Verb Learning in Non-social Contexts

Sudha Arunachalam and Leah Sheline
Boston University

1. Introduction

Acquiring word meanings is generally described as a social process involving live interaction with an interlocutor and/or joint attention to the word's referent (e.g., Baldwin & Moses, 1991; Krcmar et al., 2007; Roseberry et al., 2009; Tomasello, 2000; Tomasello & Farrar, 1986).

But for verbs in particular, access to richly informative linguistic contexts can promote acquisition of a novel verb's meaning even without interaction or visual access to the verb's referent (e.g., Arunachalam & Waxman, 2010; Arunachalam et al., 2012; Scott & Fisher, 2009; Yuan & Fisher, 2009). In these studies, toddlers were overhearers of a videotaped conversation between two adults who used the novel verb.

Here we ask whether toddlers can acquire verbs in even more socially impoverished contexts, in which syntactic information is available, but a conversational context is not. Toddlers are presented novel verbs in informative syntactic contexts, but with no discourse context, no interlocutor (neither live nor videotaped), no child-directed speech, and no visual referent for the novel verb.

We adapted a classic manipulation from the syntactic bootstrapping literature (Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990) to see whether toddlers would be able to use syntactic information in the absence of social or visual cues. Toddlers were presented with either transitive or intransitive sentences incorporating a novel verb. The sentences streamed ambiently from a speaker while toddlers watched an unrelated soundless animation or played quietly with toys. To see if they gleaned the syntactic properties of the novel verbs from this exposure, we then presented two test scenes: one depicting two actors engaged in a causative event and the other depicting two actors engaged in synchronous events, and asked them to identify the novel verb's referent. Toddlers' eye gaze was recorded and served as a measure of their interpretation of the novel verb.

2. Method

2.1. Participants

Thirty-six typically developing toddlers (18 males; mean age: 27.1 months, age range: 25.0 to 29.9 months) participated. Toddlers were recruited from Boston, MA and surrounding areas, and were acquiring English as their native language, with less than 30% exposure to another language. An additional seven toddlers were excluded from analysis due to poor tracking (> 50% of the Test Phase on three or more trials).

2.2. Apparatus

Toddlers sat in a car seat or on a caregiver's lap approximately 20 inches from a 24-inch eye-tracker monitor (Tobii T60XL), which records gaze at a rate of 60 frames/second.

2.3. Materials and Procedure

Toddlers and caregivers were welcomed into the waiting area. While the toddlers played with toys, their caregivers completed a consent form and the MacArthur-Bates short CDI checklist. The toddler and caregiver were then brought into the testing room.

Seated in front of the eye-tracker monitor, toddlers first engaged in a brief training procedure.

They viewed two training trials, each introducing two video clips playing side-by-side on the screen. On one trial, these depicted two familiar characters (Big Bird, Elmo), and on the other, a familiar character performing two familiar actions (dancing, eating). The experimenter asked the toddler about one of the scenes on each trial (e.g., "Where's Big Bird?"). No novel words were used during training.

Next, toddlers participated in four test trials, each incorporating a different novel verb (*fez, moop, lorp, biff*), and comprising two phases: Syntactic Familiarization and Test. Toddlers were randomly assigned to one of two conditions. Both conditions presented the same visual materials; the only difference was the syntactic structure of sentences heard during the Syntactic Familiarization Phase (Transitive or Intransitive).

<table>
<tr><td rowspan="2">SYNTACTIC FAMILIARIZATION</td><td colspan="3">TEST</td></tr>
<tr><td></td><td></td><td>RESPONSE WINDOW</td></tr>
<tr><td>(toddler watches an irrelevant animation or plays quietly with a shape sorter)</td><td>Causative Synchronous  </td><td>  </td><td>Causative Synchronous  </td></tr>
<tr><td>Transitive condition
The girl lorped the teacher.
Mommy was lorping the train. ...</td><td rowspan="2">Look! Wow!</td><td rowspan="2">Where's lorping?</td><td rowspan="2">Find lorping!</td></tr>
<tr><td>Intransitive condition
The girl and the teacher lorped.
Mommy and the train were lorping. …</td></tr>
</table>

Figure 1. Visual and auditory stimuli for a representative test trial.

Syntactic Familiarization Phase

During the *Syntactic Familiarization Phase*, which lasted approximately 1 min 20 sec, toddlers viewed a soundless video of animated shapes on the screen or played quietly with toys. During this time, an auditory stream consisting of 27 prerecorded sentences played from the monitor's speakers. Depending on condition assignment, the sentences were either transitive (e.g., *The girl lorped the teacher*), or intransitive with a conjoined subject (e.g., *The girl and the teacher lorped*).

The sentences were produced by a female native speaker of American English with adult-directed list-reading prosody, and included a variety of noun phrase participants in the verb's argument positions as well as a variety of tenses. The stream played as a series of unconnected sentences with approximately 1 s between them.

Test Phase

The Test Phase presented two dynamic video scenes simultaneously, side-by-side (adapted from Arunachalam & Waxman, 2010, and Arunachalam et al., 2012). These videotaped scenes depicted people and objects performing actions. In one scene, two people, or a person and an object, engaged in a causative action (e.g., boy spins girl around), and in the other, they

performed synchronous actions (e.g., both boy and girl wave one hand). (Two trials depicted two people in each scene, and the other two depicted one person and one object.)

First, the two scenes appeared for 24 seconds with no novel words. Toddlers heard prerecorded audio that had been synchronized with the visual stimuli: "Look! Wow!" Next, the two scenes disappeared and were replaced by a star positioned in the center of the display, and accompanied by the audio, "Where's *lorping*?" Finally, during the Response window, the two test scenes reappeared in their original locations, and the test query repeated: "Do you see *lorping*? Find *lorping*!"

3. Predictions and results

If toddlers were able to glean the syntactic properties (transitive or intransitive) of the novel verbs, despite that they were presented in non-social and non-interactive contexts, they should have mapped them to a corresponding event compatible with that syntax. Specifically, we expected that toddlers who had heard transitive sentences would prefer the causative event compared to toddlers who had heard intransitive sentences. Based on previous work (Arunachalam et al., 2012), we expected these diverging preferences to become evident beginning at least one 1 sec after the onset of the Response window.

Proportion of looks to the Causative Scene serves as the dependent measure. Eye gaze patterns in the two conditions begin to diverge just after 1 sec from the onset of the Response window. See Figure 2. To be consistent with previous work, we analyzed gaze from 1 to 2 sec of the Response window. As predicted, toddlers in the Transitive Condition reliably preferred the Causative Scene compared to toddlers in the Intransitive Condition in this time window ($t(34) = 2.4$, $p < .03$). (There is no reliable difference from 0 to 1 sec of Response, $t(34) = 0.06$.)

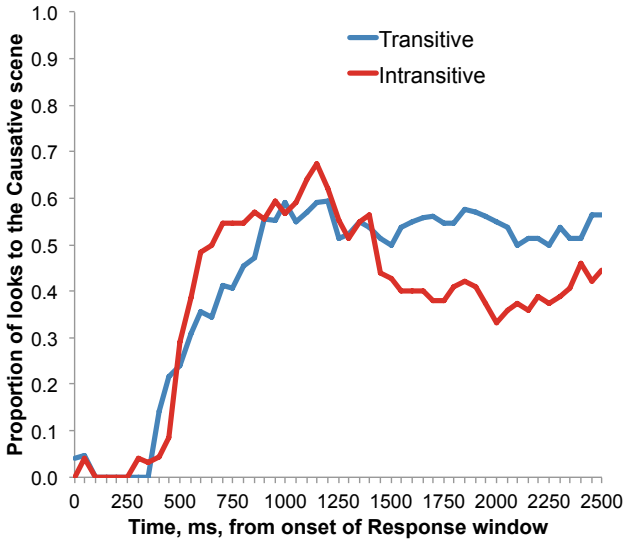

Figure 2. Proportion of looks to the causative scene over time from the onset of the Response window after the test query (e.g., *Where's lorping?*), separated by condition.

4. Discussion

The current study builds upon previous work demonstrating that by age 2, toddlers can extract information about a new verb from its syntactic context, even before viewing a relevant event (Arunachalam & Waxman, 2010; Arunachalam et al., 2012; Scott & Fisher, 2009; Yuan &

Fisher, 2009; Yuan, Messenger, & Fisher, 2011).

These findings go further to demonstrate that they can do so in an impoverished social context. Like previous studies, toddlers were presented with the novel verbs in informative linguistic contexts, but here, we presented them without visual access to the speakers, child-directed speech, or discourse context.

Our findings are also compatible with evidence that noun learning can occur in contexts in which children have visual access to the noun's referent, but have no interlocutor or no child-directed speech (e.g., Akhtar et al., 2001; Ma et al., 2011; Scofield et al., 2007). In the case of verb learning, the present study demonstrates that rich linguistic information *alone* can be sufficient for acquiring at least some aspects of verb meaning.

Of course, in most real-world learning situations, at least some social context is available, and it is well established that toddlers make use of this information when it is available. However, our findings lend insight into how toddlers may acquire aspects of verb meaning in overhearing contexts in which they may not be directly attending to the ambient speech, and in which no visual referent, discourse context, or child-directed conversation is available.

References

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*(2), 416-430. doi:10.1111/1467-8624.00287

Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. R. (2012). Out of sight but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*. *Online first,* doi:10.1080/01690965.2011.641744

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442-446. doi:10.1016/j.cognition.2009.10.015

Baldwin, D. A., & Moses, L. J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development, 10,* 309-329. doi: 10.1111/1467-9507.00168

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology*, *10*(1), 41-63.

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science*,*14*(5), 1207-1219. doi:10.1111/j.1467-7687.2011.01073.x

Ma, W., Golinkoff, R., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning And Development*,*7*(3), 185-201. doi:10.1080/15475441.2011.579839

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R.M. (2009). Live action: Can young children learn verbs from video? *Child Development*, *80*(5), 1360-1375. doi:10.1111/j.1467-8624.2009.01338.x

Scofield, J., Williams, A., & Behrend, D. A. (2007). Word learning in the absence of a speaker. *First Language*, *27*(3), 297-311. doi:10.1177/0142723707077252

Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language And Cognitive Processes*, *24*(6), 777-803. doi:10.1080/01690960802573236

Tomasello, M. (2000). Perceiving intentions and learning words in the second year of life. In: M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development*.

Cambridge University Press.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454-1463. doi:10.2307/1130423

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619-626. doi:10.1111/j.1467-9280.2009.02341.x

Yuan, S., Messenger, K., & Fisher, C. (2011, March). Syntax and selection: Learning combinatorial properties of verbs from listening. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, QC, Canada.