not occur in only one direction. As the learner changes, there are certainly "internal" effects, but there are also "external" effects: in a real sense, the *system* — the child in the world — gets smarter, not just the child. Through its own activity, the child restructures their environment as well as their internal processes and in this way fundamentally changes the learning problem and the space of possible hypotheses. For example, changing attentional biases, modulate the information from which it will learn and the scope (span) of the learning task (for an example, see Robinson and Sloutsky 2004).

How does self-organization come about? The child certainly comes to the world with an arsenal of built-in (but not necessarily language-specific) processes. Some which we know from both psychological and neuroscience research, include both basic Hebbian-style and competitive learning commonly modeled with connectionist networks and more constructive types of learning as well (e.g., as is modeled by the node recruitment employed by cascade correlation (Quartz and Sejnowski 1997). The system itself is heterogeneous and multimodal, providing lots of different "takes on the world". Internal mechanisms can become differentiated not by dedicated processes but by the cascading effects of different initial densities of cell assemblies and their intrinsic connections to other areas (usually based on proximity). Such connections can be excitatory and inhibitory and can be created anew in areas of high activity/need as well as locally strengthened or weakened by Hebbian-style learning and habituation. Brain and cognition are the emergent products of interaction between all this and the world. Given these general operating principles across all the modalities — all the heterogeneous systems — we classify most of these mechanisms as "associative" because at its core, learning is essentially about making and adjusting the strengths of *links* (connections).

2.1.2. *Not your father's behaviorism.* Claims like those in the previous two sections — e.g., that 'knowledge' is not 'in the head' but 'in the system' — often bring charges of behaviorism from theorists on both sides of the E-language / I-language spectrum. This is particularly true if the learning mechanisms are described as associative. For example, in a recent book Tomasello:

> Most people think it died with Behaviorism, but associative learning theory lives on. Smith (2000a) has argued that the essence of word learning is associating sounds with salient aspects of perceptual experience. (Tomasello 2003: 82 [the "a" was added to the Smith reference, as cited herein])

This criticism reflects a fundamental misunderstanding of both behaviorism (which denied cognition) and associative learning (which is about the formation and nature of the internal processes commonly called representations, see Smith 2000b).

However, dynamic systems approaches *are* more behaviorist than theories based on internal representations in the form of mental models or propositions in that dynamic systems theories explicitly care about the processes that link knowing to the physical/external world. Yet, it is precisely this focus on the (associative) learning mechanisms that undermines charges of behaviorism. Learning *is* about internal changes. The child's behavior is determined by processes internal to brain/mind. Indeed, one of the important contributions of the dynamic-systems can be seen as "restoring the balance" (that was tipped one way by the behaviorism revolution and then back too far the other by the cognitivism revolution) through its re-elevation of the "external" to place it *on equal footing with* the "internal."

2.2. *Situatedness: Embedded and embodied*

A system-level view of language learning requires understanding the *interactions* between individual and environment — between the traditionally internal and external. This interaction is possible — learning is possible — because individual minds are *situated* — embodied and embedded — in real time in a physical world. Situatedness allows the statistical and associative learning mechanisms that a language learner brings to the task to be engaged (see Ballard et al. 1997 for low-level details as to how). Thus, to understand how the system develops, it is crucial to understand the ramifications of situatedness. One such ramification is the elevation of *space* to a mechanistic role in language comprehension and learning, a domain in which aural information and temporal relationships have traditionally had a monopoly. This is because cognition takes place in real time in a body that is spatially oriented in the physical world. This fact is key to how we form the right associations to learn language and to how we activate and organize knowledge appropriate to the specific task at hand.

Furthermore, spatial regularities are not fixed: children learn to use the space around them to organize their cognition, for example creating spatial groupings of similar objects to make categories into *perceptual* groups (Vygotsky 1978: Chapter 2; Vygotsky 1986: Chapter 5; McGonigle and Chalmers 2003) or using space as an index into memory (Ballard et al. 1997; Richardson and Spivey 2000; Spivey and Geng 2001). Crucially, this ability *develops* as well and the interaction eventually involves language and is not one-way (Boroditsky 2001). The child gradually internalizes the skills being learned via space into their cognitive and linguistic skills. There is much recent research to this effect (e.g., Casasanto and

Boroditsky 2003, Matlock et al. 2005; and see Gentner 2001 for a relevant discussion of several ways representational systems for time and space could be related).

Consistent with these notions is evidence suggesting that gesture, speech, and thought are inherently linked, forming a single communicative system such that gesture, speech and thought mutually affect one another at multiple levels (Iverson and Thelen 1999; McNeill 1992). The embodied nature of meaning is also suggested by recent findings that some verb meanings are understood in terms of the body (Richardson et al. 2003) and even *selectively interfered with* by motor activity (Toskos et al. 2004). Glenberg and Kaschak (2002) have also presented strong evidence that the motor system is engaged when determining the meaning of several (syntactically) different types of sentences, both concrete and abstract.

3. Research

This section reviews findings — mostly in the realm of lexical learning — in support of these ideas: that language learning is both "outside-in" and "inside-out", that it is the emergent result of embodied (and thus spatial) interactions in a world replete with regularities on many scales.

3.1. *Regularities for learning*

Just as a newborn comes into existence in a physical reality with gravity, inertia, friction, etc. all of with which they can learn to exploit, they come into a world in which language is pervasive and a strong impetus to the self-organization of the system. From this immersion, the deeper principles through which the environment (E-language) is organized will "soak in" and become internalized, creating units of perception and action. We will illustrate these ideas by considering regularities at the various levels in the language system that are exploitable by associative and statistical learning mechanisms. Critically, these mechanisms when placed in a structured environment create new perceptual and conceptual "units" that can then be hierarchically applied to further speed and constrain learning.

3.1.1. *Phonemes.* Phonemic distinctions vary between languages and exposure to language plays a strong role in setting up the contrasts that comprise the phonemic system of a language. Moreover, young infants are open to learning about all possible phonetic contrasts and sometimes even discriminate them better than adults (e.g., Trehub 1976; Werker et al. 1981). However, by 12 months of age, infants lose their sensitivity to contrasts not present in their native language and show a language-specific pattern of discrimination (Werker and Tees 1984; Best et al. 1988; Bosch and Sebastian-Galles 2003). Recent evidence suggests that these developmental changes may be due to statistical learning about the distributional patterns of sub-phonetic features (Anderson and Morgan 2002; Maye et al. 2002; Maye and Weiss 2003).

In one study, Maye et al. (2002) presented 6–8 month old infants with syllables drawn from a continuum of sounds varying from *da* to *ta*. They manipulated the frequency distribution with which the individual sounds were presented. One group of infants was given exposure to a distribution in which sounds from the center of the continuum were more frequent than sounds from the ends. A second group was given exposure to a bimodal distribution in which sounds at the end were highly frequent and those in the middle were infrequent. The infants were highly sensitive to this distributional information. Following several minutes of exposure only infants in the bimodal condition discriminated the endpoint tokens, whereas those I the unimodal condition did not. Maye and colleagues interpret this result in terms of a loss of a discrimination in the unimodal group (and perhaps an enhancement of a discrimination in the bimodal case). In this way, the feature clusters (or phonemes) that comprise the primitives for phonological rules may be made in development, out of the statistical regularities in the sound patterns that children hear.

3.1.2. *Word segmentation.* One highly influential series of studies, Saffran et al. (1996), Saffran et al. (1997) and Aslin et al. (1998), showed that 8-month-old infants could discover wordlike units from as little as 2 minutes of exposure to strings of syllables. In their studies, they presented infants with a continuous two-minute stream of speech containing no pauses between words (e.g., pabikudaropitibudopabikugolatu. . . .). Within that stream they placed "words" (e.g., pabiku and daropi) that were defined only by the transitional probabilities between syllables. For example, in one study (Aslin et al. 1998), the transitional probability of one syllable following another within a "word" (e.g., between pa and bi) was equal to 1.0, and the transitional probability of one syllable following another when the two came from different "words" (e.g, ku and da) was .33. Aslin et al. found that after a brief exposure to such streams of speech, infants recognized the embedded "words" as coherent units (showing a novelty response to non-words made up of word parts, or for example, a pa followed by something other than a bi). The formation of units through the learning of statistical co-occurrences has now been demonstrated many times (see Cleeremans et al. 1998; Gomez and Gerken 2000; and Saffran 2003 for reviews). Creating units from statistical regularities appears to be a central characteristic of human (and primate) learning across speech, auditory, and visual stimuli (Kirkham et al. 2002).

Of course, these processes — evident in artificial language learning tasks — only matter if natural language presents the relevant statistical patterns. At least in the case of finding words in the English speech stream, the relevant transitional probabilities are there. Hockema (2006) recently examined phoneme transition probabilities in English-speaking parents' speech to their children and found that adjacent phoneme pairs fell virtually into two types: those occurring within a word 90% or more of the time and those occurring between words 90% or more of the time.

Furthermore, there appears to be great redundancy in the statistical cues to word boundaries. Not only are they specified by phoneme transition probabilities, but they are also specified by several other cues in the speech stream to lesser extents (including prosodic stress patterns [Cutler 1996] and allophonic variation [Jusczyk 2002]). And the information is not just limited to the speech stream. Yu et al. (2005) have utilized multi-modal information available in the combination of the raw visual and audio sensory stream to successfully extract words (and even determine their referents) as well. Apparently, there is not a poverty of the stimulus when it comes to word segmentation, but rather the opposite: words are readily recoverable from an abundance of overlapping and converging statistical information.

3.1.3. *Syntactic and semantic categories.* Higher-order units — syntactic categories such as nouns and verbs or semantic categories such as animate and inanimate — may also be formed through co-occurrence relations among words (units themselves formed from co-occurrence relations among sounds). Nearly thirty years ago, Maratsos and Chalkley (1980) argued for such a distributional approach to syntactic development. This idea was dismissed at the time on the assumption that were just too many irrelevant co-occurrences in the streams of words that a learner would hear for a simple counting of co-occurrences to arrive at the right syntactic categories (Chomsky 1965; Pinker 1987). However, recent evidence strongly indicates that this was an unwarranted concern.

Indeed, there is now a general consensus that the statistical regularities within language itself present clear evidence from which noun and verb categories can be induced (Mintz et al. 1995, 2002; Redington et al. 1998; Cartwright and Brent 1997; Elman 1990, 1998; Brill 1991; Finch and Chater 1992, 1994; Schütze 1993). The ability of statistical learning systems to identify nouns and verbs is robust under quite minimal assumptions about learning even when the learning system is sensitive to only co-occurrences among adjacent segments. The discovery of nouns and verbs is enhanced to even higher levels when other assumptions (such as learning about co-occurrence relations among nonadjacent segments or using function words as markers of phrases to be analyzed) are added (Mintz 2005). Similarly, co-occurrences among words have been shown to also yield subcategories of nouns (animates versus inanimates) as well as subcategories of verbs (transitive versus intransitives, Elman 1990; Li 2000; Burgess and Lund 2000; See also Rogers and McClelland 2004).

3.1.4. *Word meanings.* Much research has shown that some portion of a word's meaning is ascertainable from statistical regularities present in the patterns of word token usage across a variety of contexts. The most famous examples of this come from analyses of large text corpora by so-called high-dimensional vector-space approaches, such as Latent Semantic Analysis (LSA) (Landauer et al. 1998) or Hyperspace Analog to Language (HAL) (Burgess and Lund 2000) (also see Griffiths et al. 2005 for an alternative, but related, approach). What is relevant about LSA to the current discussion is that it gives an explanation for the rapid word learning — a "second vocabulary burst" — that happens in older children (10 to 12 year olds) through reading, by showing how a word seen only in one written context can be learned "deeply" via second-order correlations *through* other already-known words in its experienced context. This is an example of a form of *bootstrapping*, whereby learning can be accelerated by incorporating some of the higher-order regularities in the input.

And the regularities that children pick up on can potentially come from anywhere — either linguistic or non-linguistic. For example, another interesting line of recent research has to do with cues to word meaning embodied in the way they are pronounced, that is, in prosodic information: tone of voice, rhythm, pitch, intonation, pauses, and stress. Prosody is well known to help word segmentation (e.g., Cutler 1996) and the identification of phrase structure (e.g., Fisher and Tokura 1996), and it has also been long known that we can use it to correlate with word meaning (Kunihira 1971). For example, we might say, "It was soooooo biiiig," to emphasize or exaggerate a description of a large-sized object (see Yoshida 2003; Sasso et al. 2005). Other evidence suggests young learners also use gestures, points, and nods. These kinds of correlations in the learning environment are not yet systematically studied but may well prove crucial to understanding the embodied nature of word meaning.

3.1.5. *Acquired word-learning biases.* Language learning changes dramatically as it progresses (see Bloom 2000 for review). At about 10 months, infants show signs of comprehending a few words or phrases in highly specific contexts. On or about their first birthday, children start producing words. They add words to their productive vocabulary very slowly, adding words at a rate of one or two per week and often seem to lose ones they had produced earlier. It is not just that children have trouble producing new words at this point; word learning itself is hard at first. A study by Woodward, Markman and Fitzsimmons (1994) illustrates the point. They presented children with a novel object, for example a yellow strainer, and named it with a novel noun (e.g., "This is a toma."). The child was shown two other objects, one in the same category (e.g., another strainer) and the other a very different thing altogether. The child was asked to indicate which of these had the same name. Woodward and colleagues employed two types of trials: in one, the same-category instance was identical to the original instance; in the other, a generalization trial, the test object differed from the original instance only in color. In this task, 13-month olds succeeded when the target was identical to the original instance but chose randomly in three of four experiments when the target differed in color. (The 18-month-olds were successful in extending the name under both testing conditions.) Thus, these studies support the notion that very young word learners seem to need to hear each individual word in many contexts before they apprehend the range of the word.

Then, between 18 and 20 months, most children *become* very rapid word learners, adding new words to their vocabularies at the staggering rates of 4 to 9 *a day*. During this time, they only seem to need to hear a word used to label a single object to know the whole class of things to which the word refers (see Bloom 2000 for a review). This one-instance to whole-category learning is especially remarkable because different kinds of categories are organized in different ways. For example, animate categories are organized by many different kinds of similarities across many modalities; artifact categories are organized by shape and substance categories by material.

What developmental processes give rise to this increase in the ability of children to learn a whole category from a single instance? The evidence from both experimental studies and computational models strongly suggests that children learn how to learn words *as* they learn words. Specifically, as children slowly learn their first words, they learn the regularities that characterize different kinds of noun categories and the learning of these regularities then *creates* their ability to learn words in one trial. The nature of this learning can be characterized by four steps, illustrated

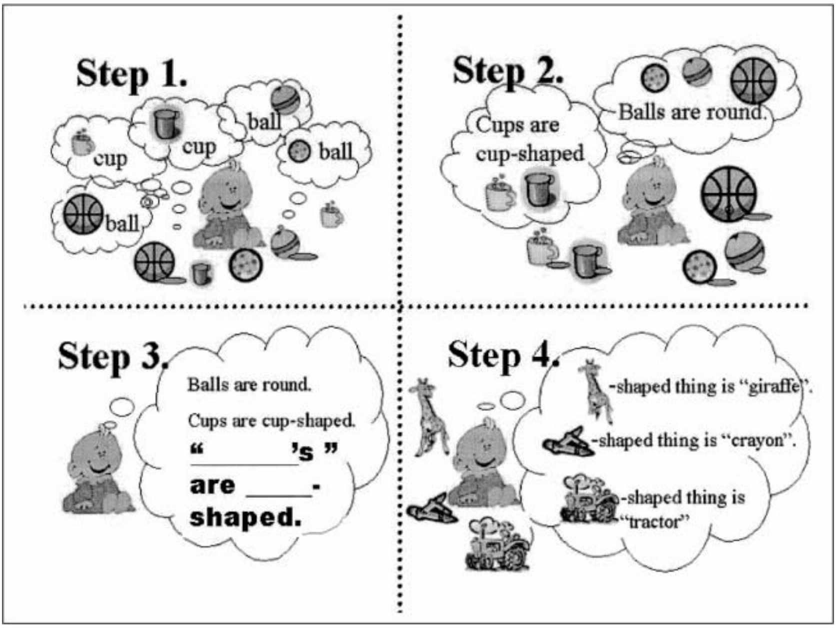

Figure 2. *The proposed four-step model by which object names and attention to shape come to be related. In Step 1, the child maps names to individual objects. First-order generalizations about the structure of individual object categories are made in Step 2. In Step 3, the child makes a higher-order generalization across learned categories about the common structure of named object categories — that is, that categories are organized by similarity in shape. Finally, in Step 4, the child has learned to attend to shape in novel name learning, leading to rapid name acquisitions (Smith et al. 2002).*

in Figure 2 (Smith et al. 2002). The figure shows just one of the regularities that children learn — that artifact categories are organized by shape. Step 1 in the learning process is the mapping of names to objects — the name "ball" to a particular ball and the name "cup" to a particular cup, for example. This is done multiple times for each name as the child encounters multiple examples. And importantly, in the early lexicon, solid, rigidly-shaped things are in categories typically well-organized by similarity in shape (Samuelson and Smith 1999). This learning of individual names sets up Step 2 — first-order generalizations about the structure of individual categories, that is, the knowledge that balls are round and cups are cup-shaped. The first-order generalization should enable the learner to recognize novel balls and cups.

Another higher-order generalization is also possible. Because most of the solid and rigid things that children learn about are named by their shape, children may also learn the *second-order* generalization that names for artifacts (solid, rigid things) in general span categories of similarly-shaped things. As illustrated in Step 3 of the figure, this second-order generalization requires generalizations over specific names and specific category structures. But making this higher-order generalization should enable the child to extend any artifact name, even one encountered for the first time, to new instances by shape. At this point, the child behaves as if it has an abstract and "variabilized" rule: For any artifact, whatever its individual properties or individual shape, form a category by shape. Step 4 illustrates the potential developmental consequence of this higher-order generalization — attention to the right property, shape — for learning new names for artifacts. The plausibility of this account has been demonstrated in simulation studies that present neural networks with the regularities that characterize common noun categories and those regularities appear sufficient to create a learner who attends to just the right properties for animals versus artifacts versus substances (Colunga and Smith 2005).

The real-world developmental consequences of these ideas have also been demonstrated in experimental studies that effectively accelerate the vocabulary acquisition function by teaching children the relevant correlations (Smith et al. 2002). In a 9-week study, children too young to show a shape bias in the generalization of artifact names were brought to the laboratory once a week and taught names for objects in four lexical categories, all well-organized by shape. At the end of the experiment, these children (but not untrained controls) were able to "learn a whole category" from hearing a single thing named, extending the name systematically by shape. More dramatically, these trained children's rate of new object name acquisitions *in the world* increased substantially, by over 350% (Smith et al. 2002; a result now replicated several times, e.g. Samuelson 2002). These results suggest that learning object names teaches children to attend to category-relevant properties, and doing so changes object-name learning — taking it from a slow process to one-trial learning.

These regularities that children learn and exploit to become fast object-name learners are between names and the *categories* of objects they label. One might ask: why do the categories have the regularities that they do? Why are instances of animate categories typically similar on many properties? Why are most artifact categories about solid things that are the same shape? Why are substance categories well-organized by material? The answer to these questions seems likely to lie in the interface among physics, biology, and human psychology. For example, artifact categories seem likely to be typically solid and defined by shape because the functional possibilities of things (their *affordances*, for example, ability to hold water or to support a sitter) are determined by their physical structure, but the functions that matter to human categories are determined by human psychology. The point is that the language-learning mechanism — children's ability to attend to just the right properties for learning the names of different kinds — is the product of regularities among words and the things to which they refer, among the physical properties of things, and in the regularities that emerge as a consequence of human goals, abilities and actions.

3.1.6. *From surface regularities to deeper regularities.* If the language learning mechanism is built from the very regularities it is learning, how does it know where to start? Where in this sea of regularities, should a learner begin? An unbiased learner would begin with — and build their language learning mechanism out of — whatever regularities are most reliable, pervasive, and salient. This leads to the possibility that learners may start with surface regularities that only roughly correspond to the deeper (higher order) regularities that characterize mature language. This idea has attracted many theorists in language acquisition from a variety of perspectives and is typically discussed under the rubric of *bootstrapping* (Landau and Gleitman 1985; Naigles 1990, 1996; Fisher et al. 1991; Pinker 1984, 1987); one insight being that easy-to-detect surface regularities may provide the learner with a partial solution that also moves the learner closer to the ultimate solution (Brent 1994; Laakso and Smith 2007). Indeed, these recent analyses of large corpora of parent speech to children suggest that relevant regularities may show up in places unexpected some 20 years ago before the possibility of computational analyses of the statistical structure of language.

One area of growing interest is the role of pronouns. Pronouns have two properties that should be relevant: frequency and a small set of surface forms. They are the most common syntactic subjects and objects in speech to children (Valian 1991; Laakso and Smith 2004, 2007) and in general discourse (Chafe 1994). Their sheer frequency makes them potentially important as statistical predictors of other aspects language. Consistent with this idea, Childers and Tomasello (2001) suggested that children acquire lexically specific frames such as "I do it" as a way into learning syntactic frames. Cameron-Faulkner et al. (2003) also observed that parents use the inanimate pronoun *it* far more frequently as the subject of an intransitive sentence than of a transitive one. Apparently in parent speech to children intransitive sentences are used more often than transitives for talking about inanimate objects, and thus the inanimate pronoun might serve as a cue to some semantic aspects of the verb.

Laasko and Smith (2007) additionally suggest that pronouns may also help learners partition verbs that express psychological attitudes toward events and states of affairs into two rough categories. They found that in speech to young children verbs that express deontic status, such as goals, purposes or intentions (*try to*), volitions or desires (*want to*), and compulsions (*have to*) tend to take infinitival complements, whereas verbs that express epistemic status, such as perceptions (*see that*), beliefs (*think that*), and knowledge (*know that*) tend to take sentential (propositional) complements (see also Tomasello 2003). Further, in the ecology of early childhood, parents tend to be the ones who *know* whereas children tend to be the ones who *need*, and Laasko and Smith found that in speech to children *you* and *I* strongly predicted the class of verb meaning. Interestingly, *I* may be a marker — a pointer — for bringing together verbs of a class, linked to the experience of who — from the child's perspective — is omnipresent.

If surface forms such as pronouns are bootstraps, they must actually take children to the deeper regularities that characterize adult language and enable productivity, etc. Laasko and Smith (2007) investigated this idea by presenting the pronoun-verb co-occurrences in parental speech to a connectionist network. The simulations both document the statistical regularities that exist between pronouns and specific verbs and also show that these regularities enable higher order generalizations that go beyond the actual co-occurrence regularities in the input data.

To summarize, there are statistical regularities in co-occurrences between pronouns and verbs in the speech that children hear from their parents. A simple statistical learner such as a connectionist network can exploit these regularities, including subtle higher-order regularities that are not obvious in a casual glance at the input data.

3.1.7. *Outside-in.* As the above examples elucidate, one of the first things one notices when considering language learning is how the task changes: from making sounds, to orienting to and learning words, then to stringing them together. The rate at which learning happens also changes — sometimes radically. As Hohenberger and Peltzer-Karpf (2005, this issue) put it, there are rapid reorganizations brought on by non-linearities in the system. In the present view, language is both a product and a cause of these changes. As children learn how language organization reflects the organization of their environment it also changes how they learn. For example, each new noun learned changes what they know about learning nouns, eventually leading English-speaking children to a shape bias for artifacts. And, in turn, learning this shape bias both accelerates the rate of word learning and changes the way infants perceive shape (Smith et al. 2002; Smith 2004; Jones 2003). In this way, language learning alters foundational cognitive processes, and it alters itself.

Three additional points about this "outside-in" notion of learning statistical regularities merit mention. First, the regularities might not always be what one expects them to be. There might be different solutions to the same problem depending on the language and environment. Indeed, as we saw in the case of word segmentation, there are abundant regularities in several different information sources in English: phoneme transition probabilities, prosodic stress patterns, and perhaps also multi-modal correlations embodied in where mother and child are looking during speech.

Second, in a sea of regularities, linguistic organization will start with "surface" regularities and "works its way in", building on itself each step of the way. The regularities the learner begins with need not be sensible or perfect. Nor does the starting point have to be exactly the same for every child. All that is necessary is for the dynamic forces (including the environment and the intrinsic dynamics of the learner) to yield a self-organizing solution that is fluency in that language. The openness of the system to different self-organizing solutions is imperative because some children end up speaking English while others end of speaking, for example, Tamil. Although there are a great deal of abstract similarities among languages (as research into universal grammar has taught us), there are also real differences among languages that (as Slobin 1985 pointed out) that can have profound and cascading consequences to the learner. There is growing evidence on this possibility that learners of different languages face fundamentally different learning tasks as in the domains of noun and verb learning (e.g., Yoshida and Smith 2001, 2003, 2005; see also description of Sethuraman 2005), and in spatial terms (Levinson 2003). Indeed, given the self-organizing nature of language learning and the fact that children are limited only to the information available to them in their native language(s), it is likely that children growing up in different linguistic environments will take different paths — utilize different information — to arrive at the solutions to the common problems they all face and solve.

We offer one illustrative example: in English the expression of a verb's arguments is obligatory; in Tamil these arguments can be (and typically are) dropped. Thus, whereas English-speaking children can use argument structure as a bootstrap to verb meaning (Naigles 1990; Ahrens 1995), Tamil-speaking children cannot and indeed appear to use learned verb meaning as a bootstrap to argument structure (Sethuraman 2005). Different languages may create different learning tasks and different developmental trajectories. However, all learners of all languages will also be constrained to similar solutions by the sameness of the physical world, by the morphology of the human body, and by shared learning mechanisms.

3.2. *Embodiment*

3.2.1. *The role of space.* In an influential series of studies, Baldwin (1991, 1994) used a clever experimental method to show that infants between 16 and 19 months could determine the intended referent for a novel label even when the referent was not being attended to by the infant (nor in view). In fact, the labeling occurred while the infant was attending to another possible, but incorrect, referent. (The correct referent was put in a bucket held by the experimenter and then labeled while the infant was playing with, and looking at, another toy.) Baldwin interpreted this to be evidence that the infants possess a word-learning strategy that "goes beyond simple associative processes such as temporal contiguity" (Baldwin 1994: 147).

In order to understand what factors were allowing the infants to make the link between label and referent, Smith (2003) extended Baldwin's method to better investigate the role being played by *space*. Infants were presented with two objects three times in a row, with the presentation of each object consistently made on the same side. For example, object A would always be presented on the left and object B on the right every time the pair was presented. After allowing the infant to play with the objects, they were taken away and the presentation then repeated with the objects again on the same side (e.g., A on the left and B on the right). After two repetitions, the experimenter put the toys out of sight and then pointed to and labeled a spot on either the left or right side of the table with a novel word (e.g., *modi*). The objects were then placed into a bucket and the infant was asked to get the *modi*. In these conditions, infants mapped the label to the referent that had been presented on the same side where the experimenter pointed. This effect was found even if the experimenter did not point to the table during the labeling, but rather snapped in the air to direct the child's attention to one side or the other.

These effects have now been replicated using three objects as well on children between 17 and 24 months (Smith et al. 2004). Interestingly, in this study, children with large vocabularies (as measured by the MacArthur-Bates communicative development inventory) were *less* successful at using spatial cues to bind a novel label to a novel referent than children with small vocabularies. This suggests that as children become better word learners, they learn to move beyond the surface cue of spatial association to deeper or more relational cues, yet another kind of bootstrapping hypothesis that is consistent with the Emergentist Account of lexical learning (see Hollich et al. 2000) as well as the evidence presented in Section 3.1 that children learn to utilize different types of information as they learn to learn words.

Taken together, these results indicate that young infants use space to bind labels to their referents, even when the referents are not in view. In fact, the power of this cue was further demonstrated by building up a memory for an object in a spatial location and then labeling another object in that same location, pitting a potential referent in plain view against a spatial memory. The effect was striking (Smith 2004). For example, if after the initial training with object A on the left and object B on the right they were switched and B was put on the left and then labeled *in full view* of the child (the experimenter pointed directly to object B and said *modi*), infants were subsequently at chance at determining the referent for the label!

It seems that infants can build up strong associations between an object and its spatial location to the point where the spatial location can act as a surrogate for the object in labeling. (What is especially noteworthy about this is that the external binding through space seems to be developmentally essential for this progress to more complicated forms of binding to be made that no longer use space.) The point is this: Our environment has a consistent spatial structure, and our cognitive processes will make use of that regularity (see also Richardson and Spivey 2000; Spivey and Geng 2001; Yu and Ballard 2007; Yu et al. 2005; Hund and Plumert 2005).

3.2.2. *The role of the body.* Space is not the only "external" medium that can subserve language learning. There is also considerable evidence that people use their own bodies to represent meanings, facilitate memory, and find higher-order structure. For example, people will look at a blank region of space where something they are trying to remember was previously displayed (Richardson and Spivey 2000; Spivey and Geng 2001), presumably because these movements help to cue recall by recreating a part of the body's state at the time of memory encoding. Furthermore, when imaging a simple story with a directional component, eye movements will tend to be in the same direction as the story (Spivey and Geng 2001). Moreover, Grant and Spivey (2003) were able to improve subjects' chances at solving an "insight problem" (a problem whose solution cannot be logically deduced) by manipulating eye movements to regions of a diagram on which fixations had previously shown to be correlated with success.

These interactions also apply to both language understanding and production as well. In terms of word meanings, it appears that the meanings of many verbs are represented in terms of the body. Toskos et al. (2004) showed that having subjects perform repetitive horizontal or vertical eye movements while listening to a list of verbs selectively hampered their recall for verbs that were previously ranked by other subjects as highly directional and/or concrete. (Also see Richardson et al. 2003.) This can be taken to indicate that the meanings of some verbs are encoded in such a way that motor activity can interfere with their processing. Meanwhile, Maouene-Cavin (2005) has shown that children (2.5 and 3.5 years old) typically encode a body part as part of the meaning of early-learned action verbs and that verbs may be progressively learned in clumps centered on one body part at a time. And, in a survey of the spatial relation terms of 125 African languages, Heine et al. (1991: 126–131) found that more than three-quarters of the terms whose etymology was known were derived from human body parts.

At the sentence level, Glenberg and Kaschak (2002) discovered what they term an "action-sentence compatibility effect". Subjects were asked to judge the sensibility of sentences and respond by pressing a button on a board that required them to either move their hand toward or away from their body. If the sentence meaning implied action in a direction opposite to the direction that the response required, subjects were significantly slower than if the sentence meaning and response direction were congruent. Glenberg and Kaschak interpret these results as evidence that the motor system was engaged during sentence comprehension, consistent with their Indexical Hypothesis (Glenberg and Robertson 1999) that meaning — even sentence meaning — is based on action. They have speculated that the meanings of sentences are composed ("meshed") out of perceptually-grounded word meanings (based on affordances) using the meanings of constructions (e.g., Goldberg 1995; Tomasello 2003) to govern the "meshing" process.

Again, the role of the body and the learner's own actions in self-organizing by processes of language learning is still promissory. But again, we suspect deep regularities that may serve as pointers to higher order meanings.

4. Discussion

Critics of this paper may appropriately offer two criticisms: First, how new is all this? How really different is it from developmental theorists such as Vygotsky (1978, 1986) or Piaget (1953) or the "whole-child" perspective of Lois Bloom (1995) on language learning? Second, the critic might add, it's mostly all promise. Although the paper points to the creation of units such as phonemes and words and attentional biases that speed up the learning of nouns, it is a far cry from a self-organized system in which the parts are integrated into a self changing whole. The critics are right on both accounts. Piaget's (1953) constructionist approach to cognitive development and Lois Bloom's "whole-child" perspective contain similar ideas to those presented here. There are others as well, particularly so-called connectionist theorists such as Rogers and McClelland (2004) who also see language learning as a self-organizing system but who perhaps have not yet paid enough attention to the role of the child's own activity — an embodied being — in creating those regularities. This, we believe, and relevant to the second criticism is where the work needs to be done. In putting the whole together — the inside, the outside, and the body which connects the two.

Received 13 May 2005
Revised version received
19 March 2008

University of Toronto
Indiana University

Note

* The first author of this article was supported through a grant from the National Institute for Child Health and Human Development (T32 HD07475). Some of the research presented in this paper was supported by grants to Linda B. Smith from the National Institute of Child Health and Human Development (RO1 HD28675) and the National Institute of Mental Health (RO1 MH60200). Correspondence address: Linda B. Smith, Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405. E-mail: smith4@indiana.edu.

References

Ahrens, Kathleen. 1995. Verbal integration: The interaction of participant roles and sentential argument structure. *Journal of Psycholinguistic Research* 22. 497–516.

Anderson, Jennifer & James L. Morgan. 2002. A statistical basis for speech sound discrimination. Poster presented at the International Conference on Infant Studies in Toronto, Canada.

Aslin, Richard, Jenny Saffran & Elissa Newport. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9(4). 321–324.

Baldwin, Dare A. 1991. Infants' contribution to the achievement of joint reference. *Child Development* 62. 875–890.

Baldwin, Dare A. 1994. Infant contributions to the achievement of joint reference. In Bloom, Paul (ed.), *Language acquisition: Core readings*, 129–153. Cambridge, MA: MIT Press.

Ballard, Dana H., Mary M. Hayhoe, Polly K. Pook & Rajesh P. N. Rao. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Science* 20. 723–767.

Best, Catherine T., Gerald W. McRoberts & Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by

English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance* 14. 345–360.
Bloom, L. 1995. *The transition from infancy to language*. New York: Cambridge University Press.
Bloom, Paul. 2000. *How children learn the meanings of words*. Cambridge, MA: MIT Press.
Boroditsky, Lera. 2001. Does language shape thought? Mandarin and English speakers' conception of time. *Cognitive Psychology* 43. 1–22.
Bosch, Laura & Núria Sebastián-Galles. 2003. Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language & Speech* 46. 217–243.
Brent, Michael. 1994. Surface cues and robust inference as a basis for the early acquisition of subcategorization frames. In Lila Gleitman & Barbara Landau (eds.), *The acquisition of the lexicon*, 433–470. Cambridge, MA: MIT Press.
Brill, Eric. 1991. Discovering the lexical features of a language. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics Berkeley, CA: Student session*, 339–340. Morristown, NJ: Association for Computational Linguistics.
Burgess, Curt & Kevin Lund. 2000. The dynamics of meaning in memory. In Eric Dietrich & Arthur B. (eds.), *Cognitive dynamics: Conceptual change in humans and machines*, 117–156. Mahwah, NJ: Lawrence Erlbaum.
Cameron-Faulkner, Thea, Elena Lieven & Michael Tomasello. 2003. A construction-based analysis of child directed speech. *Cognitive Science* 27. 843–873.
Cartwright, Timothy A. & Michael Brent. 1997. Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition* 63. 121–170.
Casasanto, Daniel & Lera Boroditsky. 2003. Do we think about time in terms of space? In Richard Alterman & David Kirsh (eds.), *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, 216–221. Boston, MA: Cognitive Science Society.
Chafe, Wallace. 1994. *Discourse, consciousness and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, IL: University of Chicago Press.
Childers, Jane B. & Michael Tomasello. 2001. The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* 37(6). 739–748.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1986. *Knowledge of language: Its nature, origins, and use*. New York, NY: Praeger.
Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
Clark, Andy. 1997. *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
Cleeremans, Axel, Arnaud Destrebecqz & Maud Boyer. 1998. Implicit learning: News from the front. *Trends in Cognitive Science* 2. 406–416.
Colunga, Eliana & Linda B. Smith. 2005. From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review* 112(2). 347–382.
Cutler, Anne. 1996. Prosody and the Word Boundary Problem. In James L. Morgan & Katherine Demuth (eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 87–99. Mahwah, NJ: Lawrence Erlbaum.
Dreyfus, Hubert L. 1997. From micro-worlds to knowledge representation: AI at an impasse. In John Haugeland (ed.), *Mind design II*, 179–211. Cambridge, MA: MIT Press.
Elman, Jeffrey. 1990. Finding structure in time. *Cognitive Science* 14(2). 179–211.
Elman, Jeffrey. 1998. Generalization, simple recurrent networks, and the emergence of structure. In Morten Ann Gernsbacher & Sharon J. Derry (eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Finch, Steven & Nick Chater. 1992. Bootstrapping syntactic categories by unsupervised learning. In *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*, 820–826. Hillsdale, NJ: Lawrence Erlbaum.

Finch, Steven & Nick Chater. 1994. Learning syntactic categories: A statistical approach. In Michael Oaksford & Glenn Brown (eds.), *Neurodynamics and psychology*, 294–321. London: Academic Press.

Fisher, Cynthia, Henry Gleitman & Lila R. Gleitman. 1991. On the semantic content of subcategorization frames. *Cognitive Psychology* 23. 331–392.

Fisher, Cynthia & Hisayo Tokura. 1996. Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development* 67. 3192–3218.

Gathercole, Susan, Elisabet Service, Graham J. Hitch, Anee-Marie Adams & Amanda J. Martin. 1999. Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology* 13(1). 65–77.

Gentner, Dedre. 2001. Spatial metaphors in temporal reasoning. In Merideth Gattis (ed.), *Spatial schemas and abstract thought*, 203–222. Cambridge, MA: MIT Press.

Glenberg, Arthur M. & Michael P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9(3). 558–565.

Glenberg, A. M. & D. A. Robertson. 1999. Indexical understanding of instructions. *Discourse Processes* 28(1). 1–26.

Gogate, Lakshimi, Arlene Walker-Andrews & Loraine E. Bahrick. 2001. The intersensory origins of word comprehension: an ecological — dynamic systems view. *Developmental Science* 4(1). 1–37.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldstone, Robert L., Andy Jones & Michael E. Roberts. 2006. Group path formation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A* 36(3). 611–620.

Gomez, Rebecca L. & LouAnn Gerken. 2000. Infant artificial language learning and language acquisition. *Trends in Cognitive Science* 4. 178–186.

Grant, E. R., & M. J. Spivey. 2003. Eye movements and problem solving: Guiding attention guides thought. *Psychological Science* 14(5). 462–466.

Griffiths, Thomas L., Mark Steyvers, David Blei & Josh Tenenbaum. 2005. Integrating topics and sytax. In Lawrence K. Saul, Yair Weiss & Léon Bottou (eds.), *Advances in Neural Information Processing Systems (NIPS) 17: Proceedings of the 2004 conference*, 537–544. Cambridge, MA: MIT Press.

Heine, Bernd, Ulrike Claudi & Frederike Hünnemeyer. 1991. From cognition to grammar: Evidence from African languages. In Elizabeth Traugott & Bernd Heine (eds.), *Approaches to grammaticalizations*, vol. 1, 149–187. Amsterdam & Philidelphia: John Benjamins.

Hockema, Stephen A. 2006. Finding words in speech: An analysis of American English. *Language Learning & Development* 2(2). 119–146.

Hollich, George, Kathy Hirsh-Pasek & Roberta M. Golinkoff. 2000. *Breaking the language barrier: An emergentist coalition model for the origins of word learning* (Monographs for the Society for Research in Child Development 262). London: Blackwell.

Hund, Alycia M. & Jodie M. Plumert. 2005. The stability and flexibility of spatial categories. *Cognitive Psychology* 50. 1–44.

Hutchins, Edwin. 1996. *Cognition in the wild*. Cambridge, MA: MIT Press.

Iverson, Jana & Esther Thelen. 1999. Hand, mouth and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies* 6(11–12). 19–40.

Jones, Susan S. 2003. Does object name learning change object perception? Paper presented at the Society for Research in Child Development 2003 Biennial Meeting in Tampa, FL, April 26.

Jusczyk, Peter. 2002. Speech perception. In Hal Pashler & Steven Yantis (eds.), *Steven's handbook of experimental psychology, volume 1*: *Sensation and perception*, 3rd edn., 493–536. New York: Wiley.

Kirkham, Natasha Z., Jonathan A. Slemmer & Scott P. Johnson. 2002. Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition* 83(2). B35–B42.

Kunihira, Shirou. 1971. The effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior* 10(4). 427–429.

Laakso, Aarre & Linda B. Smith. 2007. Pronouns predict verb meanings in child-directed speech: A corpus analysis. *Journal of Child Language* 34(4). 725–763.

Laakso, Aarre & Linda B. Smith. 2004. Pronouns predict verb meanings in child-directed speech. In Ken Forbus, Dedre Gentner, & Terry Regier (eds.), *Proceedings of the 26th annual meeting of the cognitive science society*, 767–772. Mahwah, NJ: Lawrence Erlbaum.

Landau, Barbara & Lila Gleitman. 1985. *Language and experience*. Cambridge, MA: Harvard University Press.

Landauer, Thomas, Darell Laham & Peter Foltz. 1998. Learning human-like knowledge by singular value decomposition: A progress report. In Michael I. Jordan, Michael J. Kearns & Sara A. Solla (eds.), *Advances in neural information processing systems* 10, 45–51. Cambridge, MA: MIT Press.

Levinson, Stephen. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, MA: Cambridge University Press.

Li, Ping. 2000. The acquisition of lexical and grammatical aspect in a self-organizing feature-map model. In *Proceedings of the 22nd annual meeting of the cognitive science society*. Mahwah, NJ: Lawrence Erlbaum.

Maouene-Cavin, Josita. 2005. How do children bind the meaning of verbs? Grounding verb meanings in objects, body parts and contact. Poster presented at the Society for Research in Child Development 2005 Biennial Meeting in Atlanta, GA, April 9.

Maratsos, Michael & Mary Anne Chalkley. 1980. The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In Keith E. Nelson (ed.), *Children's language*, vol. 2, 127–214. New York: Gardner Press.

Matlock, Teenie, Michael Ramscar & LeraBoroditsky. 2005. On the experiential link between spatial and temporal language. *Cognitive Science* 29. 655–664.

Maye, Jessica & Daniel Weiss. 2003. Statistical cues facilitate infants' discrimination of difficult phonetic contrasts. In Barbara Beachley, Amanda Brown & Frances Conlin (eds.), *Proceedings of the 27th annual Boston University conference on language development*, 508–518. Sommerville, MA: Cascadilla Press.

Maye, Jessica, Janet F. Werker & LeeAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82(3). B101–B111.

McGonigle, Brendon & Margaret Chalmers. 2003. Spatial representations as cause and effect: Circular causality comes to cognition. In Merideth Gattis (ed.), *Spatial schemas and abstract thought*, 247–277. Cambridge, MA: MIT Press.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Mintz, Toben H. 2005. The acquisition of grammatical categories: Insights from computational and behavioral studies. Paper presented at the Society for Research in Child Development 2005 Biennial Meeting in Atlanta, GA, April 8.

Mintz, Toben H., Newport, Elissa L. & Thomas G. Bever. 1995. Distributional regularities of form class in speech to young children. In Jill Beckman (ed.), *Proceedings of the 25th Annual Meeting of the North Eastern Linguistics Society*, vol 2. Amherst, MA: GLSA.

Mintz, Toben H., Elissa L. Newport & Thomas G. Bever. 2002. The distributional structure of grammatical categories in speech to young children. *Cognitive Science* 26. 393–424.

Naigles, Letitia. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17. 357–374.

Naigles, Letitia. 1996. The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition* 58. 221–251.

O'Grady, William D. 1997. *Syntactic development*. Chicago, IL: University of Chicago Press.

Piaget, Jean. 1953. *The origins of intelligence in children*. London: Routledge and Kegan Paul.

Pinker, Steven. 1984. Language learnability and language development. Cambridge, MA: Harvard University Press.

Pinker, Steven. 1987. *The bootstrapping problem and language development*. Cambridge, MA: Harvard University Press.

Quartz, Steven R. & Terrence J. Sejnowski. 1997. The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences* 20(4). 537–596.

Redington, Martin, Nick Chater & Steven Finch. 1998. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22. 425–469.

Richardson, Daniel & Michael Spivey. 2000. Representation, space, and Hollywood Squares: Looking at things that aren't there anymore. *Cognition* 76. 269–295.

Richardson, Daniel, Michael Spivey, Lawrence W. Barsalou & Ken McRae. 2003. Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27(5). 767–780.

Robinson, Christopher W., & Vladimir M. Sloutsky. 2004. Auditory dominance and its change in the course of development. *Child Development* 75. 1387–1401.

Rogers, Timothy T. & James L. McClelland. 2004. *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Saffran, Jennifer. 2003. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science* 12. 110–114.

Saffran, Jennifer, Richard Aslin & Elissa Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274(5294). 1926–1928.

Saffran, Jennifer, Elissa Newport, Richard Aslin, Rachel A. Tunick & Sandra Barrueco. 1997. Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science* 8(2). 101–105.

Samuelson, Larissa K. 2002. Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental Psychology* 38. 1016–1037.

Samuelson, Larissa K. & Linda B. Smith. 1999. Early noun vocabularies: Do ontology, category structure, and syntax correspond? *Cognition* 71, 1–33.

Samuelson, Larissa K. & Linda B. Smith. 2000. Grounding development in cognitive processes. *Child Development* 71(1). 98–106.

Sasso, Debora, Laura Namy, Lynn Nygaard & Kelly Chicos. 2005. Prosodic cues to word meaning. Poster presented at the Society for Research in Child Development 2005 Biennial Meeting in Atlanta, GA, April 9.

Schütze, Hinrich. 1993. Word space. In Hanson, Stephen, Jack Cowan & C. Lee Giles (eds.), *Advances in Neural Information Processing Systems 5*, 895–902. San Mateo, CA: Morgan Kaufmann.

Sethuraman, Nitya. 2005. Learning verbs in English and Tamil. Poster presented at the Society for Research in Child Development 2005 Biennial Meeting in Atlanta, GA, USA on April 9.

Slobin, Dan I. 1985. Crosslinguistic evidence for the language making capacity. In Slobin, D. (ed.), *The crosslinguistic study of language acquisition, Volume 2: Theoretical issues*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates. 1157–1255.

Smith, Linda B. 2000a. Learning how to learn words: An associative crane. In Roberta Michnik Golinkoff & Kathryn Hirsh-Pasek. (eds.), *Becoming a word learner: A debate on lexical acquisition*, 51–80. Oxford: Oxford University Press.

Smith, Linda B. 2000b. Avoiding associations when it's behaviorism you really hate. In Roberta Michnik Golinkoff & Kathryn Hirsh-Pasek (eds.), *Becoming a word learner: A debate on lexical acquisition*, 169–174. Oxford: Oxford University Press.

Smith, Linda B. 2004. Using space to bind words to objects. In Richard Alterman & David Kirsh (eds.), *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, 60–61. Hillsdale, NJ: Lawrence Erlbaum.

Smith, Linda B., Susan Jones, Barbara Landau, Lisa Gershkoff-Stowe & Larissa Samuelson. 2002. Object name learning provides on-the-job training for attention. *Psychological Science* 13. 13–19.

Smith, L. B., A. Richardson & D. Schuller. 2004. Learning words in space and time. Internal Conference on Infant Studies, Chicago, Illinois.

Smith, Linda B. & Larissa Samuelson. 2003. Different is good: Connectionism and dynamic systems theory and complementary emergentist approaches to development. *Developmental Science* 6(4). 434–439.

Smith, Linda B. & Esther Thelen. 2003. Development as a dynamic system. *Trends in Cognitive Sciences* 7(8). 343–348.

Spivey, Michael J. & Joy J. Geng. 2001. Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research* 65. 235–241.

Thelen, Esther & Linda B. Smith. 1994. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Toskos, Alexia C., Rima Hanania & Stephen A. Hockema. 2004. Eye scanpaths influence memory for spoken verbs. Poster presented at the 26th Annual Meeting of the Cognitive Science Society, Chicago, IL, August 5. Available at http://www.cogsci.northwestern.edu/cogsci2004/ma/ma367.pdf.

Trehub, Sandra E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47. 466–472.

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40(1–2). 21–81.

Vygotsky, Lev S. 1978. *Mind in society: The development of higher psychological processes*. Michael Cole, Vera John-Steiner, Sylvia Scribner & Ellen Souberman (eds.). Cambridge, MA: Harvard University Press.

Vygotsky, Lev S. 1986. *Thought and language*. Alex Kozulin (ed.). Cambridge, MA: MIT Press.

Werker, Janet F., John V. Gilbert, Keith Humphrey & Richard C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52. 349–355.

Werker, Janet F. & Richard C. Tees. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7. 49–63.

Woodward, Amanda L., Ellen M. Markman & Coleen M. Fitzsimmons. 1994. Rapid word learning in 13- and 18-month olds. *Developmental Psychology* 30. 553–566.

Yoshida, Hanako & Linda B. Smith. 2001. Early noun lexicons in English and Japanese. *Cognition* 82. B63–B74.

Yoshida, Hanako & Linda B. Smith. 2003. Shifting ontological boundaries: How Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Science* 6(1). 1–36.

Yoshida, Hanako & Linda B. Smith. 2005. Linguistic cues enhance the learning of perceptual cues. *Psychological Science* 16(2). 90–95.

Yu, Chen & Dana H. Ballard. 2007. A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing* 70(13–15). 2149–2165.

Yu, Chen, Dana H. Ballard & Richard N. Aslin. 2005. The role of embodied intention in early lexical acquisition. *Cognitive Science* 29. 961–1005.