Getting and maintaining attention in talk to young children*

BRUNO ESTIGARRIBIA AND EVE V. CLARK
Stanford University

(Received 8 December 2005. Revised 8 September 2006)

ABSTRACT

When two people talk about an object, they depend on joint attention, a prerequisite for setting up common ground in a conversational exchange. In this study, we analyze this process for parent and child, with data from 40 dyads, to show how adults initiate joint attention in talking to young children (mean ages 1;6 and 3;0). Adults first get their children’s attention with a summons (e.g. Ready?, See this?), but cease using such forms once children give evidence of attending. Children signal their attention by looking at the target object, evidence used by the adults. Only at that point do adults begin to talk about the object. From then on, they use language and gesture to offer information about and maintain attention on the target. The techniques adults rely on are interactive: they establish joint attention and maintain it throughout the exchange.

Even the most casual observation of conversation reveals that the participants rely on gaze and gesture as well as speech. In this paper, we examine a model of how gaze and gesture contribute to joint attention in parent–child exchanges. Our focus is on how adults get young children’s attention at the start of an exchange and then maintain it. Attention is a prerequisite to any communicative exchange (Brinck, 2001). With it, each participant attends and is aware that the other is attending too (H. Clark, 1996). Infants begin early on to attend to adult gaze (e.g. Moore & Dunham, 1995), and follow it from as young as 0;4 (D’Entremont, Hains & Muir, 1997).

[*] This research was supported in part by grants from the National Science Foundation (SBR97-31781), The Spencer Foundation (199900133) and the Center for the Study of Language and Information, Stanford University. We thank William Bowman, Michelle M. Chouinard, Joy C. Geren and Andrew D.-W. Wong for data collection; Caroline Bertinetti, Alegria Salaises and Elizabeth Liebert for help with transcription and analysis; and the staff of the Bing Nursery School, Stanford, and all the families for their participation. We are grateful to Herbert H. Clark, Florian Jaeger, Barbara F. Kelly and Ewart A. C. Thomas, as well as two anonymous reviewers, for helpful discussion and comments on earlier versions of this article. Address for correspondence: E. V. Clark, Department of Linguistics, Stanford University, Stanford, CA 94305-2150, USA. Email: eclark@psych.stanford.edu

799
Since visual attention is object-directed in the real world, shifts in adult gaze could be informative for identifying target objects and in resolving uncertainty about which aspect of the context to attend to (Butterworth & Jarrett, 1991; Carpenter, Nagell & Tomasello, 1998; Woodward, 2003). Infants also attend early to adult gestures (Kelly, 2001; Brand, Baldwin & Ashburn, 2002). By 0;11–1;0, they follow points to a target object without difficulty (Butterworth & Grover, 1990; Carpenter et al., 1998).

The redundancy inherent in gestures that accompany speech, given the prevalence of child-directed speech about the here-and-now, could help capture young children’s attention, and help them establish initial word-referent associations (Butler, Caron & Brooks, 2000). The parents of one-year-olds typically talk about objects and actions that are present (Snow, 1986), synchronizing gestures and speech as they do so (Messer, 1978). In fact, in playing with infants under age 2;0, mothers refer to toys they are manipulating at the time of utterance between 73% and 96% of the time. And where there is a choice of toy, the mothers' holding of the target object removes uncertainty about the intended referent (Messer, 1978). Regardless of the child’s age, mothers’ gestures appear to reinforce their speech by ‘underscoring, highlighting, and attracting attention to particular words and/or objects’ (Iverson, Capirici, Longobardi & Caselli, 1999: 72). Moreover, parents appear to use more gestures in combination with new objects and new words than with familiar objects and words (Gogate, Bahrick & Watson, 2000).

According to developmental studies, infants use gaze and gesture at several stages, alongside adults’ exaggerated gestures and modified forms of speech addressed to them, in interactions about the here-and-now. And adults rely on a small number of deictic ‘frames’ linguistically to introduce unfamiliar nouns before adding further information about the target object (Clark & Wong, 2002). Most studies of language acquisition have relied on tape-recordings but these don’t contain information about gestures or gaze. Yet gaze can signal attention in both adult and child. And gestures can both attract attention and provide non-linguistic information about properties, actions and functions.

What is the rational way to start a conversation? To understand what someone says, one has to attend to it. And one needs to attend from the onset. Indeed, there is evidence that adults start by establishing joint attention (Scheglo , 1968) and even recycle beginnings until both participants are attending (Goodwin, 1981). So adults first establish joint attention, then pursue the exchange. Do they follow the same sequence with children? If so, the start-up schema should take the following form:

1. The adult tries to get the child to attend to some object.
2. The child eventually attends to the object.
GETTING AND MAINTAINING ATTENTION

(3) The adult then introduces new information about the object.
(4) The adult maintains the child’s attention on the object.

These steps are ordered, with (1) always preceding (2), and (2) always preceding (3). Step (4) comes in whenever the child’s attention appears to waver. If the child looks away during Step (3), for example, the adult will recall the child’s attention before continuing with (3). Step (4) is therefore invoked only if the speaker sees some need to maintain the child-addressee’s attention as the exchange continues. This organization of the interactive process initiated by the adult allows for the grounding of new words, and also for the grounding of information pertinent to the meanings children assign to those words.

How else might adults and children interact? Adults might instead simply follow in on whatever the child is already attending to, and so start with Step (3), introducing information about the child’s focus of attention. In these cases, the adult is not trying to establish joint attention but simply starting from whatever the child is already attending to (Akhtar, Dunham & Dunham, 1991; Tomasello & Farrar, 1986). In the present study, our emphasis is on those cases where the adult initiates the establishment of joint attention on a specific target.

Another approach adults might adopt is to simply start talking in the expectation that children will come to attend at some point. (Adults occasionally start conversations this way with each other.) In this case, what we have called Step (3) would again be the starting point; Steps (1) and (4) could be absent, and (2) would be unordered relative to (3). In short, the ordering of (1) through (3), with intermittent reliance on (4) once (3) has been reached, is what distinguishes our proposal from other possibilities.

In this paper, we propose that when adults establish joint attention with young children, they do so in an interactive process made up of several ordered steps. We show that available evidence from adult–child interactions is consistent with these steps. In doing this, we also try to answer several basic questions. How do adults get children to attend first (Step 1)? How long does this take? Are some techniques favored over others? How do adults know that children are attending? Do they consistently look at, touch or point at the object of joint attention (Step 2)? And how do adults maintain their children’s attention as they offer information about the target (Steps 3 and 4)?

METHOD
We gave parents the task of showing some relatively unfamiliar objects to their children one-on-one, emulating a fairly common occurrence in everyday life. Children frequently encounter things that are unfamiliar, and have them named and explained by adults. In the present study, our goal
was to elicit the conversational behaviors adults normally bring to this type of exchange in order to look more closely at adult-initiated joint attention.

Procedure
We made digital video films of adult–child interactions where the parent (mother or father) was given a box of six three-dimensional objects to show to the child. Parent and child were seated side-by-side at adjacent corners of a low table in a small room in a local nursery school, as shown in Figure 1; the child sat in a small wooden chair with arms, to discourage movement around the room during the recording session. The instructions to the parent were simply to show the child the objects one at a time ‘in the way you would usually do because we are interested in how young children react’.

Each session was filmed with a Canon Optura-pi digital video camera. Before each session began, we adjusted the camera so it would capture their faces (for direction of gaze), the space in between them on the table, hand gestures and shifts in body-stance. We then left them alone in the room. Audio-recording for speech was enhanced with a directional shotgun microphone, mounted to record in front of the camera.

Participants and materials
We observed 40 parent–child dyads, drawn from middle- to upper-middle-class families, at two ages. The 20 younger children ranged in age from 1;4-7 to 1;11-13, with a mean age of 1;6-1. Half of these children (n=11) were under 1;6, mean age 1;5-10; the rest were between 1;8 and 1;11, mean age 1;9-11. This split is pertinent for some of the results. The 20
older children ranged in age from 2;8-8 to just over 3;0, 3;2-20 (mean 3;0:10). Half the children in each group were female and half male, and a little over half were first-born. All the children were acquiring American English as their first language.

The children were shown six objects, each between 7 cm and 10 cm in their most extended dimension. The objects had common names not generally known to the relevant ages (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994; Hall, Nagy & Linn, 1984; Templin, 1957), and, to be sure that as many of the objects as possible were unfamiliar, we used equivalent but different sets for the two age levels: (a) younger – an ashtray with a crenellated rim, a kitchen measuring-spoon, a pair of sunglasses, a toy dump-truck, a toy tiger and a toy crocodile; and (b) older – a strainer, a pair of salad tongs, swim-goggles, a toy tractor, a toy sting-ray and a toy gorilla. We collected the data from the younger group first, and so added a second set of items for the older group in order to maximize the probability of their being unfamiliar with the objects and the pertinent words. Each parent took the objects out one at a time, in whatever order they came to hand, from a box on the floor. In the event, a few children knew a word for some of the objects, and not all parents used the expected terms in talking about the objects.

The observation sessions varied from 4 m 21 s to 18 m 41 s depending on how talkative the dyad was. The mean time for the younger children was 8 m 20 s (median 7 m 31 s), and for the older ones was 8 m 14 s (median 8 m 7 s). Time per object (each episode within a session) averaged 1 m 23 s for the younger children and 1 m 22 s for the older ones.

Transcription and reliability

The digital video film of each session was transcribed using MediaTagger (from the Max-Planck-Institute for Psycholinguistics) with six independent tiers for the speech, gaze and gesture for each participant. Each tier was transcribed separately to avoid influence from any data already transcribed. Gestures were described in terms of trajectory and goal (e.g. ‘pick up O[bject] and display on open palm’, ‘tap index finger on O’s tail’). Gaze was noted as ‘to A[adult]’, ‘to C[hild]’, ‘to O[bject]’, or ‘to other’. Speech was transcribed orthographically. In every tier, transcriptions for each gesture, gaze and utterance were time-aligned with the video.

To assess reliability, two transcribers independently transcribed 5-minute samples from 10% of the films. Agreement on these transcriptions was high (we report percentages here since there was no finite set of categories): adult words 93% (range 91% to 96%), child words 84% (range 83% to 100%); adult gestures 90% (range 87% to 93%), child gestures 74% (range 66% to 81%); adult gazes 79% (range 64% to 86%) and child gazes 79% (range 76% to 82%).
Within each transcript we counted adult gesture types (e.g. indicating gestures like pointing, tapping, outlining), adult and child gazes (e.g. at the target object) and certain aspects of adult utterance content (e.g. use of an interjection to get attention, production of a label for the referent object). To assess reliability, we re-coded the first three and last three transcripts for 16 categories. Agreement was 95% for gestures (Cohen’s kappa 0.91), 96% for attention management (Cohen’s kappa 0.66) and 93% for verbal information (Cohen’s kappa 0.84).

RESULTS
In showing the objects to their children, all the adults labeled them and provided additional information about each one. As they talked, they also used gestures to indicate particular parts, for example, or to demonstrate how something moved or was used. And they relied on child gaze as an index of attention throughout each session. These findings are consistent with the general model we postulated for adult initiation of joint attention.

How do adults manage children’s attention?
In answering this question, we identified parental behaviors that manage children’s attention as **attention-getters** if they occurred before children’s first look at the target object, and as **attention-maintainers** if they occurred after that first look. **Attention-getting intervals** were those time intervals between the presentation of a new object and the child’s first look at it, and **attention-maintaining intervals** were the time intervals from after the child’s first look until the parent put that object away. Parents relied primarily on verbal attention-getters, along with the occasional gesture.

Parents used four different types of **verbal attention-getters**: interjections like hey and wow (including gasps); deictic terms like here, this or look; and anticipations like ready for the next one? or I have another toy. These were sometimes combined in the same turn: (gasp) look at this! combines an interjection (the gasp) with a deictic, while mommy’s gonna get something else, look combines an anticipation (something else) with a deictic (look). Lastly, they occasionally used the child’s name.

For each attention-getting and attention-maintaining interval (one of each for each of the six objects), we summed verbal attention-getters over the four categories (interjections, deictics, anticipations, and names) to get a measure of the efforts parents made to manage child attention. Table 1 shows how often adults used verbal attention-getters in the interval before children’s first look at the target compared to how often they did so in the interval after children’s first look. On 240 occasions (78 before-look
Tables 1. Number of observations with 0 to 7 verbal attention-management turns in the intervals before and after children’s first looks, summed over objects

<table>
<thead>
<tr>
<th>Number of attention-getting turns used</th>
<th>Interval before first look</th>
<th>Interval after first look</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>78</td>
<td>162</td>
<td>240</td>
</tr>
<tr>
<td>1</td>
<td>57</td>
<td>55</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>63</td>
<td>16</td>
<td>79</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

intervals and 162 after-look intervals), parents didn’t have to use any verbal attention-getters at all. On 112 occasions, they used only one, and so on.

To what extent did parental management of attention depend on how old their children were, and on how familiar their children were with the current task (i.e. whether parents were presenting the first object or subsequent ones)? And did parents make use of children’s gaze in assessing attention? Our analyses provide evidence for three major findings:

(a) Parents relied on children’s first gaze at an object as an indicator of attention.

(b) Parents made more attempts to get their children to attend to objects presented earlier than to those presented later. In effect, as children became familiar with the procedure, they attended more readily.

(c) Verbal attention-management was used more often with younger children than with older children.

(a) Child gaze. Adults could have relied on looks at either the parent or the object as an indication that children were ready for further information. All the adults in this study took their children’s gaze to the object as signaling attention. Evidence for this comes from the shift in how adults talked after the first few seconds of each episode. Parental speech was significantly more likely to be attention-oriented before children looked at the target object than after their first look. Table 2 shows the distribution of all parental speech, measured in turns, for the interval before the child looked at the target object compared to after the child looked, within the first episode. Turns were generally single utterances separated by pauses or gestures, or by child vocalizations, verbalizations or gestures. Each turn containing an attention-getter was counted as ‘attention-oriented’.

Nearly two-thirds of adult speech, 61%, was attention-oriented before children looked at the target object, but only 6% of it was attention-oriented
TABLE 2. Percentage of verbal attention-management turns (attention-oriented speech) compared to other speech before and after children’s first looks at target object 1

<table>
<thead>
<tr>
<th></th>
<th>Before first look</th>
<th>After first look</th>
<th>Overall percent of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention-oriented speech</td>
<td>61</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Non-attention-oriented speech</td>
<td>39</td>
<td>94</td>
<td>90</td>
</tr>
<tr>
<td>Total adult turns</td>
<td>258</td>
<td>3299</td>
<td>3557</td>
</tr>
</tbody>
</table>

Fig. 2. Mean number of verbal attention-getters produced in getting and maintaining children’s attention (first panel), and attention-getter means normalized per second for the getting and maintaining intervals (second panel).

after children looked. This asymmetry also shows up in the plot of mean numbers of verbal attention-managers used, shown in the first panel of Figure 2. Notice that a simple comparison of the before (getting attention) and after (maintaining attention) means doesn’t capture the fact that dyads spent only a very short time in the initial attention-getting phase (a few seconds) compared to all the time they spent interacting once the object was being attended to by both participants (over 1 m on average). If we normalize these means over time (per second), we obtain the more realistic
proportions shown in the second panel of Figure 2. As the columns here suggest, these proportions differed significantly (F(1, 452) = 95.94, p < 0.001).

These findings follow logically if parents take children’s first looks as an indicator that the children are attending to the objects. Once parents have their children’s attention, they can minimize any further efforts to get their children to attend, and instead offer them information about the objects (Clark, 2001).

Is there any difference with age in how long children take to first look at the target object? The mean time to their first look at an object is just above 8 seconds for the youngest one-year-olds, and just 4.5 seconds for the oldest ones (t(215) = 1.663, p = 0.098). Is it possible that the effects of age, object order and attention interval could be explained, at least in part, by the length of each attention interval? If younger children simply take longer to attend, can this fact alone account for the higher counts of parental verbal attention-managers? No. A Poisson-based regression (see below) showed that, after adjusting for the length of each attention interval, the other effects were still robust.

(b) Familiarity. As children became familiar with the task, parents used fewer attention-getters with each successive object, as shown in Figure 3. This finding was significant (linear regression, t(460) = 2.16, p = 0.03). The plot in Figure 3, though, shows a non-linear effect since the mean for object
6 is slightly higher than those for objects 4 and 5. This is probably because some children had become tired or bored as the end of the task approached, and their attention therefore faltered.

(c) Age. Adults used significantly more verbal attention-getters to the younger group of one-year-olds (mean = 1.03) than to the older one-year-olds or the two- to three-year-olds (combined mean = 0.76; linear regression, t(460) = 2.25, p = 0.025).

A regression model
The results so far have been based on separate statistical tests, so we modeled the data in order to take account of these factors simultaneously. Since the data were a better fit to a Poisson distribution because of the extreme skewness of the count distribution in the number of attention-getters, we used as the response variable the natural logarithm of the mean of Poisson counts (Agresti, 2002). The explanatory variables (fixed effects) were the order of presentation of the objects (a six-level factor), age (a two-level factor) and whether verbal attention-getters occurred before or after children’s first looks (a two-level factor). In the model, the mean for object 2 didn’t differ significantly from that for object 1 which provided the baseline (z = x 0.647, p = 0.52), and the mean for object 3 differed only marginally from that of object 1 (z = x 1.854, p = 0.064). The effects for subsequent objects were significantly different (for 4, z = x 2.341, p = 0.019; for 5, z = x 3.107, p = 0.002; and for 6, z = x 2, p = 0.046). The age effect was significant for the two- to three-year-olds (z = -2.672, p = 0.008). That is, they differed significantly from the baseline provided by the one-year-olds. The effect of gaze was also significant (z = x 8.672, p < 0.001): there were fewer attention-getters used after the children looked at each target object. Finally, in order to account for within-subject correlations induced by repeated measures, we included parent–child dyads as a random effect. In this mixed model (fixed and random effects), all the effects are robust, and there are no significant interactions.

The results of the regression analysis confirm the significance of the separate findings and give precise estimates of the size of the effects. The model predicts children in the young one-year-old group hear a mean of 2.03 verbal attention-getters before they look at the first object, but only 1.41 before they look at the last object, while for children in the oldest group, the means are 1.51 and 1.04 respectively. It also predicts that, for each object, the mean number of verbal attention-getters will decrease by a multiplicative factor of 0.38, after the child’s first look at the object. These predictions are all consistent with the data.

We showed that parents produce more verbal attention-getters before children look at each object, hence that they use children’s gaze as a signal
of attention. Adults also use more verbal attention-getters for earlier objects than for later objects, showing that parents are sensitive to the fact that as children become familiar with the task they attend more readily. Finally, older children in our sample attended more readily, and this was reflected in parents’ use of fewer verbal attention-getters. Parents used more verbal attention-getters with younger children because younger children, especially those under 1;6, took longer to attend.

Verbal attention-getter types

Verbal attention-getters were classified into four types – anticipations, deictics, interjections and names. Parents used more deictic terms and interjections to younger children, and more anticipatory comments about as yet unseen objects to older children. Name use was low overall. Figure 4 shows how the distribution of these types changed with the age of the child.

The two most used types were anticipations and deictics. Deictics decreased with age, whereas anticipations increased. A linear regression with response (deictic-anticipatory) and age in days as a continuous predictor confirmed this effect (p<0·01). Like deictics, interjections also decreased with age (Poisson-based regression, p=0·003), but name use was always rare.

Notice that anticipatory comments alert children ahead of time to the imminent appearance of the next object, while deictics assume the referent is already present and visible and so must either be concomitant with or immediately follow the presentation of the target object. The pattern of use by adults is consistent with older children’s becoming attuned to the task more quickly and being readier to attend than the younger ones. It is also consistent with deictics being the commonest for object 1, while
anticipations gain ground with subsequent objects. A linear regression with response (deictic-anticipatory) and object order as a continuous predictor confirmed this effect (p=0.007).

**DISCUSSION**

Establishing joint attention is an interactive process. We proposed that it consists of ordered steps: (1) the adult gets the child’s attention on X; (2) the child signals attention to X; and (3) the adult conveys information about X. Within Step (3), the adult can intervene (4) to maintain the child’s attention on X. The findings reported here support each of these steps and are captured by our regression model. In the first few seconds of an episode, adults used verbal attention-getters in 61% of their turns, together with a few gestures like pointing or displaying (too few, though, to include in the model). But after children looked at the target, adults then talked about it for the remainder of the episode (94% of turns), and made use of verbal attention-getters in only 6% of their turns (see Table 2).

**Adult–adult exchanges**

We have shown how adults use verbal attention-getters to get young children’s attention. These are the same steps that occur in adult–adult exchanges: Observers have noted that many speakers begin with a prelude to the actual exchange, often in the form of a summons (Scheglo, 1968; Scheglo & Sacks, 1973), as in (1):

1. **Ann:** Bob?
   **Bob:** Yes?

By using his name, the first speaker identifies Bob as the intended addressee, and in using a rising intonation, she asks him to respond – to show he is attending – before she continues. She could also have simply called out or said hey to get him to turn towards her. Other kinds of summons include telephone rings, doorbells, whistles, taps in the shoulder or arm and nudges – all serving the same function as the attention-getters addressed to young children.

Addressee gaze serves here too as a general signal that an addressee is attending. So if the addressee is not looking when the speaker begins to talk, the speaker may call on other techniques to attract the addressee’s gaze. One is to re-start the first turn (Goodwin, 1981), as in (2) (the asterisks mark overlapping material in the exchange):

1. **Lee:** Can you bring me [pause 0.2s]
   **Ray:** *[starts to turn his head]*
   **Lee:** *Can* you bring me here that nylon

810
Lee’s initial use of Can you bring me is to request Ray’s attention. As soon as Ray looks at him, he re-starts the utterance and this time completes it (Goodwin, 1981: 61). Or the speaker can pause in mid-utterance, until the addressee looks, and only then continue (Goodwin, 1981: 76), as in (3):

(3) Barbara: uh my kids [pause 0·8 s]
    Ethel: [starts to turn head]
    Barbara: had all these blankets, and quilts and sleeping bags.

Just as in such adult–adult exchanges, parents relied on child gaze as an indication of attention. As soon as their children looked at the target, parents began to talk about the object now in joint attention. They also used some gestures to indicate and demonstrate properties as they offered additional information (Clark & Estigarribia, in preparation; see also Zukow, 1991).

Continuing to attend

Continuing attention is largely taken for granted in adult–adult exchanges, and we know little about techniques for maintaining attention with adult interlocutors. Looking at each other is one criterion of attention – but this can be intermittent. Adult addressees also consistently track what the speaker says with back-channel forms like Mm and Uh-huh, or head gestures like nods along with shifts in facial expression. The speaker’s utterances and gestures probably serve a dual function – accumulating information and maintaining attention through voice and hand motions (see also Schnur & Shatz, 1984). The young children we studied appeared to be well aware that adults both point at and manipulate objects that are to be attended to (Woodward & Guajardo, 2002; see also Lempers, Flavell & Flavell, 1977).

To participate in joint attention, children must learn to discern the intentions of the other (Tomasello, 1995). They need to understand why the adult is trying to get their attention – that adult summonses tell them to attend to something while gaze and gesture tell them what to attend to. Part of discerning the intentions of the other may be derived from infants’ earlier participation in reciprocal exchanges with both gestures – giving and taking, showing and pulling back – and vocalizations – calls or exclamations with peek-a-boo and other hide-show games. These develop during the second half of the first year, and by 1;0 appear to be well-established (Rheingold, Hay & West, 1976). By this age, young children can also follow adult points and use pointing themselves to direct adult attention (e.g. Bates, 1976). And from 1;6, children can attend to what the adult is attending to in learning new words, even when they can’t see the target object (e.g. Baldwin, 1993). In summary, by their second year, young children can take account of adult speech, gesture and gaze as they are called on to attend jointly with adult interlocutors to specific objects.
But do parental strategies for establishing joint attention extend beyond situations where adults talk about objects? When parents read to young children, they talk about objects and about parts and properties of those objects, so the joint attention they establish is relevant for talking about both objects and their properties. Moreover, stories typically involve actions on the part of the participants, and these too are common in what adults talk about to young children. Adults also offer children demonstrations of how to do things, and here too they require joint attention for communication. Although the present study considered only a setting where parents showed children objects, these parents actually talked about parts, properties, actions and functions too (Clark & Estigarribia, in preparation). Our findings therefore seem likely to apply to a broad range of situations in which adults get their children’s attention before introducing terms for parts, properties, actions and relations, as well as terms for objects.

SUMMARY
In the parent–child exchanges explored in this paper, we focused on how parents get young children to attend. Joint attention constitutes a basic condition on communicative exchanges. We showed that adults use consistent techniques for each step in this process: they first try to get the child’s attention, and, as soon as the child looks at the target, they begin to present information about that object. Adult–child exchanges depend on a similar interactive process in starting up conversational exchanges where the addressee is not already attending (Goodwin, 1981; Kita, 2003), but just how people instantiate joint attention may vary across cultures (see e.g. Chavajay & Rogoff, 1999).

REFERENCES