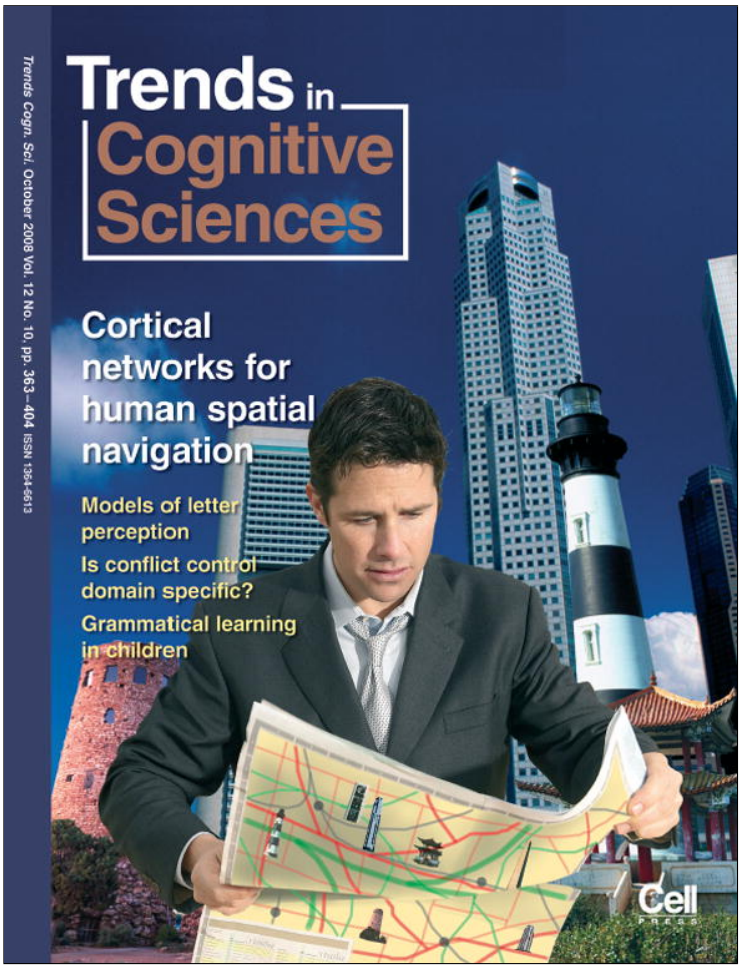

Review

How toddlers begin to learn verbs

Roberta Michnick Golinkoff[1] and Kathy Hirsh-Pasek[2]

[1] School of Education, University of Delaware, Newark, DE 19716, USA
[2] Department of Psychology, Temple University, Philadelphia, PA 19122, USA

Any theory of language must account for how children learn verbs, the gateway to grammar. Yet verbs can be difficult to learn. Building on Gentner's 'natural partitions hypothesis' we suggest that, to learn a verb, infants must conceptualize components of events and map verbs in the ambient language onto those components. Although toddlers detect and categorize at least some of the conceptual underpinnings of verb categories, the mapping of verbs onto these representations is not transparent. Mapping is a difficult problem in its own right. The Emergentist Coalition Model that has been used to explain noun learning also begins to explain how children move from perceptual to social and then to linguistic information to link verbs to actions and events.

The problem of verb learning

Verbs are the gateway to grammar. Thus, any theory of language acquisition must explain how children learn them. When expressing a thought, the choice of a verb determines what other words will be chosen (that is, what kinds of actors can perform the action the verb names) and the way in which words are assembled in sentences (that is, whether a grammatical object is needed or not). The verbs children learn can be defined syntactically and semantically. Syntactically, verbs take subjects or objects or both as in 'John [subject] blorked Mary [object]'. Semantically, verbs encode events and are a cover term for states or processes. They name visible actions (such as 'running'), invisible executive processes (such as 'thinking') [1] and are inherently 'relational' because some entity is required to carry them out. For example, the action of 'jump' (and the verb that describes it) requires someone or something to carry out the action.

As researchers have long noted [2–4], the fact that verbs label dynamic events and processes poses a special learning problem for young children. In particular, children must take the ever-changing events in the world and transform them into a categorical system represented by language. Thus, even though infants process extensive metric information about continuous events, they must rely on abstract categorical information to express these events in language [5,6]. Advances in methodology that enable the presentation of dynamic events (such as the Intermodal Preferential Looking Paradigm [IPLP] [7,8] [Box 1, Figure I]), combined with the availability of video editing technology, now empower researchers to test verb learning. This article examines how research on children's understanding of event structure and their ability to link word-to-world has boosted our understanding of verb learning [9].

In her 'natural partitions hypothesis', Gentner [2] argued that verb learning was dependent upon two capabilities: (i) a conceptual understanding of the events verbs describe and, (ii) a recognition of the way in which one's particular language expresses these events. Gentner wrote, 'It is not perceiving relations but packaging and lexicalizing them that is difficult' [2]. Gentner's hypothesis [2] (and its offshoots) fueled a new area of research that bridges event perception and early verb acquisition. The former investigates how children perceive and categorize nonlinguistic events in ways that will be relevant to language expression. The latter investigates how children learn to map their event conceptualizations into the language they are learning (see also Refs [3,4]). The sheer amount of information in an event makes mapping uncertain because verbs are not 'verbal film clips of events' [10]. For example, when children hear, 'Don't climb that!' the verb 'climb' refers to the relationship between an agent (the child) and something being ascended (the dresser). The learning problem here is twofold. First, amidst this dynamic and complex event, the child must abstract the spatial component of path (i.e. ascent). Second, the word 'climb' must be mapped to the path and not, say, to the little hands grasping the drawers.

In the face of this complexity, relational terms (verbs, prepositions and adjectives) often lag behind nouns in both studies of natural vocabulary acquisition (e.g. Refs [11,12]) and in laboratory research on verb comprehension (e.g. Refs [13–16]) (Box 2, Figure I). This general finding must be evaluated in the context of data from some Asian languages however, in which verbs appear earlier in production (e.g. Chinese and Korean) than they do in English [17,18], possibly because of how parents and caregivers speak to their children. Chinese, for example, allows verbs to appear in isolation. Thus, English requires, 'John kissed Mary' whereas in Chinese 'qin1' [kiss] is acceptable, if the listener can infer the meaning from context. Furthermore, Chinese caregivers produce both more verb types and tokens than English-speaking caregivers [17].

How children detect and categorize components of events

One reason why verb learning is challenging might be that children have difficulty extracting components of dynamic events. Several areas of research begin to tell this story. The event components that have received the most empirical attention are (i) perceptually accessible, (ii) universally represented in relational terms across languages, (iii)

Corresponding author: Golinkoff, R.M. (roberta@udel.edu).

1364-6613/$ – see front matter doi:10.1016/j.tics.2008.07.003 Available online 27 August 2008

Box 1. The Intermodal Preferential Looking Paradigm used to study verb learning

How can we study the process by which young children, notoriously poor participants, learn verbs? In the IPLP (Figure I) (e.g. Refs [8,13,39,62]), using looking time as the dependent variable, children see two simultaneous dynamic events with a linguistic stimulus that matches only one of the events. If children understand the linguistic stimulus, they look (or point when older than 30 months) at the matching scene more than at the non-matching scene. Children might see an actor dancing in one image and waving in another while they hear, 'Where's she dancing?' The prediction is that they will look more to dancing than to waving. Attention during test trials can be compared to attention during salience trials with 'neutral' language.

Figure I. The Intermodal Preferential Looking Paradigm.

Box 2. Nouns are easier to learn and extend than verbs in English, Japanese and Chinese

Nouns are apparently easier to learn and extend (apply to non-identical instances) than verbs. This is true even in 'verb friendly' languages in which verbs can appear alone without the possible confusion of surrounding noun arguments. Imai *et al.* [13] tested English, Japanese and Chinese children (ages three and five) using videotaped displays in the IPLP [7,8] (Figure I) (see also Box 1).

In the novel noun condition, children pointed in response to 'Where is the twill?'. Both age groups were successful. In both verb conditions, children were asked, 'Where is she twilling?'. Only 5-year-olds, however, successfully extended the action label, and only when the training condition matched how verbs are encoded in the target language. Thus, English-reared children succeeded in the full argument novel verb condition; Japanese children succeeded only in the bare novel verb condition, hindered by full sentences.

When Chinese children – even at the age of five – did not at first succeed, Imai *et al.* [13] reasoned that Chinese children might be relying on extralinguistic cues as Chinese uses no verb endings and permits dropping verb arguments. After a one-second segment of object-holding was removed from the beginning of each videotaped scene, Chinese children succeeded in the verb condition, demonstrating exquisite sensitivity to extralinguistic information for determining word meaning. English and Japanese children responded in the same way with and without the object-holding segment.

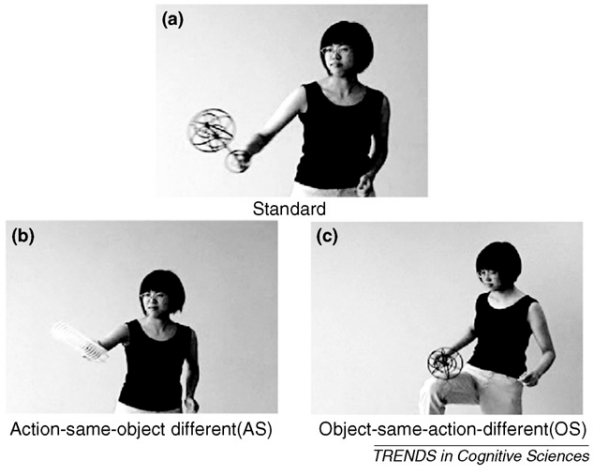

Figure I. In the standard trial **(a)**, children saw a video clip of a woman performing an unfamiliar action with an unfamiliar object. In the 'novel noun' condition, children heard, 'Look, a twill!' In the 'bare novel verb' condition they heard the equivalent of, 'Look, twilling!', whereas in the 'full argument novel verb' condition they heard a full sentence (e.g. English: 'Look, she's twilling it.'). When tested, children saw both an action-same/object-different event **(b)** and an object-same/action-different event **(c)**. Reproduced, with permission, from Ref. [13] and Oxford University Press, Inc.

encoded differently across languages and (iv) are posited to represent a set of foundational concepts from which relational terms are constructed. Among the event components studied are 'containment and support', 'path and manner' and 'source and goal' (e.g. Refs [19–22]). To show that infants are sensitive to these components of events requires evidence that they (i) discriminate these constructs as independent units within events and (ii) categorize similar event types together [23]. 'Climbing', for example, includes ascent whether done by a tiger or a child, on a mountain or a dresser.

Discriminating and categorizing containment, support and degree-of-fit

'Containment', called 'in' in English, is when an entity is in '. . .any fully or partially enclosed space. . .' [20,24]. 'On', the spatial relation 'support', occurs when a figure is supported by, attached to or encircling an external surface of a ground object. Particularly interesting is the fact that Korean verbs cross cut the English categories of 'in' and 'on' with the construct 'degree-of-fit', a concept not lexicalized in English [25,26], that captures the way two surfaces fit together. An example of tight-fit (interlocking surfaces) is the way in which Lego blocks fit snugly together; loose-fit is exemplified by the way an apple wobbles in a bowl. English-reared infants can form categories of loose-fit versus tight-fit before six months of age, despite the fact that they will not go on to learn Korean [27]. Infants categorize across categories of containment by six months of age [28], but not until 14 months across categories of support [29] (Figure 1).

Path and manner

'Path' is defined as the movement of a figure with respect to a ground object. 'Manner' is how the figure moves within an event (e.g. climbing or walking). Path and manner are represented in all languages of the world but are codified differently. Spanish speakers would probably say, '*Una mujer salió de la casa (corriendo)*' ('A woman exited the house [running]'), using a path verb [19]. English speakers would say, 'A woman ran out of the house', using a manner verb to describe how the figure moves, with the path in a

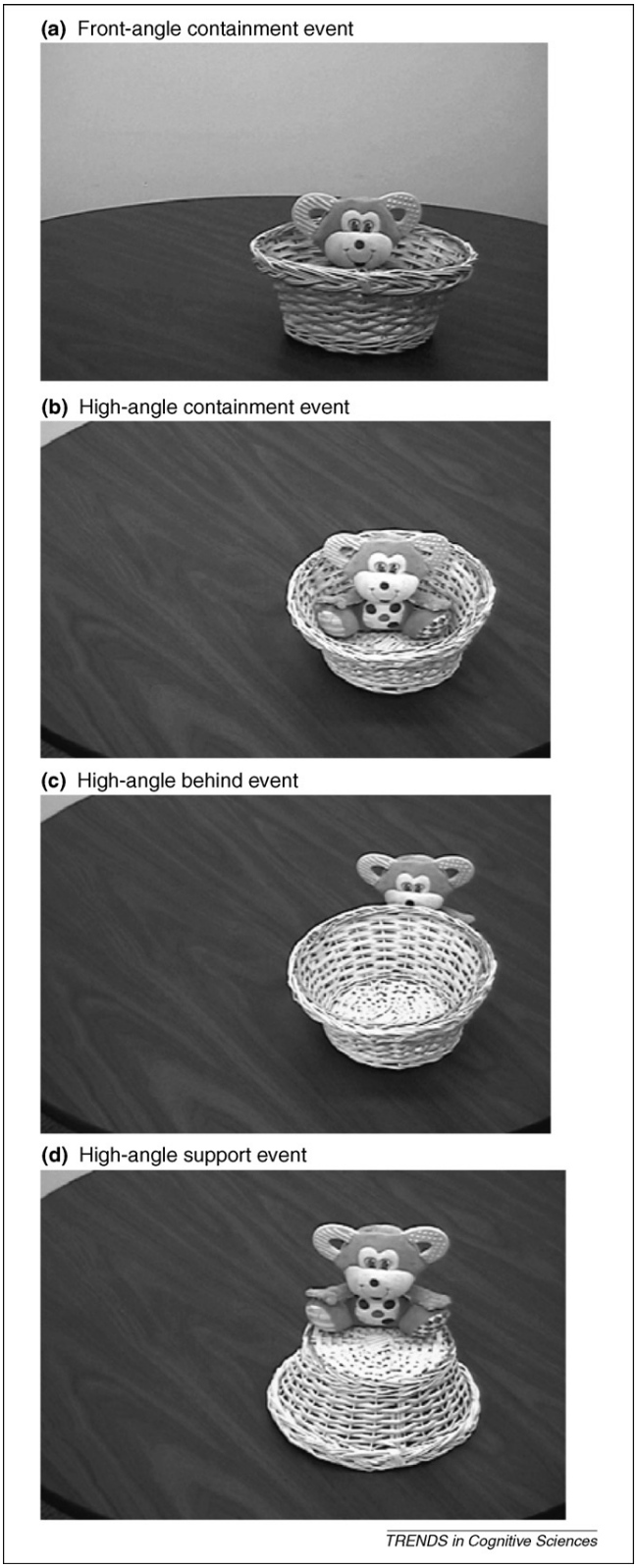

Figure 1. Stimuli used to test six-month-olds' categorization of the relation of containment. Final frames of dynamic events used to see if infants can discriminate containment from other spatial relations [28]. Infants were first familiarized to a dynamic containment event **(a)**. After familiarization, they were tested with the same familiarized containment event, in addition to a containment event filmed from a high angle so that the figure was no longer partially occluded when inserted **(b)**; a behind event in which the figure was as occluded as in the familiarization containment event **(c)**; and a support event **(d)**. Infants looked significantly longer at test events that presented a novel relation (the behind and support test events) than they did at the familiar containment event, indicating that they discriminate a containment relation from other relational events. Reproduced, with permission, from Ref. [29] and the Association for Psychological Science.

satellite prepositional phrase (viz, 'out of the house'). English uses many more manner verbs than Spanish does.

Both English- and Spanish-reared babies notice changes in paths and manners in dynamic events as early as seven months of age [23,30] and do so robustly by 14 to 17 months [30] (Figure 2). By nine months, infants categorize paths (e.g. over and under) when performed in the same way over different manners (e.g. bending or twisting); by 13 months they categorize manners [31]. Thus, path relations are easier to detect and categorize than manner relations.

Source and goal paths

The paths described do not entail goals but are 'via' paths, as the figure moves past a referent object. In 'from' paths, a figure moves from a referent object that is its source; 'to' paths are goal-oriented, involving movement towards some endpoint. The source is the starting point of the figure in an event and the goal is the endpoint of the figure. Across languages, goal paths are expressed more frequently than sources are. Homologously, 12-month-olds attend preferentially to goal paths over source paths in dynamic, nonlinguistic events [21,32]. By 14 months of age children form categories of goals, although not of sources [21,33].

In sum, although limited to a subset of the semantic components verbs encode, part of Gentner's [2] natural partitions hypothesis might be correct. Infants seem proficient at perceiving and categorizing many (although perhaps not all) of the event concepts languages encode. 'Thinking for speaking' [10] requires that infants eventually detect the characteristics of events that find expression in language, even if these event components are not accessible all at once.

How children map verbs to events

Although the conceptualizations of events by infants might not seem problematic, difficulties in verb learning (e.g. Ref. [13]) might be rooted in word-to-world mapping. Indeed, mapping verbs to referents is even difficult for adults. Yet mapping is not inherently challenging: by ten months of age, infants can link object words to referents [34]; by 15 months of age, they seem to understand verbs such as 'dance' and 'open' [35]. What, then, accounts for the generally slow progress of verb learning?

One theory of word learning, the Emergentist Coalition Model (ECM), helps unpack the complexity of verb mapping [8,36–38]. A hybrid theory, the ECM holds that children rely differentially on multiple cues over developmental time to map words onto referents – whether they are nouns or verbs. At first, infants are sensitive to perceptual cues – mapping a word to the referent that is most interesting or salient. They then use the social intent of a speaker, along with linguistic cues, to home in on word reference.

Several aspects of verb mapping might prove difficult for young learners. First, verbs are inherently relational; the actors and objects involved are often more salient than the actions themselves [39,40]. Second, as some verbs are less perceptually available (e.g. know, want), and as their

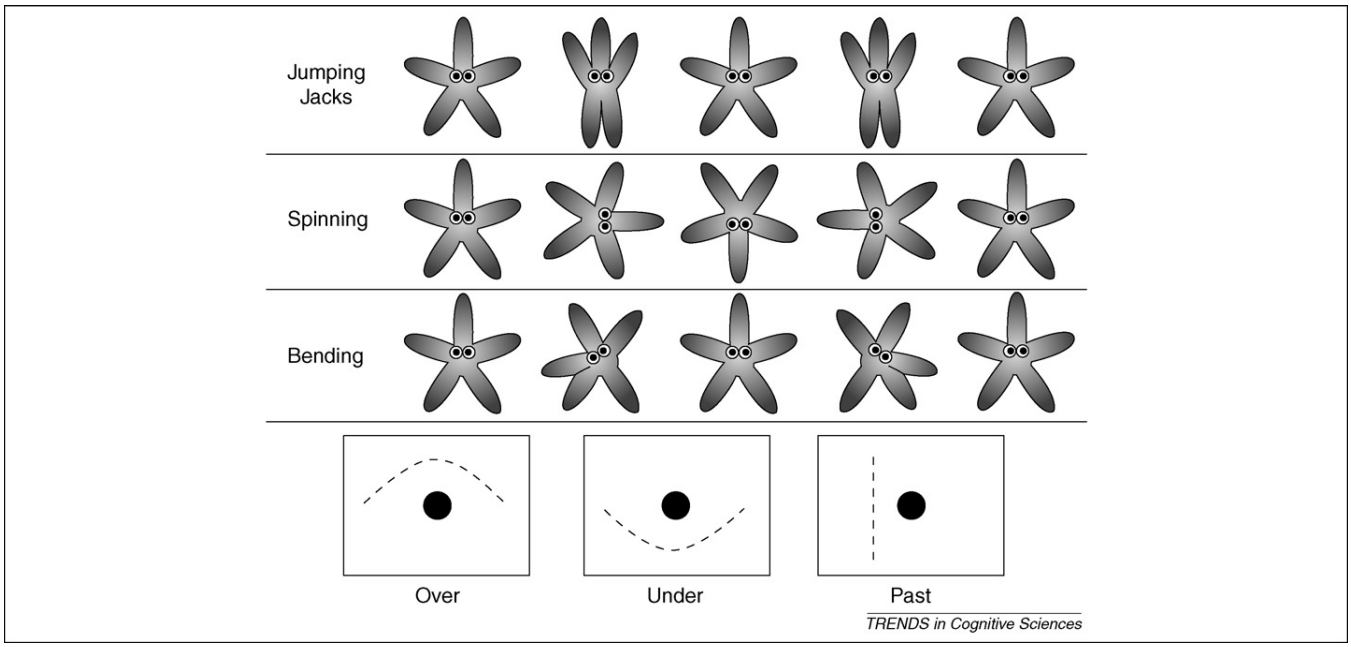

Figure 2. Stimuli used to test infants (seven months; 14–17 months) on discrimination of path and manner in dynamic events. Computer-animated motion events tested infants' discrimination of manners and paths [23,30]. In the habituation paradigm, infants saw a moving starfish character (the figure) and a stationary ball (the ground). The starfish performed an action with both a manner (jumping jacks, spinning or bending at the 'waist') and a path (over the ball, under the ball or [vertically] past the ball). Once infants became habituated to a single event (e.g. jumping jacks over), they were then tested on four different types of events: (i) a control event identical to the habituation event (e.g. jumping jacks over), (ii) an event with the same manner as the habituation event, but a different path (e.g. jumping jacks under), (iii) an event with the same path as the habituation event, but a different manner (e.g. spinning over) and (iv) an event in which both the manner and path differed from those in all other events (e.g. waist bends past). Children at both ages discriminated between the paths and the manners. Reproduced, with permission, from Ref. [30] and Oxford University Press.

meanings can turn on small differences (e.g. 'chase' versus 'flee'), children might require additional support to learn them [41]. Finally, the events that verbs name are categorized differently across languages (e.g. the Korean verbs for 'degree-of-fit' versus the 'in' and 'on' of English). Thus, verb learning might wait until children recruit not only perceptual information for word mapping, but also social and grammatical information.

One consequence of this hypothesis is that verbs – at least those that label perceptually less accessible referents – will be learned later because children cannot use social intent and grammatical knowledge in the service of word learning until ~24 months of age (e.g. Ref. [42]). Furthermore, sensitivity to grammatical information rests on children learning several nouns [43] in addition to having sufficient sentential input to note the correspondences

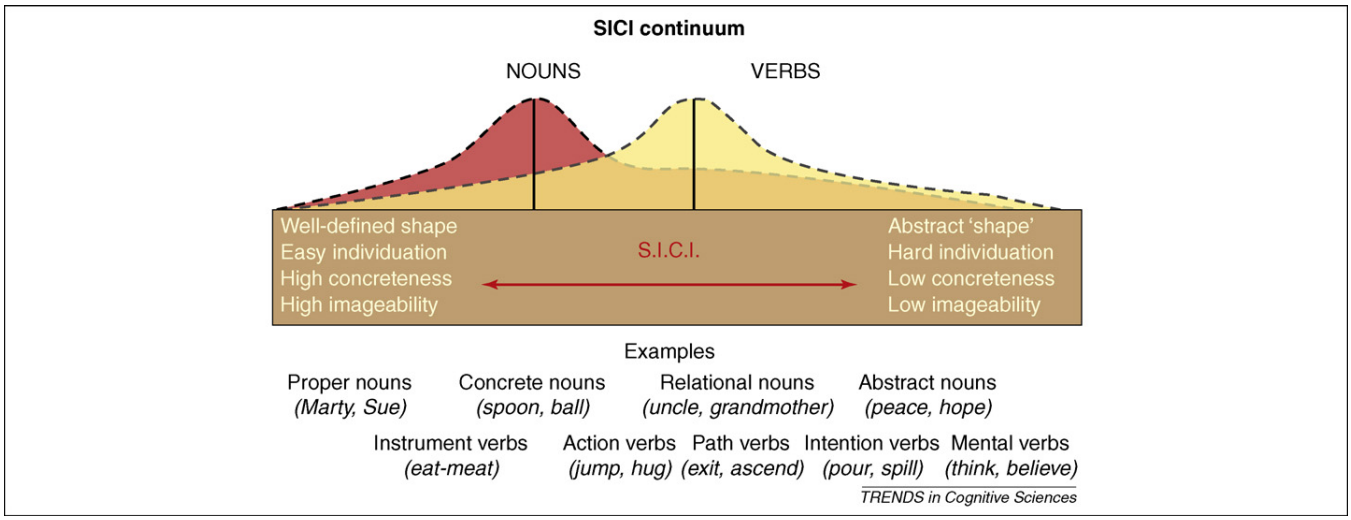

Figure 3. The SICI continuum: why some verbs appear before some nouns despite the generalization that nouns are learned before verbs in vocabulary acquisition. 'SICI' is an acronym for the factors (weights yet to be determined) [41,62–65] that contribute to the learning of nouns and verbs (shape, individuation, concreteness and imageability). Inspired by theorizing and a related chart by Gentner and Boroditsky [41], the SICI continuum represents the fact that not all nouns are learned before all verbs. The concepts words encode can be conceived as falling on a spectrum defined by the reliability and consistency of their perceptual 'shape'; the ease with which they can be distinguished from other items in the scene (individuability), whether they can be observed and are manipulable (concreteness) and how readily they yield a mental image for adults (imageability). Coupling SICI with a general word learning theory, in this case the ECM, indicates why children's earliest words will probably be at the more concrete, shape-based end of this continuum with nouns such as 'cup' learned before 'uncle' and verbs such as 'kiss' learned before a verb such as 'think.' Reproduced, with permission, from Ref. [36] and Oxford University Press.

between meanings and sentence structures [44]. The ECM makes predictions about the developmental course of verb mapping and indicates why non-relational words such as early nouns predominate in early vocabularies [36] (Figure 3). What is the evidence for this position?

Perceptual information in verb learning

If the initial strategy for word mapping is to preferentially attend to referents that are perceptually salient, then words that map to salient objects or events should appear first. Studies support this view [34,37]. With nouns, ten-month-olds assume that a label maps onto an object they find interesting (a colorful clacker) rather than onto a boring object (a beige gadget) even though the speaker is naming the latter. This results in a systematic 'mismapping' [34]. With verbs, Brandone *et al.* [45] pitted perceptual salience against speaker intention by teaching either the name for an interesting action 'with a result' versus a boring action 'without a result'. Training took place on a real object, whereas testing occurred in the IPLP (Box 1, Figure I). At 21 months, children learned the name of an action with a result (i.e. pressing a Morse code key produced a tone) but not the name for that same action when the result was disabled. Only at 33 months could children learn the name of the 'resultless' action. Consistent with this finding is the fact that children learn names for their own actions before learning the names for the actions of others [46]. Actions we carry out ourselves should be highly salient.

Visible actions such as 'running' receive higher 'imageability' ratings than unseen processes ('thinking') and are easier for adults to guess in the Human Simulation Paradigm [47] (Box 3). 'Imageability' refers to the ease with which a mental image can be generated (viz, 'apple' is easier than 'believe'). In English and Chinese, words (both nouns and verbs) for which it is easier to generate a mental image are learned earlier than words with lower imageability ratings [48]. Early-acquired Mandarin verbs are more imageable than early-acquired English verbs [48], a factor that reflects the fact that Chinese children produce more verbs early on than their English-reared counterparts [17]. Combining imageability ratings and how frequently a word appears in speech directed to children accounts for 42% of the variance in the age of acquisition for verbs in Chinese [48].

If perceptual factors dominate in early verb learning, children should also be reluctant to extend a verb to a related but unfamiliar event. Indeed, children are notorious for narrow verb extensions, with three-year-olds' interpretations of novel action verbs more specific than ten-year-olds' and adults' [49,50].

Perceptual factors are so potent for young word learners that an action performed by an unknown agent [40] or by multiple agents focuses children's attention on the agents rather than on the relation and interferes with verb learning. Maguire *et al.* [39] reported better verb learning in two-and-a-half to three-year-olds with exposure to a single actor rather than multiple actors.

Importantly, even if an event that a verb names is visually accessible and perceptually narrow, ambiguity about verb meaning still remains. In the best case scenario, perceptual data cannot go all the way towards segmenting complex events into the categorical divisions that are the stock of language.

Box 3. Findings on noun and verb learning from the Human Simulation Paradigm

The 'Human Simulation Paradigm' [47] was developed to probe a generalization about vocabulary learning: nouns are learned before verbs. In a fascinating demonstration on the relationship between event observation and word learning, researchers asked adults to guess the words used by mothers videotaped interacting with their children. Hearing a 'bleep', adults guessed the word that the mothers in the videotape were using when communicating with their 18-to 24-month-old toddlers.

With extralinguistic observation controlled by presenting the same scenes, would adults replicate children's pattern of acquisition by finding it easier to guess the nouns than the verbs? With adults, any difference between nouns and verbs could not be attributed to conceptual factors. Varying the linguistic information offered along with the scene allowed exploration of how accompanying language influenced target word identification [43,63].

Adults correctly guessed the missing nouns in 45% of the cases; their score for verbs was only 15%, and for mental verbs, 0% [47]. When language was systematically added [63], ranging from giving adults a list of the nouns that accompanied the verb to a full sentence with the verb omitted, a step function emerged in adults' ability to accurately guess the verb. Verb identification improves dramatically as more language is offered.

Snedeker *et al.* [66] conducted the same experiment with Mandarin infant-directed speech. Mandarin speakers guessed the words from the English and the Mandarin samples, and Americans guessed the words in Mandarin samples. Both Taiwanese and American adults duplicated the patterns found in children's acquisition: better identification of nouns than verbs from the English samples and equally good identification of nouns and verbs from the Mandarin samples. The source of these differences is not clear.

Social information in verb learning

As Tomasello and colleagues argue [51,52], language acquisition is embedded in the nexus of social interaction. Verb learning rests on social information in at least three ways, (i) in discerning which referent action the speaker is naming, (ii) in discerning the actor's intent (did she *mean* to kill Cock Robin or was it an accident?) and (iii) inferring the speaker's interpretation of the event (e.g. 'die' versus 'kill') [53].

A demonstration of the importance of social information in word learning is offered by the case of autistic children. Notoriously poor at interpreting the intentions of others, 67% of the variance in the size of autistic children's vocabularies was accounted for by their sensitivity to speaker intent [54]. Using a novel verb for a novel action – even one that looks accidental (dropping beads outside a cup) – influences typical two-year-olds to believe that the action was done on purpose [55]. Finally, there is evidence that toddlers are sensitive to whether an event was intentional in assigning verb meaning. By 24 months, toddlers can infer that a speaker meant to label an unseen action [56]. By 27 months, children can learn different verbs for similar actions (e.g. 'knock over' versus 'topple'), if offered nonverbal information for whether the action was accidental or intentional [53].

In sum, attention to perceptual cues for verb meaning precedes children's attention to social intentional cues.

Box 4. Outstanding questions

- Is there a relationship between children's abilities to analyze the subcomponents of events that will be encoded in verb meaning and their ability to learn novel verbs? Because part of the difficulty of verb learning is discovering how particular languages encode events in their verbs, future research should link the perception of events to the discovery of verb meaning.
- Are there individual differences in verb learning that are predictive of later language development? That is, are children who are better at conceptualizing nonlinguistic events also better at learning verbs?
- What are the challenges faced by second and bilingual language learners given that languages package verb meaning differently?
- Can a Bayesian model of word learning incorporate the interactive components of the ECM along with verb frequency to offer a simulation of real verb learning?
- What factors underlie delayed verb learning? Given that the ECM emphasizes the use of perceptual, social and linguistic cues, perhaps we can determine which of these factors are stumbling blocks for individual children. Should such patterns emerge, how might we craft interventions for children with atypical language development?

However, as extralinguistic context is insufficient for interpreting verb meaning, perceptual and social cues alone leave the verb learner lost in a sea of possible meanings.

Linguistic cues to verb meaning

Consider the case of invisible verb referents such as 'think' and the fact that identical scenes can be described in various ways (i.e. Mary chases John; John flees from Mary). Syntactic structure serves as a 'zoom lens' [57] highlighting a speaker's perspective and constraining verb meaning. 'Syntactic bootstrapping' [57] is the theory that children use the number and arrangement of arguments in the sentence to compute verb meaning [43]. For example, 'John spoodled Mary the crick' suggests that 'spoodled' must be about actual or metaphorical transfer. Analogous to 'put', 'spoodled' takes three arguments: a subject (John), an indirect object that is spoodled (the crick) and a direct object (Mary). Fisher [58] illustrated this point with two-year-olds, showing them a complex event of one woman pushing another woman forward and back in a red wagon. Hearing either (i) an intransitive sentence ('*She* pilks back and forth!') or (ii) a transitive sentence ('*She* pilks *her* back and forth!'), children were asked to point to static images in response to test questions. As the training sentences only differed in the number of arguments surrounding the verb, children used this syntactic information to decide whether the verb described the agent or the receiver of the action. Syntactic bootstrapping can even be used with languages that allow some argument dropping [59], and even overrides regularity in morphology for signaling verb meaning [60]. It also seems to work in reverse: shown an event with a single participant, toddlers expect to hear an intransitive sentence. [61].

Conclusions

Children's difficulty with verb learning could lie in their inability to detect and categorize the components of dynamic events or in the mapping process or both. Although infants seem sensitive to the components of events, this research is still in its infancy (Box 4). An interesting question is why some event components (such as path or containment) become available before their counterparts (manner and support, respectively).

The factors that influence mapping seem to undergo developmental change as well. Learners face a challenge in translating dynamic, continuous events into the categorical units required by language. Verb learning is first governed by perceptual factors, with interesting events preferred as verb referents. Even if strong perceptual cues are available, however, verb meaning is still ambiguous and hence requires additional scaffolds in the form of social and linguistic information. One model, the ECM, provides a framework for organizing the course of early verb learning and offers an explanation that invokes the multiple cues that children must coordinate to learn verbs. Why children shift in their use of these cues and how they interact for determining verb meaning are projects for further research.

Acknowledgements

This work was supported by an NICHD grant (5R01HD050199) and by an NSF grant (BCS-0642529) to both authors. We thank our wonderful graduate students Julia Parish, Sarah Roseberry, Weiyi Ma, Wilkey Wong, Tilbe Goksun and Lulu Song, in addition to Rebecca Seston and David Strober for their help in preparing this manuscript.

References

1 Frawley, W. (1992) *Linguistic Semantics,* Lawrence Erlbaum Associates
2 Gentner, D. (1982) Why are nouns learned before verbs: linguistic relativity versus natural partitioning. In *Language Development: Language, Thought and Culture* (Vol. 2) (Kuczaj, S.A., ed.), In pp. 301–334, Lawrence Erlbaum Associates
3 Tomasello, M. (1992) *First Verbs: A Case Study of Early Grammatical Development,* Cambridge University Press
4 Gleitman, L.R. and Gleitman, H. (1992) A picture is worth a thousand words, but that's the problem: the role of syntax in vocabulary acquisition. *Cur. Dir. Psychol. Sci.* 1, 31–35
5 Kosslyn, S.M. (2006) You can play 20 questions with nature and win: categorical versus coordinate spatial relations as a case study. *Neuropsychologia* 44, 1519–1523
6 Regier, T. and Carlson, L.A. (2001) Grounding spatial language in perception: an empirical and computational investigation. *J. Exp. Psychol. Gen.* 130, 273–298
7 Golinkoff, R.M. *et al.* (1987) The eyes have it: lexical and syntactic comprehension in a new paradigm. *J. Child Lang.* 14, 23–45
8 Hirsh-Pasek, K. and Golinkoff, R.M. (1996) *The Origins of Grammar: Evidence From Early Language Comprehension,* MIT Press
9 Hirsh-Pasek, K. and Golinkoff, R.M., eds (2006) *Action Meets Word: How Children Learn Verbs*, Oxford University Press
10 Slobin, D.I. (2003) Language and thought online: cognitive consequences of linguistic relativity. In *Advances in the Investigation of Language and Thought* (Gentner, D. and Goldin-Meadow, S., eds), pp. 157–191, MIT Press
11 Gentner, D. (2006) Why verbs are hard to learn. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 544–564, Oxford University Press
12 Bornstein, M. *et al.* (2004) Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Dev.* 75, 1115–1140
13 Imai, M. *et al.* (2008) Novel noun and verb learning in Chinese, English, and Japanese children: universality and language-specificity in novel noun and verb learning. *Child Dev.* 79, 979–1000
14 Golinkoff, R.M. *et al.* (1996) Lexical principles may underlie the learning of verbs. *Child Dev.* 67, 3101–3119
15 Childers, J.B. and Tomasello, M. (2006) Are nouns easier to learn than verbs? Three experimental studies. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 311–335, Oxford University Press
16 Werker, J.F. *et al.* (1998) Acquisition of word-object associations by 14 month old infants. *Dev. Psychol.* 34, 1289–1309

17 Tardif, T. (2006) But are they really verbs? In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 477–498, Oxford University Press
18 Choi, S. and Bowerman, M. (1991) Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition* 41, 83–121
19 Talmy, L. (1985) Lexicalization patterns: semantic structure in lexical forms. In *Language Typology and Syntactic Description* (Shopen, T., ed.), pp. 57–149, Cambridge University Press
20 Mandler, J.M. (2004) *The Foundations of Mind: Origins of Conceptual Thought.* Oxford University Press
21 Lakusta, L. *et al.* (2008) Conceptual foundations of spatial language: evidence for a goal bias in infants. *Lang. Learn. Dev.* 3, 179–197
22 Jackendoff, R. (1983) *Semantics and cognition.* MIT Press
23 Pulverman, S. *et al.* (2006) Conceptual foundations for verb learning: celebrating the event. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 134–159, Oxford University Pres
24 Mandler, J.M. (2006) Actions organize the infant's world. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 111–133, Oxford University Press
25 Choi, S. (2006) Preverbal spatial cognition and language-specific input: categories of containment and support. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 191–207, Oxford University Press
26 Gentner, D. and Bowerman, M. Why some spatial semantic categories are harder to learn than others: the typological prevalence hypothesis. In *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin* (Guo, J. et al., eds), Lawrence Erlbaum Associates (in press)
27 Hespos, S.J. and Spelke, E.S. (2004) Conceptual precursors to spatial language. *Nature* 430, 453–456
28 Casasola, M. *et al.* (2003) Six-month-old infants' categorization of containment spatial relations. *Child Dev.* 74, 679–693
29 Casasola, M. (2008) The development of infants' spatial categories. *Cur. Dir. Psychol. Sci.* 17, 21–25
30 Pulverman, R. *et al.* Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition* (in press)
31 Pruden, S.M. *et al.* (2008) Current events: how infants parse the world events for language. In *Understanding Events: From Perception to Action* (Shipley, T. and Zacks, J., eds), pp. 160–192, Oxford University Press
32 Wagner, L. and Carey, S. (2005) Twelve-month-old infants represent probable endings of motion events. *Infancy* 7, 73–83
33 Lakusta, L. and Landau, B. (2005) Starting at the end: the importance of goals in spatial language. *Cognition* 96, 1–33
34 Pruden, S.M. *et al.* (2006) The birth of words: ten-month-olds learn words through perceptual salience. *Child Dev.* 77, 266–280
35 Fenson, L. *et al.* (1994) Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59, 1–189
36 Maguire, M. *et al.* (2006) A unified theory of word learning: putting verb acquisition in context. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 364–391, Oxford University Press
37 Hollich, G.J. *et al.* (2000) Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monogr. Soc. Res. Child Dev.* 65, 1–135
38 Golinkoff, R.M. and Hirsh-Pasek, K. (2006) Baby wordsmith: from associationist to social sophisticate. *Cur. Dir. Psychol. Sci* 15, 30–33
39 Maguire, M. *et al.* (2008) Focusing on the relation: fewer exemplars facilitate children's initial verb learning and extension. *Dev. Sci.* 11, 628–634
40 Kersten, A.W. and Smith, L.B. (2002) Attention to novel objects during verb learning. *Child Dev.* 73, 93–109
41 Gentner, D. and Boroditsky, L. (2001) Individuation, relativity and early word learning. In *Language Acquisition and Conceptual Development* (Bowerman, M. and Levinson, S., eds), pp. 215–256, Cambridge University Press
42 Naigles, L.R. and Swensen, L.D. (2007) Syntactic supports for word learning. In *The Handbook of Language Development* (Hoff, E. and Shatz, M., eds), pp. 212–231, Blackwell
43 Gleitman, L.R. *et al.* (2005) Hard words. *Lang. Learn. & Dev.* 1, 23–64
44 Fisher, C. *et al.* (1991) On the semantic content of subcategorization frames. *Cognit. Psychol.* 23, 331–392
45 Brandone, A. *et al.* (2007) Action speaks louder than words: young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Dev.* 78, 1322–1342
46 Smiley, P. and Huttenlocher, J. (1995) Conceptual development and the child's early words for events, objects, and persons. In *Beyond Names for Things: Young Children's Acquisition of Verbs* (Tomasello, M. and Merriman, W.E., eds), pp. 21–61, Lawrence Erlbaum Associates
47 Gillette, J. *et al.* (1999) Human simulations of vocabulary learning.. *Cognition* 73, 135–176
48 Ma, W. *et al.* Imageability predicts age of acquisition of verbs in Chinese children. *J. Child Lang.* (in press)
49 Behrend, D.A. (1990) The development of verb concepts: children's use of verbs to label familiar and novel events. *Child Dev.* 61, 681–696
50 Forbes, J.N. and Farrar, M.J. (1995) Learning to represent word meaning: what initial training events reveal about children's developing action verb concepts. *Cognit. Dev.* 10, 1–20
51 Akhtar, N. and Tomasello, M. (2000) The social nature of words and word learning. In *Becoming a Word Learner: A Debate on Lexical Acquisition* (Golinkoff, R.M. and Hirsh-Pasek, K., eds), pp. 115–135, Oxford University Press
52 Tomasello, M. (2003) *Constructing a Language.* Harvard University Press
53 Poulin-Dubois, D. and Forbes, J.N. (2002) Toddlers' attention to intentions-in-action in learning novel action words. *Dev. Psychol.* 38, 104–114
54 Parish-Morris, J. *et al.* (2007) Children with autism illuminate the role of social intention in word learning. *Child Dev.* 78, 1265–1287
55 Behrend, D.A. and Scofield, J. (2006) Verbs, actions, and intentions. In *Action Meets Word: How Children Learn Verbs* (Hirsh-Pasek, K. and Golinkoff, R.M., eds), pp. 286–307, Oxford University Press
56 Akhtar, N. and Tomasello, M. (1996) Twenty-four-month-old children learn words for absent objects and actions. *Br. J. Dev. Psychol.* 14, 79–93
57 Gleitman, L.R. (1990) The structural sources of verb meanings. *Lang. Acquisition* 1, 3–55
58 Fisher, C. (2002) Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Dev. Sci.* 5, 55–64
59 Lee, J. and Naigles, L.R. (2005) Input to verb learning in Mandarin Chinese: a role for syntactic bootstrapping. *Dev. Psychol.* 41, 529–540
60 Lidz, J. *et al.* (2003) Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition* 87, 151–178
61 Brandone, A. *et al.* (2005) One-for-one and two-for-two: anticipating parallel structure between events and language. In *Proceedings of the 30th Boston University Conference on Language Development* (Bamman, D. *et al.*, eds), pp. 36–47, Cascadilla Press
62 Golinkoff, R.M. *et al.* (2002) Young children can extend motion verb labels to point-light displays. *Dev. Psychol.* 38, 604–614
63 Snedeker, J. and Gleitman, L. (2004) Why is it hard to label our concepts? In *Weaving a Lexicon* (Hall, D.G. and Waxman, S.R., eds), pp. 255–293, MIT Press
64 Bird, H. *et al.* (2000) 'Little words'-not really: function and content words in normal and aphasic speech. *J. Neuroling.* 15, 209–237
65 Druks, J. and Masterson, J. (2003) Editorial. *J. Neuroling.* 16, 59–65
66 Snedeker, J. *et al.* (2003) Cross-cultural differences in the input to early word learning. In *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society* (Alterman, R. and Kirsh, D., eds), pp. 1094–1099, Lawrence Erlbaum Associates