Cognition 129 (2013) 569–573

Contents lists available at ScienceDirect

Cognition

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

Brief article

Two-year-olds can begin to acquire verb meanings in socially impoverished contexts

Sudha Arunachalam *

Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA 02215, United States

ARTICLE INFO

Article history:
Received 23 January 2013
Revised 19 August 2013
Accepted 21 August 2013
Available online 17 September 2013

Keywords:
Language acquisition
Word learning
Syntactic bootstrapping

ABSTRACT

By two years of age, toddlers are adept at recruiting social, observational, and linguistic cues to discover the meanings of words. Here, we ask how they fare in impoverished contexts in which linguistic cues are provided, but no social or visual information is available. Novel verbs are presented in a stream of syntactically informative sentences, but the sentences are not embedded in a social context, and no visual access to the verb's referent is provided until the test phase. The results provide insight into how toddlers may benefit from overhearing contexts in which they are not directly attending to the ambient speech, and in which no conversational context, visual referent, or child-directed conversation is available.

1. Introduction

Learning new words is a critical task of early childhood. It is typically thought of as a social process in which the child apprehends the intentions of an interlocutor and/or engages in joint attention with the interlocutor to the word's referent (e.g., Baldwin, 1993; Baldwin & Moses, 2001; Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009; Tomasello, 2000; Tomasello & Barton, 1994; Tomasello & Farrar, 1986). But many potential learning situations do not present such rich social or visual information. Imagine a toddler playing quietly in the living room while her parents talk in the kitchen. The toddler may overhear many words for which she does not have meanings, but in this situation she lacks visual access to the referents of those words, and the speech she is hearing is not child-directed and is not embedded in a child-centered conversation or interaction. Can toddlers benefit from such socially impoverished learning situations by positing at least basic meanings for some of the new words they hear?

Recent studies have begun to address this question by closing in on the kinds of information, at a minimum, that toddlers require to establish a new word's meaning. For example, by age two, toddlers can acquire the meaning of a novel noun by watching two adults interact with and talk about its referent, even if they are not directly spoken to and do not themselves interact with the object (e.g., Akhtar, Jipson, & Callanan, 2001; Gampe, Liebal, & Tomasello, 2012; O'Doherty et al., 2011; Scofield, Williams, & Behrend, 2007; Shneidman, Buresh, Shimpi, Knight-Schwarz, & Woodward, 2009). They can also acquire novel nouns introduced in adult-directed speech rather than child-directed speech (Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011).

For verbs, too, there is evidence of learning in the absence of joint attention. Shown a videotaped interaction between two adults who are animatedly engaged in conversation, and using a novel verb in informative linguistic contexts, toddlers can assign the verb a broad meaning, even though the event it describes is not visually available (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2012; Arunachalam & Waxman, 2010; Scott & Fisher, 2009; Yuan & Fisher, 2009). The critical information provided to toddlers in these studies is syntactic. Toddlers hear the novel verb in syntactic contexts that provide information about the verb's meaning, and can use this information to set

* Address: Department of Speech, Language, and Hearing Sciences, 635 Commonwealth Ave., Rm. 323, Boston, MA 02215, United States. Tel.: +1 617 353 7491.
E-mail address: sarunach@bu.edu

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.08.021

down a basic lexical representation for the verb, an ability known as syntactic bootstrapping (Gleitman, 1990; Landau & Gleitman, 1988). In Arunachalam and Waxman (2010), for example (following on results from Naigles, 1990; see also Hirsh-Pasek, Golinkoff, & Naigles (1996)), toddlers heard adults conversing using novel verbs in either transitive sentences (e.g., *The girl lorped the boy*) or intransitive sentences (e.g., *The girl and the boy lorped*). These two sentence types describe fundamentally different kinds of relations between the event participants: transitive verbs typically describe causative events, in which one actor acts on another (e.g., a girl pushes a boy), and intransitive verbs typically describe non-causative events (e.g., a girl and a boy wave). Even though in Arunachalam & Waxman's task toddlers were only overhearing this syntactic information, and were not ostensibly shown the verb's referent, they nevertheless mapped novel transitive verbs, but not intransitive verbs, to causative events.

But is social context, such as an animated conversation between two adults, *required* to trigger toddlers' abilities to use linguistic context to acquire verb meanings? For some other kinds of learning from auditory stimuli, social context is not required. Statistical learning studies demonstrate that even young infants can glean patterns from auditory input presented in socially impoverished contexts in which the infant hears ambient auditory streams (e.g., Saffran et al., 1996; Marcus et al., 1999). However, these patterns are devoid of meaning. Identifying coherent forms from an auditory stream may differ fundamentally from assigning those forms a referential meaning (Naigles, 2002), with the latter requiring social context. In fact, toddlers have difficulty assigning meaning to isolated word forms if they do not perceive them as referential (Fennell & Waxman, 2010).

Nevertheless, we suggest that assigning basic aspects of meaning from syntactic cues is different from other referential mapping tasks, and that it does not require social context. Because syntactic cues convey aspects of meaning in themselves without identification of a specific visual referent in the world, cues available in social interaction, such as eye gaze, may not be necessary. In fact, we hypothesize that the socially impoverished conditions typically used in statistical learning studies will be sufficient. After all, syntax provides distributional information about where different kinds of words can and cannot appear in a sentence, and it stands to reason that humans' skills at discerning patterns in auditory input will serve them well in discerning that an unfamiliar transitive verb is both preceded and followed by a noun phrase. Syntactic information may thus provide a different inroad to acquisition, one that does not require the rich social contexts that other kinds of mapping tasks do.

Therefore, in the current study we asked whether toddlers could acquire the meaning of a novel verb from severely socially impoverished situations: we did not provide child-directed speech, conversational context, interaction with or eavesdropping on people speaking (neither live nor videotaped), or visual access to the event the verb describes. The only information provided was linguistic: toddlers heard the novel verbs in informative sentences (either transitive or intransitive). These sentences were presented in a maximally non-social context: they were recorded in adult-directed speech and were not embedded in a discourse, and streamed ambiently from a speaker while the child engaged in an unrelated activity (Lany & Saffran, 2011). At test, the toddlers saw two candidate referents for the novel verb: a causative event and a synchronous event, and were asked, e.g., "Find lorping." Our goal was to determine whether toddlers, on hearing the transitive or intransitive sentences, could use their syntactic content alone to map the novel verbs to meaning, even though no social or visual information was available.

2. Methods

2.1. Participants

Forty typically-developing toddlers (age range: 25.0–29.9 months, mean: 27.2) were included in the final sample. All were recruited from Boston, MA and surrounding communities, and were acquiring English as their native language, hearing other languages less than 30% of the time. Caregivers completed the MacArthur-Bates Communicative Development Inventory Short Form Level II A (Fenson et al., 2000). Toddlers' production vocabulary ranged from 38 to 100 words (mean 75 words) and did not differ between conditions. An additional eight toddlers were excluded from analysis due to inattentiveness and one due to parental interference.

2.2. Materials

Participants participated in four trials, each consisting of a Familiarization and a Test phase.

2.2.1. Visual stimuli

During Familiarization, the visual stimuli consisted of animated shapes moving silently on the screen, and during Test, they consisted of digitized video recordings of live actors performing actions.

2.2.2. Auditory stimuli

During Familiarization, the auditory stimuli consisted of 27 sentences playing from the monitor's speakers. These sentences were produced by a female native speaker of American English with 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. Speech was recorded in a sound-attenuated booth. During Test, toddlers heard two attention-getting phrases (*Look! Wow!*) and the test query (e.g., *Where's lorping? Find lorping!*), also recorded by a female native speaker of American English in a sound-attenuated booth, and synched with the visual stimuli.

2.3. Apparatus and procedure

Toddlers played with toys while the caregiver signed a consent form and completed the MacArthur-Bates checklist. The toddler and caregiver were then brought into the