

Twelve-Month-Olds Understand Social Intentions Based on Prosody and Gesture Shape

Núria Esteve-Gibert

CNRS

LPL UMR 7309

Aix Marseille Université and

Department of Translation and Language Sciences

Universitat Pompeu Fabra

Pilar Prieto

Department of Translation and Language Sciences

Universitat Pompeu Fabra and

ICREA (Institució Catalana de Recerca i Estudis Avançats)

Ulf Liszkowski

Department of Developmental Psychology

University of Hamburg and

MPI for Psycholinguistics

Max Planck Research Group Communication before Language

Infants infer social and pragmatic intentions underlying attention-directing gestures, but the basis on which infants make these inferences is not well understood. Previous studies suggest that infants rely on information from preceding shared action contexts and joint perceptual scenes. Here, we tested whether 12-month-olds use information from act-accompanying cues, in particular prosody and hand shape, to guide their pragmatic understanding. In Experiment 1, caregivers directed infants' attention to an object to request it, share interest in it, or inform them about a hidden aspect. Caregivers used distinct prosodic and gestural patterns to express each pragmatic intention. Experiment 2 was identical except that experimenters provided identical lexical information across conditions and used three sets of trained prosodic and gestural patterns. In all conditions, the joint perceptual scenes and preceding shared action contexts were identical. In both experiments, infants reacted appropriately to the adults' intentions by attending to the object mostly in the sharing interest condition, offering the object mostly in the imperative condition, and searching for the referent mostly in the informing condition. Infants' ability to comprehend pragmatic intentions

based on prosody and gesture shape expands infants' communicative understanding from common activities to novel situations for which shared background knowledge is missing.

Theories of language acquisition posit that understanding others' communicative intentions emerges prior to and is causally related to the acquisition of meaningful language use (Tomasello, 2003). In support, experimental research shows that around the time of their first birthdays, infants understand the pragmatic meaning of nonverbal communicative acts (Tomasello, Carpenter, & Liszkowski, 2007). Less is known about the cognitive processes underlying these abilities, that is, *how* infants understand others' communicative acts and what sources of information they rely on. Most studies have addressed infants' understanding of referential intentions (*what* to attend to), typically in cases of referentially arbitrary requests (e.g., at 12–14 months: Liszkowski, Carpenter, & Tomasello, 2008; Moll, Richter, Carpenter, & Tomasello, 2008; Saylor & Ganea, 2007; at 17 months: Southgate, Chevallier, & Csibra, 2010). These studies have manipulated the shared action contexts preceding a communicative act, that is, the shared activities or “common grounds” of participants in an interaction. The findings show that infants bind information from preceding shared action contexts and joint visual scenes to a communicative act to interpret its referent.

However, to fully understand a communicative act and react appropriately, one also needs to understand the social intention underlying the act (*why* somebody directs one's attention). Only a few studies have addressed whether infants distinguish among different types of social intentions (Aureli, Perucchini, & Genco, 2009; Behne, Liszkowski, Carpenter, & Tomasello, 2012; Camaioni, Perucchini, Bellagamba, & Colonesi, 2004; Liebal, Behne, Carpenter, & Tomasello, 2009). For example, in Liebal et al. (2009), an experimenter played a cleanup game with 14-month-old infants, then surreptitiously left one object in the middle of the room and left. After that, either the same or a new experimenter entered the room and pointed to the object. Infants interpreted the point as a request to clean up the remaining object and reacted appropriately only when the person was the experimenter with whom they just had shared the activity of cleaning up. Thus, similarly to studies on referential understanding, these studies on social intention understanding suggest that infants bind information from preceding shared action contexts and joint visual scenes to a communicative act in order to react appropriately.

One problem with these paradigms, however, is that it has remained unclear whether infants are inferring a social intention underlying the communicative act or are simply taking the next relevant step in a shared action sequence (e.g., continuing with the cleanup game). For example, infants might have cleaned the object away had they found it by themselves, or the mere presence of the cleanup partner may have sufficed to instigate the next step in the sequence. Differentiating between these two possibilities is crucial to determine whether infants indeed infer social and pragmatic intentions underlying communicative acts before they acquire language. Further, in many accounts, attention-directing deictic acts are conceptualized as inherently arbitrary and do not mean anything by themselves: They require resorting to information about shared action contexts outside the communicative act (e.g., Tomasello, 2008). The problem, however, is that shared knowledge is not always readily available or easily established, especially for infants who are novices at most

tasks. In quotidian communication, however, social and pragmatic meaning is also expressed in the co-occurring features of a communicative act, like its prosody and gesture shape (e.g., Kendon, 2004). For example, when an attention-directing act is accompanied by speech, the prosodic pattern used by speakers signals their physical, mental, or emotional states and instigates pragmatic implicatures (Gussenhoven, 2004; Ladd, 1996; Pierrehumbert & Hirschberg, 1990; Prieto, 2015; Wilson & Whar-ton, 2006). Older children use prosodic cues to infer linguistic meaning (Cutler & Swinney, 1987; Ito, Bibyk, Wagner, & Speer, 2014), but it has remained unknown whether young infants on the cusp of language acquisition can use vocal characteristics or gesture shape to distinguish social intentions. One recent study found that 14-month-olds distinguish intentional from accidental actions based on the actor's intonational expressions (Sakkalou & Gattis, 2012); it is unclear, however, whether these abilities extend to infants' pragmatic understanding of communicative actions. Unraveling the kinds of cues infants pick up on when making inferences about others' communicative acts is crucial to understanding the cognitive nature and development of the human capacity to infer others' communicative intentions.

Infant-directed interactions typically involve specific prosodic and gestural cues. Caregivers produce longer pauses, more prosodic repetitions, higher mean pitch, and wider pitch range when interacting with infants, and they use gestures to disambiguate and emphasize the accompanying speech (e.g., O'Neill, Bard, Linnell, & Fluck, 2005; Saint-Georges et al., 2013). While these characteristics help infants detect *communicative* intent, that is, the fact *that* they are being addressed and that something relevant will follow (Senju & Csibra, 2008), the question has been whether caregivers also use different prosodic characteristics to express different types of *social* intentions (e.g., requests, offers, or comments) and whether prelexical infants can pick up on these act-accompanying cues to infer the different social intentions and react appropriately.

A recent seminaturalistic study of 12-month-old infants in their home environments suggests that caregivers use distinct prosodic patterns and gestural hand shapes when drawing infants' attention to a referent for different reasons: caregivers mostly used high pitch range, long syllables, and index-finger pointing when sharing interest in exciting events; moderate pitch range, short syllables, and palm-up whole-hand gesture when requesting objects; and narrow pitch range, moderate short syllables, and index-finger pointing when informing about hidden toys (Esteve-Gibert, Liszkowski, & Prieto, 2016). That study, however, did not address conclusively whether infants understood parents' communicative acts appropriately, and in particular, whether infants relied on the features accompanying the acts as cues pointing to the intended meaning, or on the act-preceding shared activities, which also provided disambiguating information regarding the speaker's intentions. Naturalistic observation studies cannot easily tease apart information from shared action contexts and information from the act itself, and "in the wild," it is notoriously difficult to discern whether infants base their communicative inferences on lexical, prosodic, and gestural cues accompanying the act, preceding shared activities, or any combination of these factors.

The current study aimed to establish whether 12-month-old infants can distinguish among different types of social intention by reading the prosodic characteristics of point-accompanying vocalizations and gesture shapes. We designed a laboratory-based procedure in which an adult directed an infant's attention to a novel referent while

expressing one of three types of social intention (Tomasello et al., 2007): expressive (sharing an attitude about an object with the infant), imperative (asking the infant to give them an object), and informative (helping the infant to find something relevant). Across these three conditions, we equalized the preceding shared action contexts and perceptual co-presence (sources of information, which disambiguate meaning, see Clark & Marshall, 1981), the spatial layout, and the response opportunities to exclude the possibility that infants could react appropriately based on information sources other than the style of the communicative act. Infants were tested at the age of 12 months because at this age they are already able to produce gesture and prosodic cues to communicate intentionally (Esteve-Gibert & Prieto, 2013; Grünloh & Liszkowski, 2015) and because they comprehend referential acts to hidden entities within appropriate preceding shared action contexts and joint visual scenes (Behne et al., 2012).

In Experiment 1, we invited caregiver–infant dyads and analyzed caregivers’ various attention-directing acts in terms of accompanying prosody and gesture shape as they conveyed the three aforementioned social intentions to determine relevant differences in the information infants were exposed to. We then tested whether infants would react appropriately to their caregivers’ different communicative acts. If infants had some categorical understanding of accompanying characteristics as cues to pragmatic intentions, they should react appropriately to these distinctly expressed meanings even when the preceding shared activity and perceptual co-presence alone would not suffice to disambiguate the meaning of the gestures. That is, when caregivers expressed their interest in an object, infants should mostly share attention with them; when caregivers requested an object, infants should mostly offer it to them; and when caregivers provided information about a hidden item, infants should search for it. In Experiment 2, we used the same paradigm but controlled for the lexical content of speech. To this end, we used trained experimenters who produced gestures with the specific prosodic patterns and hand shapes observed in caregivers in Experiment 1. Experiment 2 thus aimed to replicate the findings from Experiment 1 while constraining the interpretation by excluding the possibility that the 12-month-olds could perhaps base their responses on a semantic understanding of the lexical content.

EXPERIMENT 1

Methods

Participants

Twenty-two caregiver–infant dyads participated in the study. Four dyads were excluded because of parental procedural error ($N = 1$) or infants’ refusal to participate ($N = 3$). The final sample included 18 caregiver–infant dyads (nine girls). Infants’ mean age was 12 months and 17 days (range: 12 months and 7 days–12 months and 28 days), and they all came from families with mid-to-high socio-economic status. All dyads were recruited from a Dutch database of caregivers from a middle-sized city in the Netherlands who expressed interest in participating in research with their infant.



Figure 1 Setup of the test room. Left: screenshot; right: schematic sketch of the setup.

Setup and materials

Two tables were arranged in an L shape (see Figure 1). An infant chair was attached to the table on the inside of the L. The caregiver's chair was placed at 90° to the infant chair so that the adult would be facing the side of the infant. Under the table and in front of the caregiver's chair, there was a small bench with two toys on it. A black cardboard occluder was placed along the lower part of the L-shaped table, behind the infant's line of sight, to hide the experimenter (E, henceforth) and stimuli. The cardboard occluder had a small opening at table height through which small objects could be passed. Two room dividers also blocked the infant's view of E during the experiment. Caregiver and infant were recorded by a camera connected to a video monitor placed behind the cardboard occluder so that E could observe their interactions from behind the screen.

A total of eight empty paper cupcake cups were used as stimuli, one for each trial. The cups were all of different colors and looked appealing to the infant. They were presented in front of the infant on a black stick which E passed through the opening in the cardboard occluder. The black stick had a round plate-like platform affixed to its end onto which the cups were placed. A round colored sticker was attached to the center of this platform such that it would be hidden under the cup. The same setup and materials were used across all three conditions.

Procedure

Infants were randomly assigned to one of three conditions in which the caregivers were instructed to act in one of three ways, resulting in 6 caregiver–infant dyads per condition. There were eight test trials per condition. The general procedure in the warm up, play phase, and test phase was identical in the three conditions.

Warm up. Before entering the test room, caregiver, infant, and E played with some toys together for 5 min as E first explained the experiment to the caregiver in a general way and then gave specific instructions on what s/he was to do during the test phase. These instructions differed across conditions. After these instructions were completed, all three went into the test room. E helped the caregiver to accommodate the infant in the infant seat and instructed the caregiver to sit down in the caregiver's

chair. E briefly reiterated the instructions to the caregiver and then hid behind the black cardboard occluder.

Play phase. In each trial, caregiver and infant played with the two toys (which were out of sight on the bench under the table) for about a minute, with the explicit instruction that they could only have one toy at a time on the table. This play period was crucial to distract the infant before the adult directed the infant's attention toward the target object at the other end of the table. Otherwise, the infant could have focused on the target object by chance independently of the adult's communication.

Test phase. E made sure that the infant played with the caregiver and did not attend to the occluder, and then, she passed the upside-down cup on the black stick inconspicuously through the opening of the occluder in the screen until it was now 10 cm from the infant's seat. There was a white mark behind the occluder and on the stick to signal how far E had to push the stick. E could also monitor the location of the cup by means of the camera and adjust it depending on how far the infant could reach: For example, if, after the first trial, E saw that at 10 cm, the cup was too easy for the infant to reach; the stick was left a bit further away. Importantly, the distance between the caregiver and the cup had to be big enough to induce caregivers to use a deictic gesture to direct the infants' attention toward the cup. We chose not to have the target object in infants' view from the beginning of the trial to prevent infants from exploring or taking the cupcake cup before the caregiver actually directed his/her attention toward it. Once the cup was situated in front of the infant, the caregiver directed the infant's attention to it with an expressive, imperative, or informative motive, following the respective specific instructions.

Expressive condition: Caregivers were instructed to direct the infant's attention to the cup as if wanting to share their interest in the cup with the infant. The explicit instruction given to the caregiver was (in Dutch): *Deel je interesse voor de beker met je zoon/dochter. Gebruik gerust woorden of gebaren als u wilt. Het enige is dat u het object zelf niet aan mag raken* "Share your interest in the cup with your son/daughter. Feel free to use words or gestures if you want to. The only thing is that you should not touch the object yourself."

Imperative condition: Caregivers were instructed to direct the infant's attention to the cup in order to get the infant to give them the cup. The explicit instruction given to the caregiver was as follows: *Vraag je zoon/dochter om jou de beker te geven. Gebruik gerust woorden of gebaren als u wilt. Het enige is dat u het object zelf niet aan mag raken* "Ask your son/daughter to give you the cup. Feel free to use words or gestures if you want to. The only thing is that you should not touch the object yourself."

Informative condition: Caregivers were instructed to direct the infant's attention to the cup in order to inform the infant that there was a sticker under it. The explicit instruction given to the caregiver was as follows: *Informer je zoon/dochter dat er iets onder de beker verstopt is. Gebruik gerust woorden of gebaren als u wilt. Het enige is dat u het object zelf niet aan mag raken* "Inform your son/daughter that there is something hidden under the cup. Feel free to use words or gestures if you want to. The only thing is that you should not touch the object yourself."

No explicit instruction was given to the caregiver regarding the gesture and speech strategies they should use. However, E told the caregiver to direct the infants' attention toward the cup only once or twice per trial to prevent them from being too insistent. In total, the cup was in the infant's field of vision for 20 sec. If the infant took the cup before the 20 sec was over, the caregiver placed the cup back on the stick and E retracted the stick again. However, if the infant had not shown any reaction, E retracted the stick with the cup again behind the occluder. If the 20 sec had passed and the infant was still playing with the cup, E shook the stick to signal the caregiver to place the cup on the stick again, and retrieved the cup. Then, the play phase of the next trial started. Each trial involved a differently colored cup and sticker.

Data coding

The data were first coded in terms of the caregiver's use of speech and gesture and then with regard to the infant's behavior.

Caregivers' speech and gesture. We coded the caregivers' use of gesture and speech. First, the hand shape of pointing gestures (either accompanied by speech or not) was classified as *index-finger pointing gesture* when the arm was extended and the index finger was directed at a specific location, or as *palm-up whole-hand pointing gesture* when the arm was extended toward the object or infant and the hand was open with the palm oriented upwards. Second, the prosodic features of caregivers' speech (with or without accompanying gestures) were transcribed with the help of Dutch experts using Praat sound-analysis software (Boersma & Weenink, 2012). Prosodic analyses were conducted on the utterance conveying the target intended meaning in each condition. When caregivers produced more than one utterance to convey the target intended meaning within a trial, they were analyzed as different data points and averaged. Three features were coded: the intonation pattern, the mean syllable duration, and the pitch range. These three features were chosen following previous research on their relevance in early prosodic comprehension and production (e.g., Papaeliou & Trevarthen, 2006). Intonation pattern was annotated using the well-established ToDI system for the transcription of Dutch intonation (Gussenhoven, 2005). Mean syllable duration was calculated by dividing the total duration of the sentence (in ms) into the number of syllables it contained. The pitch range of the utterance was coded by locating the maximum and minimum pitch points in the F0 line and subtracting the minimum from the maximum. Finally, the lexical content of the caregivers' speech was analyzed using the following data-driven categories: "Look" (e.g., *Kijk eens* "Look!"); "Look" + Adjective (e.g., *Kijk eens! Mooie, die is mooi he!* "Look! Nice, it is nice!"); "Look" + Location (e.g., *Wat is dat nou? Zit daar iets onder?* "What's that? Do you see what is under it?"); "Take" + Object (e.g., *Pak het bekertje* "Take this little cup!"); or "Give" + Object + Person (e.g., *Geef maar aan mama* "Give it to mama").

The caregivers' fidelity in acting as instructed and their naturalness in doing so was judged for 20% of the data set by one additional rater observing the video recordings. Caregivers were judged as acting as instructed (i.e., conveying the appropriate social intention) in 93.3% of the trials, no condition showing more faithful behavior than the others ($\chi^2(2, 30) = 4.286, p = .117$). Caregivers were also judged as

acting naturally in 100% of the trials: on a 1–5 Likert scale (1 = very artificial, 5 = very natural), 33.3% of the trials obtained a 4 and 66.7% obtained a 5, with no effect of condition, $F(2, 30) = 1.016$, $p = .375$, $\eta^2 = .070$. We also ensured that between-trial intervals did not include information about the social intention of the interaction. In 20% of the between-trial intervals, a naïve rater was asked to guess what would happen next in the interaction, with four options: The caregiver will share interest about the cup, the caregiver will request the cup, the caregiver will inform about something hidden in the cup, or “I don’t know”. The “I don’t know” option was chosen in 100% of the cases.

An inter-rater reliability test for the coding of the caregivers’ speech and gesture was conducted with 10% of the data for each condition ($N = 21$) by two independent trained coders who were unaware of the purpose of the study. The reliability of ToDI intonation coding was good, yielding 75.4% agreement and a Cohen’s $\kappa = 0.71$,¹ and excellent for gesture shape (100% agreement, Cohen’s $\kappa = 1.0$).

Infants’ behaviors. Infants’ behavior after each trial was coded using ELAN. Four categories were used: (1) *offering cup*, when the infant took the cup from the stick and gave it to the caregiver; (2) *attending cup*, when the infant looked at the cup ostensively, pointed at it, or picked it up and played with it; (3) *attending sticker*, when the infant took the cup off the stick and looked at, pointed at or played ostensively with the sticker or the black stick under the cup; and (4) *no reaction*, when the infant did not show any of these reactions within the 20 sec during which the cup was placed in front of him/her. When more than one of these reactions was observed within the 20 sec of a trial, the primary, predominant reaction of the infant was coded, that is, the reaction that was most salient, longest, or not a consequence of an unprompted discovery of the cup or the sticker. Thus, if the infant explored the cup for a few seconds and then gave it to the caregiver, the most salient behavior was *offering cup* and the trial was coded as such. By the same token, if the infant picked up the cup, handled it for a couple of seconds, but later put it down to pay attention to the sticker and this second behavior lasted longer than his/her handling of the cup, it was coded as *attending sticker*. Finally, if the infant attended the sticker because (s)he discovered it by chance after having handled the cup, this was considered incidental behavior and the primary behavior was still classified as *attending cup*.

Inter-rater reliability was conducted with 20% of the data for each condition ($N = 26$) by two independent coders who were unaware of the purpose of the study. Overall, the reliability in coding of the infants’ behavior as offering cup, attending cup, or attending sticker was very good (Cohen’s $\kappa = 0.83$). When considering cases in which infants produced only one response, agreement was perfect (Cohen’s $\kappa = 1$), and in cases in which infants produced more than one response, agreement was substantial (Cohen’s $\kappa = 0.70$).

Regarding our data analyses, we used mixed ANOVAs when analyzing within-subjects dependent variables with several levels (caregivers’ intonation contours, infants’ behaviors) as a function of between-subjects independent variables (condition); and we used one-way ANOVAs for within-subjects variables with one level (pitch range, speech rate) as a function of between-subjects independent variables (condition). When main effects or interactions were observed, these were further explored with simple

¹These results are consistent with the intertranscriber agreement results found in studies on intonation transcription using the ToBI system (see Escudero, Aguilar, Vanrell, & Prieto, 2012, for a review).

effects based on the overall ANOVA (Fisher's LSD) to know at which level that variable was significant. If significances were obtained, direct comparisons based on mean differences were performed to determine the exact nature of that effect. Most dependent variables were normally distributed. Regarding infants' behaviors, nonparametric tests (Kruskal–Wallis; planned Mann–Whitney U comparisons) essentially yielded the same pattern of significances as the reported comparisons based on ANOVAs.

Results

Caregivers' speech and gesture

One aim of this experiment was to identify the specific gesture shape and prosodic strategies that the adults used to convey imperative, expressive, and informative intentions. From the total amount of data (144 trials), 14 trials of six individual infants were excluded from the analysis because of parental procedural error (six trials of one individual infant) or the infant's refusal to participate (e.g., when they tore the cup or threw it away; eight trials of five individual infants). Thus, a total of 130 valid trials were used for subsequent analyses (see Table 1, for a distribution of valid trials across conditions and act modality). The proportion of valid trials differed significantly across conditions, $F(2, 143) = 4.012$, $p < .05$, $\eta^2 = .054$, with more valid trials in the imperative than the informative condition ($p < .01$), and all other comparisons being non-significant. This difference, however, does not reflect distinct levels of difficulty for the infants because it was mainly driven by parental procedural errors. There was no relation between the condition and the caregivers' use of a specific modality ($\chi^2(4, N = 130) = 2.78$, $p = .596$).

The analysis of the caregivers' use of different gesture shapes across conditions revealed that caregivers always used index-finger pointing gestures during the expressive and informative conditions (100% of the cases in both conditions), while they always used palm-up whole-hand pointing gestures directed at the object or the infants during the imperative condition (100% of the cases).

The intonation analysis showed that four contours were the most commonly observed (representing 75% of the total number of utterances): the fall (40.4%, $N = 84$), the fall-rise (14.4%, $N = 30$), the half-completed fall (12.1%, $N = 25$), and the low rise (8.2%, $N = 17$). A mixed ANOVA with these four intonation contours as dependent within-subjects variable and condition as between-subjects variable (three

TABLE 1
Distribution of Valid and Invalid Trials Across Conditions and Act Modalities

	<i>Expressive</i>		<i>Imperative</i>		<i>Informative</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Valid trials						
Gesture only	1	2	4	8.3	2	4.2
Speech only	13	27	10	20.8	12	25
Gesture + speech	32	66.7	33	68.8	25	52.1
Total valid	44	91.7	47	97.9	39	81.3
Invalid trials	4	8.3	1	2.1	9	18.7

levels: expressive, imperative, and informative) revealed a main effect of intonation, $F(3, 45) = 8.307$, $p < .001$, $\eta^2 = .356$, no main effect of condition, $F(2, 15) = 1.441$, $p = .268$, $\eta^2 = .161$, and no interaction between intonation and condition, $F(6, 45) = 1.660$, $p = .153$, $\eta^2 = .181$. Thus, the fall contour appeared most frequently in all conditions.

Exploratory follow-up analyses for each intonation contour separately showed that only the fall-rise intonation differed across conditions, $F(2, 15) = 4.133$, $p < .05$, $\eta^2 = .355$. Direct comparisons based on the mean differences revealed more fall-rise in the expressive condition than in the other two conditions (both p 's $< .05$). For each condition separately, simple effects based on the overall ANOVA showed that intonation varied significantly within the expressive condition, $F(3, 13) = 7.399$, $p < .01$, $\eta^2 = .631$. Direct comparisons based on the mean differences within the expressive condition showed that the fall contour was more frequent than both the half-fall and the rise contours ($p < .01$ for all comparisons), and the fall-rise was more frequent than the rise ($p < .05$).

Analysis of the mean syllable duration yielded significant differences between conditions, one-way ANOVA, $F(2, 214) = 26.894$, $p < .001$, $\eta^2 = .201$. Follow-up comparisons showed that the mean syllable duration differed significantly between the expressive and imperative conditions, $F(1, 180) = 37.824$, $p < .001$, $\eta^2 = .174$, and between the expressive and informative conditions, $F(1, 170) = 19.430$, $p < .001$, $\eta^2 = .103$, but not between the imperative and informative conditions, $F(1, 78) = 3.078$, $p < .083$, $\eta^2 = .038$ (see Table 2, for mean values).

Analysis of the pitch range revealed significant differences between conditions, one-way ANOVA, $F(2, 214) = 4.115$, $p < .05$, $\eta^2 = .037$. Follow-up comparisons revealed a statistically significant difference between pitch range in the expressive condition compared to the informative condition, $F(1, 170) = 6.710$, $p < .01$, $\eta^2 = .038$. Pitch range in the expressive condition did not differ from that seen in the imperative condition, $F(1, 180) = 2.651$, $p = .105$, $\eta^2 = .015$, and the imperative and informative conditions did not differ significantly from one another, $F(1, 78) = 1.154$, $p < .286$, $\eta^2 = .015$ (see Table 2, for mean values).

Analysis of the lexical content in caregivers' utterances (see Table 3) showed that "Look" + Adjective was the most frequent type of utterance in the expressive condition, "Give" + Object + Person was the most frequent in the imperative condition, and "Look" + Location was the most frequent in the informative condition. Caregivers thus tended to use specific lexical cues alongside their prosodic and gesture strategies to convey their intentions.

TABLE 2
Caregivers' Use of Pitch Range and Syllable Duration Across the Three Pragmatic Situations

	<i>Syllable duration (ms)</i>		<i>Pitch range (Hz)</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Expressive	252.36	96.98	215.12	132.32
Imperative	155.51	72.78	179.38	112.61
Informative	178.86	33.82	151.60	117.45

TABLE 3
Distribution of the Lexical Content of the Caregivers' Speech Across the Three Conditions

	"Look"		"Look" + Adjective		"Look" + Location		"Take" + Object		"Give" + Object + Person	
	N	%	N	%	N	%	N	%	N	%
Expressive	8	18.2	34	77.2	1	2.3	1	2.3	0	0
Imperative	2	4.3	1	2.1	0	0	5	10.6	39	83
Informative	4	12.1	0	0	28	84.8	1	3.1	0	0

Bold indicates the Most Frequent Pattern in Each Condition.

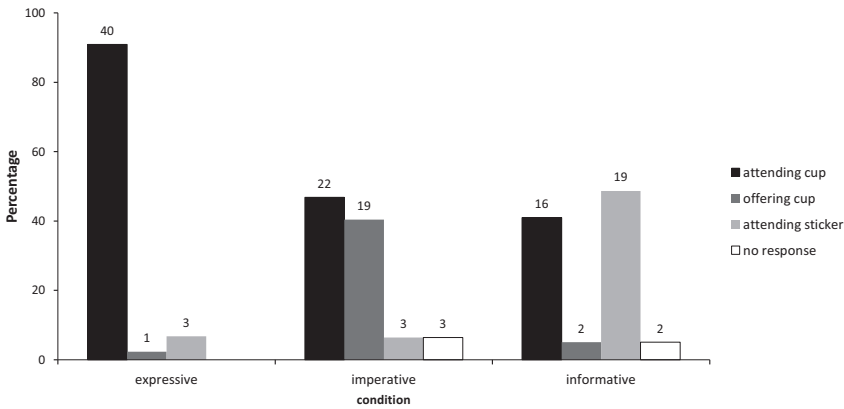


Figure 2 Percentage of infant's behaviors across conditions in Experiment 1. Numbers over the bars represent numbers of trials in which the behavior was shown.

Infants' behaviors

A 3 (behavior: attending cup, offering cup, attending sticker) \times 3 (condition: expressive, imperative, informative) mixed ANOVA was applied to the total number of valid trials. It revealed significant differences across behaviors, $F(2, 30) = 15.959$, $p < .001$, $\eta^2 = .515$, no effect of condition, $F(2, 15) = 1.995$, $p = .171$, $\eta^2 = .210$, and an interaction between behavior and condition, $F(4, 30) = 8.235$, $p < .001$, $\eta^2 = .523$ (see Figure 2). For each behavior separately, simple effects based on the overall ANOVA showed that all infants' behaviors varied depending on the condition, attending cup: $F(2, 15) = 7.230$, $p < .01$, $\eta^2 = .491$; offering cup: $F(2, 15) = 11.123$, $p < .01$, $\eta^2 = .597$; attending sticker: $F(2, 15) = 6.305$, $p < .05$, $\eta^2 = .457$. Direct comparisons based on the mean differences revealed more attending cup behavior in the expressive condition compared to the imperative and informative conditions (respectively, $p < .05$ and $p < .01$), more offering behavior in the imperative condition compared to the expressive and informative conditions (both p 's $< .01$), and more "attending sticker" behaviors in the informative condition compared to the expressive and imperative conditions (both p 's $< .01$).

For each condition separately, simple effects based on the overall ANOVA showed that infants' behaviors varied significantly within all conditions, expressive: $F(2, 14) = 17.749$, $p < .001$, $\eta^2 = .717$; imperative: $F(2, 14) = 4.673$, $p < .05$, $\eta^2 = .400$; informative $F(2, 14) = 6.029$, $p < .05$, $\eta^2 = .463$. Direct comparisons on the mean differences revealed that within the expressive condition, infants attended the cup more than they offered it or attended the sticker (respectively, $p < .001$, $p < .001$), within the imperative condition infants attended the cup and offered it equally often, but they offered the cup more than they attended the sticker (respectively, $p = .67$ and $p < .01$), and within the informative condition infants attended the sticker more than they offered the cup ($p < .01$).

Thus, infants attended the cup a lot in all conditions but did so most often in the expressive condition, they offered the cup across all conditions but most often in the imperative one, and they attended the sticker across all conditions but most often in the informative condition (see Figure 2).

TABLE 4
Number of Infants Showing the Predominant Behavior at Least Once in Each Condition and Experiment

	<i>Attending cup</i>	<i>Offering cup</i>	<i>Attending sticker</i>
Exp.1			
Expressive	6	1	0
Imperative	6	6	3
Informative	5	1	5
Exp. 2			
Expressive	10	0	5
Imperative	10	6	4
Informative	10	4	9

Bold indicates the predicted behavior in each condition.

There was no evidence that the infants learned over trials. There was no significant difference between the mean proportion of trials with the appropriate response in the first half of the trials (55.6%) compared to the second half of the trials (52.8%), as shown by the results of a paired-samples *t*-test: $t(17) = 0.287$, $p = .777$, $d = .139$. A 2 (appropriate response: in the first half of the trials, in the second half of the trials) \times 3 (condition: expressive, imperative, informative) mixed ANOVA revealed no effect of appropriate responses, $F(1, 15) = .204$, $p = .658$, $\eta^2 = .013$, and no interaction between condition and appropriate responses, $F(2, 15) = .663$, $p = .530$, $\eta^2 = .081$. We also analyzed how many of the infants showed the expected behavior for each condition at least once during the experiment. First, all infants participating in the expressive condition attended the cup at least once during the experiment trial ($N = 6$). Second, in the imperative condition all infants offered the cup to the adults at least once during the experiment ($N = 6$). Third, in the informative condition, all infants except for one attended to the sticker at least once during the experiment ($N = 5$) (see Table 4).

Discussion

Experiment 1 showed that 12-month-old infants can infer the different social and pragmatic intentions underlying caregivers' attention-directing acts. Infants reacted to caregivers' acts of requesting, sharing, or informing about an object, respectively, mostly by offering it, sharing interest in it, or searching for something hidden by it. Our new paradigm revealed that infants can make these inferences even when the preceding shared action context does not disambiguate the meaning of caregivers' attention-directing acts. It shows that infants can base their inferences on information from accompanying speech and gesture shapes, while excluding several alternatives.

Importantly, across all experimental conditions infants played with their parent using the same toys; their attention was not focused on the target object before the communicative act; and the spatial layout, seating arrangement, and response period were identical across conditions. Therefore we can exclude the possibility that forms of common ground like preceding action contexts and perceptual co-presence, which have been shown previously to guide infants' interpretations of communicative acts, differentially influenced the infants' behavior across the three conditions. One could argue that the infants differentially learnt over trials and accumulated some form of common ground over repeated trials. However, caregivers were instructed to always put the cup back on the stick, which would be a rather unexpected reaction to a fulfilled

communicative act in the imperative and informative conditions. More importantly, a direct test of learning over trials revealed no increase in target behaviors, so we can also exclude the possibility that information from repeated trials differentially influenced infants' responses in each condition.

Experiment 1 showed that caregivers express different social and pragmatic intentions not only through lexical means but also through distinct patterns of gesture shape and prosody, which is in line with recent findings from home observations (Esteve-Gibert et al., 2016). Research has often established the meaning of prosody as a function of the context within which it is used, but, in the current experiment, adults were instructed to convey different intentions, so that the prosodic realization was a direct consequence of that instruction. The expressive condition had the most varied contours, widest pitch range and longest syllable durations, a pattern reminiscent of infant-directed speech characteristics which have been related to affective talk (Bryant & Barrett, 2007; Trainor, Austin, & Desjardins, 2000) and likely reflect the sharing of an affective state with the infant. In the informative condition, by contrast, the fall contour with narrow pitch range and short syllables—the most neutral way of producing a broad focus statement in Dutch (Gussenhoven, 2005)—predominated, presumably reflecting a less emotion-laden content, and highlighting the informational element. The imperative condition was mostly characterized by a falling intonation which was produced with wide pitch range and small syllable duration, a pattern that confirms previous findings on imperative intonation in Dutch (Van Heuven & Kirsner, 1999). Another strong cue to the requesting in contrast to the informing intention was the upward-oriented palm shape. Expressive and informative pointing gestures shared the same index-finger hand shape, suggesting that a common primary function of these declarative acts is to signal an epistemic state (e.g., attention to something) in contrast to imperative pointing gestures, which primarily signal a motivational state (e.g., a desire to get something).

These results suggest that 12-month-old infants understand others' social intentions not just thanks to extensive preceding action contexts and attentional scenes, as previous research has suggested, but also because they have some understanding of the form of the communicative acts, that is, their accompanying features. However, one limitation of this study is the small sample size, which requires replicating the effect of infants' appropriate responses. Another limitation is that we cannot conclusively rule out the possibility that infants understood caregivers' intentions based on the lexical–semantic information contained in the caregivers' speech. In order to control for lexical information and test exclusively the role of prosody and hand shape, we conducted Experiment 2, in which we excluded the disambiguating information present in lexical–semantic speech.

EXPERIMENT 2

Three experimenters were trained to replicate the caregivers' three strategies identified in Experiment 1 regarding hand shape and prosodic cues but had them express the same lexical content across conditions. Otherwise, the paradigm was the same as in Experiment 1. Based on the results of Experiment 1, our predictions were that (1) the infants would attend to the cup more in the expressive condition than in the other two conditions; (2) the infants would offer the cup more in the imperative condition than

in the other two conditions; and (3) the infants would explore the sticker more in the informative condition than in the other two conditions.

Methods

Participants

Thirty-six 12-month-old infants participated in the study. Six infants were excluded from the sample because they became fussy in more than half of the trials ($N = 3$), did not want to play ($N = 2$), or because of mother interference ($N = 2$). The final sample included 30 infants (nine girls). The infants' mean age was 12 months and 12 days (range: 12 months and 3 days–12 months and 26 days) and they all came from mid-to-high SES families. All infants were recruited from a Dutch database of caregivers from a middle-sized city in the Netherlands who expressed interest in participating in research with their infant. None of them had participated in Experiment 1.

Setup and materials

The setup in Experiment 2 was identical to Experiment 1, with the only differences that (1) an experimenter sat where the caregiver had previously been seated, and (2) there was a chair behind the infant on which caregivers sat during the experiment.

Procedure

Infants were randomly assigned to one of three conditions, resulting in 10 dyads per condition. The procedure was the same as in Experiment 1 except that (1) an experimenter rather than the caregiver drew the infants' attention toward the objects; (2) specific gesture and prosodic strategies were used by the experimenter for each condition; and (3) the same lexical information was conveyed verbally across conditions.

Three experimenters were trained to use gesture and prosodic strategies modeled on those used by the caregivers in Experiment 1 to convey the three different types of social intentions under study (see Figure 3). The training sessions were administered by the first author and consisted of watching a video with the target gesture and prosodic strategies and then practicing how to produce them. The gesture and prosodic strategies were repeated to ensure that the experimenters could all reproduce them properly and uniformly. To reduce a potential effect of experimenter in the infants' behavior, all experimenters tested the same number of infants per condition. In addition, during each trial, the first author of this study monitored the E's performance from behind the occluder to check whether the experimenter performed properly and provide feedback if necessary. E's body posture was controlled so that it did not differ across pragmatic conditions: moving slightly forward during the pointing gesture and then quickly returning to her initial position. E's gaze patterns were also controlled by having her shift her gaze between cup and infant twice during the attention-directing act and then fixing her gaze on the infant for the rest of the trial. In all three conditions, E produced one of two pointing gestures while speaking the same words (in Dutch): "*Hey! Die! Die!*" "*Hey! This! This!*" with a falling intonation contour (H*L L %). The specific gesture and prosodic strategies in each condition were as follows:

Imperative condition: E produced a palm-up whole-hand pointing gesture, palm tilted slightly upwards, and used a wide pitch range and short syllables.

Expressive condition: E produced an index-finger pointing gesture and used a wide pitch range and long syllables.

Informative condition: E produced an index-finger pointing gesture and used a narrow pitch range and short syllables.

The experimenter produced the pointing gesture accompanied by speech only once and then returned to her initial position, without interacting with the infant during the rest of the trial except for cases when the infant offered the cup to her. In those cases, the experimenter took the cup and placed it back on the stick. As for facial gestures, we decided that the experimenters should not act artificially and violate the naturally co-occurring facial expressions of the different acts. We therefore asked them to have a smiling face during the expressive condition and a friendly but slightly less enthusiastic facial expression during the imperative and informative conditions.

Data coding

The data were coded as in Experiment 1. A manipulation check was conducted with 100% of our data to check for potential differences in experimenters' behaviors between conditions in terms of body posture, facial gestures, and gaze alternations. In 100% of the trials, the experimenters produced the body posture and gaze alternations as

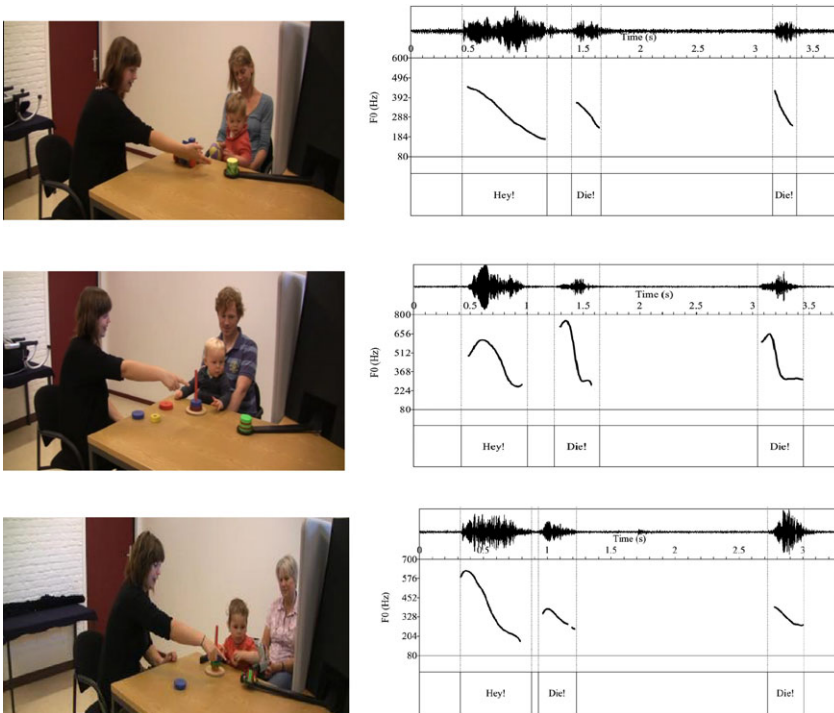


Figure 3 Stimuli used in Experiment 2: top panel, stimuli in the imperative condition; middle panel, stimuli in the expressive condition; bottom panel, stimuli in the informative condition.

trained. In 89% of the trials, the experimenters produced the intended facial expression. The experimenters' fidelity in acting as instructed and their naturalness in doing so was judged for 20% of the data. Experimenters used the instructed hand shape in 100% of the trials, the instructed intonation contour in 100% of the expressive and imperative trials and in 98.7% of the informative trials, and they also consistently varied syllable durations and pitch range depending on the condition, one-way ANOVAs: $F(2, 88) = 15.141, p < .001, \eta^2 = .256$ for syllable duration, and $F(2, 76) = 39.149, p < .001, \eta^2 = .507$ for pitch range. As expected, expressive trials had significantly longer syllable durations than imperative and informative trials (both p 's $< .001$, imperative and informative trials not differing from each other). Expressive trials had wider pitch range than imperative and informative trials (both p 's $< .001$), while imperative and informative trials did not differ in terms of pitch range. Experimenters used the stipulated prosodic and gesture features and, although they were less good at reproducing the appropriate pitch range in imperative trials, infants were probably guided by the distinctive palm-up whole-hand attention-directing gesture in this condition.

In the majority of trials the experimenter was judged as natural (no cases were judged "very artificial" or "quite artificial"; 26.5% of the cases were judged "a bit natural", 34.7% "quite natural", and 38.8% "very natural"). Condition influenced the degree of naturalness, $F(2, 46) = 6.876, p < .01, \eta^2 = .23$, expressive trials being judged more natural than imperative and informative trials (both p 's $< .01$), the latter two being judged equally natural ($p = .808$).

Inter-rater reliability was conducted as in Experiment 1 with 20% of the data from each condition ($N = 47$) by two independent coders who were unaware of the purpose of the study. The reliability was excellent (Cohen's $\kappa = 0.93$). When considering only the cases in which infants produced only one response, reliability was perfect (Cohen's $\kappa = 1$), and in cases where the infant produced more than one response, reliability was substantial (Cohen's $\kappa = 0.75$).

Results

From the total amount of data (240 trials), two trials were excluded from the analysis because of experimenter error, so a total of 238 valid trials were used for analysis. Figure 4 shows an overview of the reactions across conditions. A 3 (behavior: attending cup, offering cup, attending sticker) \times 3 (condition: expressive, imperative, informative) mixed ANOVA revealed a main effect of behavior, $F(2, 54) = 52.253, p < .001, \eta^2 = .659$, an effect of condition, $F(2, 27) = 4.362, p < .05, \eta^2 = .244$, and an interaction between behavior and condition, $F(4, 54) = 6.564, p < .001, \eta^2 = .327$. For each behavior separately, simple effects based on the overall ANOVA revealed that the three behaviors occurred at significantly different rates within each condition, attending cup: $F(2, 27) = 7.195, p < .01, \eta^2 = .348$; offering cup: $F(2, 27) = 3.828, p < .05, \eta^2 = .221$; attending sticker: $F(2, 27) = 6.688, p < .01, \eta^2 = .331$. In line with our predictions and results from Experiment 1, we conducted three planned contrasts to test whether each behavior followed the predicted pattern across the three conditions. The first planned contrast revealed that infants attended the cup significantly more often in the expressive condition than in the other two conditions, $t(27) = 3.118, p < .01, d = 1.200$. The second planned contrast revealed that infants offered the cup to the adult significantly more often in the imperative condition than in the other two conditions, $t(27) = 2.502, p < .05, d = .963$. The third planned contrast revealed that infants

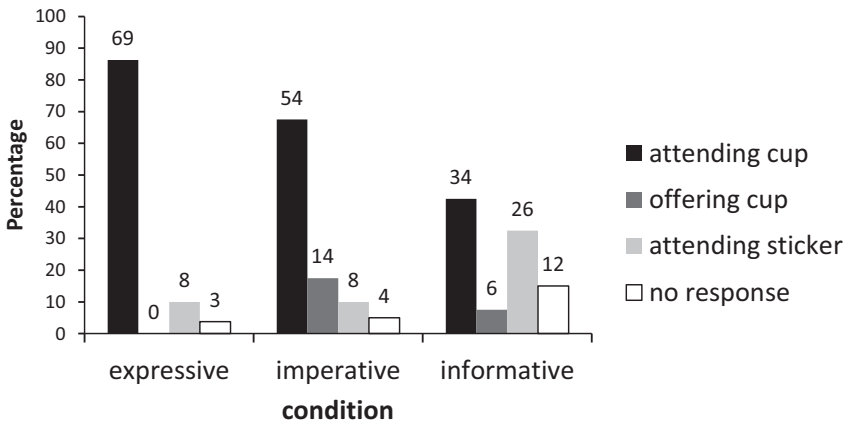


Figure 4 Percentage of infant's reactions across conditions in Experiment 2. Numbers over the bars represent number of trials in which that reaction was observed within each condition.

attended to the sticker significantly more often in the informative condition than in the other two conditions, $t(27) = 3.657$, $p < .001$, $d = 1.407$.

For each condition separately, simple effects based on the overall ANOVA showed that the three behaviors occurred at significantly different rates within each condition, expressive: $F(2, 26) = 33.863$, $p < .001$, $\eta^2 = .723$; imperative: $F(2, 26) = 11.511$, $p < .001$, $\eta^2 = .470$; informative: $F(2, 26) = 11.289$, $p < .001$, $\eta^2 = .465$. Direct comparisons based on the mean differences revealed that in the expressive condition, infants attended the cup more than they offered it or attended the sticker (both p 's $< .001$). In the imperative condition, infants offered the cup slightly more than they attended the sticker, but this difference did not reach significance, and they attended the cup more than they offered it or attended the sticker (both p 's $< .001$). In the informative condition, infants attended the cup and the sticker equally often, but rarely offered the cup compared to the other behaviors (both p 's $< .01$).

There was no evidence that infants learned over trials. There was no significant difference between the mean proportion of trials with the appropriate response in the first half of the trials (45%) compared to the second half of the trials (47.5%), as shown by the results of a paired-samples t -test, with $t(29) = 0.462$, $p = .647$, $d = .172$. A 2 (appropriate response: in the first half of the trials, in the second half of the trials) \times 3 (condition: expressive, imperative, informative) mixed ANOVA revealed that the appropriate responses in the first half of the trials compared to the second half of the trials did not vary, $F(1, 27) = .591$, $p = .449$, $\eta^2 = .021$, and no interaction between appropriate responses in the first and second half and condition, $F(2, 27) = .732$, $p = .490$, $\eta^2 = .051$. We also calculated the number of infants showing the expected behavior at least once during the experiment to be sure that the results did not come from only one infant being better than the others. First, all infants participating in the expressive condition attended the cup in at least one trial during the experiment. Second, in the imperative condition, more than half of the infants offered the cup to the experimenter at least once during the experiment ($N = 6$). Third, in the informative condition, all infants except for one attended the sticker at least once during the experiment ($N = 9$) (see Table 4).

Discussion

Experiment 2 replicated the main findings of Experiment 1, namely that infants used the act-accompanying speech and gesture features to interpret the caregivers' social intent. Importantly, Experiment 2 controlled for the lexical information conveyed across conditions, thus ruling out the possible alternative interpretation of Experiment 1 that infants reacted only based on lexical-semantic information. As in Experiment 1, our analyses of learning over trials revealed no change in behavior when comparing the first half of the trials with the second half of the trials across conditions, thus excluding the possibility that infants used differently accumulating information in the different conditions. The three conditions were exactly the same in terms of spatial layout, action context, and lexical information. The only available sources of information that we manipulated experimentally were the gesture shape and prosodic cues of pitch range and syllable duration. Thus, infants' different interpretations of the attention-directing gestures were driven neither by differences in preceding action context, perceptual co-presence, nor the lexical content of speech. Instead, the experimental results reveal that infants reacted appropriately by relying on the shape of the pointing gesture combined with the prosodic cues of duration and pitch range.

Future work could address the relative contribution of each of these act-accompanying characteristics to the infants' interpretation of social intentions and attempt to determine whether infants rely more on one cue than another at different developmental stages. In the current study, infants extracted information from an adult's prosodic and gesture patterns in the absence of disambiguating information from a shared action context.

In Experiment 2, sharing attention was the infants' most frequent behavior across conditions. One possible explanation is that sharing attention is a default interpretation for referential acts to new things, perhaps because the relevance is guided bottom-up (gloss: "oh, something new"). In the absence of pre-established common ground and lexical information, it is presumably easiest and most natural for infants to interpret a pointing gesture as an expression of affective interest, because this motive is rooted in communicative (nonreferential) exchanges that develop soon after birth (Trevarthen, 1979). In contrast, interpreting informative and imperative pointing gestures requires a more complex understanding of others' goals, and in the absence of pre-established common ground and lexical cues, this should be more difficult and perhaps requires deeper processing of the style of the act. This could explain, among other things, why the appropriate response rates in the imperative and informative conditions were overall a bit lower than in the expressive condition, suggesting that interpretations of imperative and informative pointing gestures rely more on information from preceding common ground and social context than interpretations of expressive pointing acts.

GENERAL DISCUSSION

Social and pragmatic accounts suggest that the acquisition of language and perhaps the emergence of social understanding more generally are rooted in early social interactions and communicative exchanges. Accordingly, from a developmental point of view, the challenge has been to show that infants communicate meaningfully before they

engage in verbal communication (Tomasello, 2003) and explicit theory-of-mind reasoning (Liszkowski, 2013). From a cognitive point of view, the question is: How do infants do this? One approach is to investigate what information infants use to guide their pragmatic understanding of others' communicative acts. The few available studies on social intention comprehension have suggested that infants interpret pragmatic intentions based on information from preceding shared action contexts and perceptual co-presence (Aureli et al., 2009; Camaioni et al., 2004; Liebal et al., 2009).

The current findings show that when information from preceding shared contexts is ambiguous or underspecified, 12-month-olds can use another source of meaning to comprehend social intentions. This additional information source derives from the quotidian concurrent characteristics of communicative acts, such as prosody and gesture shape, through which social intentions are expressed. This finding is important in at least two respects. First, it provides clearer evidence that prelexical infants indeed understand the social intentions underlying a communicative act, because our paradigm excluded the possibility that infants only inferred a next step in a sequence. Second, it accounts for a more flexible understanding of communication in novel situations that go beyond routinized ritual activities.

The current study did not address the emergence of infants' comprehension of prosody and gesture shape before 12 months of age. We propose that this kind of comprehension builds on skills of understanding communicative acts within contextualized activities. Infants likely first learn about act-accompanying characteristics because they repeatedly co-occur with acts that are embedded in extensively shared, meaningful action contexts, like rituals and routines. For example, at 4 months infants already anticipate others' actions in known routines (e.g., being picked up; Reddy, Markova, & Wallot, 2013). These kinds of interpersonal activities are accompanied by rich additional cues in prosody, gesture, and posture, thus providing infants with ample opportunity to learn about act-accompanying characteristics and how they map onto the ongoing activity. Just like children need some understanding of the concept of focus to benefit from accented words (Cutler & Swinney, 1987), infants need an understanding of social intentions within shared activities to benefit from their characteristic prosodic and gestural expressions. In this respect, infants' understanding of accompanying characteristics presumably originates in simpler forms of action understanding and social engagement in the first year of life and reflects a broad social-cognitive achievement, not simply a precursor to meaningful language use.

ACKNOWLEDGMENTS

We thank Marloes van der Goot, Margret van Beuningen, and Mireille Hassemer for helping in running the experiments, Thomas Grünloh for his comments on the methodology, Aoju Chen and Yiya Chen for their help in transcribing Dutch intonation, Rafèu Sichel-Bazin and Pauline Welby for their help with Praat scripts, and Alfonso Igualada and Santiago González-Fuente for participating in the inter-rater reliability tests. We are really grateful to the children and the children's caregivers for voluntarily taking part in this study. This research has been funded by a research grant awarded by the Spanish Ministry of Science and Innovation (FFI2012-31995 "Gestures, prosody and linguistic structure"), by a grant awarded by the Generalitat de Catalunya (2014 SGR-925) to the *Grup d'Estudis de Prosòdia*, by a traveling grant

awarded by the Department of Translation and Language Sciences at the Universitat Pompeu Fabra, and by the Max-Planck-Society.

REFERENCES

- Aureli, T., Perucchini, P., & Genco, M. (2009). Children's understanding of communicative intentions in the middle of the second year of life. *Cognitive Development*, 24(1), 1–12.
- Behne, T., Liszkowski, U., Carpenter, N., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, 30, 359–375.
- Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* (Version 5.3.04) [Computer program]. Retrieved from www.praat.org
- Bryant, G. H., & Barrett, H. C. (2007). Recognizing Intentions in Infant-Directed Speech: Evidence for Universals. *Psychological Science*, 18(8), 746–775.
- Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnese, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy*, 5(3), 291–308.
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press.
- Cutler, A., & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language*, 14, 145–167.
- Esteve-Gibert, N., Liszkowski, U., & Prieto, P. (2016). Prosodic and gesture features distinguish the pragmatic meaning of pointing gestures in child-directed communication. In M. E. Armstrong, N. Henriksen & M. M. Vanrell (Eds.), *Approaches to intonational grammar in Ibero-Romance* (pp. 251–275). Amsterdam, NL: John Benjamins.
- Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: Evidence from Catalan-babbling infants. *Journal of Child Language*, 40, 919–944.
- Escudero-Mancebo, D., Aguilar, L., Vanrell, M. D. M., & Prieto, P. (2012). Analysis of inter-transcriber consistency in the Cat_ToBI prosodic labelling system. *Speech Communication*, 54(4), 566–582.
- Gussenhoven, C. (2004). *The phonology of tone and intonation (research surveys in linguistics)*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Grünloh, T., & Liszkowski, U. (2015). Prelinguistic vocalizations distinguish pointing acts. *Journal of Child Language*, 42(6), 1312–1336. doi:10.1017/S0305000914000816.
- Ito, K., Bibyk, S., Wagner, L., & Speer, S. R. (2014). Interpretation of contrastive pitch accent in 6- to 11-year-old English Speaking children (and adults). *Journal of Child Language*, 41, 84–110.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, 12, 264–271.
- Liszkowski, U. (2013). Using theory of mind. *Child Development Perspectives*, 7(2), 104–109.
- Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108, 732–739.
- Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what “we” have shared in a special way. *Infancy*, 13(1), 90–101.
- O'Neill, M., Bard, K. A., Linnell, M., & Fluck, F. (2005). Maternal gesture with 20-month-old infants in two contexts. *Developmental Science*, 8, 352–359.
- Papaeliou, C., & Trevarthen, C. (2006). Prelinguistic pitch patterns expressing ‘communication’ and ‘apprehension’. *Journal of Child Language*, 33(1), 163–178.
- Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, USA: MIT Press.
- Prieto, P. (2015). Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6, 371–381.
- Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS ONE*, 8, e65289.

- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., . . . Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS ONE*, *8*, e78103.
- Sakkalou, E., & Gattis, M. (2012). Infants infer intentions from prosody. *Cognitive Development*, *27*(1), 1–16.
- Saylor, M. M., & Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, *43*, 696–704.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668–671.
- Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*, 907–912.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: The MIT Press.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*, 705–722.
- Trainor, L. J., Austin, C. M., & Desjardins, N. (2000). Is Infant-Directed Speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*, 188–195.
- Trevarthen, C. (1979). Communication and cooperation in early infancy. *A description of primary intersubjectivity*. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 321–347). London: Cambridge University Press.
- Van Heuven, V. J., & Kirsner, R. S. (1999). Interaction of grammatical form and intonation: Two experiments on Dutch imperatives. In R. Bezooijen, & R. Kager (Eds.), *Linguistics in the Netherlands* (pp. 81–96). Amsterdam: John Benjamins.
- Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*, 1559–1579.