Infant Behavior & Development 39 (2015) 42–52

Contents lists available at ScienceDirect

Infant Behavior and Development

Language development at 18 months is related to multimodal communicative strategies at 12 months

Alfonso Igualada[a], Laura Bosch[b,d], Pilar Prieto[c,a,*]

[a] Universitat Pompeu Fabra, Spain
[b] Universitat de Barcelona, Spain
[c] Institució Catalana de Recerca i Estudis Avançats, ICREA, Spain
[d] Institute for Brain, Cognition and Behaviour, UB, Spain

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form
23 December 2014
Accepted 10 February 2015

Keywords:
Multimodal communication
Joint attention
12-month-old infants
Language measures
Pointing task

ABSTRACT

The present study investigated the degree to which an infants' use of simultaneous gesture–speech combinations during controlled social interactions predicts later language development. Nineteen infants participated in a declarative pointing task involving three different social conditions: two experimental conditions (a) available, when the adult was visually attending to the infant but did not attend to the object of reference jointly with the child, and (b) unavailable, when the adult was not visually attending to neither the infant nor the object; and (c) a baseline condition, when the adult jointly engaged with the infant's object of reference. At 12 months of age measures related to infants' speech-only productions, pointing-only gestures, and simultaneous pointing–speech combinations were obtained in each of the three social conditions. Each child's lexical and grammatical output was assessed at 18 months of age through parental report. Results revealed a significant interaction between social condition and type of communicative production. Specifically, only simultaneous pointing–speech combinations increased in frequency during the available condition compared to baseline, while no differences were found for speech-only and pointing-only productions. Moreover, simultaneous pointing–speech combinations in the available condition at 12 months positively correlated with lexical and grammatical development at 18 months of age. The ability to selectively use this multimodal communicative strategy to engage the adult in joint attention by drawing his attention toward an unseen event or object reveals 12-month-olds' clear understanding of referential cues that are relevant for language development. This strategy to successfully initiate and maintain joint attention is related to language development as it increases learning opportunities from social interactions.

1. Introduction

Gesture–speech integration is an important feature of human communication. As McNeill (1992) noted, in human languages gesture and speech modalities are coordinated not only at the temporal and phonological levels (i.e., the most prominent part of the gesture is typically aligned with the most prominent part of speech), but also at the semantic

* Corresponding author at: Department of Translation and Language Sciences, Universitat Pompeu Fabra, Roc Boronat, 138, 08018 Barcelona, Spain. Tel.: +34 935421114; fax: +34 93 93 542 16 17.
E-mail addresses: alfonso.igualada@upf.edu (A. Igualada), laurabosch@ub.edu (L. Bosch), pilar.prieto@upf.edu (P. Prieto).

http://dx.doi.org/10.1016/j.infbeh.2015.02.004
0163-6383/

and pragmatic levels (i.e., the two components can share similar semantic functions). Infants begin to use simultaneous gesture–speech combinations intentionally near the end of the first year of life, a few months after the onset of canonical babbling and typically preceding the beginning of the one-word production stage (Butcher & Goldin-Meadow, 2000; Carpenter, Mastergeorge & Coggins, 1983; Esteve-Gibert & Prieto, 2014). The presence of these combined, multimodal communicative behaviors have been taken as an indicator of intentional communication, representing a step further on the way toward linguistic communication (Bates, Camaioni, & Volterra, 1975; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Wetherby & Prizant, 1989). But research is still needed concerning the prevalence and pragmatic function of simultaneous gesture–speech combinations in specific socio-communicative contexts and their potential predictive value for later language development.

The developmental pathway of simultaneous gesture–speech combinations was studied in Esteve-Gibert and Prieto (2014). The study showed that at 11 months infants already produced simultaneous gesture–speech combinations, but pointing without speech still occurred more frequently. In their longitudinal sample they also found a significant increase in gesture–speech productions by 15 months of age. These multimodal productions mostly involved pointing and reaching gestures with a declarative communicative purpose, and when combined with speech, the two modalities were temporally coordinated in an adult-like way. The use of simultaneous gesture–speech combinations may serve to provide redundant information about the same referent through multimodal means, thereby highlighting a particular piece of information and minimizing joint effort in a communicative context (see Wagner, Malisz, & Kopp, 2014, for a review). In other words, infants may intentionally use multimodal strategies to mark a prominence in their communicative productions, a behavior that favors joint attention processes.

There is a considerable body of evidence that infants' joint attention abilities are linked to later language development (Kristen, Sodian, Thoermer, & Perst, 2011; Laakso, Poikkeus, Katajamäki, & Lyytinen, 1999; Mundy & Gomes, 1998; Tomasello & Farrar, 1986; Tomasello, 1988). Studies have provided evidence that caregivers' contingent interactions (e.g., those that follow on the infant's focus of attention) tend to elicit more pointing and speech combinations by infants (e.g., Miller & Gros-Louis, 2013; Miller & Lossia, 2013) and also lead to better language abilities later in development (McGillion et al., 2013; Rollins, 2003; Tamis-LeMonda, Bornstein, & Baumwell, 2001; Tomasello & Farrar, 1986). These results provide indirect evidence about the potential relationship between an infant's multimodal communicative ability to initiate joint attention (i.e., to communicate and influence an adult's attention regarding an intended referent) on the one hand and the infant's later language abilities on the other.

Literature addressing early infants' communication abilities has typically focused on separate analyses of either gestures (and mainly pointing gestures) or speech modality but not both. For example, the ability to use pointing gestures has been regarded as a clear and powerful non-verbal strategy to also initiate joint attention between the infant and the adult with regard to an object or event (Tomasello, Carpenter, & Liszkowski, 2007). Likewise, research on infants' gesture production has shown that communicative gestures (e.g., iconic and pointing gestures) signal intentional communication (Bates et al., 1979; Bavin et al., 2008; Caselli, Rinaldi, Stefanini, & Volterra, 2012) and that pointing gestures with a declarative intention are a good predictor of the emergence of verbal language (Colonnesi, Stams, Koster, & Noom, 2010). On the other hand, literature on speech development has also documented that acoustic measures of early infants' vocalizations vary depending according to their communicative intentionality (Esteve-Gibert & Prieto, 2014) and that vocalizations coordinated with gaze directed at the referent affect adult-infant social interactions and support language learning (Goldstein, Schwade, Briesch, & Syal, 2010; Gros-Louis, West, & King, 2014). While some studies with slightly older infants have shown that one particular use of supplementary gesture–speech combinations – that in which the gesture modality conveys a different meaning than the one conveyed by speech – predicts the onset of grammatical development (Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson & Goldin-Meadow, 2005; Özçaliskan & Goldin-Meadow, 2005; Pizzuto, Capobianco, & Devescovi, 2005; Rowe & Goldin-Meadow, 2009), the emergence of simultaneous gesture–speech combinations (i.e., gesture co-occurring with speech to express the same meaning) and their relation to later language development has not been analyzed in detail.

In this study we are interested in exploring the link between the early ability to intentionally produce simultaneous pointing–speech combinations in specific communicative contexts and later language development. To our knowledge only two studies have explored the predictive role of early simultaneous gesture–speech combinations on later language development. In Murillo and Belinchón (2012), a sample of eleven parent-infant dyads were recorded interacting in a semi-structured play context at three longitudinal moments, namely at 9, 12, and 15 months. The results showed that the use of pointing gestures at 12 months, especially when accompanied by vocalizations and directed gaze on the part of the infant, correlated positively with vocabulary development at 15 months of age. In a recent study, Wu and Gros-Louis (2014) analyzed the spontaneous interactions of 10- to 13-month-old infants with their mothers in fifty-one dyads and showed that the infants' combinations of vocalization and pointing, and especially those produced when mothers were not attending to the target event, were related to the infants' subsequent comprehension skills at 15 months. It should be noted that both of the studies mentioned above are based on the analysis of spontaneous interactions, where it is difficult to behaviorally control for two important aspects of early communicative patterns, namely, (a) the pragmatic intention or motive behind children's use of pointing gestures to comment on an event or object; and (b) the social interaction gaze patterns used by the adult during the communication. In this study we will attempt to address this issue by controlling for these two factors. To do so, we will examine pointing gestures that express a declarative intention (i.e., the communicator engages with the recipient to share information with him/her about something) by using a task that was specifically designed to elicit this behavior in infants, namely the declarative pointing task (Carpenter, Nagell, & Tomasello, 1998).

Liszkowski, Albrecht, Carpenter, and Tomasello (2008) showed that the communicative behaviors of 12- and 18-month-olds were affected by the patterns of adult attention to both the child and the event of reference. The authors measured children's communicative responses in two experimental social interaction conditions involving differences in an adult's availability in relation to the infant. The behavioral procedure used in the study consisted of a declarative pointing task which took into account different social conditions in order to control for the adult's joint attention patterns (Carpenter et al., 1998; Matthews, Behne, Lieven, & Tomasello, 2012). In the baseline condition the adult jointly engaged with the infant's event of reference, while in the critical conditions, the adult either looked at the infant but not at the object of reference (available condition), or was not visually attending to either of them (unavailable condition). The results of the study revealed that infants pointed significantly more, and produced more vocalizations during and after the infant's first point in the available and unavailable conditions than in the baseline condition. Moreover, the adult's social interaction patterns during the unavailable condition triggered less pointing behavior than during the available condition. Therefore, infant communicative responses changed depending on adult attention behaviors. Similarly, Gros-Louis and Wu's (2012) analysis of 12-month-old infants' interactions with their mothers showed that the children were more likely to combine vocalizations with pointing when mothers were not looking at the target event. These studies suggest that infants use pointing gestures and speech intentionally in specific communicative situations, and that they seem to efficiently adapt their communicative behavior to the adults' availability for joint attention.

The current study investigates the predictive value of simultaneous pointing–speech combinations in different social interaction contexts. We will test whether the use of multimodal cues to attract an adults' attention in a communicative context is an important ability related to later language development. Following Liszkowski et al. (2008) and using a similar procedure (i.e., a declarative pointing task carried out under social conditions that differ according to adult joint attention patterns), we aim more specifically to investigate the role of adult interaction patterns in the integrated use of pointing gesture and speech by 12-month-old infants. As already mentioned, the declarative pointing task is especially suited to our purposes for two main reasons: (a) it elicits the production of declarative pointing in a situation in which infants are likely to initiate joint attention with the adult about an event of reference (Carpenter et al., 1998; Matthews et al., 2012); and (b) it controls for the adult's joint attention patterns, thus increasing the child's opportunity to produce simultaneous gesture–speech combinations in more demanding social conditions where the adult is either available (but not jointly attending to the child's object of reference) or unavailable. According to previous research, we expect a greater use of early multimodal productions during the available condition compared to baseline, but also differences in the use of gesture and speech combinations between the available and unavailable conditions. Liszkowski et al. (2008) reported measures of the vocalizations produced during and after the first pointing behavior. Since in their study the vocalizations produced during the first point (multimodal productions) and after the first point (unimodal productions) were grouped together, their results do not reflect a clear measure of the use of pointing–speech combinations which are the focus of the present research. The first goal of our study is thus to replicate and extend prior findings on the role of different social conditions in triggering multimodal communicative strategies, such as simultaneous pointing–speech combinations.

The second goal of the study is to explore the degree to which an infant's early ability to use multimodal, simultaneous pointing–speech combinations at 12 months of age predicts subsequent vocabulary acquisition, with measures at 18 months of age. By using an experimental task which favors a specific declarative intention from the child and which controls for the adult's patterns of responses, we aim at extending results elicited in research by Murillo and Belinchón (2012) and Wu and Gros-Louis (2014) through an experimental task where infants are not interacting with their mothers. This controlled scenario will allow us to more thoroughly analyze the connection between infants' communicative strategies and their language outcomes 6 months later (measures in those studies were obtained just 3 months later, at 15 months of age). It is important to point out that analyzing a child's interaction with an unfamiliar adult (as opposed to mothers or other habitual caregivers) can provide a stronger assessment of infants' communicative abilities because their behavior in the task will not be influenced by prior experience or shared routines in infant-caregiver interactions. In line with previous research showing that the use of combinations of gestures and words are good predictors of both lexical and grammatical development (Iverson & Goldin-Meadow, 2005; Özçaliskan & Goldin-Meadow, 2005; Rowe & Goldin-Meadow, 2009), we expect that the early ability to use simultaneous multimodal combinations at 12 months of age in communicatively demanding social interaction situations will be positively correlated with measures of language growth at 18 months of age. It is suggested that infants' ability to successfully engage the adult in joint attention using a combined, multimodal strategy can increase language-learning opportunities from social interaction contexts.

2. Methods

2.1. Participants

A final sample of nineteen infants ($N = 19$; 12 boys and 7 girls) participated in this study. They had all been born at term, were healthy, and had normal hearing. They were followed longitudinally, first being tested at 12 months of age (mean age: 12; 12; range: 11; 23–12; 27) on a task involving different interaction conditions and then contacted again at 18 months of age to obtain language outcome measures from parental reports. An additional seven infants were initially recruited and tested but had to be excluded from the final sample because of oral habits which interfered with the pointing activity

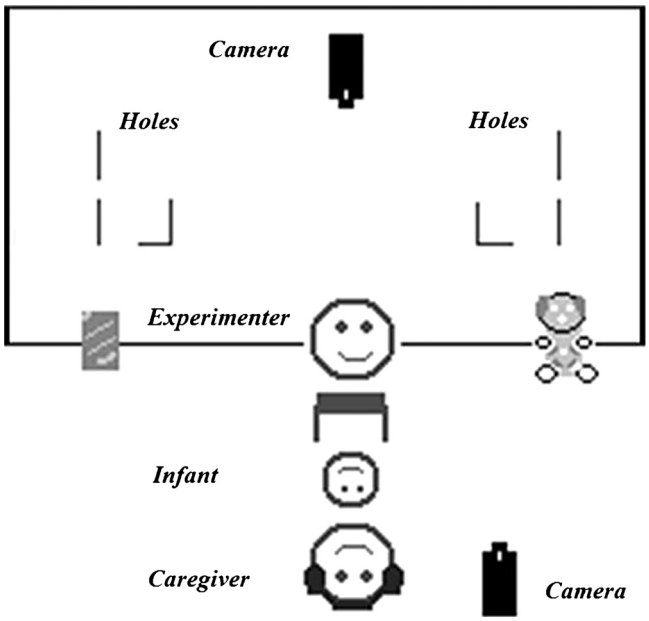

Fig. 1. Schematic representation of the central area within the testing room. The setting includes a curtain with six openings, three on each side, where six of the objects manipulated by an assistant hidden behind it were presented, two cameras (frontal and back position) and two additional objects placed in front of the curtain, on the floor, to the left and right of the experimenter. Locations of the experimenter, child, and caregiver are also indicated.

(2) or crying (1), or because follow-up data on their language outcomes at 18 months was unavailable (4). All participants were raised in monolingual Spanish-speaking homes: six of them were recruited in a monolingual Spanish-speaking area (Albacete) from public nurseries and thirteen participants were recruited from the APAL Infant Lab database in Barcelona. Results from the language exposure questionnaire (Bosch & Sebastián-Gallés, 2001) administered to the caregivers ensured that even the infants recruited in Barcelona, a linguistically mixed city, were predominately spoken to by their parents in Spanish on a daily basis. Exposure to a different language, if present, was sporadic and restricted to encounters with people outside the home environment (percentage of overall exposure to Spanish: median: 100%; mean 90.4%; SD: 12.2). Since the minimum exposure to a second language at home required to qualify infants as bilingual is 25% (Bosch & Sebastián-Gallés, 2001), participants in the present study could be safely qualified as monolinguals. When initially contacted, all families reported that their infant had already begun to point at objects, an essential eligibility criterion for participation in this experiment.

2.2. Experimental setting and materials

The experimental setting was based on Liszkowski et al. (2008). Experimental sessions took place in a 2 m × 3.5 m distracter-free testing room where the portable set-up could be easily placed. A large opaque curtain divided the room into two unequal areas, a central one where the caregiver and experimenter were seated and a small one behind the curtain, where an assistant, hidden from the infants' view, manipulated the objects to be presented during the task. The infant sat on his or her caregiver's lap in the central area at a distance of 2 m from the curtain, facing the experimenter, who had the curtain behind him. A small table was placed between the experimenter and the caregiver (see Fig. 1). The caregiver wore a pair of earphones which continuously played music to distract him or her from the activity and avoid interference with the child's spontaneous behavior.

A video camera was positioned so that it could record the child's reactions through a large opening in the upper center of the curtain while a second camera was positioned at the back of the room to record the sequence of events as seen from the child's perspective. The curtain had three lateral openings on each side through which the puppets were made visible to the child, one at a time (see Fig. 1). These openings (four of them at a distance of 60 cm from the floor and two of them at 100 cm) were symmetrically positioned at about 30° and 25° respectively to the left and right of the infant's direct frontal view. A total of ten different stimuli were manipulated by an assistant hidden behind the screen. These stimuli followed the same characteristics as in Liszkowski et al. (2008), namely six similar hand puppets (a cat, a frog, a cow, a rooster, a sun, and a snail), two different hand puppets (an articulated mouth and a grandmother), and two electronic stimuli (a dancing pig and a light). The latter two were located on the floor approximately 30° to the infant's left and right. Both electronic devices had switches which allowed the assistant to activate and deactivate them from behind the curtain. These electronic stimuli remained inactive except for the trials in which they were used. The labels of all these objects are included in the Spanish version of MacArthur's Communicative Development Inventory for children aged 8–15 months (López-Ornat

et al., 2005). A moveable bead toy and a pair of books were used between conditions to return the infant's attention to the experimenter.

2.3. Procedure

Liszkowski et al.'s (2008) procedure was adapted to elicit a range of infant communicative behaviors (i.e., pointing–speech combinations, pointing-only, and speech-only productions) through an enjoyable event, by presenting puppets or toys from behind the experimenter. The procedure involved social interaction in two experimental conditions (i.e., available and unavailable) and a baseline condition, these three conditions differing in terms of the adult's joint attention patterns as in Liszkowski et al.'s study:

(a) In the *baseline condition*, the experimenter jointly engaged with the infant and the stimulus. First, the stimulus was activated and the experimenter looked at the infant, ignoring the stimulus until the infant had pointed to it. After this first pointing the experimenter reacted with joint attention (i.e., by looking back and forth between the stimulus and the infant's face), pointed to the object, and said things like, "Oh. . ., look, it's a cat!" "Look! It's saying hi to you!" "Oh, a cat!"
(b) In the *available condition*, the experimenter looked at the infant but did not look at the stimulus. First, the stimulus was activated and the experimenter looked at the infant, ignoring the stimulus until the infant had pointed to it. After this first pointing the experimenter maintained eye contact with the infant and did not look at the stimulus, while saying "Hmm? What? What's there? Hmm?"
(c) In the *unavailable condition*, the experimenter attended to neither the infant nor the object. When the stimulus was activated the experimenter's attention was directed at neither the infant nor the stimulus, but rather at the book. Even when the infant pointed to the stimuli, the experimenter continued looking at the book while saying "Hmm? What? What's there? Hmm?"

The testing session was organized in the following way. First, caregivers were informed about the experimental procedure and permission to record was obtained. Then they received general instructions on how to behave during the task: they were instructed to hold the child gently but firmly on their lap in order to maintain the infant's position constant but to avoid all interaction with the child during testing and to avoid looking at the curtain and the objects that would be appearing there. They were encouraged to sit calmly while listening to music through the headphones.

After instructions were given, the warm-up period began. This took place in a separate room and consisted of an enjoyable play activity (lasting from 5 to 20 min) between the experimenter and the infant. Then, accompanied by the caregiver, they moved to the testing room, where the pointing task was carried out. Before baseline trials, there was a short play period with the bead toy to keep the infant interested in the experimenter as a social partner. When the experimenter judged that the infant was relaxed and attentive, he gradually withdrew from the interaction and signaled to the assistant by means of snapping his fingers out of the infant's sight that the first stimulus could be activated. The infant had 20 s within which to initiate a pointing gesture. When the infant pointed for the first time, the stimulus continued to be activated for another 20 s or until he/she showed that he/she was no longer interested in the object by ceasing to look at it by more than 10 s. If no communicative behavior was produced in reaction to the stimuli (i.e., no gestures, vocalizations, or any combination thereof), the stimulus was withdrawn after the first 20 s had elapsed. In all cases, the assistant, who could see the behavior of the infant from one of the holes in the curtain, monitored the duration of the trial and signaled to the experimenter when the trial was finished by clucking her tongue. Before experimental trials, the experimenter and the infant shared a book activity until the infant was relaxed and attentive; then the experimenter gradually withdrew the activity, signaled to the assistant by means of a finger snap to activate the next stimulus while continuing to look at the book, which he held at the opposite side of the infant's field of view relative to where the stimulus was going to appear in front of the screen.

The within-subjects experimental design was organized as sequences of three different types of trials, always starting with a baseline condition, followed by the available and unavailable conditions in a counterbalanced order across participants (i.e., ten participants were tested in the Baseline-Available-Unavailable order and nine participants were tested in the Baseline-Unavailable-Available order). Each sequence was repeated five times, so that each child completed a total of 15 trials. The right or left side where the first stimulus appeared was also presented in a counterbalanced order across participants. Side presentation was alternated from right to left in successive trials until 15 trials were completed. The electronic stimuli were placed at infant's left side (light stimuli) and the right side (dancing pig stimuli), they were only used in those trials when the assistant alternated the position to the left or right side, respectively. The other experimental stimuli (handheld puppets) were randomly protruded by the assistant through one of the three holes of each side. Five of the 10 stimuli were used twice in order to complete 15 trials. The order of presentation of the stimuli and the stimuli to be used for a second time were also randomly chosen by the assistant. The full experimental session lasted approximately 18 min. Following the session, parents were given instructions on how to complete the Spanish version of the MacArthur-Bates Communicative Development Inventories, Words, and Sentences section of the 16–30 months CDI (López-Ornat et al., 2005). They were contacted again and asked to fill out and return the form 6 months later, when their child was 18 months of age.

2.4. Coding and reliability

Coding was performed using ELAN software (Lausberg & Sloetjes, 2009), which is especially well suited for video annotations. Measures of communicative modality were separately obtained for baseline, available, and unavailable conditions. Behaviors corresponding to three different modalities were registered, namely, pointing-only, speech-only, and pointing–speech combinations. In what follows, the specific criteria used for coding the infants' behavior are described.

Pointing-only. Only instances of pointing directed at the target stimulus of the trial were coded, while other communicative gestures (e.g., waving the hand to say "hello" or clapping hands) were not taken into account.

We followed Liszkowski et al.'s (2008) coding of pointing gestures (isolated or in combination with speech), in which pointing gestures were coded when the infant extended the arm (either fully or slightly bent) and index finger or open hand downwards (similarly to Brooks & Meltzoff, 2008; Cartmill, Demir, & Goldin-Meadow, 2012).

Speech-only. This category included any vocalization produced by the infant except infants' fixed signals (e.g., cries, shouts, laughs or groans) or vegetative sounds (e.g., sneezes or burps) (Nathani & Oller, 2001; Oller et al., 2010). Following these authors, we coded vocalizations as independent utterances when they were separated by a silence longer than 300 ms. Then, following Goldstein et al. (2010) and Gros-Louis et al. (2014), we coded infants' gaze direction as pertaining to one of the following categories: stimulus-directed (looking at the target stimulus presented in the trial), experimenter-directed (looking at the experimenter after seeing the target stimulus), looking-caregiver (looking at the caregiver in the moment after seeing the target stimulus), looking-other (looking at other objects in of the room like the books or the bead toy, to a stimulus visually present in the room like the pig or light, or to the caregiver/experimenter when they were not previously looking at the target stimulus). We only included as vocalizations those that clearly referred to the target stimulus, that is, were stimulus-directed, and also those that were experimenter-directed after the child had seen the stimulus.

Pointing–speech combinations. Simultaneous pointing–speech combinations are defined as sharing all pragmatic function, semantic content, and phonological temporal cues (Butcher & Goldin-Meadow, 2000; McNeill, 1992). In the latter regard, the stroke phase of the gesture must coincide with the interval of maximum effort in the gesture. We therefore classified communicative productions as simultaneous pointing–speech combinations by looking at their temporal alignment, so that vocalizations which overlapped with at least some portion of the stroke of the pointing gesture were coded as simultaneous (following Esteve-Gibert & Prieto, 2014; Gros-Louis & Wu, 2012; McNeill, 1992). Such combinations were counted as such only when they were clearly directed at the target stimulus.

After coding, the number of occurrences of each communicative modality (speech-only, pointing-only, or pointing-speech combinations) per trial was obtained and their frequency was computed. Inter-rater reliability was assessed by two observers who had been trained for 2 h in the coding procedure. Observers assessed a total number of 61 trials, which corresponds to 21.4% of the trials across conditions. Agreement for presence/absence of communicative productions in each trial was very high: overall agreement was 96% and the fixed-marginal kappa statistic was 0.90. Observers assessed a total number of 141 infant productions, which corresponded to 43% of the data. The overall agreement for the classification of communicative acts (141 items) into one of the three categories (namely, *speech-only*, *gesture-only* and *pointing–speech combinations)* was 95% and the fixed-marginal kappa was 0.94, indicating that there was substantial agreement among independent coders. Overall agreement for the classification by coders of infant gaze behaviors into one of the four categories (namely, *stimulus-directed, experimenter-directed, caregiver-directed*, and *other-directed*) was 88% and the fixed-marginal kappa statistic was 0.78, indicating that there was substantial agreement among independent coders.

Finally, to control for consistency of the experimenter's behavior within a given condition, we assessed whether the experimenter's expected speech and gesture performance in the three conditions was as defined in the procedure section. We did this by monitoring the experimenter' gaze, gestures, and speech across conditions. The results showed that the experimenter used speech and gesture behaviors as defined in the procedure in 100% and 93.8% of cases, respectively. The following descriptive information shows that in the baseline condition the number of gazes directed at the stimuli after the infant's first pointing (Median = 4; mode = 4; Mean = 3.78; SD = 1.25) and the number of pointing gestures directed at the stimuli (Median = 2; Mode = 2; Mean = 2.14; SD = 0.787) was consistent across participants. In the experimental trials, the gaze behavior before and after infants' first point was correctly performed by the experimenter in 97.9% and 100%, respectively. In trials in which there was no pointing gesture, the experimenter directed his gaze correctly in 98.9% of the trials (that is, at the infant in the baseline and available conditions and at the book in the unavailable condition). Thus the experimenter consistently performed according to the prescribed behaviors in each condition.

3. Results

The results section is divided in two subsections, which correspond to the two main goals of this research: (1) the effects of social conditions on the production of simultaneous pointing–speech combinations at 12 months of age and (2) the predictive value of early pointing–speech combinations with regard to language development measures at 18 months.[1]

[1] The application of generalized linear mixed models (GLMM) is especially suitable for our data because this technique extends the linear model so that the target is linearly related to the factors and covariates via a specified link function, thus allowing the target to have a non-normal distribution and the observations to be correlated (e.g., Nouri, 2010; West, Welch, & Galecki, 2007) All statistical analyses were performed using SPSS Statistics 19.0 (SPSS Inc., Chicago IL).

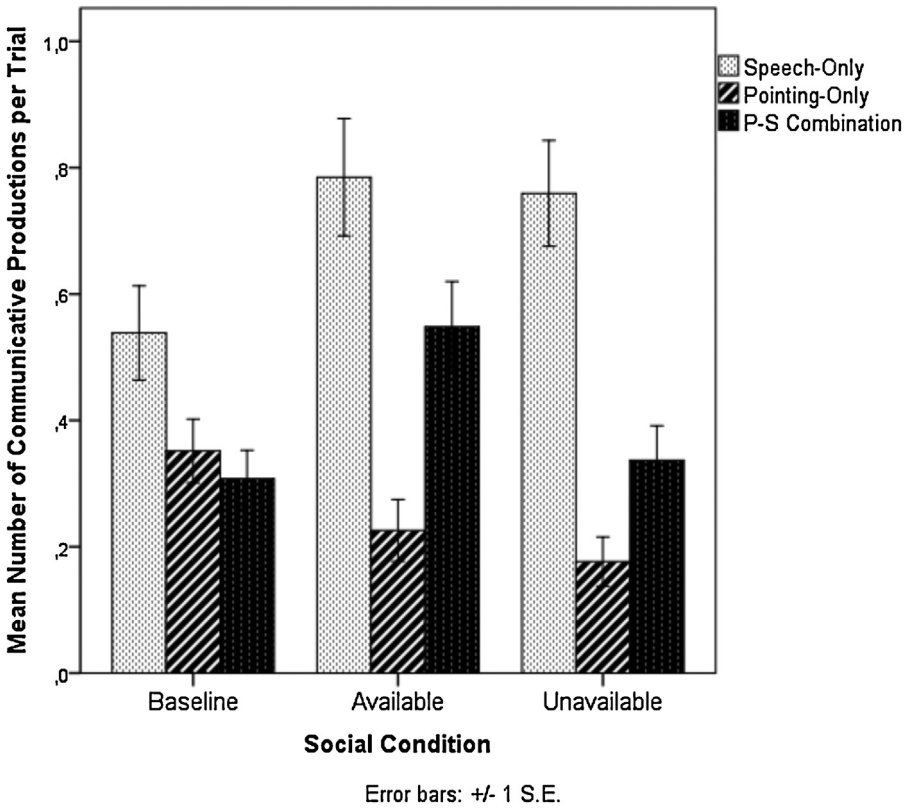

Fig. 2. Mean number of occurrences of each type of communicative production (speech-only, pointing-only, and pointing–speech combinations) per trial as a function of social condition (baseline, available, and unavailable conditions). Error bars ±1 S.E.

First, to check for the potential effects of the stimuli used on the infants' communicative productions, three GLMMs were run with number of communicative productions as a dependent variable. The results of the first analysis show that the number and modality of infants' communicative productions were not affected by the frequency of appearance of each stimulus type ($F(18, 807) = 1.243$, $p = 0.22$). A second analysis revealed that infants' communicative productions were not influenced by stimuli which were presented once (new) or twice (given) ($F(1, 835) = 0.287$, $p = 0.593$). The third analysis revealed that infants' communicative productions were not affected by the side of the screen in which the stimuli were activated ($F(1, 817) = 0.001$, $p = 0.973$). This shows that the variation in the number and location of the stimuli used in the procedure did not influence the number of infants' communicative productions.

3.1. Effects of social condition on the use of pointing–speech combinations

In order to assess the effects of the different social conditions on the number of occurrences per trial of speech-only, pointing-only, and pointing–speech combinations that were elicited in the baseline and the two experimental conditions (i.e., available and unavailable), we conducted a Generalized Linear Mixed Model (GLMM) with number of communicative productions (three levels: speech-only, pointing-only, and pointing–speech combinations) as the dependent variable (Poisson distribution, log link); social conditions (three levels: baseline, available, and unavailable conditions), communicative modality (three levels: speech-only, pointing-only, and pointing–speech combinations), and all their possible interactions as fixed factors; and subject, trial, and social conditions as random factors. Bonferroni paired post hoc tests were carried out for the significant main effects and interactions.

A total of 322 communicative behaviors were coded, including speech-only productions ($N = 174$), pointing-only productions ($N = 76$), and pointing–speech combinations ($N = 72$). The following infants' behaviors were excluded from analysis: communicative behaviors not directed at the stimuli activated/protruded during the trial (speech-only $N = 90$; pointing-only $N = 23$; and pointing–speech combinations $N = 24$) (e.g., books, bead toy, pig, light, holes at the opposite side of the screen, or other points in the room); vegetative and fixed vocal signals like cries or laughter produced during the trial ($N = 49$); and other communicative but non-stimulus-oriented gestures produced during the trial ($N = 62$) (e.g., greeting gestures, clapping gestures, or beat gestures accompanying speech).

The results of the GLMM analysis showed a main effect of communicative modalities ($F(2, 846) = 16.772$, $p < 0.001$), with a greater production of speech-only productions than the other two communicative modalities, that is, pointing–speech combinations and pointing-only productions ($F(2, 846) = 14.833$, $p < 0.001$). The mean number of communicative productions were the following. Baseline trials (speech-only: 0.46 [SD = 0.85]; pointing-only: 0.39 [SD = 0.69]; pointing–speech combination: 0.21 [SD = 0.52]); available trials (speech-only: 0.71 [SD = 1.21]; pointing-only: 0.26 [SD = 0.53]; pointing–speech combination: 0.41 [SD = 0.87]); unavailable trials (speech-only means = 0.61 [SD = 1.09]; pointing-only means = 0.17 [SD = 0.54]; pointing–speech combination means = 0.13 [SD = 0.36]) (see Fig. 2).

With respect to the distribution of communicative modalities in relation to the social conditions presented in the task, there was a significant main interaction between social condition and communicative modality, ($F(4, 846) = 3.754$, $p < 0.01$), indicating that communicative productions behave differently depending on the social condition. More interestingly for our purposes, post hoc analyses of the effect of social condition on the distribution of communicative modalities showed significant effects of simultaneous pointing–speech combinations, with a greater number of productions in the available condition than in the baseline or unavailable conditions, ($F(2, 846) = 6.292$ $p < 0.01$). By contrast, speech-only and pointing-only productions did not significantly differ between social conditions, respectively ($F(2, 846) = 2.816$, $p = 0.06$ and $F(2, 846) = 0.947$, $p = 0.388$). Thus, although speech-only productions

Table 1
CDI scores of infants at 18 months as reported by parents.

	Median	Mean	SD	Min	Max
Expressive vocabulary section	44	65.42	82.82	3	367
Morphosyntax section	6	11.21	15.07	0	67

Table 2
Multiple regression analyses of the most effective models predicting infants' vocabulary and morphosyntax measures at 18 months based on early communicative productions at 12 months during a specific social condition. R^2 statistics and p-values are reported for each model.

Social condition	Communicative production included in the model	R^2 statistic (%)	β-Typified
Model with expressive vocabulary measure at 18 months as dependent variable			
Available	P–S combinations	30.2*	0.550*
Model with morphosyntactic measure at 18 months as dependent variable			
Available	P–S combinations	29.5*	0.543*

noticeably increased from available to baseline condition, only simultaneous pointing–speech combinations increased in number in the available condition with respect to the baseline condition.

Post hoc analyses revealed that communicative modalities did not show significant differences among themselves in frequency of production in the baseline condition (baseline $F(2, 846) = 2.111$, $p < 0.01$). pointing–speech combinations were more frequently produced than pointing-only in available condition, and pointing-only was more frequently produced than pointing–speech combinations in the unavailable condition (available $F(2, 846) = 9.853$, $p < 0.001$; unavailable $F(2, 846) = 11.800$, $p < 0.001$). Again, speech-only productions occurred more frequently than the other communicative productions in available and unavailable conditions, but they did not differ significantly between conditions.

Additionally, taking into account that there were two subgroups exposed to different orders of social condition, i.e., ten infants in order 1 (Baseline-Available-Unavailable trial order), and nine infants in order 2 (Baseline-Unavailable-Available trial order), the possibility of order effects has been analyzed. Infants' communicative productions were significantly different depending on social condition's order of presentation ($F(2, 837) = 5.342$, $p = 0.005$). Post hoc analyses revealed that (a) children tested with order 2 produced a greater amount of speech-only production than those tested with order 1 ($F(1, 837) = 11.594$, $p = 0.001$), (b) infants tested with order 1 produced a greater amount of pointing–speech combinations than infants tested with order 2 ($F(1, 837) = 6.515$, $p = 0.010$), and (b) infants tested with both orders produced similar amounts of pointing-only productions ($F(1, 837) = 0.338$, $p = 0.561$).

3.2. Predictive value of simultaneous pointing–speech combinations for expressive language outcomes at 18 months

The results of the Spanish version of the MacArthur-Bates CDI that were obtained at 18 months of age (as reported by parents) are provided in Table 1. The expressive vocabulary section measures the total number of words that each child produced at that age, while the other section measures the child's ability to produce morphosyntactic features in their utterances. Table 1 shows the median, average, standard deviation (SD), minimum, and maximum CDI scores on vocabulary and grammar section for our sample of nineteen infants at 18 months. By that time, all participants had already begun to produce words, but morphosyntactic markers were still absent in some of them as reported by their parents (note the ranges in the last two right-hand columns). We did not include word-ending measures in the analysis because such responses were very infrequent (median = 0; mean = 1.63; SD = 2.4; min = 0; max = 8).

For each participant the dependent variables, defined as the total number of communicative productions (separated into speech-only, pointing-only, and pointing–speech combination productions) in each of the three social conditions (baseline, available, and unavailable conditions) were obtained.

In order to analyze the predictive value of earlier pointing–speech combinations in different social conditions, one multiple regression analysis was run with each language measure (expressive vocabulary and morphosyntax) as dependent variables. Thus total of two multiple regression analyses were run with communicative productions (i.e., speech-only, pointing-only, and pointing–speech combination) uttered during different social conditions (i.e., baseline, available, and unavailable conditions) as independent variables, using a step-wise method. Table 2 shows the results of the two most effective models of communicative production uttered in trials with different social conditions (independent variables) for the prediction of different language measures at 18 months (dependent variables).

As can be seen in the table, the results of the first regression model, which included communicative productions elicited in the baseline, available, and available conditions, showed significant effects of pointing–speech combinations during the available condition; specially, the model indicated that the number of pointing–speech combinations in the available condition at 12 months explained 30.2% of the expressive vocabulary variance ($R^2 = 0.302$, $F(1, 18) = 7.365$, $p < 0.05$). It was found that simultaneous pointing–speech combinations produced during the available condition were the best predictor of vocabulary measures at 18 months ($\beta = 0.55$, $p < 0.05$). The best model relating communicative productions to morphosyntactic measures reported significant differences in the measures of pointing–speech combinations in the available condition. The results of the regression model indicated that the number of pointing–speech combinations found in the available condition at 12 months explained 29.5% of the morphosyntax section variance at 18 months ($R^2 = 0.295$ $F(1, 18) = 7.365$, $p < 0.05$). In sum, it was found that the number of simultaneous pointing–speech combinations produced during the available condition at 12 months was the best predictor of both vocabulary measures ($\beta = 0.55$, $p < 0.05$) and morphosyntax measures at 18 months ($\beta = 0.543$, $p < 0.05$).

4. Discussion and conclusions

This study set out to investigate the effects of specific contexts of social interaction on the elicitation of multimodal communicative abilities by 12-month-olds, and more importantly, to investigate the predictive value of the integrated use of speech and gesture during these social interaction contexts for later language abilities. We used a declarative pointing task (Carpenter et al., 1998) with controlled social interactions (as used in Liszkowski et al., 2008) to measure infants' communicative productions (i.e., speech-only, pointing-only, and pointing–speech combinations). The first set of results showed significant effects of social condition (baseline, available, and unavailable conditions) on the type of communicative productions produced by children. Infants used pointing–speech combinations in the available condition, i.e., the condition in which the adult looked at the infant and did not look at the event of reference, more frequently than in the baseline condition, i.e., the condition in which the adult showed gaze engagement between the infant and the event of reference.

By contrast, the frequency of use of pointing-only productions did not change significantly between baseline condition and either available or unavailable condition. And though speech-only productions occurred more frequently than pointing-only and pointing–speech combinations, the frequency of oral productions did not differ between baseline condition and either experimental condition. Thus, simultaneous pointing–speech combinations were the only type of communicative production that was significantly more frequent in the available condition than in the baseline condition.

Despite the observed interaction between order of presentation and communicative productions (see Section 3.1), related to infants ability to adapt to the adult responses, the overall predominance of pointing–speech combinations in the available condition remains unaffected.

Interestingly, the ability to fully integrate gesture and speech did not increase when the adult failed to show joint attention in response to a communicative intent of the infant (i.e., in the unavailable condition). There was no increase in any communicative modality between the baseline and unavailable conditions. This result differs from results reported by Liszkowski, Carpenter, Henning, Striano, and Tomasello (2004); Liszkowski et al. (2004, 2008), which showed an increase in the number of pointing productions when the adult was not available to communicate about a referent (unavailable condition) in comparison to trials where the adult shared joint attention with the infant about a referent (baseline condition). In their studies, the experimental conditions were tested between participants, so that trials were presented in the baseline and available conditions for one group of participants and in the baseline and unavailable conditions for a different group. In our study we tested trials belonging to three social conditions within participants (baseline and both experimental conditions). Therefore, the infants' social interaction opportunities to communicate with adults in our study differed from those in previous studies. It is possible that testing the three conditions within participants (e.g., two trials in which the adult was available and one more trial in which the adult was not available to communicate) might have changed infants' communicative strategies such that they would employ predominately pointing behavior when the adult was ready to communicate. This could explain why the frequency of pointing by infants in this study was lower when the adult was not available than in previous studies. Nevertheless, our main results clearly show evidence that infants' use of pointing–speech combinations is dependent on adult attention patterns, with an increase in the number of multimodal productions when the adult is available to communicate but does not look at the referent of the infants' interest.

Our results extend those of Liszkowski et al. (2008) in confirming that the ability to use simultaneous multimodal combinations is employed by children as a repairing strategy to reinforce information related to their communicative goal when the adult does not share attention to the referent but is available to communicate with the infant. On the other hand, the difference between the amount of speech-only productions in the available and baseline conditions only approached but did not reach significance. This result differs from the one related to pointing behavior, whereby the difference in the number of pointing–speech combinations across the two conditions was significant. This difference may be explained by the fact that the experimenter responded to pointing gestures in the baseline and experimental conditions but ignored vocalizations (e.g., the experimenter started to communicate only after the infant's first point). This might have reinforced and promoted the use of pointing gestures. Nevertheless, these results go in the same direction as those yielded by Gros-Louis and Wu (2012) and Wu and Gros-Louis (2014) through naturalistic observations, since they also noted an increase in simultaneous gesture–speech combinations when the adult was available to the infant but not attending to the event or object of interest. Therefore, the boosting effect of multimodal gesture–speech combinations in the available condition may be interpreted as a signal of the intentional ability of the child in a situation in which the adult has yet not seen the object but is crucially expressing an interest by looking at the child. This ability to deploy multimodal means might thus be a good reflection of a better ability to intentionally pursue a communicative goal (Bates et al., 1975, 1979; Liszkowski et al., 2008; Wetherby & Prizant, 1989).

Importantly, the second set of results revealed a predictive relationship between the capacity to produce early multimodal communicative productions at 12 months and language measures at 18 months. Two multiple regression analyses were run to test whether the type and frequency of multimodal communicative productions expressed during a specific social condition significantly predicted later expressive vocabulary and sentence measures at 18 months. The use of simultaneous pointing–speech combinations elicited during the available condition at 12 months was the variable most predictive of expressive vocabulary and morphosyntactic measures at 18 months. Overall, results showed that these two measures of language development were significantly related to early multimodal productions: a total of 30.2% and 29.5% of the variance of vocabulary and morphosynctactic development at 18 months, respectively, could be explained by the frequency of use of simultaneous pointing–speech combinations obtained during the available condition at 12 months. Therefore, the ability to produce pointing–speech combinations was positively correlated to later language measures extracted from parental reports.

The results of this study support the hypothesis that pointing gestures synchronized with speech constitute evidence of a powerful joint engagement ability for infants which is related to later language development. In line with our results, the observation of naturalistic interactions in Wu and Gros-Louis (2014) also revealed that an infant's ability to produce multimodal utterances when an adult looked at the infant but not at an object of interest predicted later language abilities. We have thus extended previous results with the use of an experimental task which controls for social interactions with an unfamiliar adult and favors a given pragmatic intention from the child during the task (e.g., a declarative intention). One of the positive outcomes of our experiment is to show that a set of infants with different family communicative backgrounds react in similar ways to specific social conditions. The fact that infants were interacting with an unfamiliar adult neutralized the possible influence of prior caregiver-child routines and shaped interaction. Moreover, with regard to the potential effects

of the use of multimodal productions on later language development, our results extend the results yielded by Murillo and Belinchón (2012) and Wu and Gros-Louis (2014) to the age of 18 months, when early production of gesture–speech combinations are found to correlate with lexical and grammatical output at that stage of development.

Our first finding on the positive effects of the available condition for the use of simultaneous pointing–speech combinations backs up the hypothesis that infants' sensitiveness to the common conceptual ground of the interlocutor is expressed through multimodal cues (deRuiter, 2000; Holler & Stevens, 2009; Tomasello, 2008). That is, infants at an early stage of pointing–speech multimodal development are able to adjust their response to their interlocutor's knowledge of their shared space. Moreover, this finding also relates to a substantial body of literature on how adult-infant joint attention processes affect infants' communication and language abilities (Carpenter et al., 1998; Hoff, 2006; Tomasello & Farrar, 1986). Yu and Smith (2012) noticed that cooperation between adults and infants favors the creation of optimal visual moments of language learning which reduce distracters from the scene. Likewise, recent studies have found that adult contingent responses (i.e., adult communication following the infant's focus of attention) are linked to the ability to produce simultaneous gesture–speech combinations (Miller & Gros-Louis, 2013; Miller & Lossia, 2013). We still do not how the adults' contingent interactions might affect and the infant's ability to persist with their communicative goal,as well as his or her multimodal abilities. Future studies could test whether or not caregiver-infant contingent interactions are related to the infants' abilities to achieve their communicative goals and their successful use of simultaneous gesture–speech combinations. The fact that the recipient's auditory and visual sensory channels are activated to share attention with the adult may serve as a strategy to reduce the number of distracters from the context. Our interpretation is that simultaneous gesture–speech combinations may work as an effective communicative strategy to highlight a piece of information and reduce ambiguity (see Wagner et al., 2014, for a review). As Goldin-Meadow, Goodrich, and Iverson (2007) suggested, pointing gestures reinforce speech by singling out the referent indicated by the accompanying speech.

Importantly, the results of the present study have shown that the ability to use gesture–speech multimodal integration as a communicative strategy at 12 months is related to later language development. Though firmer conclusions could be drawn on the basis of a greater sample, the results of this study provide important information about early language precursors. Gesture–speech integration may be an early indicator of intentional communicative efficiency in those situations where an infant intends to highlight a piece of information and draw an adult's attention toward an object. This research has shown the importance of an infant's capacity to convey meaning simultaneously in two distinct modalities, gesture and speech, as a precursor of language development. That is, pointing in combination with early speech may be an important signal of intentional communication, in which semantic, pragmatic, and phonological information is integrated for the first time.

Acknowledgments

This research was funded by a Recercaixa 2013–2015 project to Pilar Prieto and two projects from the Spanish MINECO (BFU2012-31995 to Pilar Prieto and PSI-2011-25376 to Laura Bosch). We would like to thank all GrEP and APAL Lab members for their support and comments to earlier versions of this paper, and especially Jorgina Solé for help in sample recruitment and testing assistance, and Joan Borràs-Comes for his help with the statistical analysis of the data. We are grateful to all the subjects and their families for having participated in this longitudinal study.

References

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly: 21.*, 205–226.
Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergency of symbols: Cognition and communication in infancy*. New York: Academic Press.
Bavin, E. L., Prior, M., Reilly, S., Bretherton, L., Williams, J., Eadie, P., et al. (2008). The early language in victoria study: Predicting vocabulary at age one and two years from gesture and object use. *Journal of Child Language: 35.*, (3), 687–701.
Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy: 2.*, (1), 29–49.
Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language: 35.*, (1), 207–220.
Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one-to-two word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–258). New York: Cambridge University Press.
Capirci, O., Iverson, J. A., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language: 23.*, (3), 645–673.
Carpenter, R. L., Mastergeorge, A. M., & Coggins, T. E. (1983). The acquisition of communication intentions in infants eight to fifteen months of age. *Language and Speech: 26.*, 101–116.
Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society of Research in Child Development: 63.*, (4), 1–143 (Serial No. 255).
Cartmill, E. A., Demir, O. E., & Goldin-Meadow, S. (2012). Studying gesture. In E. Hoff (Ed.), *Research methods in child language: A practical guide*. Blackwell Publishing.
Caselli, M. C., Rinaldi, P., Stefanini, S., & Voterra, V. (2012). Early action and gesture vocabulary and its relation with word comprehension and production. *Child Development: 83.*, (2), 526–542.
Colonnesi, C., Stams, G. J., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review: 30.*, (4), 352–366.
deRuiter, J. P. (2000). The production of gestures and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge University Press.
Esteve-Gibert, N., & Prieto, P. (2014). Infants temporally coordinate gesture–speech combinations before they produce their first words. *Speech Communication: 57.*, 301–316.
Goldin-Meadow, S., Goodrich, W. G., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science: 10.*, (6), 778–785.

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy: 15.*, (4), 362–391.
Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy: 19.*, (4), 385–408.
Gros-Louis, J., & Wu, Z. (2012). Twelve-month-olds' vocal production during pointing in naturalistic interactions: Sensitivity to parents' attention and responses. *Infant Behavior and Development: 35.*, (4), 773–778.
Hoff., E. (2006). How social contexts support and shape language development. *Developmental Review: 26.*, (1), 55–88.
Holler., J., & Stevens, R. (2009). The effect of common ground on how speakers use gesture. *Journal of Language and Social Psychology: 26.*, (1), 4–27.
Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science: 16.*, (5), 367–371.
Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology: 47.*, (5), 1207–1219.
Laakso, M. L., Poikkeus, A. M., Katajamäki, J., & Lyytinen, P. (1999). Early intentional communication as predictor of language development in young toddlers. *First Language: 19.*, (56), 207–231.
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, and Computers: 41.*, (3), 841–849.
Liszkowski, U., Albrecht, K., Carpenter, K., & Tomasello, M. (2008). Infants' visual and auditory communication when a partner is or is not visually attending. *Infant, Behavior and Development: 31.*, (2), 157–167.
Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science: 7.*, (3), 297–307.
López-Ornat, S., Gallego, C., Gallo, P., Karousou, A., Mariscal, S., & Martínez, M. (2005). *Inventario del desarrollo MacArthur: Versión española*. Madrid: TEA.
Matthews, D., Behne, T., Lieven, E., & Tomasello, M. (2012). Origins of the human pointing gesture: A training study. *Developmental Science: 15.*, (6), 1–14.
McGillion, M. L., Herbert, J. S., Pine, J. M., Keren-Portnoy, T., Vihman, M. M., & Matthews, D. (2013). *IEEE Transactions on Autonomous Mental Development: 5.*, (3), 240–248.
McNeill, D. (1992). *Hand and Mind*. Chicago: The Chicago University Press.
Miller, J. L., & Gros-Louis, J. (2013). Socially guided attention influences infants' communicative behavior. *Infant Behavior and Development: 36.*, (4), 627–634.
Miller, J. L., & Lossia, A. K. (2013). Prelinguistic infants' communicative system: Role of caregiver social feedback. *First Language: 33.*, (5), 433–448.
Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development: 21.*, (3), 469–482.
Murillo, E., & Belinchón, M. (2012). Gestural-vocal coordination. Longitudinal changes and predictive value on early lexical development. *Gesture: 12.*, (1), 16–39.
Nathani, S., & Oller, D. K. (2001). Beyond ba-ba and gu-gu: Challenges and strategies in coding infant vocalization. *Behavior Research Methods, Instruments, and Computers: 33.*, (3), 321–330.
Nouri, M. (2010). . *IBM statistics advanced statistics : Vol. 19.* (Vol. 19) Prentice Hall.
Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., et al. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America: 107.*, (30), 13354–13359.
Özçaliskan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition: 96.*, (3), 101–113.
Pizzuto, E., Capobianco, M., & Devescovi, A. (2005). Gestural–vocal deixis and representational skills in early language development. *Interaction Studies: 6.*, (2), 223–252.
Rollins, P. R. (2003). Caregivers' contingent comments to 9-month infants: Relationships with later language. *Applied Psycholinguistics: 24.*, (2), 221–234.
Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language development. *Developmental Science: 12.*, (1), 182–187.
Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development: 72.*, (3), 748–767.
Tomasello, M. (2008). *The origins of human communication*. MIT Press.
Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development: 78.*, (3), 705–722.
Tomasello, M. (1988). The role of joint attentional processes in early language development. *Language sciences: 10.*, (1), 69–88.
Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development: 57.*, (6), 1454–1463.
Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech interaction: An overview. *Speech Communication: 57.*, 209–232.
West, B., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. New York: Chapman & Hall/CRC.
Wetherby, A. M., & Prizant, B. M. (1989). The expression of communicative intent: Assessment guidelines. *Seminars in Speech and Language: 10.*, (1), 77–91.
Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language: 34.*, (1), 72–90.
Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition: 125.*, (2), 244–262.