Cognition 114 (2010) 442–446

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Meaning from syntax: Evidence from 2-year-olds

Sudha Arunachalam *, Sandra R. Waxman

Department of Psychology, Northwestern University, Evanston, IL 60208, United States

ARTICLE INFO

Article history:
Received 29 June 2009
Revised 16 October 2009
Accepted 20 October 2009

Keywords:
Language acquisition
Word learning
Syntactic bootstrapping

ABSTRACT

When toddlers view an event while hearing a novel verb, the verb's syntactic context has been shown to help them identify its meaning. The current work takes this finding one step further to reveal that even in the absence of an accompanying event, syntactic information supports toddlers' identification of verb meaning. Two-year-olds were first introduced to dialogues incorporating novel verbs either in transitive or intransitive sentences, but in the absence of any relevant referent scenes (see Yuan & Fisher, 2009). Next, toddlers viewed two candidate scenes: (a) two participants performing synchronous actions, (b) two participants performing a causative action. When asked to "find mooping", toddlers who had heard transitive sentences chose the causative scene; those who had heard intransitive sentences did not. These results demonstrate that 2-year-olds infer important components of meaning from syntactic structure alone, using it to direct their subsequent search for a referent in a visual scene.

Intuitively, we assume that to learn the meaning of a novel word, children must hear the word while simultaneously observing its referent, for example, hearing /kæt/ in the presence of a cat, or /kik/ in the presence of a kicking action (e.g., Augustine, 398/1992; Locke, 1690; Pinker, 1984). But visual access to potential referents varies across contexts and language learners. For example, blind children lack visual access to potential referents, yet they arrive at virtually the same meanings for novel words as do sighted children. Landau and Gleitman (1985) have observed that they do so in part by using the syntactic contexts in which words appear. Of course, sighted children too recruit syntactic information to glean broad aspects of word meaning. For example, they expect that verbs taking sentential complements (e.g., the boys [verb] that the tiger will pounce) refer to mental states rather than activities (e.g., Asplin, 2002; Gleitman, 1990; Papafragou, Cassidy, & Gleitman, 2007). But it is an open question how much learners can glean about verb meaning from syntactic information alone.

In word learning tasks, sighted children can coordinate their sensitivity to syntactic information with their observation of the visual world. This coordination is especially important in verb learning (e.g., Fisher, 2002; Gleitman, 1990; Naigles, 1990), where syntactic information inherent in the sentence structure provides a "'linguistic zoom lens' to help the learner detect what is currently being expressed about an ongoing event or a state or relation" (Gleitman & Fisher, 2005: p. 132). In this 'zoom lens' metaphor, syntactic information, available concurrently with the visual scene, focuses learners' attention on the part of the observed scene that is most compatible with the meaning conveyed by the sentence.

Recent evidence has gone further, demonstrating that simply hearing a novel verb in a sentence, without an accompanying scene, helps toddlers to infer some information about its meaning, and to use that information later to find an appropriate referent when a visual scene becomes available (Yuan & Fisher, 2009). Yuan and Fisher first introduced 2-year-olds to two actors, engaged in a dialogue in which a novel verb was mentioned either in an intransitive sentence (e.g., the boy mooped) or a transitive sentence (e.g., the boy mooped the girl). Toddlers then viewed two scenes, one depicting an event with one actor, the other an event with two actors. When asked to "find mooping", toddlers looked longer at the two-actor event in the transitive, than in the intransitive, condition. Performance in a

* Corresponding author. Tel.: +1 847 467 0737.
E-mail address: s-arunachalam@northwestern.edu (S. Arunachalam).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.10.015

control condition revealed that this preference for the two-actor scene reflected their interpretation of the transitive verb, *per se*, and not the presence of the two nouns alone. Thus, even before seeing an event, when toddlers heard a novel verb in a sentence, they associated the number of nouns in the sentence with the novel verb, and brought this into correspondence with the number of actors in the subsequently presented visual scene.

Toddlers' spontaneous ability to match the number of nouns with the number of participants in the event described by the verb is impressive; to succeed, they had to have used linguistic information alone to determine the number of participants to associate with the verb. But recent work also reveals that toddlers can glean more from syntactic context than the number of likely participants. For example, when a novel verb is presented in conjunction with a visual scene, 26-month-olds can determine whether a verb taking two noun phrases refers to a causative scene (e.g., a duck pushing down on a rabbit's head, forcing him to squat) or synchronous scene (e.g., a duck and bunny each twirling one arm in circles) by noting whether the verb appears in a transitive (e.g., the duck is gorping the bunny) or intransitive sentence (e.g., the duck and the bunny are gorping) (Naigles, 1990; see also Bunger & Lidz, 2004; Fisher, 2002; Naigles & Kako, 1993). Thus, when a novel verb is introduced in conjunction with visual scenes, toddlers use the syntactic structures in which two noun phrases occur to infer the relation between two participants.

But can toddlers use syntactic structure to zoom in on the relation between participants even when the verb is introduced without accompanying visual information? The current experiment addresses this issue directly. Following Yuan and Fisher, we presented the linguistic stream before providing any visual information, but following Naigles, we introduced sentences with two nouns and subsequently tested learners' interpretations by showing them visual scenes with two participants. At issue is whether toddlers can use syntactic information in absence of an event, to form a representation of the novel verb that allows them to determine not just the number of participants involved, but also the relation in which the participants will stand to each other.

First, toddlers heard a novel verb, presented in either transitive (e.g., the boy is going to moop the lady) or intransitive sentences (e.g., the boy and the lady are going to moop). Only later were they given relevant visual information. This consisted of two test scenes, presented side-by-side, each involving two actors. In one test scene, the actors were engaged in a causative event (e.g., a boy spins a girl in a chair); in the other, the same actors were engaged in a synchronous event (e.g., a boy and a girl each wave one hand in circles). Crucially, while the test scenes were presented, no syntactic or semantic information was available to help the toddler infer which scene depicted "mooping". Therefore, their choice of test events had to be guided by the syntactic information they had heard before exposure to the events.

We reasoned as follows: if toddlers can use the syntactic structure in which a novel verb is presented to infer meaning, then those who had heard the novel verbs in transitive sentences should choose the causative event when asked to "find mooping" more often than those who had heard intransitive sentences. Those who had heard intransitive sentences should choose either scene. See Naigles and Kako (1993) for evidence that although toddlers have a clear preference to interpret transitive sentences as referring to causative events, they accept both synchronous and causative events as referents of intransitive sentences, a pattern that likely reflects the fact that intransitive sentences are compatible with both specifically synchronous meanings (e.g., "waving hand in circles") and more general meanings (e.g., "playing").

1. Methods

1.1. Participants

Forty typically-developing toddlers (20 males; mean age of 27.3 months, ranging 24.9–29.9) were included in the final sample. We focused on 27-month-olds because at this age, toddlers actively add verbs in their own spontaneous speech and have demonstrated success in similar experimental tasks (e.g., Naigles, 1990; Yuan & Fisher, 2009). Toddlers were recruited from Evanston, IL and surrounding areas, and were acquiring English as their native language, with less than 25% exposure to another language. Parents completed the MacArthur long form vocabulary checklist: words and sentences (Fenson et al., 1993). Mean production vocabulary was 442 words (range: 50–681); there were no differences between conditions in vocabulary. To be included in the final sample, toddlers had to (a) point correctly on at least two (of four) pointing games and training trials, and (b) point clearly on at least one test trial. Fourteen toddlers who failed to meet these criteria were excluded; another 13 were excluded due to fussiness, and 2 to experimenter error.

2. Materials

2.1. Visual stimuli

In the Dialogue phase, toddlers viewed digitized video recordings of two live actors speaking. In the test phase, they viewed videos of different actors performing actions; half of the actions involved two human actors, and half involved one human actor and one inanimate object. Videos were presented on a 20-in. television screen.

2.2. Auditory stimuli

A female native speaker of American English produced the speech stimuli (described in Table 1) using child-directed speech. Speech was recorded in a sound-attenuated booth, synchronized with the visual stimuli, and presented on a speaker centered below the visual display.

2.3. Apparatus and procedure

Toddlers played freely with toys while the caregiver signed a consent form and completed the MacArthur

Table 1
Representative set of stimuli.

	Dialogue phase		Test	
			Baseline (24 s)	Response (24 s)
Linguistic stream	Transitive condition A: You know what? B: What? A: The lady mooped my brother B: Really? The lady mooped your brother? A: And the boy is going to moop the girl B: Oh yes. He is going to moop her	Intransitive condition A: You know what? B: What? A: The lady and my brother mooped B: Really? The lady and your brother mooped? A: And the boy and the girl are going to moop B: Oh yes. They are going to moop	Look! Wow!	Where's mooping?
Observational stream			 	

checklist. The toddler and caregiver were then brought into an adjoining room where the toddler was seated in an infant seat, 16 in. from the television screen. The caregiver sat behind the toddler and was requested not to talk or otherwise interact with her during the session. One experimenter controlled the experimental procedure from behind a curtain; another sat next to the toddler to elicit responses. We asked toddlers to indicate their choice of scenes by pointing. We chose pointing as a measure because it is a more active behavioral response than traditional looking-time methods, has been successfully used at this age in word learning tasks (Bernal, Lidz, Millotte, & Christophe, 2007; Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008), and converges well with looking-time (Arunachalam & Waxman, 2009). Pointing was recorded with a video camera centered above the screen.

Toddlers first participated in a warm-up game designed to encourage them to point to the screen. Two video clips of Sesame Street characters were presented on the screen, side-by-side, and the experimenter asked the toddler to point, once to a particular character (e.g., Elmo), and once to a particular action (e.g., dancing). If a toddler was reluctant to point or pointed incorrectly, the experimenter demonstrated the correct response and encouraged the toddler to do the same.

Next, each toddler participated in six trials, each featuring a different verb. Two training trials (involving familiar verbs, sleep and hug) were followed by four experimental trials (involving novel verbs). Because the training trials were designed to familiarize toddlers with the experimental procedure, the structure of the training and experimental trials was identical. Each trial included two phases: dialogue and test. For experimental trials, toddlers were randomly assigned to either the transitive or intransitive condition. Toddlers in both conditions saw exactly the same video scenes, but heard different auditory stimuli see Table 1. The four experimental trials were presented in one of two random orders, balanced across conditions. The left–right positions of the two types of test scene were counterbalanced across trials.

2.3.1. Dialogue phase

Each trial began with a scene of two women having a conversation in which they used either a known verb (on the two training trials) or a novel verb (on the four experimental trials). One training trial involved an unergative intransitive verb (sleep), and the other involved a transitive verb (hug). The experimental trials involved either transitive sentences, e.g., "the boy is going to moop the lady" (transitive condition) or conjoined-subject intransitive sentences, e.g., "the boy and the lady are going to moop" (intransitive condition). Each dialogue consisted of two six-sentence video clips, averaging 34 s, including eight mentions of the verb. Dialogue videos appeared in the center of the screen.

2.3.2. Test phase

Toddlers then saw the two test scenes side-by-side. On the training trials, the event type was held constant across test scenes. On the sleep trial, both test scenes depicted one participant: (1) a woman sleeping, and (2) the same woman crying. On the hug trial, both scenes depicted two participants: (1) a woman hugging a toy, and (2) the same woman lifting a box. On each experimental trial, both test scenes depicted the same two participants (e.g., a man and a woman), with the event type differing across the two test scenes: (1) a synchronous event (e.g., man and woman each wave one of their own hands in circles), and (2) a causative event (e.g., the same man spins the same woman around). On all four experimental trials, each test scene depicted two moving participants.[1]

For both training and experimental trials, the test phase began with a 24 s inspection period, during which toddlers heard "Look! Wow!" and had an opportunity to inspect the test scenes, both of which were novel to them. The screen

[1] On half of the experimental trials, one participant was animate and the other inanimate; on the others, both were animate. On animate–inanimate trials, the animate participant was always the agent of the causative action. Because there were no effects of animacy, we collapse across this factor in the analysis.

then went blank for 1.5 s, during which time the novel verb was presented. Importantly, at this point, the verb was presented with no syntactic information to indicate its meaning: In both conditions, toddlers heard, e.g., "Where's mooping?" Next, the test scenes reappeared for 24 s, and infants heard: "Do you see mooping? Find mooping!" The experimenter seated next to the toddler repeated the question and encouraged her to point (e.g., "Can you show me?"). Neutral feedback (e.g., "Good pointing!") was given on all trials, regardless of the child's response. Notice that if toddlers' choices differ as a function of condition, these differences must be attributable to the syntactic information (i.e., transitive or intransitive sentences) provided before the test scenes appeared.

2.3.3. Coding

All pointing responses were verified by a second condition-blind coder from the video recordings. Agreement between coders was 100%. We analyzed toddlers' first point on a given trial, with one exception: one toddler spontaneously corrected his first response on one trial, explicitly saying, "No, that's mooping", pointing to the other scene.

2.4. Results and discussion

The results, depicted in Fig. 1, provide the first documentation that even in the absence of a visual scene, 27-month-olds can use syntactic information to zoom in on a particular relation between event participants. As predicted, toddlers in the transitive condition were more likely to point to the causative scene ($M = .67$) than those in the intransitive condition ($M = .43$), $F(1, 38) = 10.06$, $p < .01$, $d = 1.00$. In addition, toddlers in the transitive condition pointed to the causative scene significantly more often than would be expected by chance ($t(1, 19) = 3.19$, $p < .005$, $d = 0.71$). In contrast, those in the intransitive condition did not differ from chance ($t(1, 19) = 1.34$, $p = .20$). Although this null effect must be interpreted with caution, it is worth noting that this outcome for the intransitive condition accords with reports that toddlers accept both causative and non-causative events when introduced to intransitive sentences (Naigles & Kako, 1993; Yuan & Fisher, 2009). Recall that intransitive sentences are compatible with both specific synchronous meanings like "waving hand in circles" and more general meanings like "playing". Clearly, then, by 27 months, toddlers successfully extract information about a new verb's meaning from the syntactic context in which it is introduced, even in the absence of visual information, and recruit this information later when they are shown candidate events. Toddlers' ability to cull an initial representation of a novel verb's meaning in the dialogue phase, on the basis of syntactic context alone, is striking. It reveals an early ability to do more than count noun phrases and match them to event participants (Yuan & Fisher, 2009). Counting noun phrases has been argued to be a central ability in early verb learning (Fisher, 2002; Lidz, Gleitman, & Gleitman, 2003; Yuan & Fisher, 2009), but the current results reveal that toddlers use syntactic information to encode more than just participant number; they also use syntactic information to home in on the kind of relation the verb describes.

This interpretation warrants careful consideration. After all, it is also possible that toddlers formed no representation of the verbs' meaning during the dialogue phase, instead simply remembering the sentences and waiting until they encountered the verb again in the presence of a visual scene before constructing a representation of its meaning. However, decades of research on 'fast mapping' cast doubt on this possibility. The evidence reveals that infants and young children form at least a preliminary representation of a new word's meaning, using whatever information is available rather than waiting until they have established its meaning fully (e.g., Carey & Bartlett, 1978; Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996). In fact, this ability to establish an initial, if incomplete, representation of a word's meaning is, at its core, the phenomenon of 'fast mapping'.

We suggest that in the current experiment, toddlers are 'fast mapping' verb meanings in the dialogue phase, using syntactic information alone to establish a representation of verb meaning. This capacity, while striking, may be instru-

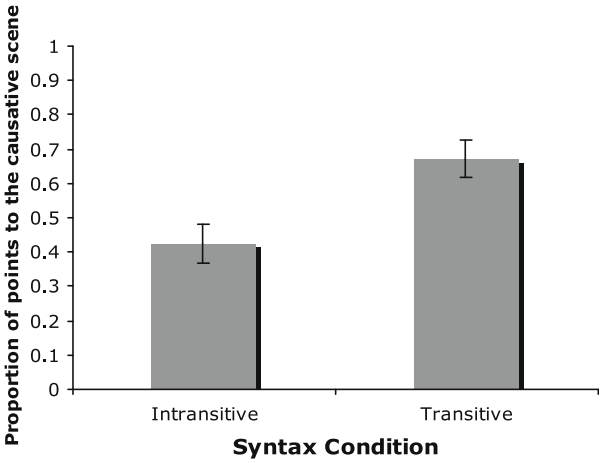

Fig. 1. Mean proportion of points to the causative scene, expressed as a function of syntactic condition.

mental in the natural course of verb learning. After all, verbs are often introduced in the absence of the events to which they refer (e.g., "let's find your shoes") (Gleitman & Gleitman, 1992; Tomasello & Kruger, 1992). The evidence reported here reveals that under such circumstances, toddlers are capable of more than just associating a novel verb with a subsequently-presented action (Ambalu, Chian, & Print, 1997; Tomasello & Kruger, 1992): they can use the verb's syntactic structure to construct a representation of its meaning—including the kind of relation it describes—and to query this representation later when shown candidate events. Syntactic information itself is therefore a powerful and precise support for word learning.

Acknowledgements

This research was supported by National Institutes of Health (NIH) Grant HD30410 to the second author. A portion of this data was presented at the 33rd Annual Boston University Conference on Language Development (November 8, 2008; Boston, MA). We are grateful to the infants and caretakers who participated in this study, and to C. Fisher, E. Leddon, S. Yuan, and the reviewers for helpful comments.

References

Ambalu, D., Chian, S., & Print, T. (1997). When is it best to hear a verb? The effects of the timing and focus of verb models on children's learning of verbs. *Journal of Child Language, 24*, 25–34.

Arunachalam, S., & Waxman, S.R. (2009). 24-Month-old infants' use of syntactic context in noun and verb learning. Paper presented at the society for Research in Child Development Biennial Meeting, Denver, CO.

Asplin, K.N. (2002). Can complement frames help children learn the meaning of abstract verbs? Unpublished Doctoral Dissertation, University of Massachusetts, Amherst.

Augustine (398/1992). *Confessions*. New York: Oxford University Press.

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development, 3*, 325–341.

Bunger, A., Lidz, J. (2004). Syntactic bootstrapping and the internal structure of causative events. In *Proceedings of the 28th Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17–29.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. S. (1993). *User's guide and technical manual for the MacArthur Communicative Development Inventories*. San Diego, CA: Singular Press.

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*, 55–64.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 1–55.

Gleitman, L. R., & Fisher, C. (2005). Universal aspects of word learning. In J. A. McGilvray (Ed.), *The Cambridge companion to Chomsky*. Cambridge: Cambridge University Press.

Gleitman, L., & Gleitman, H. (1992). A picture is worth a thousand words: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science, 1*, 31–35.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, R. N. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28*, 99–108.

Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development, 67*, 3101–3119.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Lidz, J., Gleitman, H., & Gleitman, L. R. (2003). Understanding how input matters: The footprint of Universal Grammar on verb learning. *Cognition, 87*, 151–178.

Locke, J. (1690). *An essay concerning human understanding*. Cleveland: Meridian Books, 1964.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*, 628–634.

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

Naigles, L., & Kako, E. (1993). First contact: Biases in verb learning with and without syntactic information. *Child Development, 64*, 1665–1687.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition, 105*, 125–165.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*, 311–333.

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*, 619–626.