Journal of Memory and Language 88 (2016) 104–116

Contents lists available at ScienceDirect

Journal of Memory and Language

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

A new experimental paradigm to study children's processing of their parent's unscripted language input

Sudha Arunachalam

Dept. of Speech, Language, & Hearing Sciences, 635 Commonwealth Ave., Boston University, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 17 September 2015
revision received 1 February 2016

Keywords:
Children
Referential expressions
Eye-tracking
Methodology
Language acquisition
Language processing

ABSTRACT

This paper introduces a new experimental paradigm for studying children's real-time language processing of their parents' unscripted speech. Focusing on children's processing of referential expressions, or the phrases that parents used to label particular objects, we engaged dyads in a game in which parents labeled one of several objects displayed on a screen, and the child was to quickly identify it as their eye gaze was tracked. There were two conditions; one included a competitor object (e.g., the target was a striped umbrella and the display also included an umbrella with polka dots), while the other one did not (e.g., only one umbrella was present). The results revealed evidence of children's incremental processing of their parents' referential expressions. They also showed faster processing of postnominally-modified as compared to prenominally-modified referential expressions. Parents tended to produce postnominally-modified referential expressions in the more difficult experimental condition in which there was a competitor object, suggesting either that these expressions are also easier for them to produce, or that they accommodate their children by producing more easily processed expressions. We discuss the potential of this paradigm for advancing theories of the relationship between child-directed language input and children's language processing.

Introduction

Children's knowledge about the words, syntactic structures, and discourse-pragmatic properties of their language develops in tandem with their very ability to process and understand these elements. Children are not born with adult-like language processors; rather, their processing abilities increase over development and are shaped by features of the ambient language input. In turn, as they parse the input, they acquire new words and structures and their language competence increases. One domain in which the parallel development of language knowledge and parsing skill is evident is in children's understanding of referential expressions, or the linguistic expressions used to refer to entities in the world. Children's abilities to identify the referents of these expressions begin early in life and develop rapidly. By 6 months of age, infants asked, for example, to "Look at the apple" from a display depicting an apple and a mouth prefer to look at the apple, indicating that they have processed to some degree the noun "apple" and can shift their gaze to its referent (Bergelson & Swingley, 2012). By 24 months of age, toddlers look to a named image within 500 ms of the noun's onset (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998), and they process referential expressions incrementally, as indicated by the fact that they launch eye movements to potential referents even before the noun is complete (e.g., Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999). As children get older and processing speed increases, they are able to use prenominal information from determiners and modifiers to narrow down the space of possible referents (e.g.,

E-mail address: sarunach@bu.edu

http://dx.doi.org/10.1016/j.jml.2016.02.001

Fernald, Thorpe, & Marchman, 2010; Lew-Williams & Fernald, 2007), and they show recognition of pragmatic and discourse factors that influence their interlocutor's choice of referential expression (e.g., Allen, Skarabela, & Hughes, 2008; Clancy, 1993, 1997; Clark & Amaral, 2010; Graham, Sedivy, & Khu, 2014; Hughes & Allen, 2015; Matthews, Lieven, & Tomasello, 2010).

This developing ability to process even complex referential expressions is of course important in itself, allowing children to identify entities that are the topic of conversation, but it is also important because it allows children to take advantage of opportunities to increase their lexical and grammatical knowledge. Identifying the referent of one expression will in many situations facilitate the assignment of meaning to other linguistic elements later in the utterance. For example, Fernald, Marchman, and Hurtado (2008) found that the faster 3-year-old children are to parse through a modified determiner phrase (e.g., "the red car"), the more likely they are to acquire the meaning of a novel noun downstream in the utterance (as in "The red car is on the *deebo*"). Studies of verb learning, too, suggest that the ability to process referential expressions efficiently is essential. For example, 2- and 3-year-olds more easily acquire novel verb meanings when the verbs are preceded by a short unmodified description (e.g., "The man is *pilking*") than when preceded by a longer description (e.g., "The nice tall man is *pilking*") (He, Kon, & Arunachalam, in preparation; Kon, Goksun, Bagci, & Arunachalam, 2016). Given the importance of quick and efficient comprehension of referential expressions for acquisition of new vocabulary, it will be helpful to understand precisely how children's online processing of these expressions develops.

Importantly, which referential expressions a child will hear depends on properties of the referential world, such as how many potential referents are co-present, as well as properties of the dyad, such as the goals of the communication (e.g., parents may seek to educate or entertain in addition to achieving successful reference) and the parent's perceptions of the child's developmental level. Parents are known to tailor their speech according to their child's level (e.g., Bellinger, 1980; Bornstein, Hendricks, Haynes, & Painter, 2007; Hoff-Ginsberg, 1994; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007; Newport, Gleitman, & Gleitman, 1977; Pan, Rowe, Singer, & Snow, 2005; Snow, 1972, 1977) as evident, for example, in the higher pitch and slower tempo characteristic of infant-directed speech (Fernald & Simon, 1984; McRoberts & Best, 1997), as well its redundancy and shorter and simpler grammatical structures (e.g., Snow, 1972).

Because of these individual and situational differences, it is important to study parents' speech in naturalistic situations. But studying what parents produce offers only an incomplete picture of what children will understand from that input. After all, child-directed speech is unlikely to be taken up veridically by the child's developing language processing system (e.g., Harris, 1992; Lidz & Gagliardi, 2015; Omaki & Lidz, 2015). (Most obviously, syntactic structures that the child parses incorrectly will be represented differently in the child's mind than in the speaker's.) Understanding how children processes the input in real-time will offer insight into what linguistic representations they are likely to form. New approaches are thus needed to reveal not only the linguistic input that children are exposed to, but how they comprehend this input, and in turn, whether parents, too, are aware of the limitations of children's intake and tailor their child-directed speech accordingly.

To study these issues, we developed a paradigm to integrate the study of unscripted productions from parents and the study of children's language processing. We took advantage of recent advances in portable eye-tracking by using a small eye-tracker with a tablet; the tablet setup allowed parents and children to sit together relatively naturally and to feel like they were playing a game, while yielding eye-tracking evidence about the children's online language processing. We presented the dyads with a finding game in which parents labeled one of six pictures on the screen for their child, and the child had to guess which picture was intended. On each trial the dyads viewed an array of six clipart images of animals, objects, and people. We first numbered the six locations for the parent and subsequently referred to the images by their numbered location. We indicated a specific image to the parent on each trial by whispering a number to them. Parents were not told what to say; they were only told that they could talk, but not point, to help their child find the correct picture as quickly as possible. From the recordings of these interactions and the child's gaze coordinates as recorded by the eye-tracker's software, we asked what kinds of referential expressions parents used to label the pictures, and how quickly these expressions were understood by the children as they sought to identify the target on each trial.

We designed two trial types. On trials in the Same condition, the target images were ones that had to be described with a modifier in order to disambiguate them from one of the distractors, such as a striped umbrella in an array that also included an umbrella with polka dots. In the other trial type, the Different condition, only one umbrella was present in the array, and so no modifier was necessary; it would thus be *over*informative for parents to produce, "a striped umbrella" or "an umbrella with stripes" on that trial. We analyzed children's eye gaze to determine whether their latency to look to the target image differed by type of referential expression.

One might expect parents to produce "good" referential expressions that support their children's comprehension. After all, adult speakers engage in *audience design* (Clark & Murphy, 1982), and generally speaking, their referential expressions are adapted to the knowledge state of their interlocutor, both in adult- and child-directed speech (e.g., Ariel, 1990; Chafe, 1987; Clancy, 1993, 1997; Du Bois, 1987; Gundel, Hedberg, & Zacharski, 1993; Hughes & Allen, 2015; Prince, 1985; Rohde & Frank, 2014; Skarabela, 2006).[1] The fact that parents' child-directed

[1] At least in some cases, what appears to be audience design may be minimization of the speaker's own processing load (e.g., Ferreira & Dell, 2000; Horton & Gerrig, 2005; Horton & Keysar, 1996; Wardlow Lane & Ferreira, 2008; Wardlow Lane & Liersch, 2012). We return to this possibility in the General Discussion; our question here is simply to what degree parents' choice of referential expression converges with what children find easy to process.

speech is attuned to their child's developmental level (e.g., Bellinger, 1980; Bornstein et al., 2007; Hoff-Ginsberg, 1994; Huttenlocher et al., 2007; Newport et al., 1977; Pan et al., 2005; Snow, 1972, 1977) suggests that child-directed speech may be particularly well adapted in this sense.

But audience design may also be particularly difficult in the context of dyadic parent–child interactions, given that children's needs and internal knowledge states are likely quite difficult to model; these change rapidly over development, and children do not always provide clear feedback about their understanding. How are parents to know what referential expressions will best support their child's comprehension in any given situation? We suspect that this is not trivial; psycholinguists, too, have difficulty characterizing what a "good" referential expression looks like for a particular context, even in the context of adult–adult interactions. For example, one might expect a helpful speaker to follow Grice's (1975) second maxim of Quantity ("Do not make your contribution more informative than is required"). In the present context, that would mean not providing modifiers in the Different condition when a basic-level object noun suffices. But not only do adult speakers often overmodify (e.g., Deutsch & Pechmann, 1982; Engelhardt, Bailey, & Ferreira, 2006; Ferreira, Slevc, & Rogers, 2005; Gann & Barr, 2014; Mangold & Pobel, 1988; Sedivy, 2003; Sonnenschein, 1984), in some cases, overinformativeness appears to *benefit* adult listeners: Rather than being slowed by the Gricean violation, they identify referents more quickly (e.g., Arts, Maes, Noordman, & Jansen, 2011; Davies & Katsos, 2010; Levelt, 1989). In other cases, however, overinformative expressions have been found to be harmful, slowing down processing and comprehension (e.g., Engelhardt, Demiral, & Ferreira, 2011; Engelhardt et al., 2006; Grodner & Sedivy, 2011; Sedivy, 2003, 2007). Naturally, the particulars of the situation matter (Engelhardt et al., 2011).

For children, there does not appear to be evidence that overinformativeness hinders their comprehension. Although Sonnenschein (1982) found that 5-year-olds did not benefit from overinformativeness, Davies and Katsos (2010) found that neither did 5-year-olds reject overinformative expressions when asked to judge whether a speaker spoke in a good or bad way (though they did exhibit sensitivity to overinformativeness when given a graded rather than binary judgement task). Further, Thorpe and Fernald (2006) found that children can "listen through" unfamiliar or unnecessary adjectives, suggesting that they may simply be able to ignore overinformative modifiers if they are not helpful. (However, note that overinformative referential expressions may be harmful if they are to be used in the service of acquiring a novel word occurring later in the utterance, e.g., Kon et al., 2016; Lidz, Bunger, Leddon, Baier, & Waxman, 2009.) If parents are sensitive to the possibility that overinformativeness does not hinder their children's comprehension, and if their tendency toward redundancy in infant-directed speech governs their choices of referential expressions as well, parents may be inclined to provide extra information to give the child more descriptive information with which to identify the intended referent.

But another factor that should affect children's comprehension of referential expressions is how easy or difficult they are to process. If parents provide more information than is strictly necessary to support their child's identification of the referent, they will likely produce longer, more detailed referential expressions. More information usually poses a bigger processing burden. And as children's processing abilities are continuing to develop throughout early childhood, an increased processing load may be insurmountable. This processing issue should be most evident in the Same condition, in which the parent must provide more than a single noun to disambiguate the intended referent.

There are two salient choices for how to realize this modifying information: prenominally or postnominally within the referential expression. Prenominal modifiers (e.g., "striped umbrella"), provide an earlier disambiguation than postnominal modifiers (e.g., "umbrella with stripes"), and so if speakers are interested in efficiency, they may choose the prenominal modifier—in the current study we attempted to encourage efficiency by telling dyads that the game involved finding the pictures as quickly as possible. On the other hand, the evidence on children's processing of referential expressions suggests that they sometimes have difficulty integrating prenominal modifiers. Sekerina and Trueswell (2012) found that 6-year-old Russian speakers hearing an expression like "red butterfly," given a display containing two red referents and two butterflies, did not show anticipatory eye movements to red referents before the noun, suggesting that they were not making use of the prenominal information. Huang and Snedeker (2013) found that 5-year-olds did show evidence of incremental processing of prenominal scalar adjectives (e.g., "Point to the big coin" in a display including a big and a small coin), but they were slower than adults to do so (see also Nadig, Sedivy, Joshi, & Bortfeld, 2003). In simpler tasks, involving only one distractor, children as young as 3 years of age perform better (e.g., Fernald et al., 2010; Thorpe, Baumgartner, & Fernald, 2006, Experiment 2; Tribushinina & Mak, 2015). Fernald et al. (2010) found that 36-month-olds incrementally interpreted a prenominal adjective given a display with a blue car and red car and the instruction, "Can you find the blue car?" Thirty-month-olds, however, appeared to wait, failing to make use of the prenominal adjective until after hearing the noun. Tribushinina and Mak (2015) found that 3-year-olds who heard "a soft pillow" looked to a pillow rather than a book more quickly than when they heard, "a nice pillow." But the objects were not evidently either "soft" or "nice" in the images, and so children may have relied on conceptual knowledge about things that can be soft or the collocational experience they have had with these adjective-noun pairings. They could also have simply listened to the adjective and not integrated it with the noun at all, because only one of the objects fit the description.

These studies suggest that prenominal modifiers pose challenges for young learners, particularly in visual displays that present more than one distractor. (No studies that we are aware of have explicitly compared English learners' success with prenominal and postnominal modifiers, but see Ninio, 2004 for evidence from Hebrew, and Weisleder & Fernald, 2009, for evidence from Spanish—both languages that place modifiers postnominally.) Thus, if parents are aiming to maximize their child's comprehension, rather than point them to the target as early in the referential expression as possible, they may avoid prenominal modifiers which young children may have difficulty integrating. Postnominal modifiers (or other strategies, like two separate utterances) may be easier because the child can first identify the object category and then narrow down the precise referent from among the available choices.

Our paradigm allows us to ask whether parents' referential expressions tend to be supportive of their children's comprehension because we target both parents' productions and children's processing. To determine whether these two converge, we ask, with respect to parents' choice of referential expressions:

(1) How often do parents provide overinformative referential expressions, that is, how often do they produce expressions containing modifiers on Different condition trials? We predict they will, at least some of the time, because modifiers may not hinder children's processing and may be helpful.
(2) When parents produce modifiers, do they more often produce them prenominally or postnominally? If the modifier is necessary for identifying the referent, as in the Same condition, we predict that parents will tend to produce postnominal modifiers, despite that postnominal modifiers result in later disambiguation, because children may have difficulty integrating prenominal modifiers.

And with respect to children's gaze:

(1) How are children's latencies to look to the target object affected by properties of the referential expression the parent produces? If prenominal modifiers are difficult to integrate, we expect to see shorter latencies when parents use postnominal modifiers.

Materials and method

Participants

Thirty-two typically-developing children (16 girls, 16 boys) ranging in age from 3;2 to 4;11 years (mean 3;11 years) and one of their parents (28 mothers, 4 fathers) were included in the final sample. Dyads were recruited from Boston, MA and surrounding areas, and were acquiring English as their native language, hearing other languages less than 30% of the time. An additional nine dyads were excluded from the final analysis because the child talked continuously while the parent was speaking on most trials ($N = 1$), the child was a twin of another child who had already participated with that parent ($N = 1$), there was excessive track loss such that the child contributed data on fewer than 40% of trials ($N = 2$),[2] the parent provided almost all "roundabout" descriptions such as animal sounds instead of referential expressions ($N = 1$), or experimenter error, that is, failure to provide appropriate instructions to the parent ($N = 2$) or failure to record the session ($N = 2$). Parents provided informed consent on behalf of themselves and their children.

Materials

The stimuli consisted of a .pdf file with several pages, each depicting six images. See Fig. 1. The first page depicted six numbers, used to explain the procedure to parents. On Same condition trials (Fig. 1B), there were two pictures depicting the same basic level object category (e.g., an umbrella). The distinguishing feature was a salient aspect of the object's appearance or location, for example whether an umbrella had stripes or polka dots or whether a book was open or closed. One member of the pair was designated the target object and the other was the distractor. On Different condition trials (Fig. 1C), there was only one picture depicting any given basic level category. Participants viewed five trials from each condition. We also included one initial warm-up trial and five filler trials. Four of the filler trials included pairs of objects like Same condition trials, but the target was not a member of the pair. Two lists were created such that Same condition trials in List 1 were converted to Different condition trials in List 2 by replacing the distractor with an object from a different basic-level object category, and List 1 Different condition trials were similarly converted to List 2 Same condition trials. Dyads were randomly assigned to a list.

Apparatus

The eye-tracking device, Tobii X2-30 (sampling at 30 Hz),[3] was attached to an EyeMobile bracket with a Windows Surface Pro 2 tablet running Tobii Studio 3.2. These sat on a child-sized table in front of the child, who sat in a child-sized chair. The parent sat next to the child. While in many developmental eye-tracking studies, parents wear blindfolds to prevent their gaze from being tracked, for this setup it is critical that parents can see clearly while still not being tracked. The head movement box for the X2-30, as outlined in the device's technical specification, is approximately 20″ and a height of 14″ at 70 cm. In our setup, parents sitting next to their child were unlikely to cross into this space unless they leaned over considerably. Nevertheless, to be

[2] We were surprised at the low track loss rates in this sample, given that our setup permitted children much more freedom to move than more traditional paradigms in which children are often seated in a chair that restricts or discourages movement. Naturally we reminded children on occasion to "sit back in your chair" and "don't touch the screen," though anecdotally this did not seem to occur moreso than in our other studies with this age group. It may be that the design of our game, in which children had to remain quiet in order to hear their parents' utterances and respond quickly, encouraged them to sit relatively still.

[3] We conducted a timing test with this portable eye-tracker using procedures recommended by Tobii (http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii_Eye_Tracking_Timing_whitepaper.pdf). This yielded similar performance to the large stand-alone monitor T60XL model that we also have in our lab.

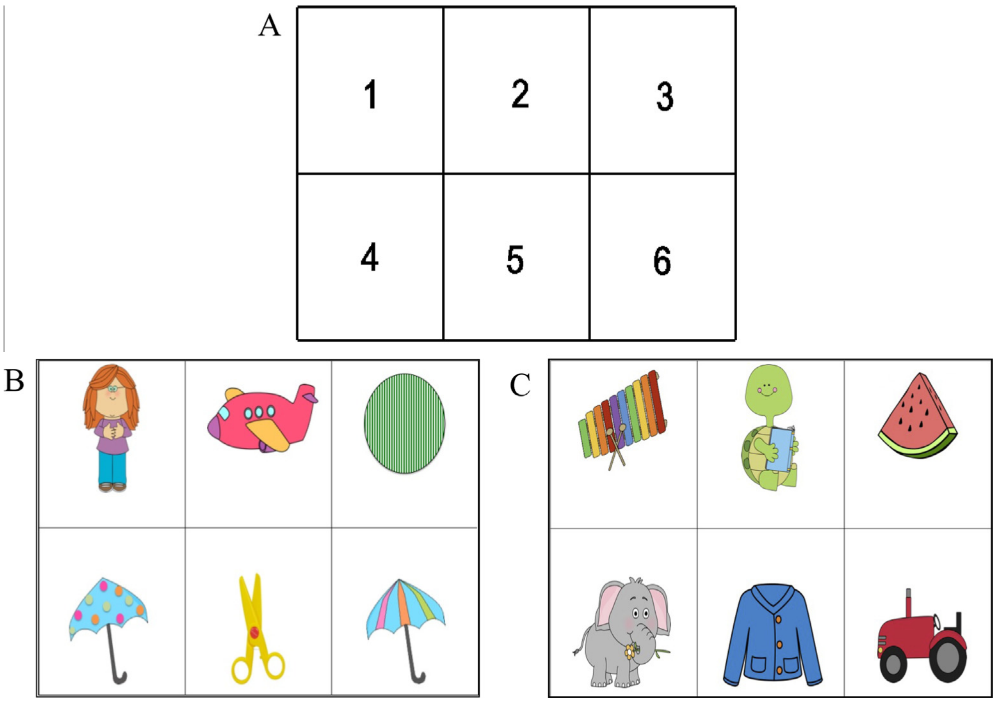

Fig. 1. A. Numbered grid used during instructions. B. Representative grid from the Same Condition. C. Representative grid from the Different Condition.

Fig. 2. A child participating with his mother.

sure that parents could move around comfortably and engage naturally with their child, they wore laser goggles that blocked the near-infrared wavelengths that the Tobii X2-30 uses but not shorter wavelengths (we found these for under $50). The parent was able to see the screen very well, similarly to wearing sunglasses. See Fig. 2.

Procedure

The child and parent were first welcomed into our playroom, where the child played with an experimenter while the parent read and signed the consent form. The parent and child were then seated in front of the experimental apparatus as described above, and the experimenter explained the game. The parent first saw the array of six boxes containing the numbers 1 through 6 (Fig. 1A). Parents were informed that each trial would display six images, one in each box, and that the experimenter would refer to the boxes by number on each trial. Their goal was to get their child to point to the image in the box we named. Children were also told that they would see pictures in the boxes and play a finding game with their parent, and that their job was to point to the correct picture as quickly as possible. We embedded the task in a finding game to provide motivation and to ensure that the location of the target object was random and initially unknown to the child and parent.[4]

[4] While we initially experimented with open-ended games and picture book contexts in which the parent was provided no instruction on what to label (in addition to no instruction on how to label it), we worried that children's gaze might align with what the parent was labeling simply because both were scanning the scene similarly, or were interested in the same objects. The finding game, for which we provided the target on each trial, eliminated these possibilities.

With the child seated approximately 65 cm from the eye-tracker, the child's gaze was calibrated using Tobii Studio, and recording began. The recording was set up using the Screen Recording stimulus type in Tobii Studio. The experimenter navigated to the stimuli file and reminded the parent of the instructions. On each trial, the experimenter whispered a number into the parent's ear on each trial, out of the child's earshot. We opted for this low-tech method of prompting the parent to keep the casual gameplay aspect at the forefront. The parent responded at his or her own pace, and the trial concluded only when the child indicated a response. The experimenter provided encouraging feedback, focusing on the dyad's teamwork to ensure that the child did not feel she was taking a test (e.g., "You and Daddy make a great team!"). The experimenter advanced to the next trial by swiping to the next page of the .pdf or using a keyboard connected via Bluetooth.

Results and discussion

This paradigm yields two types of data: the referential expressions produced by the parents, and the children's eye gaze. We first extracted the referential expressions from the tablet's webcam recording of the session using Praat software (Boersma & Weenink, 2011), and identified the onset, offset, and disambiguation point—the first point of each expression at which a listener could *in principle* identify a unique referent (e.g., onset of *stripes* in *umbrella with stripes*), and the point at which adult listeners would be expected to direct gaze to that referent (e.g., Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). These time stamps were then used to identify where the child was looking during and after the referential expression from the gaze data file exported from Tobii Studio.

We first discuss the parents' referential expressions, and then children's eye gaze in the context of the different types of expressions parents produced. For the parents' speech, we first report on their speech rate and the productions that were excluded from further analysis. We then report on the main issues of interest: how informative parental expressions were in each of the Different and Same conditions, and where parents placed modifiers within the referential expression when they did produce them. For children, in addition to discussing trials that were excluded from analysis, we report on their latencies to their first look to the target image in each condition, separated by the presence and placement of modifiers within the referential expression.

Parents' productions

Speech rate

We first assessed parents' speech rate across all trials to determine if we would have to normalize speech rate in our analyses. Parents might be expected to speak more slowly to younger than older children, or on Same condition trials as compared to Different conditions, given that the former present a more difficult task. After all, slower speech rate can support children's language comprehension (e.g., Haake, Hansson, Gulz, Schötz, & Sahlén, 2014) and parents who are accommodating their children might well speak more slowly in more difficult contexts. However, speech rate per referential expression (calculated as the number of syllables per second) differed minimally across conditions (Same: 0.27, Different: 0.26, $t(31) = 0.10$, $p = .92$), and did not correlate with age (Pearson's $r = .084$). The standard deviation in speech rate across participants was quite small as well (0.15 syll/s) and so we did not analyze speech rate further.

Excluded trials

For the subsequent analyses, we excluded trials on which parents did not produce a clear referential expression to label the target. Some parents on some trials produced descriptions related to a recent personal experience (e.g., "What did we use this morning with Daddy?" for the target *hammer*), made animal sounds (e.g., "Hoo, hoo" for the target *owl*), or produced roundabout guessing game descriptions (e.g., "We take it with us when we go outside on a rainy day. It keeps the rain off our faces," etc.). This was the case on only 4% of trials, including filler and warmup trials, not counting the one dyad that was excluded altogether because the parent produced mostly expressions of this type. We also excluded 13% of trials in the Same condition, from 15 different parents, on which the parent did not provide sufficient information for the child to identify the target—this occurred, as far as we can tell, because the parent did not initially notice the distractor or recognize it as a competitor. We were not specifically interested in repairs in this study, and so these trials were excluded.

Informativeness

On all included trials, parents labeled the target object in a single referential expression that included a basic-level noun (e.g., "umbrella"); when a modifier occurred it was also part of the noun or determiner phrase. We therefore assessed informativeness by asking whether parents included modifiers. Because we excluded Same condition trials on which the parent was underinformative, modifiers were produced 100% of the time in the Same condition in the final data set. In the Different condition, we asked whether parents included a modifier at all—recall that in the Different condition no modifiers were necessary because there were no basic-level competitors. Across all Different condition trials, parents provided a modifier (that is, an adjective or prepositional phrase in addition to an optional determiner and a noun) 21% of the time; the proportion of trials including a modifier did not correlate with children's age ($r = -.079$, $p = .67$).

Modifier placement

Within the modified referential expressions, we also examined whether the parent modified prenominally or postnominally (e.g., striped umbrella vs. umbrella with stripes). Modifier placement differed by condition, with modifiers appearing prenominally only 21% of the time in the Same condition, but 60% of the time in the Different condition.[5] The proportion of pre- vs. postnominal modifiers was not correlated with children's age in either condition (Different: $r = .076$, $p = .72$; Same: $r = .054$, $p = .79$).

Why might more postnominal modifiers occur more frequently in the more difficult Same condition? The prenominal modifier permits earlier disambiguation (because there was only striped object in the display), and so if parents were simply interested in efficiency—recall that we told dyads the goal of the game was for the child to point to the target as quickly as possible—we would expect them to prefer prenominal modifiers. We suspect that parents instead preferred to produce postnominal modifiers because children find it easier to process them; we return to this below.

Children's eye gaze

Excluded trials

We performed analyses on the subset of trials for which we had useable gaze data. We excluded trials on which, in the window beginning with the onset of the parent's referential expression and extending for 2 s after the offset of this expression, there was more than 25% track loss (12%). We also excluded trials on which the child happened to be looking at the target at the referential expression's onset (3%), and trials on which the child never looked at the target (1%) (as all children pointed to the target eventually, these latter few trials may have been ones on which the child noted the target peripherally without fixating on it, or were ones on which calibration quality had deteriorated over the course of the session). Note that no trials were excluded because the child did not point correctly; in all cases the child identified the target initially or, on a few trials, did so shortly afterward, with a verbal correction, e.g., "oh, I mean this one."

Same condition

Because the Same condition presented two similar referents, we first examined the data coarsely to ascertain whether children were considering both possible referents during the referential expressions and narrowing down their hypotheses over time. Thus, we plotted children's gaze to the target, the pair distractor, and the other four images over time from the onset of the parent's referential expression, excluding frames during which children were not looking at any of the six pictures. See Fig. 3. The mean referential expression duration was 1771 ms ($SD = 820$ ms); the figure depicts gaze for 3000 ms after referential expression onset.

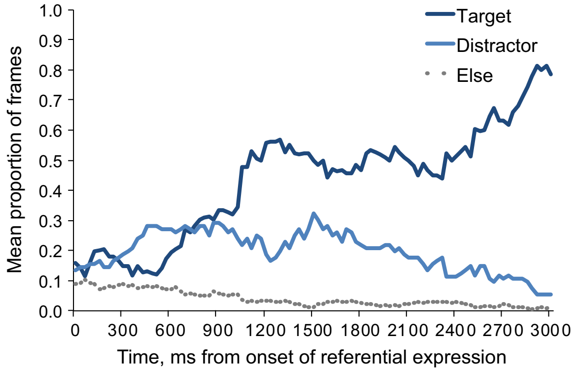

Fig. 3. Children's gaze to the images over time in the Same condition.

Children began by looking at the target, distractor, and other images at roughly chance levels (because there are six images, chance is ~17%). Looks to the target and distractor then increase over time, compared to the other four images, and finally looking to the distractor drops off while looking to the target continues to increase. This provides only a coarse picture, as the particular referential expression to which children are responding, and therefore the timeline along which these changes are occurring, necessarily differs not only for each trial but also for each participant, depending on what the parent chose to produce and their speech rate. Nevertheless, inspection of Fig. 3 suggests that children are indeed experiencing competition between the target and pair distractor as the referential expression unfolds.

To understand how children's eye gaze related to the type of referential expression they heard, we analyzed latencies to look to the target image. Because referential expressions varied in length, we calculated latencies from the offset of the referential expressions to make interpretation easier, although looks before the offset of the expression (i.e., negative latencies) were included. We divided the trials into prenominal and postnominal categories, although note that because these were not experimentally assigned, the number of trials in each category differs. For prenominal modifiers, the mean latency to look to the target was 276 ms, and for postnominal modifiers it was −256 ms. See Fig. 4.

We fit these latencies to a mixed-effects model with subject and item as random effects and Modifer Placement (before vs. after) as a fixed effect, using the lmer() function in the lme4 package in R (R v. 2.14.2) (Bates, Maechler, Bolker, & Walker, 2015; R Development Core Team, 2012). The parameter estimates, listed in Table 1, indicate a significant effect of modifier placement, with postnominal modifiers showing significantly shorter latencies than prenominal modifiers. We hypothesized that children might get faster to identify the target with age, but we did not find a significant negative correlation of latency with children's age ($r = -.21$, $p = .28$).

Given that the mean latency from referential expression offset was negative when a postnominal modifier was present, it could be that children were looking to the target (and distractor) simply on the basis of the noun, without

[5] We were concerned that the distribution of pre- to post-nominal modifiers was distorted in the Same condition by the plausibility of prenominal modifiers for the objects depicted. While on most trials, pre- and post-nominal modifiers were both possible (e.g., striped umbrella vs. umbrella with stripes), on three of the ten trials we thought it much more likely that an adult would only use a postnominal modifier given the nature of the property that distinguished the target and distractor—for example, on one trial the target was a rabbit on a table, with the distractor a rabbit under a table (see Edwards & Chambers, 2011). Excluding these trials from analysis, the percentage remained exactly the same—21% of the time on trials that permitted either pre- or postnominal modifier descriptions, parents used postnominal modifiers.

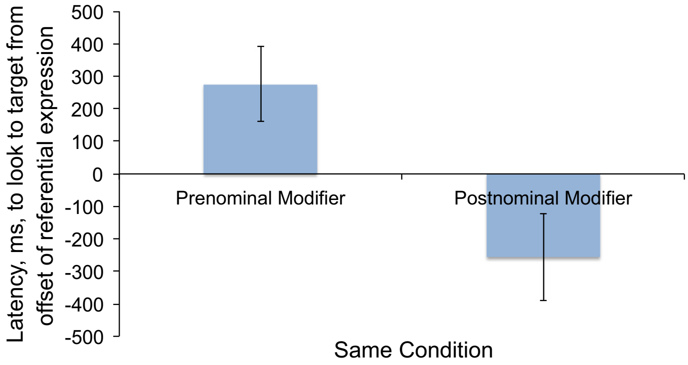

Fig. 4. Children's latencies to look to the target in the Same condition, by type of referential expression. Error bars represent standard error of subject means.

Table 1
Fixed effects from model of latencies to look to the target in the Same condition.

Effect	Estimate	*SE*	*t*-value
Intercept	272.2	228.4	1.19
Modifier Placement (after vs. before)	−563.4*	241.2	−2.34

* $p < .05$ (on normal distribution).

having integrated the modifier information.[6] However, reassessing these latencies from the disambiguation point of the referential expression instead of its offset (e.g., the onset of "stripes" in "umbrella with stripes") reveals that the mean latency to look to the target from the disambiguation point is 489 ms, well into this modifying information. In fact, only on 26% of trials were children looking to the target within 200 ms after the disambiguation point (the 200 ms offset is to account for the time it likely takes to launch an eye movement, Matin, Shao, & Boff, 1993). It is likely, then, that on most trials children's gaze to the target was driven by the information available in the modifier rather than sheer chance in looking to the target over the distractor after hearing the object category label.

The fact that children had difficulty with "striped umbrella" relative to "umbrella with stripes" suggests that linguistic integration of the adjective and noun is still a challenge for preschoolers. This result is consistent with Sekerina and Trueswell (2012), although our design, unlike theirs, presented only one object in the scene bearing the relevant property. Our design also did not permit the use of conceptual or collocational knowledge as Tribushinina and Mak (2015) did, because the properties of the objects were not intrinsic to the categories named (e.g., softness is a very typical property of a pillow, but stripedness is not of an umbrella).

Different condition

In the Different condition, we investigated latencies according to whether the trials had no modifier, a prenominal modifier, or a postnominal modifier. See Fig. 5. Two trials on which the (same) parent produced one modifier in each position (e.g., "a little turtle that's reading") were included in both groups. Latencies were longer when no modifier was produced (240 ms) than when a modifier was produced (prenominal: −121 ms, postnominal: −614 ms). This is to be expected given that the referential expressions were very short when there was no modifier, and thus children required time to direct their eye gaze toward the target after the offset of the expression. The fact that the presence of modifiers did not slow children down, despite that they were overinformative, is in line with previous literature finding no evidence that overinformativeness is harmful for children's comprehension at this age (e.g., Davies & Katsos, 2010; Thorpe & Fernald, 2006).

We analyzed latencies in the Different condition to examine the effect of modifier placement as we did for the Same condition (excluding trials on which no modifier was produced). See Table 2. This analysis indicated, as in the Same condition, significantly shorter latencies with postnominal modifiers than prenominal modifiers. There was a non-reliable trend toward shorter latencies with increased age ($r = -.35$, $p = .083$).

The results from the Different condition thus parallel those from the Same condition. The interpretation of these differences could be somewhat different, though, as in the Different condition postnominal modifiers were simply extra information; children could uniquely identify the referent without them. Prenominal modifiers, of course, added linguistic material before the identifying noun and thus lengthened the delay between the onset of the referential expression and when the referent was labeled. Despite that in many cases the parent's modifier did in principle provide enough information to identify the intended referent even in the Different condition, it is perhaps not surprising that children did not use this information; after all, given the referential scene, in which each box presented a unique object type, there was no reason to expect a modifier and thus to attempt to integrate it. Taken together with performance in the Same condition, however, these results are suggestive toward our hypothesis that preschoolers find it difficult to incrementally use prenominal information to identify named referents (at least with a complex display including six images rather than two, cf. Fernald et al., 2010).

[6] Thanks to an anonymous reviewer for raising this hypothesis.

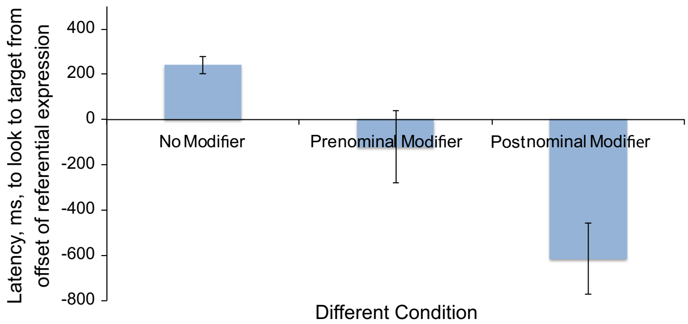

Fig. 5. Children's latencies to look to the target in the Different condition, by type of referential expression. Error bars represent standard error of subject means.

Table 2
Fixed effects from model of latencies to look to the target in the Different condition on trials on which parents produced a modifier.

Effect	Estimate	*SE*	*t*-value
Intercept	−50.72	149.4	−0.34
Modifier Placement (after vs. before)	−568.06*	206.4	−2.75

* $p < .05$ (on normal distribution).

General discussion

We have introduced a new experimental paradigm to study children's online comprehension of their own parent's unscripted speech. It goes beyond conventional paradigms by linking the child's online processing to the specific linguistic expressions produced by their parents. The unscripted nature of the parent speech reveals the choices parents make about what expressions to use, and the gaze data from children allows for inferences about how these expressions are processed. In the current study, we used this paradigm to explore parents' productions of, and children's online comprehension of, referential expressions labeling objects in a visual display. More specifically, we examined parents' production of modifiers and modifier placement when the display either did or did not warrant modifier inclusion—that is, there either was or was not a basic level object competitor. Our goal was to see if the types of referential expressions that parents most often produced would match the types that children found easiest to process.

On Different condition trials, in which there was no competitor, parents overmodified about 20% of the time. This is not surprising given that overmodification occurs even in adult-directed speech (e.g., Deutsch & Pechmann, 1982; Engelhardt et al., 2006; Ferreira et al., 2005; Gann & Barr, 2014; Sedivy, 2003; Sonnenschein, 1984). In fact, it is perhaps surprising that parents did not overmodify more often; we suspected that parents would be particularly prone to provide redundant or overinformative expressions in an effort to provide the child with more descriptive information with which to identify the referent, especially because children may not be hindered by the presence of extra information (e.g., Davies & Katsos, 2010; Thorpe & Fernald, 2006). Just as in the adult literature, research with a variety of situations and tasks will be necessary to determine when, how, and why parents choose to or choose not to provide overinformative or redundant information. When parents did provide modifiers, we found that these occurred prenominally a little over half of the time. Children showed shorter latencies in looking to the target image when their parents produced postnominal modifiers than when their parents produced prenominal modifiers, though we do not know whether this was because children struggled to integrate prenominal modifiers or whether they were faster with postnominal modifiers because they could simply ignore them.

In the Same condition, in which there were two potential referents from the same basic-level object category, parents did, as expected, produce disambiguating modifiers in their referential expressions. However, these modifiers much more often appeared postnominally than prenominally, and just as in the Different condition, children were faster to identify the referent when the modifier was postnominal. It could be that the object names are more familiar, more robustly represented, and more easily accessible for young children than the modifier labels, and so they find it easier to begin their search for the referent from this starting point—just as hearing modifiers postnominally can help children acquire the modifiers themselves (Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010).

The correspondence between parents' choice of postnominal modifiers in the more difficult Same condition and children's shorter latencies with postnominal modifiers is interesting. Could it be that parents produced more postnominal modifiers in the more difficult condition because they were aware of their children's processing preferences and were trying to support their children's comprehension? This would mean that parents not only accommodate their children's language levels by producing less complex speech overall (e.g., Huttenlocher et al., 2007), but that they also accommodate very specific (and perhaps surprising, given adults' skillful integration of prenominal modifiers) features of children's language processing—namely, their relative difficulty with online integration of prenominal as opposed to postnominal modifiers with their noun complements.

Another possibility, however, is that parents chose these expressions not to support their children's comprehension, but because they were easier for the parents themselves to produce. Just as modifier-noun integration introduces difficulties for children, it may increase cognitive load for the speaker. After all, in addition to being parents, in these interactions parents were also speakers, and the psycholinguistic literature on audience design indicates that while on the whole, speakers appear to model a hearer's internal knowledge state and adapt their productions accordingly (e.g., respecting common ground) (e.g., Clark & Marshall, 1981; Clark & Wilkes-Gibbes, 1986; Fussell & Krauss, 1989), there are limitations on these abilities, such that the information speakers include in an expression also depends on how accessible the information is to the speaker himself, thus minimizing demands for the speaker (e.g., Ferreira & Dell, 2000; Horton & Gerrig, 2005; Horton & Keysar, 1996; Wardlow Lane & Ferreira, 2008; Wardlow Lane & Liersch, 2012).

To begin to explore this possibility, we conducted a small follow-up study in which 17 adult participants provided referential expressions for an adult confederate in the same task. A similar pattern obtained, though attenuated; when speaking to other adults, our adult participants produced prenominal modifiers 48% of the time that they produced modifiers at all in the Different condition, compared to 32% of the time in the Same condition (recall that the patterns were 60% in the Different condition and 21% in the Same condition in child-directed speech). This suggests that ease or accessibility of a particular referential expression for the speaker, rather than a desire to be easily comprehended by the child, may partially (but not completely), account for parents' postnominal modifier preference in the Same condition,[7] and the difference in modifier position between the easy Different and the difficult Same condition suggests that this variation likely only surfaces when the cognitive load for the listener and/or speaker is high. It may thus be that parents' choice of linguistic expression is determined not only by their interest in supporting their child's comprehension but also by their own needs as speakers.

We see the new paradigm we have developed as particularly promising for improving theories of the relationship between parental input and children's language outcomes. Much attention has been paid to the now robust findings that the quantity and especially quality of the input children hear predict their language skill (e.g., Cartmill et al., 2013; Fernald & Morikawa, 1993; Goldfield, 1993; Hart & Risley, 1995; Hoff, 2003; Hoff & Naigles, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Huttenlocher et al., 2007; Naigles & Hoff-Ginsberg, 1998; Newport et al., 1977; Pan et al., 2005; Rowe, 2008, 2012; Weisleder & Fernald, 2013). With our experimental approach, one can zoom in on specific elements that constitute low- or high-quality input and examine how children process that input. Fernald and colleagues (Hurtado, Marchman, & Fernald, 2008; Weisleder & Fernald, 2013) have argued that language processing is a mediator in the relationship between the amount of child-directed speech children hear and their expressive vocabulary. Indeed, we have noted that there has been increasing focus in the literature on not only what input is directed to children, but children's *intake* of that input, which is necessarily different because the input is "filtered" through the child's language processing system. It is likely, therefore, that examining children's processing of particular elements of child-directed speech will lend further specificity to these mediation models, revealing not only what features of parental input best predict increases in processing skill across development, but also what aspects of processing skill are enhanced, and in turn, the extent to which parents are aware of their children's processing limitations and skills and accommodate their needs accordingly.

Methodologically, we found strong similarities to existing literature on children's processing of referential expressions in more controlled studies, even with the increased ecological validity gained by studying parents' speech instead of pre-recorded speech. The similarities are naturally reassuring; they suggest that children's processing of their own parent's speech does not qualitatively differ from their processing of experimentally designed or pre-recorded stimuli. However, because we could not predict exactly what parents would choose to say, the auditory stimuli that children were processing was necessarily—and interestingly—diverse as compared to a typical experiment.

This paradigm also has considerable flexibility for exploring language development beyond the research questions we pursued here. Within our own data set we observed quite a bit of variation in features we did not explicitly study, like determiner choice and prosodic cuing to signal contrastive information. The stimuli could be adapted to study a variety of other syntactic, semantic, and pragmatic phenomena as well. Given the open-ended nature of the task, any such investigation would likely yield a rich data set. More generally, this experimental setup could be used to see how children follow along in real-time during book-reading. Though two recent studies have used eye-tracking to examine book-reading, they involved either two eye-trackers on two different screens (Guo & Feng, 2013), or a head-mounted eye-tracker that young children may not tolerate (Evans & Saint-Aubin, 2005).

Another advantage of our combination of controlled experimental and naturalistic methods is that the same dyads are used to study both parental speech and children's comprehension, permitting analyses of individual differences. The importance of individual differences is increasingly recognized in the study of language comprehension, particularly as it has implications for detection and treatment of language impairment (e.g., Fernald & Marchman, 2012). In fact, the current experimental method may be especially useful for informing parent training interventions for children whose language processing skills are likely to be poor (e.g., children with specific language impairment) by identifying features of parent speech that their child can or cannot easily process.

[7] This raises the possibility that children struggle with prenominal modifiers because parents disprefer them and produce them less often. We think this is unlikely because some common modifiers like color names almost always occur prenominally in English (Thorpe & Fernald, 2006). It is likely that children hear many prenominal modifiers, and that parents only choose between prenominal and postnominal placement with some kinds of modifiers.

Our experimental method has corollaries in adult psycholinguistic research and in infant research on visual attention. With adults, eye-tracking has been used to study how adults achieve collaborative reference in naturalistic conversations between naïve participants (Brown-Schmidt & Tanenhaus, 2008; Gergle & Clark, 2011). Brown-Schmidt and Tanenhaus used a game setup modeled closely after previous work using a confederate and a naïve participant, except that both individuals were naïve participants and both eye-tracked. They found this to be a valuable method; the controlled game context meant that sufficient useable data was generated, but participants were not explicitly told what to say, and potential problems associated with using confederates were avoided (Kuhlen & Brennan, 2013). The current study uses a similar design, but in a more practical setup for young children, and without gaze data from the parent.

With infants, head-mounted eye-tracking has recently been used to determine where infants look during free play with their mothers (Franchak, Kretch, Soska, & Adolph, 2011). Headcams (small cameras attached to the infant's forehead with a headband) allow for recording of what the child sees, and reveal how the infant's visual perspective might differ from an adult's (Smith, Yu, & Pereira, 2011; Yoshida & Smith, 2008). The focus of this work is on what infants see and interact with. But to determine how infants process *language* addressed to them in real-time, this method would likely be too uncontrolled to yield useful data. Our paradigm allows for a more constrained visual environment as well as more constraints on what parents are likely to say.

In sum, we see this new paradigm as a tool for bridging two currently disparate areas of study: the language input directed to children by their parents, and the language skills—including knowledge of vocabulary and syntax as well as language processing ability—of the children themselves. We expect that future work using this tool will help to refine models of the relationships between input, intake, and outcomes in language development.

Acknowledgments

This work was funded by NIH K01 DC013306. Thanks to the reviewers for helpful comments on this manuscript, to the research assistants at the Boston University Child Language Lab for assistance with data collection, to the audience at the Symposium on Research in Child Language Disorders for helpful feedback, to the parents and children who participated, and to the family whose games of "I Spy..." during a long airplane ride inspired the task design.

References

Allen, S., Skarabela, B., & Hughes, M. (2008). Using corpora to examine discourse effects in syntax. *Trends in Language Acquisition Research, Corpora in Language Acquisition Research, 6*, 99–138.
Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification facilitates object identification. *Journal of Pragmatics, 43*(1), 361–374.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8*. <http://CRAN.R-project.org/package=lme4>.
Bellinger, D. (1980). Consistency in the pattern of change in mothers' speech: Some discriminant analyses. *Journal of Child Language, 7*(3), 469–487.
Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253–3258.
Boersma, P., & Weenink, D. (2011). *Praat: Doing phonetics by computer [Computer program]. Version 5.2.42*. <http://www.praat.org/> Retrieved on 18 September 2011.
Bornstein, M. H., Hendricks, C., Haynes, O. M., & Painter, K. M. (2007). Maternal sensitivity and child responsiveness: Associations with social context, maternal characteristics, and child characteristics in a multivariate analysis. *Infancy, 12*(2), 189–223.
Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-time investigation of referential domains in unscripted conversation: A targeted language game approach. *Cognitive Science, 32*(4), 643–684.
Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National academy of Sciences of the United States of America, 110*, 11278.
Chafe, W. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 21–51). Amsterdam: Benjamins.
Clancy, P. M. (1993). Preferred argument structure in Korean acquisition. In *Proceedings of the 25th annual child language research forum* (pp. 307–314).
Clancy, P. M. (1997). Discourse motivations for referential choice in Korean acquisition. In H. Sohn & J. Haig (Eds.). *Japanese/Korean linguistics* (Vol. 6, pp. 639–659). Stanford, Calif: Published for the Stanford Linguistics Association by the Center for the Study of Language and Information.
Clark, E. V., & Amaral, P. M. (2010). Children build on pragmatic information in language acquisition. *Language and Linguistics Compass, 4*(7), 445–457.
Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding*. Cambridge: Cambridge University Press.
Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology, 9*, 287–299.
Clark, H. H., & Wilkes-Gibbes, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.
Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua, 120*(8), 1956–1972.
Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11*(2), 159–184.
Du Bois, J. W. (1987). The discourse basis of ergativity. *Language, 63*, 805–855.
Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24*(6), 409–436.
Edwards, J. D., & Chambers, C. G. (2011). It's not what you said, it's how you said it: How modification conventions influence on-line referential processing. In E. Gibson & N. J. Pearlmutter (Eds.), *The processing and acquisition of reference* (pp. 219–237). Cambridge: MIT Press.
Engelhardt, P. E., Bailey, K. G., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of quantity? *Journal of Memory and Language, 54*(4), 554–573.
Engelhardt, P. E., Demiral, Ş. B., & Ferreira, F. (2011). Over-specified referring expressions impair comprehension: An ERP study. *Brain and Cognition, 77*(2), 304–314.
Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science, 16*, 913–920.
Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*(1), 203–222.
Fernald, A., Marchman, V. A., & Hurtado, N. (2008). Input affects uptake: How early language experience influences processing efficiency and vocabulary learning. In *Proceedings of the 7th IEEE international conference on development and learning* (pp. 37–42). IEEE.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*, 637–656.
Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*(3), 228–231.
Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), 104.
Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*, 1003–1015. http://dx.doi.org/10.1111/1467-8624.00331.
Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology, 60*, 190–217. http://dx.doi.org/10.1016/j.cogpsych.2009.12.002.
Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*(4), 296–340.
Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition, 96*(3), 263–284.
Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development, 82*(6), 1738–1750.
Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology, 25*(3), 203–219.
Gann, T. M., & Barr, D. J. (2014). Speaking from experience: Audience design as expert performance. *Language, Cognition and Neuroscience, 29*(6), 744–760.
Gergle, D., & Clark, A. T. (2011, March). See what I'm saying? Using Dyadic Mobile Eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 conference on computer supported cooperative work* (pp. 435–444). ACM.
Goldfield, B. A. (1993). Noun bias in maternal speech to one-year-olds. *Journal of Child Language, 20*, 85–99.
Graham, S. A., Sedivy, J., & Khu, M. (2014). That's not what you said earlier: Preschoolers expect partners to be referentially consistent. *Journal of Child Language, 41*(01), 34–50.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 41–58). New York: Academic Press.
Grodner, D., & Sedivy, J. (2011). The effect of speaker-specific information on pragmatic inferences. In E. Gibson & N. J. Pearlmutter (Eds.), *The processing and acquisition of reference* (pp. 239–272). Cambridge: MIT Press.
Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*(2), 274–307.
Guo, J., & Feng, G. (2013). How eye gaze feedback changes parent–child joint attention in shared storybook reading? In Y. I. Nakano, C. Conati, & T. Bader (Eds.), *Eye gaze in intelligent user interfaces* (pp. 9–21). London: Springer.
Haake, M., Hansson, K., Gulz, A., Schötz, S., & Sahlén, B. (2014). The slower the better? Does the speaker's speech rate influence children's performance on a language comprehension test? *International Journal of Speech-Language Pathology, 16*(2), 181–190.
Harris, M. (1992). *Language experience and early language development: From input to uptake*. Hove: Lawrence Erlbaum Associates.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.
He, A. X., Kon, M., & Arunachalam, S. (2016). When information hurts: Informativity and processing cost in verb learning (in preparation).
Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368.
Hoff, E., & Naigles, L. (2003). How children use input to acquire a lexicon. *Child Development, 73*, 418.
Hoff-Ginsberg, E. (1994). Influences of mother and child on maternal talkativeness. *Discourse Processes, 18*(1), 105–117.
Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition, 96*, 127–142.
Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*(1), 91–117.
Huang, Y. T., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental Psychology, 49*(6), 1090.
Hughes, M. E., & Allen, S. E. (2015). The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. *Lingua, 155*, 43–61.
Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*, F31.
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236.
Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to young children. *Developmental Psychology, 43*, 1062.
Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J. L., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*, 343–365.
Kon, M., Goksun, T., Bagci, A., & Arunachalam, S. (2016). Verb acquisition in English and Turkish: The role of processing. In J. Scott & D. Waughtal (Eds.), *BUCLD 40 online proceedings supplement*. <www.bu.edu/bucld/supplementvol40/>.
Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20*(1), 54–72.
Levelt, Willem J. M. (1989). *Speaking: From intention to articulation*. Cambridge/London: MIT Press.
Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198.
Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. (2009). *When one cue is better than two: Lexical vs. syntactic cues to verb learning*. Unpublished manuscript. <http://ling.umd.edu/labs/acquisition/papers/2009/LidzBungerLeddonBaierWaxman2009.pdf>.
Lidz, J., & Gagliardi, A. (2015). How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics, 1*(1), 333–353.
Mangold, R., & Pobel, R. (1988). Informativeness and instrumentality in referential communication. *Journal of Language and Social Psychology, 7*(3–4), 181–191.
Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: Information processing time with and without saccades. *Perceptual Psychophysics, 53*, 372–380.
Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology, 46*(4), 749.
McRoberts, G. W., & Best, C. T. (1997). Accommodation in mean f 0 during mother–infant and father–infant vocal interactions: A longitudinal case study. *Journal of Child Language, 24*(03), 719–736.
Nadig, A., Sedivy, J., Joshi, A., & Bortfeld, H. (2003). The development of discourse constraints on the interpretation of adjectives. In *Proceedings of the 27th annual Boston University conference on language development* (pp. 568–79).
Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*, 95.
Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. *Talking to Children: Language Input and Acquisition*, 109–149.
Ninio, A. (2004). Young children's difficulty with adjectives modifying nouns. *Journal of Child Language, 31*(02), 255–285.
Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition, 22*(2), 158–192.
Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76*, 763.
Prince, E. (1985). Fancy syntax and 'shared knowledge'. *Journal of Pragmatics, 9*(1), 65–81.
Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label order and their implications for symbolic learning. *Cognitive Science, 34*, 909–957.
Rohde, H., & Frank, M. C. (2014). Markers of topical discourse in child-directed speech. *Cognitive Science, 38*(8), 1634–1661.
Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*, 185.
Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*, 1762.
Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research, 32*(1), 3–23.

Sedivy, J. C. (2007). Implicature during real time conversation: A view from language processing research. *Philosophy Compass, 2*(3), 475–496.

Sekerina, I. A., & Trueswell, J. C. (2012). Interactive processing of contrastive expressions by Russian children. *First Language, 32*(1–2), 63–87.

Skarabela, B. (2006). *The role of social cognition in early syntax: The case of joint attention in argument realization*. Unpublished Doctoral Dissertation. Boston, MA: Boston University.

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science, 14*(1), 9–17.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, 549–565.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language, 4*(01), 1–22.

Sonnenschein, S. (1982). The effects of redundant communications on listeners: When more is less. *Child Development, 53*, 717–729.

Sonnenschein, S. (1984). The effects of redundant communications on listeners: Why different types may have different effects. *Journal of Psycholinguistic Research, 13*(2), 147–166.

Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Team R Development Core (2012). *A language and environment for statistical computing. Vienna: R-Foundation for Statistical Computing*.

Thorpe, K., Baumgartner, H., & Fernald, A. (2006). Children's developing ability to interpret adjective-noun combinations. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th annual Boston University conference in language development* (pp. 631–642). Somerville, MA: Cascadilla Press.

Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds "listen through" ambiguous adjectives in fluent speech. *Cognition, 100*, 389–433.

Tribushinina, E., & Mak, W. M. (2015). Three-year-olds can predict a noun based on an attributive adjective: Evidence from eye-tracking. *Journal of Child Language*, 1–17.

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1466.

Wardlow Lane, L., & Liersch, M. J. (2012). Can you keep a secret? Increasing speakers' motivation to keep information confidential yields poorer outcomes. *Language and Cognitive Processes, 27*(3), 462–473.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*(11), 2143–2152.

Weisleder, A., & Fernald, A. (2009). Real-time processing of postnominal adjectives by Latino children learning Spanish as a first language. In J. Chandlee, M. Franchini, S. Lord, & G. Rheiner (Eds.), *Proceedings of the 33th annual Boston University conference on language development* (pp. 611–621). Somerville, MA: Cascadilla Press.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy, 13*(3), 229–248.