Pragmatic inferences in context: learning to interpret contrastive prosody*

CHIGUSA KURUMADA

University of Rochester, USA

AND

EVE V. CLARK

Stanford University, USA

ABSTRACT

Can preschoolers make pragmatic inferences based on the intonation of an utterance? Previous work has found that young children appear to ignore intonational meanings and come to understand contrastive intonation contours only after age six. We show that four-year-olds succeed in interpreting an English utterance, such as "It LOOKS like a zebra", to derive a conversational implicature, namely [but it isn't one], as long as they can access a semantically stronger alternative, in this case "It's a zebra". We propose that children arrive at the implicature by comparing such contextually provided alternatives. Contextually leveraged inferences generalize across speakers and contexts, and thus drive the acquisition of intonational meanings. Our findings show that four-year-olds and adults are able to bootstrap their interpretation of the contrast-marking intonation by taking into account alternative utterances produced in the same context.

INTRODUCTION

Children's interpretation of contrast-marking prosody

Although many aspects of language have been mastered by the age of five or six, other aspects, such as pragmatic understanding of the meaning of intonational contours, have longer developmental trajectories. Many researchers have argued that young children do not rely on intonation in inferring the speaker's pragmatic intention, affect, or emotional state (e.g. Aguert, Laval, Le Bigot & Bernicot, 2010; Capelli, Nakagawa & Madden, 1990; Cruttenden, 1985; Cutler & Swinney, 1987; Hornby & Hass, 1970; Morton & Trehub, 2001; Quam & Swingley, 2012; Solan, 1980; Van Der Meulen, Janssen & Den Os, 1997; Wells, Peppe & Goulandris, 2004; Winner & Leekam, 1991). Moreover, the ability to interpret an intonation contour that signals contrast is said to develop late, after age six, and slowly (Ito, Bibyk, Wagner & Speer, 2012; Ito, Jincho, Minai, Yamane & Mazuka, 2012; Speer & Ito, 2009). For example, Cruttenden (1985) tested British English-speaking ten-year-olds and adults to determine whether they could reliably identify the referents of utterances such as those in (1a) and (1b), given a choice of three pictured referents [1]–[3].

(1) a. It's a very nice **garden**.
                          [fall–rise]
      b. It's a very nice **garden**.
                          [fall]

   Pictures

[1]   a nice garden but a house falling down
[2]   a garden and a house, both very nice
[3]   a garden overgrown but a very nice house

   (1a) has the fall–rise pitch accent on the word *garden*, which results in a slight rise at the final position of the sentence. This is typically associated with reservation (Cruttenden's [1985] term) or contrast in meaning. This reservation is expected to match the picture in [1], in which the nice garden is contrasted with the not-nice house. (1b), however, does not convey such a contrast and, hence, could describe [1] or [2]. Cruttenden found that 70% of the adult listeners chose picture [1] for (1a) but that only 20% of the ten-year-olds did so; 70% of them chose [2] instead. The majority in both groups consistently selected [2] when they heard (1b). These results and those of other studies strongly suggest that comprehension of the contrast-marking function of an intonation contour has yet to be acquired by ten-year-olds (see also Cutler & Swinney, 1987; Speer & Ito, 2009).

851

Children's difficulty in understanding intonation contours cannot be attributed solely to a limited ability to PERCEIVE pitch contours. Much younger children are acutely sensitive to prosodic information in the input (e.g. Fernald 1985; Morgan & Demuth, 1996; Sakkalou & Gattis, 2012). In addition, in a few limited domains, children under six years of age do show some understanding of pragmatic functions marked by intonation contour (e.g. Armstrong, 2014; Grassmann & Tomasello, 2007, 2010; Sekerina & Trueswell, 2012), and they PRODUCE distinct contours that depend on the pragmatic function of their utterance (Cutler & Swinney, 1987; Hornby & Hass; 1970; Ito, 2014; Thorson & Morgan, 2015; Wells et al., 2004; Wieman, 1976). At the same time, young children appear to be insensitive to intonational meanings in situations in which adult listeners derive specific pragmatic inferences effortlessly. This raises the question: What causes these young children difficulty?

To investigate the development of the ability to interpret intonational meanings, consider example (1) again. Adult-like interpretation of intonational meanings involves the following three steps (Pierrehumbert & Hirschberg, 1990). First, listeners need to identify a particular utterance as a variant of (1a) or (1b). Second, they need to remember that an intonation contour such as (1a) is typically associated with the speaker's highlighting a contextual contrast. Third, they reason about why the speaker used a particular intonation in context in arriving at a specific interpretation of the target utterance (e.g. [It's a very nice garden, but the house is not as nice]). In particular, this process requires listeners to derive a contextually plausible alternative expression (namely, what the speaker could have said, e.g. (1a) vs. (1b)) and make inferences about why the speaker chose a particular expression (Grice, 1975).

These three steps are not necessarily acquired in order, and their developmental trajectories are most likely intertwined. Nonetheless, the separation of intonation interpretation into these three steps is useful, as each one taps a different aspect of children's linguistic knowledge. The first step depends on representations of meaningful intonational contours in English, as well as cognitive and attentional resources, to process a sequence of speech as it unfolds (Ito et al., 2012; Speer & Ito, 2009). The second includes knowledge specific to English intonational contours, as languages differ in how much of the speaker's pragmatic intention is conveyed prosodically and how this is accomplished (see Büring & Gutiérrez-Bravo, 2001; Ladd, 2008; Jun, 2005, 2014; Vallduví, 1992). The third step requires the ability to make use of contextual information to derive a conversational implicature. Researchers have argued that preschoolers lag behind adults in using contextual information to constrain possible interpretations of an utterance (Snedeker & Trueswell, 2004). Indeed, children's intonational interpretations could differ from those of

852

adults in any one of these steps and so result in difficulty in arriving at the intonational meaning intended.

Two important questions here are: Which step constitutes an obstacle? and How do children achieve adult-like interpretations? To answer these questions, we first need to take a brief look at another domain of development with a similarly late, and protracted, development in children: scalar implicature. Preschoolers typically fail to derive scalar implicatures based on quantifiers (e.g. *none, some*, and *all*) or logical words such as *and* vs. *or* (e.g. Chierchia, Crain, Guasti, Gualmini & Meroni, 2006; Musolino, 2006; Musolino & Lidz, 2006; Noveck, 2001; Papafragou, 2006; Papafragou & Musolino, 2003; Papafragou & Tantalou, 2004). Their difficulty could be due to the level of their general processing capacity, language-specific lexical knowledge (including knowledge about relevant scales), or pragmatic ability in deriving the appropriate implicatures.

One method for sorting out these components is to provide extra scaffolding for one component and to determine whether this improves children's performance. Recent studies have shown that, with such an approach, the difficulty that children have with scalar implicature often lies in identifying a relevant scale. In fact, preschoolers can manage pragmatic inferences when they are shown visually represented ad-hoc scales (Stiller, Goodman & Frank, 2015), familiar scales such as numbers, or even explicit descriptions. So, when shown a picture of three animals sleeping, four-year-olds correctly reject *Only the cat and the cow are sleeping* (Barner, Brooks & Bale, 2011). The availability of contrasting prenominal adjectives also helps children to make contrastive inferences (Horowitz & Frank, 2012, 2014). Together, these results support the hypothesis that children's difficulty with scalar implicature lies in their limited ability to call to mind relevant alternatives in response to words such as *some*, rather than in their cognitive ability to compute contextually enriched pragmatic meanings.

If we apply the same logic to the interpretation of intonational contours, we can test whether preschoolers understand a contrastive intonation contour better, as in (1a), when there is a linguistic alternative available in the same context. In other words, the non-adult-like performance observed in earlier research may stem from children's difficulty in arriving at relevant alternatives based on an intonational cue alone. We define an alternative as an expression that is (a) lexically, syntactically, or prosodically different from a target utterance, and (b) is strongly linked to one of the possible meanings that the speaker could have expressed in that context. If a likely alternative expression is explicit in context, preschoolers may be able to bootstrap their inference by reasoning about why the speaker used one particular expression and not another. This type of reasoning follows from adherence to the principle of contrast (Clark, 1990), known for its facilitative roles in the early stages of children's word

853

learning. The current study was designed to test this prediction in the domain of intonational meanings.

We also test an extended version of this prediction. We asked: If the availability of alternatives facilitates children's contextual inferences, does successful comprehension contribute to the LEARNING of an intonational meaning? By learning, we mean acquisition of the ability to identify an intended meaning based on a particular prosodic contour, even in the absence of an explicit alternative. Contextually derived meanings may be short-lived or strongly tied to a particular setting, and thus unlikely to generalize across speakers and contexts (e.g. Carey, 1978; Goodman, McDonough & Brown, 1998; Horst & Samuelson, 2008; Mervis & Bertrand, 1994; Wilkinson & Mazzitelli, 2003). Alternatively, young children could learn to associate a particular intonation contour with some pragmatic meaning through contextually supported inferences.

In what follows, we present findings from three studies of four-year-olds. We show that these children succeed in the appropriate pragmatic interpretation of contrast-marking intonation contours when they can leverage their interpretations against alternatives in context. We also show, in a two-day-long study, that contextually leveraged inferences generalize across speakers and contexts, potentially contributing to longer-term acquisition of intonational meanings.

A case in point: "It looks like an X"

The current study focuses on the construction *It looks like an X* (e.g. *It looks like a zebra*). (We use italics for example words or sentences abstracted away from acoustic detail, e.g. *It is raining outside* can be said as "It's RAINING outside!", "It IS raining outside!", and other variations. We use double-quotation marks for quoted speech, with phonetic and prosodic specification; capital letters for prosodic emphasis; and square brackets [ ] for intended interpretations.) Depending on the listener's construal of the intonation contour, this construction can convey either an affirmative 'hint' (e.g. [(It looks like a zebra) and I think it is one]) or, with a different contour, a negative 'warning' (e.g. [(It looks like a zebra) but it actually isn't one]) (Kurumada, Brown & Tanenhaus, 2012, submitted; Kurumada, Brown, Bibyk, Pontillo & Tanenhaus, 2014). Both interpretations are attested in spontaneous conversations between adults and children.

Hansen and Markman (2005), in their analysis of language uses pertinent to the appearance/reality distinction, found examples of *look(s) like* in 478 recording sessions from eight children between 2;0 and 3;11 and their parents (in 226,629 turns). Of these, 56 instances (12%, virtually all from adults) were used to indicate the identity of the object (e.g. Child: *What's*

854

that? – Adult: *It looks like a bottle opener*); and 71 (15%, again, adult uses mainly) were used to refer to the similarity in appearance of the referents mentioned (e.g. Adult: *Yeah, that rock looks like a beehive*). By four years of age, then, children clearly have been exposed to instances of the *It looks like an X* construction, with an affirmative (It is an X) and comparative interpretations in context, usually offered in response to children's *What's that* questions. Nevertheless, Hansen and Markman's study did not examine the prosodic features of utterances, and thus does not provide any information on how often adults contradict the child's inappropriate reference by means of "It LOOKS like X (but it isn't one)."

Other researchers have documented adult speakers' uses of *looks like* as well as *is like* to convey one of two different speaker intentions marked by differences in prosodic contour (Clark & Wong, 2002). In (2) below, the adult speaker (Karen) uses *They look like cows to me* as a hint to categorize the referents as cows. In (3), however, Adam's mother uses *It's like* to give feedback on an erroneous word–object mapping, with an alert or warning. Clark and Wong suggested that the lengthened vowel (here transcribed as 'li::ke') emphasizes this function of the utterance by highlighting the reading [It is like a rope but it's actually *not* a rope].

(2)   Abe (2;7·26; Kuczaj 26:87, from Clark and Wong (2002),
       Karen is an adult)

ABE:	What's those?
KAREN:	What do they look like?
ABE:	I don't know.
KAREN:	They look like cows to me, don't they look like cows
	to you?

(3)   Adam (3;0.lo; Brown/Adam 20:1671, from Clark and Wong (2002)

ADAM:	What kind o(f) rope is dat?
MOTHER:	It's not a rope.
	It's li::ke a rope.
	It's a cord.

Figure 1 provides intonational contours that a native speaker of American English uses when asked to express these two meanings. In THE NOUN-FOCUS PROSODY (on the left), the sentence bears a canonical accent placement, with sentential stress on the last constituent (i.e. *X* in *It looks like an **X***), and the overall pitch contour ends with a low–falling boundary tone. We follow the ToBI convention, in which L and H to represent a low tone and a high tone respectively, and % indicates an utterance-final boundary tone (Beckman & Ayers, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005; Silverman *et al.*, 1992). An asterisk (*) means that the tone is aligned with a pitch
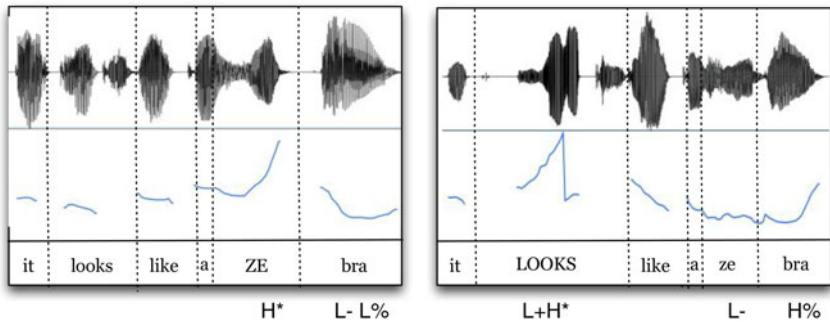
855

Fig. 1. Waveforms (top) and pitch contours (bottom) of the utterance "It looks like a zebra". The affirmative interpretation *It is a zebra* is typically conveyed by the contour on the left, and the negative one, *It is not a zebra,* by the contour on the right.

accent. The noun-focus prosody is therefore annotated as a contour with an H* on the final noun followed by a falling boundary tone L-L%. This pattern typically evokes the affirmative interpretation (i.e. [It looks like an X and it is one]) usually produced by adults to provide a hint as to the identity of the referent. In THE VERB-FOCUS PROSODY (on the right), the vowel in the verb *looks* is lengthened and emphasized with a contrastive accent $(L + H^*)$, and the sentence typically ends with a rising, L-H%, boundary tone. This pattern biases individuals' interpretation toward the negative interpretation (i.e. [It looks like an X but it is not one]). This provides a warning. Studies with adult speakers of English corroborate this distinction between hint and warning (Kurumada *et al.*, 2012, submitted; Kurumada *et al.*, 2014). Although adult judgments are not always categorical and can be modulated by contextual factors, such as the preceding utterance, an expected level of speaker expertise, and distributional information about the speaker's prosodic production, they reliably interpret these contours as [It is an X] (hint) and [It is not an X] (warning) (see Kurumada *et al.*, 2012).

With this construction, we ask whether four-year-olds can make a contrastive inference based on prosodic information. To call to mind that "it LOOKS like an X … " (verb-focus prosody) means [It is not an X], one needs to make inferences based on two prosodic representations: a fall–rise pitch accent $(L + H^*)$ and a rising boundary tone (L-H%). The L $+ H^*$ accent on the verb evokes a set of alternatives, in this case, any alternative predicates, including the semantically stronger *It is an X*. The utterance-final boundary tone signals that the propositional content is incomplete, which invites an additional inference from the listener (e.g. [… but it is actually something else]). This is similar to the inference

856

expected in (1a), "It's a very nice garden …" [but the house is not nice] (Cruttenden, 1985).

In Experiments 1–3, we test whether four-year-olds can, in fact, engage in such an inference if they have access to an alternative expression (e.g. "It is an X" vs. "It LOOKS like an X …"). In an everyday conversational context, as in (2) and (3) above, adult speakers tend to provide alternative expressions through rephrasing (e.g. "It's not a rope"; "It's like a rope"; "It's a cord"). We predict that young children will succeed in making a contrastive inference from "It LOOKS like an X" if they receive such information in context. We then ask whether such contextually supported inferences can be extended to understanding the same intonation contour when an alternative structure is not present. We end with a discussion of the role of such contextual inferences in intonational development.

EXPERIMENT 1

METHOD

Participants

We recruited and tested twelve children who were acquiring English as their first language (five girls, seven boys; mean age 4;2, age range 3;8–4;7) at a nursery school in Stanford, California. We also collected data from twenty-four adults on-line, using Amazon's Mechanical Turk service (https://www.mturk.com/mturk/). Adult participants were all self-identified as native speakers of American English who currently resided in the United States. Data from three adult participants were excluded because their participation time was two standard deviations below the mean. Adults were paid $1 to participate in this task.

Stimuli

First, we embedded sixteen high-frequency animal names in the sentence frame *It looks like an X* (e.g. *It looks like a zebra*). All the items were recorded twice by a female native speaker of American English: once with noun-focus prosody (e.g. "It looks like a ZEBRA!") and once with verb-focus prosody (e.g. "It LOOKS like a zebra …"). We created two experimental lists and counterbalanced the stimuli across these two lists: items pronounced with noun-focus prosody in List one had verb-focus prosody in List two, and vice versa. The mean duration (in milliseconds) and fundamental frequency of the words used (i.e. *it, looks, like, a, noun*) are summarized in Table 1 and Figure 2.

We chose sixteen animal pictures to visually represent the animal terms. We then chose sixteen more animal pictures to form pairs whereby the animals resembled each other perceptually (e.g. a zebra and an okapi;

857

TABLE 1. *Mean fundamental frequency values and mean word duration in Sally's utterances used in Experiments 1–3*

	Mean fundamental frequency (Hz) [standard deviation]				
	It	looks	like	a	noun
Noun-Focus	247·26 [6·67]	196·71 [7·12]	239·63 [9·4]	292·57 [3·58]	299·29 [11·17]
Verb-Focus	262·01 [12·56]	366·55 [22·32]	244·06 [19·6]	212·65 [12·03]	204·6 [2·3]

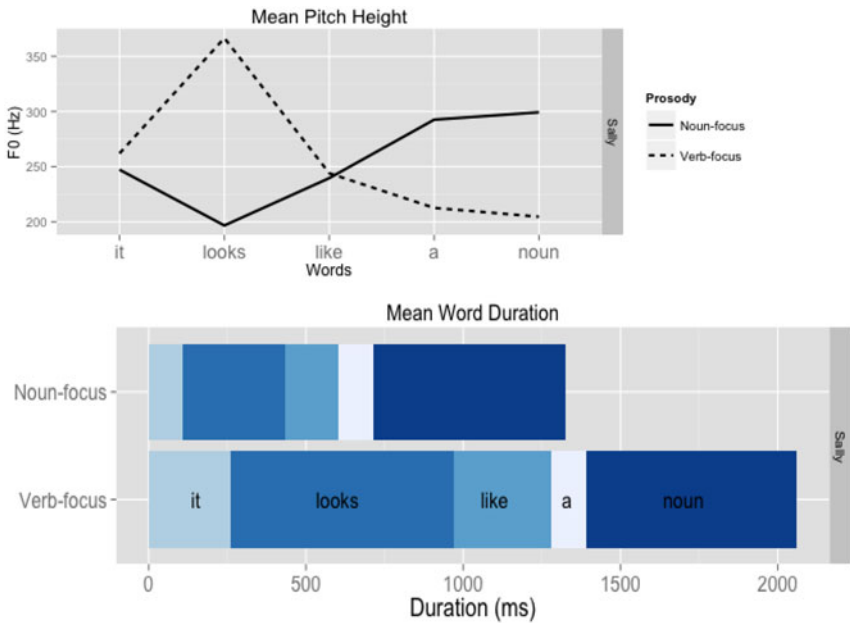|	Mean word duration (milliseconds) [standard deviation]				
	It	looks	like	a	noun
Noun-Focus	107·51 [18·05]	326·42 [12·06]	171·21 [18·17]	109·5 [40·23]	610·12 [93·86]
Verb-Focus	262·47 [47·18]	709·01 [83·72]	308·1 [50·73]	110·07 [51·17]	671·61 [119·78]

Fig. 2. Mean fundamental frequency (Hz) and mean word duration (in milliseconds) in the female puppet's (Sally) utterances.

Figure 3). In each pair, the target named in the input sentence (e.g. *It looks like a zebra*) was the more frequent of the two and was expected to be more familiar to the children who were being tested. Hereafter, the target named in a sentence (e.g. a zebra) is referred to as the animal MENTIONED, and the

858

paired animal (e.g. an okapi) is the UNMENTIONED animal. The animals in each pair served as target referents for one or the other of the two prosodic contours used in the task (e.g. a zebra as the target referent for "It looks like a ZEBRA" [and it is one] and an okapi for "It LOOKS like a zebra …" [but it's not one]).

Why did we use this asymmetry in the familiarity of the two items in each pair? For our study, we needed to ensure that, for each pair, children were more familiar with the animal mentioned than with the unmentioned animal. If the children were equally familiar with both animals, e.g. horse and zebra, no adult speaker would say, "It LOOKS like a zebra …" to refer to a horse; instead, the speaker would simply say, "It's a horse". In other words, for the verb-focus prosody to be used felicitously, one animal has to be unfamiliar and less likely to be referred to by name. To make sure that preschoolers registered this asymmetry, we did two surveys. The first one was an informal interview, for which we showed all the depictions of animals to eight children and had them name them one by one. Most could name all the mentioned animals (e.g. horse, zebra, elephant) or else recognized the names once the adult interviewer mentioned them (e.g. beetle, tadpole). At the same time, none of the children could correctly name all the unmentioned animals (e.g. okapi, tapir, bison). These children did not participate in the actual experiment. We also showed all of the stimuli to the teachers at the nursery school to confirm that the unmentioned animals were, indeed, less familiar to the children tested.

In the second survey, we checked that preschoolers could correctly associate a given mentioned name (e.g. zebra) to a target picture when presented with the two options of zebra and okapi. We did this to reject the possibility that children would accept both pictures (a zebra and an okapi) as equally good referents for the word used in an instruction (e.g. *It looks like a zebra*). We showed our picture pairs to five four-year-olds, who did not participate in the experiments, and asked them to point to the picture that matched best with the name of the animal. They answered correctly, e.g. choosing a zebra over an okapi almost all the time (96%). This suggests that four-year-olds indeed recognized one picture in each pair as a better referential fit for the noun mentioned. This licensed adult uses of "it LOOKS like an X …" in referring to the unfamiliar item in each pair.

The task

The experiment took place in a quiet room with a low table and child-sized chairs. The experimenter sat across from the child and placed experimental equipment (a puppet, a box, and a file binder displaying all the two choice options) on the table. Participants took part in a two-alternative

859

forced-choice task with an experimenter who was a native speaker of American English. The task comprised sixteen trials (two practice trials and fourteen critical trials) and lasted approximately 20 minutes.

The child participant was first introduced to the puppet named Sally. A mini portable speaker was attached to the puppet as a means to play the audio stimuli, and the experimenter who manipulated the puppet controlled the audio stimuli with a small mp3 player. Children first took part in a picture-naming task, for which they labeled eight animals, one by one (Picture-naming Task 1). These animals were later used as 'mentioned' animals in the immediately following block of the guessing game (Guessing-game Phase 1). Then the child and the puppet repeated the procedure with eight more critical items. Children were randomly assigned to one of two item lists with different presentation orders of the animal pairs.

In the guessing game, children were first presented with a box and told that it contained pictures of many different animals. In each trial, the child was shown two pictures – the target and a distractor (e.g. a zebra [mentioned] versus an okapi [unmentioned]), as shown in Figure 3. The pictures were presented in a red frame (left) and a blue frame (right), and the location of the mentioned and the unmentioned pictures was counterbalanced across items. The puppet was allowed to peek inside the box and to give the child a clue in the form of "It looks like an X" vs. "It LOOKS like an X …". The experimenter controlled the puppet's movements so that there was no other extralinguistic cue (such as the puppet's gesture or gaze direction) for the child. Following the puppet's utterance, the child was asked to point to the picture of whichever animal was hidden inside the box. When the child's point was ambiguous, the experimenter followed up and asked, "Is it the one in the red box or the blue box?" until the child provided a clear verbal or non-verbal response. After the child's response, the experimenter took a picture card from the box, showed it to the child, and said, "Oh, it was this one!" This served as feedback about which animal the puppet-speaker had intended to identify.

All sessions were recorded on a camcorder for later review and coding. Two coders annotated the children's responses. All of the responses were made either non-verbally or verbally (i.e. choosing the red or the blue box). When a child changed his or her mind, we always noted that the one that he or she chose last was the chosen animal. There was no disagreement between coders on coding decisions.

To equate children's initial experience with the task, we gave all of them an identical set of example items. The first practice trial used pictures of a horse and a donkey as choice options, and the children heard an utterance with noun-focus prosody ("It looks like a horse"). The second practice trial used pictures of a butterfly and a moth, and they heard an utterance with
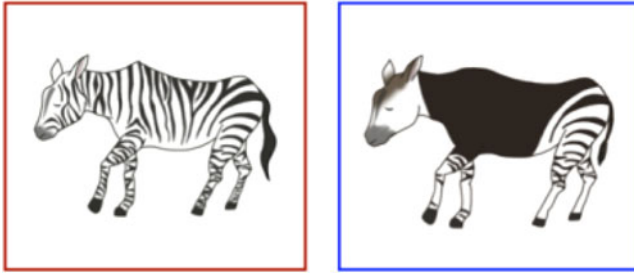
860

Fig. 3. Examples of two choice options: a mentioned animal (zebra) on the left and the unmentioned animal (okapi) on the right.

verb-focus prosody ("It LOOKS like a butterfly ..."). Children's responses for these practice trials were analyzed separately.

To collect data from twenty-four adult speakers, we conducted the same experiment as an Internet-based survey through Amazon Mechanical Turk. We used the web-based platform for the adult data as a means to collect the data efficiently and economically. To ensure that they had access to audio clips, prior to the experiment, we asked all of the participants to play a demo sound clip and to type in a word that they heard. We presented the task as "a language experiment targeted for preschoolers", and we described, with written texts and pictures, how the guessing game was conducted with children. The participants were instructed to listen to each sentence only once and to answer two-alternative forced-choice questions by clicking on a picture of the intended referent.

As in the live experiment with child participants, all of the adult participants were exposed to the same two example items (i.e. "It looks like a HORSE" and "It LOOKS like a butterfly ... ") and then received the fourteen critical items in a pseudo-randomized order. They received feedback after each trial. The adult task was set up so that sound clips were auto-played at the start of each trial, and participants could replay the sound as many times as they needed. Only two participants used this replay function, twice each. They could not proceed to the next trial without choosing at least one item. As in the child version of the task, they could change their mind and reselect an animal in each trial. They could, however, not go back to previous trials once they had proceeded to the next trial.

We decided a priori to remove datapoints from any participant who showed no variability in responses (i.e. always choosing a mentioned or an unmentioned animal). No participants, however, followed that strategy in the current dataset.

861

RESULTS AND DISCUSSION

The children's and adults' responses are presented in Figure 4. As shown in Panels (a) and (c), children and adults provided similar responses to the first two practice trials: in response to the noun-focus contour (i.e. "It looks like a HORSE"), the animals mentioned were chosen about 70% of the time. In response to the verb-focus prosody (i.e. "It LOOKS like a butterfly …"), however, they chose the two pictures (i.e. a butterfly and a moth) at an approximately equal rate. In addition, adults and children differed significantly in how they incorporated the feedback given after these first two trials. As can be seen in Panel (d), adults quickly associated the prosodic contours with the intended interpretations and proceeded to provide near-categorical responses for the fourteen critical trials. Children's responses, however, remained at chance level for both prosodic contours. On average, they chose a mentioned animal (a zebra when the input was *It looks like a ZEBRA*) 46% of the time for the noun-focus prosody and 51% of the time for the verb-focus prosody.

Do the current results simply reflect the fact that adults are better than children at learning an abstract form–meaning pairing quickly? The strongest version of this view would predict that adults should quickly learn any prosody–meaning mapping. Experimental evidence shows, however, that adults are NOT willing to learn a new pairing if the relation between the contours and the meanings are unexpected or unpredictable compared to what they are familiar with. Heeren, Bibyk, Gunlogson, and Tanenhaus (2015), for instance, demonstrated that adult listeners could not easily learn to remap rising/falling utterance-final boundary tones and assertion vs. question interpretations, even with constant feedback. This suggests that, while cognitive flexibility and better understanding of the task probably facilitated the form–meaning mappings in adults, the results must also reflect existing prosodic knowledge.

The adults' response patterns were compared with those in a previously published study that used a very similar set of visual and audio stimuli. Kurumada *et al*. (2014) conducted a laboratory-based eye-tracking experiment using the current construction "it looks like an X!" and "it LOOKS like an X …" while, unlike the current study, providing no feedback after each trial. They found that adult participants selected the mentioned animal on 66% of critical trials with noun-focus prosody, but only 26% of trials with verb-focus prosody. Adult participants in the current experiment selected the mentioned animal on 74% of the two practice trials and on 98% of the fourteen critical trials with noun-focus prosody, and on 52% of the practice trials and 7% of the critical trials with verb-focus prosody (Figure 4). Taken together, these results support the view that constant feedback helps adult listeners derive clearly
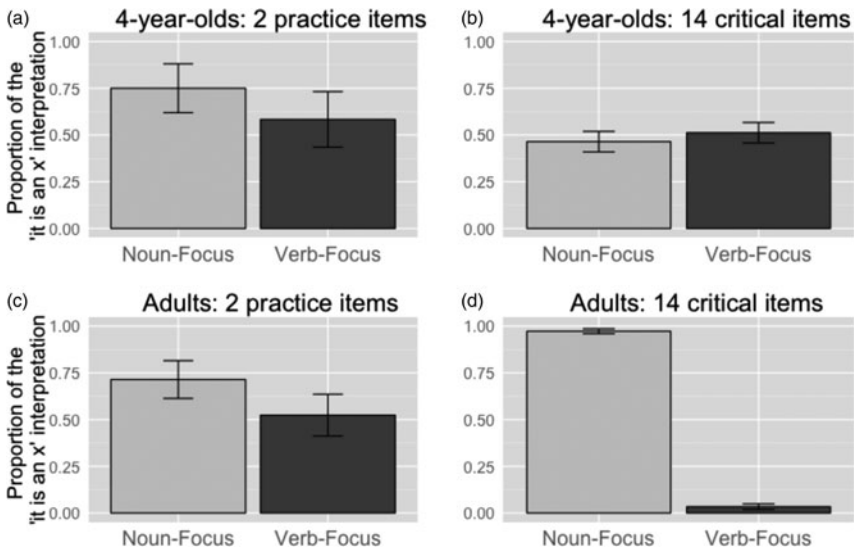
862

Fig. 4. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in the responses from four-year-olds (top) and adults (bottom). Left-hand panels summarize responses to the two practice trials, and right-hand panels summarize responses to the fourteen critical items. Error bars represent the standard error of the mean.

distinguished interpretations based on noun-focus versus verb-focus intonation contours. Children, however, do not seem to use the feedback as effectively as adults do.

We conducted a mixed-effects regression analysis with the response data for critical trials produced by the four-year-olds (Gelman & Hill, 2006). The model included the input prosody (i.e. noun-focus vs. verb-focus), item order, the children's gender and age in months as fixed effects, and both children and items as random effects. The main effect of the input prosody was not significant ($p > \cdot 4$), and there were no significant effects of gender or age for the children. The model also suggested that there was no effect of order for trials. Thus, despite the feedback provided after each trial, the children did not learn the contour contrast within the task itself.

Next, to take a closer look at individual children, we plotted response patterns for the fourteen critical items for each child (Figure 5). There was a lot of variability in their responses, and almost all of their responses deviated from the pattern observed in adults. Some children, such as E1_F, E1_G, E1_H, and E1_J (where E1 stands for Experiment 1), categorically chose either the animal mentioned (e.g. zebra) or unmentioned (e.g. okapi), regardless of the prosodic input. Other children showed less variability, but their responses did not distinguish significantly
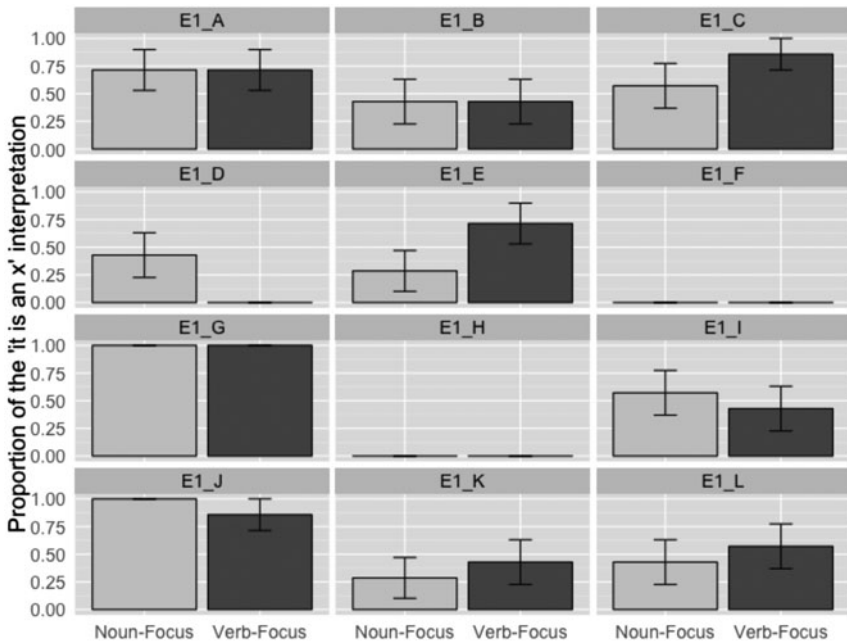
863

Fig. 5. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in the responses from the twelve individual children. Error bar represents the standard error of the mean.

between noun-focus and verb-focus prosody. These results lend support to earlier observations that four-year-olds are not sensitive to contrastive prosodic contours in the absence of other contextual information.

The results of Experiment 1 suggest that adults and four-year-olds differ in how they integrate feedback and determine the relevant associations between the prosodic input (i.e. noun-focus vs. verb-focus) and the interpretation intended (i.e. [It is an X] vs. [It is not an X]). In the next experiment, we sought to determine whether the presence of an alternative expression would facilitate four-year-olds' identification of these associations. The contrastive inference based on verb-focus prosody requires reasoning that the verb phrase "it LOOKS like an X" should be chosen over other expressions, such as "It is an X", when the target is NOT an X. Based on earlier findings (Barner *et al.*, 2012; Horowitz & Frank, 2012, 2014), we expected that explicit mention of an alternative would help children derive the contrastive inference signaled by the prosody. If this prediction is supported, we can argue that four-year-olds have the basic ability to attend to intonational contours and compute the pragmatic interpretation that is intended by the speaker. Children's insensitivity to

864

intonational meanings, attested in Experiment 1, can then be attributed to a still underdeveloped ability to identify linguistic alternatives that could constrain their interpretation of a particular intonation contour. When the alternatives are made explicit, children may be able to make use of the intonation contour to derive a contrastive interpretation when given verb-focus prosody.

EXPERIMENT 2

METHOD

Participants

We recruited and tested twenty-four children who were acquiring English as their first language (fourteen girls, ten boys; mean age 4;6, age range 4:2–4;8) at the same nursery school as used in Experiment 1. They were randomly assigned to one of two conditions (Forms-only and Combined) described below. We also collected data from forty adults on-line, using Amazon Mechanical Turk. Data from three adult participants were excluded based on the same criterion as in Experiment 1.

Stimuli

The visual stimuli and experimental settings were identical to those in Experiment 1. The audio stimuli were changed as follows: in the CONSTRUCTION-ONLY CONDITION, the puppet said "It's an X" when the target animal was indeed an X (e.g. "It is a ZEBRA" when the target picture depicted a zebra), and "It looks like an **X**," again, with a focus on the final noun, as a warning, when the picture was not an X (e.g. "It looks like a ZEBRA" when the target picture depicted an okapi).

In the COMBINED CONDITION, the puppet used "It's an X" with the noun-focus prosody (i.e. a nuclear accent on X with a rising boundary tone), to convey the [It is an X] interpretation, and verb-focus prosody (e.g. "It LOOKS like an X …") for the [It is not an X] interpretation. The pattern of clues to meaning by condition is summarized in Table 2. We used a between-subjects design, with each child participating in just one condition.

The Combined condition was designed to determine whether children could associate verb-focus prosody and the meaning of [but it is not one]. The association should be more accessible to children in the Combined condition than in Experiment 1 because the linguistic signals are distinguished by both prosody and syntactic construction. The statement type *It is an X* is unambiguously associated with the intention to assert that [It is an X]. Casillas and Amaral (2011) showed that four-year-olds could infer core vs. peripheral category membership of objects based on a

865

TABLE 2. *The between-subject manipulation of Experiments 1 and 2. White cells identify sentence patterns used for identifying a target as the animal mentioned. Shaded cells identify sentence patterns used for identifying the hidden animal as* not *being the mentioned animal.*

	Experiment 1 (Prosody-only)	Construction-only	Combined
It's an X		It's a ZEBRA	It's a ZEBRA
Noun-focus	It looks like a ZEBRA	It looks like a ZEBRA	
Verb-focus	It LOOKS like a zebra …		It LOOKS like a zebra …

contrast in linguistic construction (e.g. "it's a butterfly" vs. "It's sort of a butterfly"). If the contrast in construction (i.e. "It's a ZEBRA" vs. "It looks like a ZEBRA") is sufficient for a contrastive inference, and children show no awareness with contrastive prosody, the results in the Construction-only and Combined conditions should be identical. In other words, a significant difference between the results of the Construction-only and the Combined conditions would support the view that the children can make use of prosodic information on top of syntactic information to make the relevant pragmatic inferences.

RESULTS AND DISCUSSION

Figure 6 provides a summary of children's responses to the first two practice trials. Across all conditions, children first received the input type associated with the [It is an X] interpretation (i.e. "It's an X" in both the Form-only and Combined conditions). They then received the input type associated with the [It is not an X] interpretation (i.e. "It looks like an X" in the Construction-Only condition and "It LOOKS like an X …" in the Combined condition).

All responses in the fourteen trials from children and adults are summarized in Figure 7. In Experiment 1 (henceforth, the Prosody-only condition), children's responses to verb-focus did not differ from chance. With the same acoustic input as in Experiment 1, however, children in the Combined condition successfully interpreted verb-focus prosody ("It LOOKS like an X …"), treating it as [It is not an X] 89% of the time. This strongly suggests that the presence of the syntactic form "It's an X"
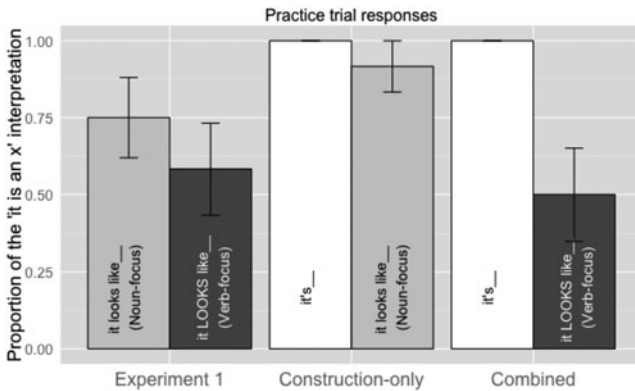
866

Fig. 6. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in the two practice trials, given first, in Experiments 1 and 2.

facilitated children's integration of the feedback that they received during the first two practice trials. In the Construction-only condition, children also showed sensitivity to the contrast in syntactic form between the two sentence types. They chose the mentioned animal 79% of the time when they heard, "It's an X", but only 46% of the time when they heard, "It looks like an X", also with noun-focus prosody. That their response to the latter was at chance suggests that the structural contrast alone was not sufficient for children to associate "it looks like an X" with the interpretation [It is not an X]. Notice, however, that the adults in the Construction-only condition also chose a mentioned animal about 35% of the time when they heard "it looks like an X". This may reflect an overall bias for the noun-focus prosody to be associated with an [It is an X] interpretation (Kurumada *et al.*, 2012; Kurumada, Brown & Tanenhaus, unpublished observations). Alternatively, this result could arise from the conflicting cues of contrasting predicates ("is" vs. "looks like"), along with prosodic focus always on "X" (e.g. zebra).

These results suggest that the presence of the alternative expression "It is an X" facilitated children's integration of contextual feedback, which led, in turn, to adult-like performance on the fourteen critical items in the Combined condition. The difference between the Construction-only and Combined conditions excludes the possibility that children's improved performance was simply due to the formal contrast between "It is an X" and "It looks like an X". Four-year-olds are able to make the intended inference from the verb-focus prosody when the input expression is explicitly contrasted with a semantically stronger alternative expression.

Do such contextually supported inferences have an effect on the long-term learning of prosodic interpretations? Carey and Bartlett (1978), in the
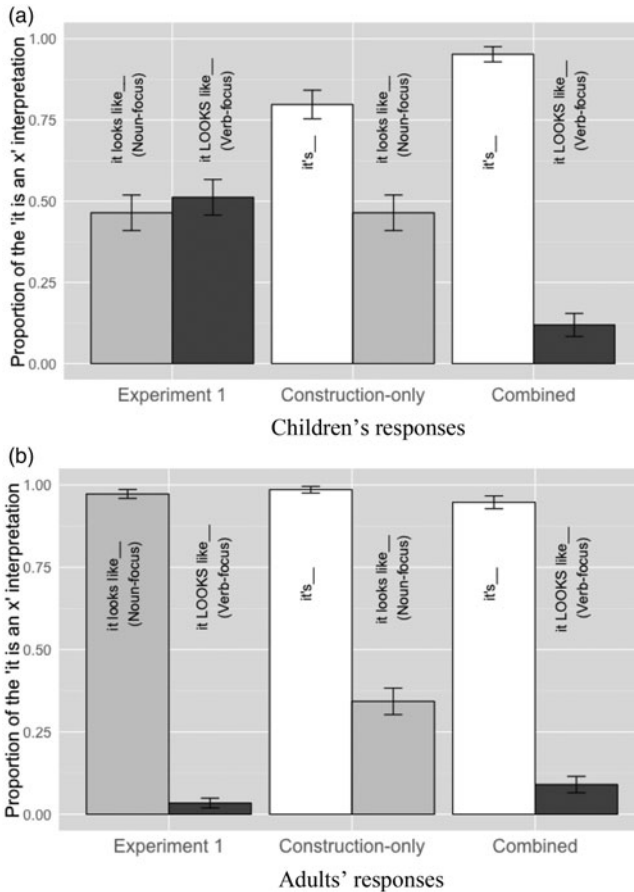
867

Fig. 7. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in Experiments 1 and 2: (a) in children's responses for target items, (b) in the adult control study. Error bars represent the standard error of the mean.

domain of word learning, proposed that the process of learning a new word could be separated into two phases. The first phase is FAST-MAPPING, by which the child associates a newly encountered word with a contextually constrained word meaning. The second phase is EXTENDED-MAPPING, by which a newly learned referent–label mapping is maintained over several successive uses. In this second phase, some meaning for the word is abstracted from properties of particular referents and becomes stored as part of a new lexical entry. In prosodic learning as well, children need to extract properties of prosodic contours that can be generalized across speakers and contexts.

868

In Experiment 3, we conducted a two-day-long study in which we exposed children to the Combined condition and then to a Prosody-only condition. If the children have learned the mapping between verb-focus prosody and the [It is not an X] interpretation, then this should improve their performance in the Prosody-only condition.

EXPERIMENT 3

The goal of Experiment 3 was to determine whether experience with the contextual inferences in the Combined condition would lead to a more adult-like interpretation of the verb-focus prosody in the absence of any formally contrasting alternative. We conducted a two-day experiment in which children were first tested on the Combined condition with contrasting syntax and prosody and then on the Prosody-only condition (as in Experiment 1) with the two sessions run one day apart. We addressed two questions here: (a) Does the contextually supported interpretation of contrastive intonation contour generalize to a situation where there is no such support? and (b) If so, does the effect also generalize to a new speaker?

METHOD

Participants

We recruited and tested twenty-four children who were acquiring English as their first language (ten girls, fourteen boys; mean age 4;5 age range 4;1–4;11) at a nursery school in Rochester, New York, and in the Baby Lab at the University of Rochester.

Stimuli

The animal pictures used in this experiment were identical to those in Experiments 1 and 2. In this experiment, we introduced a male puppet (Dave) in addition to the female puppet (Sally) from Experiments 1 and 2. A male speaker of American English recorded the stimuli for the Dave puppet. Mean pitch values and segment durations of Dave's utterances are provided in Table 3 and Figure 8. We created two between-subject conditions. In the same-speaker condition, children heard Sally on both days. In the different-speaker condition, children heard Dave in the Combined condition on Day 1 and Sally in the Prosody-only condition on Day 2 (Table 4). The audio stimuli for Sally's speech were identical to those in the Prosody-only and Combined conditions from Experiments 1 and 2.

The goal of this experiment was to determine whether the advantage observed in the Combined condition could be maintained and transferred

869

TABLE 3. *Mean fundamental frequency values and mean word duration in Dave's (male puppet) utterances used in the different speaker condition in Experiments 3*

	Mean fundamental frequency (Hz) [standard deviation]				
	It	looks	like	a	noun
Noun-Focus	148·14 [108·14]	124·96 [8·6]	122·47 [7·3]	151·94 [8·4]	214·83 [40·47]
Verb-Focus	182·92 [117·33]	184·73 [19·06]	119·26 [4·14]	112·14 [4·5]	125·57 [28·08]

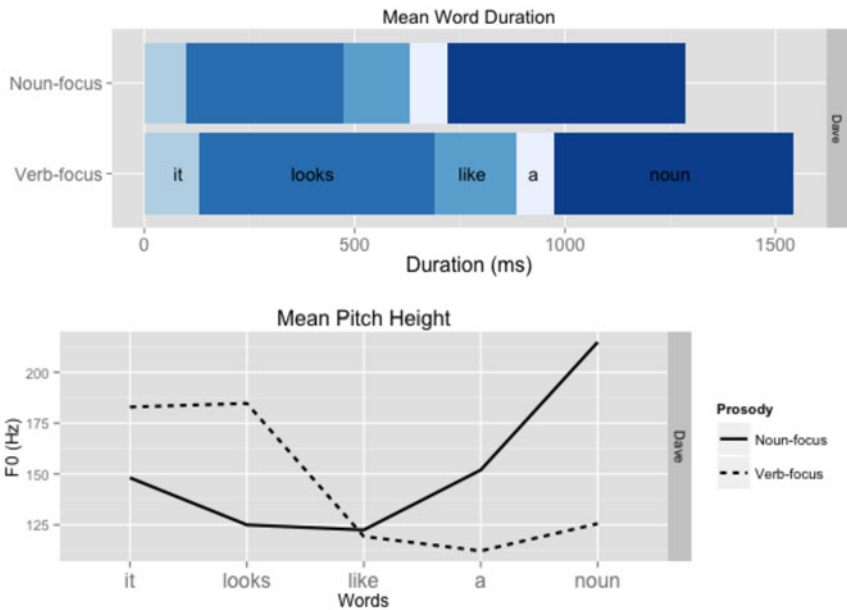|	Mean word duration (milliseconds) [standard deviation]				
	It	looks	like	a	noun
Noun-Focus	99·32 [15·87]	374·47 [32·93]	157·29 [9·19]	88·90 [38·31]	565·83 [137·65]
Verb-Focus	131·49 [24·37]	559·28 [18·08]	193·56 [18·08]	90·2 [31·1]	569·45 [96·68]

Fig. 8. Mean fundamental frequency (Hz) and mean word duration (in milliseconds) in the male puppet's (Dave) utterances.

to the Prosody-only condition, which had yielded a null result in Experiment 1. We introduced the different-speaker condition in this experiment for two reasons. First, we wanted to determine whether the contextually supported prosodic interpretation generalized to a new speaker. Using a speaker of

870

TABLE 4. *Items used in Experiment 3. White cells identify sentence patterns used for identifying a target as the animal mentioned. Shaded cells identify sentence patterns used for identifying the hidden animal as* not *being the mentioned animal.*

Same speaker condition		Different speaker condition	
Day 1 (Combined)	**Day 2** (Prosody-only)	**Day 1** (Combined)	**Day 2** (Prosody-only)
It's a ZEBRA		It's a ZEBRA	
	It looks like a ZEBRA		It looks like a ZEBRA
It LOOKS like a zebra …	It LOOKS like a zebra …	It LOOKS like a zebra …	It LOOKS like a zebra …

the opposite sex with a different voice quality offers one test of this. As can be seen in Tables 1 and 3 and Figures 2 and 8, the male and the female speakers' utterances had different prosodic profiles, which made it possible for us to test whether children were simply tracking the mapping between acoustic patterns of speech and different interpretations or whether they were learning more abstract templates of prosodic contours applicable across speakers and speech conditions.

Second, we introduced the condition to tease apart two possible mechanisms of intonational learning. One possibility is that children simply memorize the observed association between the verb-focus prosody "It LOOKS like an X" and the feedback [It is not an X] on Day 1. They then use it to distinguish the noun-focus and the verb-focus prosody in the Prosody-only condition on Day 2. If memorization were used, children in the same- and the different-speaker conditions should behave similarly on Day 2. Indeed, the benefit of the Combined condition might be greater in the same-speaker condition, as the children would hear exactly the same voice across two days.

Alternatively, we would see a different result if children's learning involves additional assumptions beyond memorized mappings between prosodic contours and their interpretations. If children are learning the mapping of

871

which linguistic expression the speaker could use to convey a particular meaning, interacting with the same speaker over two days might, in fact, cause confusion. Children might find it odd for the speaker to use "It's an X" to convey the [It is an X] interpretation on Day 1 and then switch to "It looks like an X" to express the same interpretation on Day 2. Previous studies of children's sensitivity to referential precedents show that preschoolers expect a speaker to adhere to referential pacts. They are surprised when a given speaker switches referential expressions without any clear contextual justification. Nevertheless, they accepted a new expression for a previously mentioned object when a novel speaker produced it (Graham, Sedivy & Khu, 2013; Matthews, Lieven & Tomasello, 2010). This led us to predict that, if children learn to process intonational meanings through reasoning along the lines of "What would the speaker say if she meant X?", those in the different-speaker condition should perform better on Day 2. The encounter with a new speaker justifies the introduction of a new expression, "It looks like an X" (noun-focus prosody), for the interpretation that was previously expressed by "It's an X."

We altered the order of the items from Day 1 to Day 2. In addition, the input prosody was flipped for two-thirds of the items to exclude the possibility that children could simply answer questions by memorizing the target animals from Day 1.

The task

The task was identical to those in Experiments 1 and 2.

RESULTS AND DISCUSSION

The children's responses in the example trials were similar to what we observed in Experiment 1. Only responses from the fourteen critical trials were included in the following analysis. Children in both conditions reliably distinguished the two prosodic contours on Day 1, and thus replicated the results of the Combined condition in Experiment 2 (see Figure 9). This replication highlights the robustness of the effect of the Combined condition because Experiment 3 was conducted in a different geographical location (Experiment 2: Stanford, California; Experiment 3: Rochester, New York) and included a new speaker as well as the original speaker.

On Day 2, in the Prosody-only condition, children differed in their performance, depending on whether they heard the same speaker as on Day 1 or a different speaker from that of Day 1 (see Figure 9). With the same speaker (right-hand-side of Panel a, Figure 9), children's responses for both the noun-focus and verb-focus versions were equally biased toward the "It is not an X" interpretation. With a different speaker
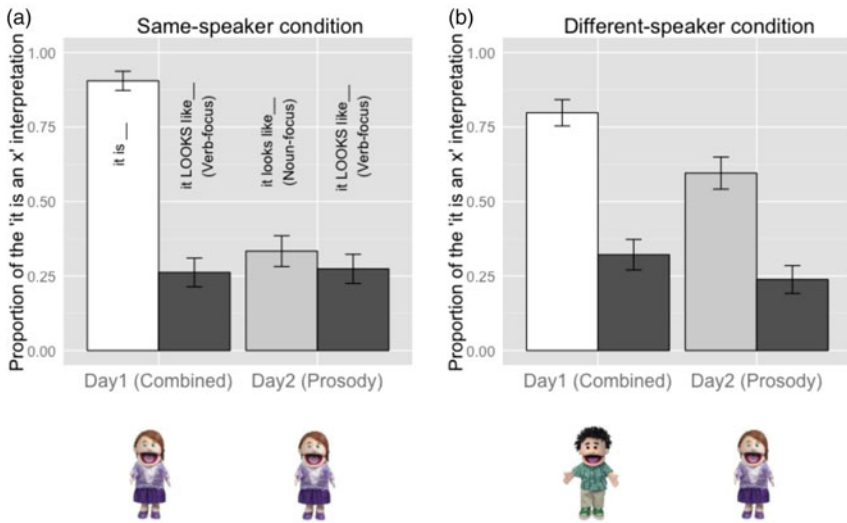
872

Fig. 9. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in Experiment 3. Error bars represent the standard error of the mean.

(right-hand-side of Panel b, Figure 9), however, children successfully assigned different responses to the noun-focus and verb-focus versions. This difference, though, was a little smaller than the difference the same children assigned on Day 1 (compare the left- and right-hand-sides of Panel b, Figure 9).

We constructed a mixed model to fit the data collected on Day 2. The first model included the prosodic input (i.e. noun-focus vs. verb-focus), the condition (e.g. same- or different-speaker), item order, children's gender and age as fixed effects, and children and items as random effects. Item order, gender, and age were dropped in subsequent models due to lack of significance in model comparisons. The final model revealed significant effects of prosodic input ($\beta = 0.89$, $p < .01$), condition ($\beta = 1.1$, $p < .04$), and a marginal interaction between them ($\beta = 0.92$, $p < .06$). The interaction term suggests that children derived different interpretations, depending on the two kinds of the prosodic input, and did so more in the different speaker condition. Children in the same-speaker condition, in contrast, were likely to interpret both types of input, "It looks like an X" and "It LOOKS like an X", as [It is an X].

Thus, in the different-speaker condition, prior exposure to the Combined condition supported a successful distinction between the two prosodic contours in the Prosody-only condition. The size of the distinction, however, varied across individuals. Figure 10 provides a summary of the responses from each child in Experiment 3. Four of the twelve children (E3_A, B, C, D)
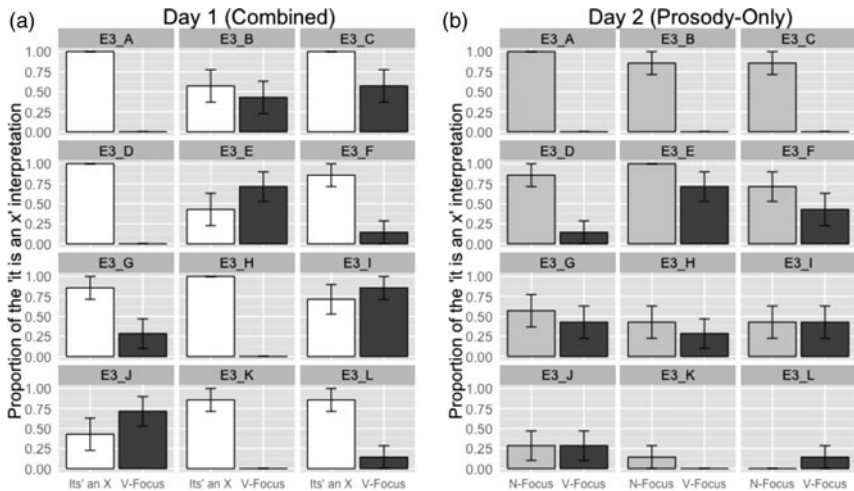
873

Fig. 10. Proportion of the 'it is an X' interpretation (choice of a mentioned animal) in the responses from the twelve individual children in the different-speaker condition. Error bars represent the standard error of the mean.

showed categorical or near-categorical judgment patterns on Day 2. Interestingly, two of the four (E3_B and E3_C) showed significant improvement from Day 1 to Day 2 in their interpretation of verb-focus prosody. Responses from four children (E3_E, F, G, H) showed the expected numerical trend but contained much more variability. Responses from the remainder of the children (E3_I, J, K, L) were biased toward the [It is not an X] interpretation, just as were those in the same-speaker condition. In short, there appear to be significant individual differences in how exposure on Day 1 was generalized to Day 2. Further, some children showed similar response patterns to those we found in the same-speaker condition.

GENERAL DISCUSSION

Studies of children's acquisition of intonational meanings have long posed a puzzle in that preschoolers appear unable to interpret contrastive prosody despite their general sensitivity to prosodic features of speech. We hypothesized that young children's difficulty may stem in part from their weaker expectations about possible alternative expressions with which to contrast whatever the speaker has explicitly mentioned. We examined four-year-olds' interpretations of the construction *It looks like an X* produced with two different intonation contours: one with canonical accent placement that indicated the speaker's intention to provide a hint (e.g. "It looks like a ZEBRA" [and I think it is one]; noun-focus prosody)

874

and the other with a contrastive pitch accent on the verb that provides a warning (e.g. "It LOOKS like a zebra …" [but it isn't one]; verb-focus prosody). Our main objective was to determine whether the presence of a semantically strong alternative, namely "It is an X", would facilitate contrastive inferences.

The three experiments reported here yielded support for the facilitative role of contextual alternatives. In the absence of "It's an X" (Experiment 1), four-year-olds' performance in the current task replicated previous findings in that they appeared insensitive to contrast-marking intonation contour. Despite their receiving feedback in each trial, some children consistently chose either the mentioned or the unmentioned animals, regardless of the prosodic input. Others gave more variable responses, but none consistently showed the adult-like responses of associating "It looks like an X" with [It's an X] and "It LOOKS like an X" with [It is not an X].

In Experiment 2, we demonstrated that four-year-olds could successfully associate verb-focus contour with the [It is not an X] interpretation when it contrasted with the same speaker's use of the form "It is an X" (that is, in the Combined condition). In the Construction-only condition, the children responded differentially to the forms of "It's an X" and "It looks like an X" (where the latter also had noun-focus prosody), but the difference was smaller than in the Combined condition. The difference between these two conditions suggests that they do, in fact, treat the noun-focus and verb-focus prosody differently. Taken together, the results support our proposal that children are able to derive contrastive inferences based on verb-focus prosody when the speaker's intended contrast is made explicit with alternative expressions.

These results provide insight into how young children may discover intonational meanings in the input. As is seen in word-learning (e.g. Carey & Bartlett, 1978; Clark, 1990; Markman & Wachtel, 1988), scalar implicature (Barner et al., 2011; Stiller et al., 2015), and contrasting adjectives (Horowitz & Frank, 2012, 2014), children appear to be able to leverage their knowledge about an unfamiliar intonation contour with contextually provided contrast. That is, they interpret intonational contours conditionally, depending on alternative expressions used by the same speaker to convey interpretations warranted by the context. Such contextual bootstrapping may allow young learners to associate a particular intonation contour and a contextually supported interpretation. Further, the results from the different-speaker condition support the idea that the contextually facilitated intonation interpretations can be retained and generalized across speakers, leading to LEARNING of intonational meanings.

It is also important to note that a presence of an alternative expression may have contributed to better comprehension of alternative intonational meanings by aiding memory for intonational contours. In the Combined

875

condition, verb-focus prosody ("It LOOKS like an X …") differed from its contextual alternative ("It's an X") in two ways: syntactic structure and prosodic contour. This may have made it easier for children to notice that the speaker was, in fact, using two distinct linguistic structures, helping them to integrate the feedback from the experimenter and to better remember the mapping between the structures and the meanings that they conveyed. In the Prosody-only condition (i.e. "It looks like an X" vs. "It LOOKS like an X …") and the Construction-only condition (i.e. "It is an X" vs. "It looks like an X"), in contrast, the contrasts in the linguistic information that the children heard were more subtle, possibly making it more difficult for them to remember how the two linguistic forms mapped onto the speaker's intentions.

We cannot, however, straightforwardly attribute the results of Experiment 3 to improvements in memory alone. In the same-speaker condition, memorized representations of the verb-focus prosody "It LOOKS like an X" could have helped children better distinguish it from noun-focus prosody. Instead, what we found was a strong tendency in children in this condition to generalize the knowledge along the dimension of the construction. Their responses were strongly biased toward [It is not an X] regardless of the prosodic input. As seen in Figure 10, there were a few children in the different-speaker condition who showed the same preference. In both cases, this was likely due to perseveration on the construction–interpretation mapping heard on Day 1. There, the structural contrast (*It's an X* vs. *It looks like an X*) distinguished the two interpretations. If they carried this assumption over from Day 1, they should interpret both noun-focus and verb-focus prosody on Day 2 as [It is not an X] because they both have the construction *It looks like an X*.

The different degree of perseveration between the two conditions, we believe, supports the idea that children's intonational interpretations occur as part of a more general pragmatic inferential process similar to their referential comprehension. They construct a speaker-specific expectation – a pact – as to how a particular interpretation has been encoded (Graham *et al.*, 2013; Matthews *et al.*, 2010). They then use any formal (lexical, prosodic, and syntactic) deviation from an expected expression as a signal of differences in interpretations. In this process, they appear to rely more on lexically or structurally encoded contrast than on intonational contrast, presumably because structural contrasts are registered categorically, and hence more reliably mapped onto distinct interpretations, compared to highly variable prosodic information. In the different-speaker condition, most of the children seemed to successfully block generalization of the pact with the previous speaker, picking up on the differences between the noun-focus and the verb-focus prosody produced by the new speaker as meaningful contrasts.

876

Our results thus highlight the importance of assessing children's prosodic interpretations within discourse, in which they can make use of all available alternative form–meaning mappings produced in the same context. Previous studies have examined how children interpret 'intonational' minimal pairs (such as (1a) and (1b)) to determine whether they could derive different meanings. Such an approach, however, tends to overlook other types of alternative expressions that could help bootstrap children's understanding of a target utterance. The current results suggest that both children and adults constantly rely on contextual alternatives and feedback to fine-tune their expectations about the semantic, syntactic, and prosodic devices that the speaker uses to express a particular meaning. In turn, those contextual inferences, over time, guide their learning of intonational meanings.

We leave for future studies the question of whether exposure to the Combined condition could facilitate children's interpretation of a construction other than *it LOOKS like an X*. For this, one could substitute a different verb (e.g. "it SOUNDS like a fire engine …") or use unrelated sentences produced with the same intonation contour (e.g. "It's a very nice GARDEN …" as in Cruttenden (1985), or "She HAD a bell … [but she no longer has one]" as in Dennison (2010) and Dennison and Schafer (2010)). Indeed, future research should consider whether form-based inferences also facilitate comprehension of a wider range of prosodic representations (e.g. accented vs. unaccented referential expressions to signal given vs. new discourse status; see Arnold, 2008). Such integrative approaches will advance our understanding of the process of prosodic development in relation to children's lexical, syntactic, and pragmatic abilities.

REFERENCES

Aguert, M., Laval, V., Le Bigot, L. & Bernicot, J. (2010). Understanding expressive speech acts: the role of prosody and situational context in French-speaking 5- to 9-year-olds. *Journal of Speech, Language, and Hearing Research* **53**, 1629–41.

Armstrong, M. E. (2014). Child comprehension of intonationally-encoded disbelief. In W. Orman & M. J. Valleau (eds), *Proceedings of the 38th Annual Boston University Conference on Language Development*, 25–38. Somerville, MA: Cascadilla Press.

Arnold, J. (2008). THE BACON not the bacon: how children and adults understand accented and unaccented noun phrases. *Cognition* **108**, 69–99.

Barner, D., Brooks, N. & Bale, A. (2011). Accessing the unsaid: the role of scalar alternatives in children's pragmatic inference. *Cognition* **118**, 84–93.

Beckman, M. E. & Ayers, G. E. (1997). *Guidelines for ToBI labelling, version 3.0.* Manuscript and accompanying speech materials, Ohio State University. Online: <http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/>.

Beckman, M. E., Hirschberg, J. & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (ed.), *Prosodic typology: the phonology of intonation and phrasing*, 9–54. Oxford: Oxford University Press.

Büring, D. & Gutiérrez-Bravo, R. (2001). Focus-related constituent order variation without the NSR: a prosody-based crosslinguistic analysis. In J. McCloskey (ed.), *SASC 3:*

877

Syntax and Semantics at Santa Cruz (pp. 41–58). Santa Cruz, CA: Linguistics Research Center, University of California, Santa Cruz.

Capelli, C. A., Nakagawa, N. & Madden, C. M. (1990). How children understand sarcasm: the role of context and intonation. *Child Development* **61**, 1824–41.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. A. Miller (eds), *Linguistic theory and psychological reality*, 264–93. Cambridge, MA: MIT Press.

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

Casillas, M. & Amaral, P. (2011). Learning cues to category membership: patterns in children's acquisition of hedges. In C. Cathcart, I.-H. Chen, G. Finley, S. Kang, C. S. Sandy & E. Stickles (Eds.), *Annual Meeting of the Berkeley Linguistics Society 37th Annual Meeting* **37**(1), 33–45. Available from Linguistic Society of America, eLanguage platform. Online: <http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/836>.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A. & Meroni, L. (2006). The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L. Dominguez. & A. Johansen (eds), *Proceedings of the 25th Annual Boston University Conference on Language Development*, 157–68. Somerville, MA: Cascadilla Press.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language* **17**, 417–31.

Clark, E. V. & Wong, A. (2002). Pragmatic directions about language use: words and word meanings. *Language in Society* **31**, 181–212.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language* **12**, 643–61.

Cutler, A. & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language* **14**, 145–67.

Dennison, H. Y. (2010). *Processing implied meaning through contrastive prosody*. Unpublished PhD dissertation, Department of Linguistics, University of Hawaii, Manoa.

Dennison, H. Y. & Schafer, A. J. (2010). Online construction of implicature through contrastive prosody. *Proceedings of Speech Prosody 2010*. Online: <http://speechprosody2010.illinois.edu/papers/100338.pdf>.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development* **8**, 181–95.

Gelman, A. & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Goodman, J. C., McDonough, L. & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development* **69**, 1330–44.

Graham, S. A., Sedivy, J. & Khu, M. (2013). That's not what you said earlier: preschoolers expect partners to be referentially consistent. *Journal of Child Language* **41**, 1–17.

Grassmann, S. & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language* **34**, 677–87.

Grassmann, S. & Tomasello, M. (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics* **42**, 3098–105.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (eds), *Syntax and semantics, vol.3*, 41–58. New York: Academic Press.

Hansen, M. B. & Markman, E. M. (2005). Appearance questions can be misleading: a discourse-based account of the appearance–reality problem. *Cognitive Psychology* **50**, 233–63.

Heeren, W. F., Bibyk, S. A., Gunlogson, C. & Tanenhaus, M. K. (2015). Asking or telling: real-time processing of prosodically distinguished questions and statements. *Language and Speech* **58**(4), 474–501.

Hornby, P. A. & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* **13**(2), 395–9.

Horowitz, A. & Frank, M. C. (2012). Learning from speaker word choice by assuming adjectives are informative. In N. Miyake, D. Peebles & R. P. Cooper (eds), *Proceedings of*

the 34th Annual Conference of the Cognitive Science Society, 473–8. Austin, TX: Cognitive Science Society.

Horowitz, A. & Frank, M. C. (2014). Preschoolers infer contrast from adjectives if they can access lexical alternatives. In P. Bello, M. Guarini, M. McShane & B. Scassellati (eds), *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 625–30. Austin, TX: Cognitive Science Society.

Horst, J. S. & Samuelson, L. (2008). Fast mapping but poor retention in 24-month-old infants. *Infancy* **13**, 128–57.

Ito, K. (2014). Children's pragmatic use of prosodic prominence. In D. Matthews (ed.), *Pragmatic development in first language acquisition*, 199–218. Amsterdam: John Benjamins.

Ito, K., Bibyk, S. A., Wagner, L. & Speer, S. R. (2012). Interpretation of contrastive pitch accent in six- to eleven-year-old English-speaking children (and adults). *Journal of Child Language* **41**, 1–27.

Ito, K., Jincho, N., Minai, U., Yamane, N. & Mazuka, R. (2012). Intonation facilitates contrast resolution: evidence from Japanese adults and 6-year-olds. *Journal of Memory and Language* **66**, 265–84.

Jun, S.-A. (2005). *Prosodic typology: the phonology of intonation and phrasing.* Oxford: Oxford University Press.

Jun, S.-A. (2014) *Prosodic typology II: the phonology of intonation and phrasing.* Oxford: Oxford University Press.

Kurumada, C., Brown, M., Bibyk, S., Pontillo, D. & Tanenhaus, M. K. (2014). Is it or isn't it: listeners make rapid use of prosody to infer speaker meanings. *Cognition* **133**, 335–42.

Kurumada, C., Brown, M. & Tanenhaus, M. K. (2012). Pragmatic interpretation of contrastive prosody: it looks like speech adaptation. In N. Miyake, D. Peebles & R. P. Cooper (eds), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 647–52. Austin, TX: Cognitive Science Society.

Ladd, D. R. (2008). *Intonational phonology*, 2nd ed. Cambridge: Cambridge University Press.

Markman, E. M. & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* **20**, 121–57.

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**, 749–60.

Mervis, C. B. & Bertrand, J. (1994). Acquisition of the novel name/nameless category (N3C) principle. *Child Development* **65**, 1646–62.

Morgan, J. & Demuth, K. (1996). *Signal to syntax: bootstrapping from speech to grammar in early acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Morton, J. B. & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development* **72**, 834–43.

Musolino, J. (2006). On the semantics of the subset principle. *Language Learning and Development* **2**, 195–218.

Musolino, J. & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics* **44**, 817–52.

Noveck, I. (2001). When children are more logical than adults. *Cognition* **78**, 165–88.

Papafragou, A. (2006). From scalar semantics to implicature: children's interpretation of aspectuals. *Journal of Child Language* **33**, 721–57.

Papafragou, A. & Musolino, J. (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition*, **86** 253–282.

Papafragou, A. & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition* **12**, 71–82.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan & M. E. Pollack (eds), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press.

Quam, C. & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development* **83**, 236–50.

Sakkalou, E. & Gattis, M. L. (2012). Infants infer intentions from prosody. *Cognitive Development* **27**, 1–16.

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

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., and Hirschberg, J. (1992). TOBI: a Standard for labeling English prosody. *Proceedings of the 1992 International Conference on Spoken Language Processing*, Vol. 2, 867–70. Banff, Canada.

Snedeker, J. & Trueswell, J. C. (2004). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* **49**, 238–99.

Solan, L. (1980). Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research* **23**, 688–98.

Speer, S. R. & Ito, K. (2009). Prosody in first language acquisition: acquiring intonation as a tool to organize information in conversation. *Language and Linguistics Compass* **3**, 90–110.

Stiller, A., Goodman, N. D. & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language, Learning, and Development* **11**, 176–90.

Thorson, J. C. & Morgan, J. L. (2015). Acoustic correlates of information structure in child and adult speech. In E. Grillo & K. Jepson (eds), *Proceedings of the 39th Annual Boston University Conference on Language Development*, 411–23. Somerville, MA: Cascadilla Press.

Vallduví, E. (1992). *The informational component*. New York: Garland Press.

Van Der Meulen, S., Janssen, P. & Den Os, E. (1997). Prosodic abilities in children with specific language impairment. *Journal of Communication Disorders* **30**(3), 155–69.

Wells, B., Peppe, S. & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language* **31**, 749–78.

Wieman, L. A. (1976). Stress patterns of early child language. *Journal of Child Language* **3**, 283–6.

Wilkinson, K. M. & Mazzitelli, K. (2003). The effect of 'missing' information on children's retention of fast-mapped labels. *Journal of Child Language* **30**, 47–73.

Winner, E. & Leekam, S. (1991). Distinguishing irony from deception: understanding the speaker's second-order intention. *British Journal of Developmental Psychology* **9**(2), 257–70.