

Journal of Child Language

<http://journals.cambridge.org/JCL>

Additional services for *Journal of Child Language*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants

NÚRIA ESTEVE-GIBERT and PILAR PRIETO

Journal of Child Language / Volume 40 / Issue 05 / November 2013, pp 919 - 944
DOI: 10.1017/S0305000912000359, Published online: 10 August 2012

Link to this article: http://journals.cambridge.org/abstract_S0305000912000359

How to cite this article:

NÚRIA ESTEVE-GIBERT and PILAR PRIETO (2013). Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. *Journal of Child Language*, 40, pp 919-944 doi:10.1017/S0305000912000359

Request Permissions : [Click here](#)

Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants*

NÚRIA ESTEVE-GIBERT

Universitat Pompeu Fabra, Spain

AND

PILAR PRIETO

ICREA – Universitat Pompeu Fabra, Spain

*(Received 9 November 2010 – Revised 3 August 2011 – Accepted 21 June 2012 –
First published online 10 August 2012)*

ABSTRACT

There is considerable debate about whether early vocalizations mimic the target language and whether prosody signals emergent intentional communication. A longitudinal corpus of four Catalan-babbling infants was analyzed to investigate whether children use different prosodic patterns to distinguish communicative from investigative vocalizations and to express intentionality. A total of 2,701 vocalizations from 0;7 to 0;11 were coded acoustically (by marking pitch range and duration), gesturally, and pragmatically (by marking communicative status and specific pragmatic function). The results showed that communicative vocalizations were shorter and had a wider pitch

[*] An earlier version of this article was presented at the VI Conference on Language Acquisition (Barcelona, 8–10 September 2010). We would like to thank participants at that meeting, and especially S. López-Ornat, J. Trueswell, and L. Bosch. We are grateful to the editor, the action editor and the two reviewers for their comments, which have been very helpful to us in revising the text. We also thank M. Armstrong, J. Borràs-Comes, and S. Berends for the reliability coding, M. M. Vanrell for her help with statistics, and Paolo Roseano for his help with Praat figures, all of them members of the Grup d'Estudis de Prosòdia. Finally, we thank the children and the children's parents for voluntarily taking part in this study. This research has been funded by three research grants awarded by the Spanish Ministerio de Educación y Ciencia, FFI2009-07648/FILO 'The role of tonal scaling and tonal alignment in distinguishing intonational categories in Catalan and Spanish', by the Consolider-Ingenio 2010 (CSD2007-00012) Program, and by a grant awarded by the Generalitat de Catalunya to the Grup d'Estudis de Prosòdia (2009SGR-701). Address for correspondence: Universitat Pompeu Fabra – Departament de Traducció i Ciències del Llenguatge, Edifici Roc Boronat 138, Barcelona – 08018, Spain. tel: 935422409; e-mail: nuria.esteve@upf.edu

range than investigative vocalizations and that these patterns in communicative vocalizations depended on the intention of the vocalizations: requests and expressions of discontent displayed wider pitch range and longer duration than responses or statements. These results support the hypothesis that babbling children can successfully use a set of prosodic patterns to signal intentional speech.

INTRODUCTION

A number of studies have investigated early prosodic patterns in babbling infants. Some of them have focused on the presence or absence of language-specific prosodic patterns in terms of contour direction, metrical bias, or syllable duration (Davis, MacNeilage, Matyear & Powell, 2000; Engstrand, Williams & Lacerda, 2003; Kent & Murray, 1982; Levitt & Utman, 1992; Lieberman, 1967; Mamepe, Friederici, Christophe & Wermke, 2009; among others). Although it is well known that adults use prosody to express communicative intentions, attitudes, and meanings, this first group of studies investigated prosodic development irrespective of the potential differences in the pragmatic meaning of the vocalizations. A second group of studies did not incorporate intentionality as a factor in their analysis of prosodic development but, when discussing results, they stated that the differences they found in contour direction could be due to communicative purposes (Whalen, Levitt & Wang, 1991) or to a dynamic relationship between physiological constraints and emotional experience (Snow, 2006; Snow & Balog, 2002).

A third group of studies, however, investigated the emergence of communicative intention in relation to prosody. Many of them have analyzed children at the one-word stage, finding that at this stage children produce adult-like prosodic contours to express distinct pragmatic intentions (Astruc, Prieto, Payne, Post & Vanrell, *in press*; Balog & Brentari, 2008; Balog, Roberts & Snow, 2009; Flax, Lahey, Harris & Boothroyd, 1991; Furrow, 1984; Furrow, Podrouzek & Moore, 1990; Galligan, 1987; Marcos, 1987; Prieto, Estrella, Thorson & Vanrell, 2012; Vihman & DePaolis, 1998; Vihman, DePaolis, & Davis, 1998). In a longitudinal study from the babbling stage to the one-word and two-word, Halliday (1975) analyzed his son's early pitch contours from 0;9 to 2;6 and discovered that different vocal expressions were able to convey distinct functions. Halliday found that his child produced mid falling tones when interacting with other people but low falling tones with narrower range when he was interested in the modification of an object. Also, he found that at 1;0 his son produced requests with rising tones.

Within this last group of studies investigating prosodic development with respect of intentionality, only a few of them have analyzed infants during

the pre-babbling and babbling periods. D'Odorico and Franco (1991), for instance, acoustically analyzed the vocalizations produced by five Italian-learning children from 0;4 to 0;11, in terms of mean *f₀* values, maximum and minimum pitch, melody type structure and units of vocalizations in a prosodic unit, and mean duration. As for context types, vocalizations were classified as vocalizations during infant manipulation of a toy (VIM), vocalizations during shared experience (VSE, i.e. manipulating a toy but looking at the adult), vocalizations during adult manipulation of a toy (VAM), and vocalizations during exchanges with the adult (VEA, i.e. neither of them is manipulating the toy but they are both looking at each other). Results offered support for a 'selective production hypothesis' whereby different types of vocalizations were produced in different communication contexts until children were 0;9. Thus, children at 0;4-0;6 used different contour directions when producing a VIM and a VSE; at 0;6-0;8 children assimilated categories VSE and VAM; and at 0;8-1;0 VIM vocalizations could not be distinguished from the other vocalizations. The authors hypothesized that a child's ability to acoustically distinguish between categories tends to disappear as age increases. Therefore, until 0;9 but not thereafter children show a selective production hypothesis, i.e. different patterns of non-segmental features characterize sounds produced in different contexts. Because their results revealed many individual differences among their infant subjects, the authors concluded that they had failed to capture communicative differences across contexts.

Papaeliou, Minadakis and Cavouras' (2002) study represented a step forward in identifying the prosodic cues that children use in the babbling period to express intentionality. They examined the acoustic patterns of six English-speaking infants from 0;7 to 0;11 and acoustically analyzed vocalizations expressing either emotions or communicative functions. According to Trevarthen (1990), vocalizations expressing emotions identify the quality of communication, whereas vocalizations expressing communicative functions identify the direction and purpose of communication. They analyzed the following features in the vocalizations: duration; initial, final, peak, lowest, and mean *f₀* values; range of *f₀*; standard deviation of *f₀*; ratio of standard deviation of *f₀*; and duration of the vocalization. The meaning of the vocalizations was assigned by interviewing mothers about the meaning they would attribute to their infant's vocalizations, a system that, according to the authors, simulates the natural conditions of communication. They found that prosodic patterns were different when vocalizations conveyed communicative functions from when they expressed emotions: vocalizations carrying communicative functions were shorter, with lower *f₀* values, and had greater intensity than vocalizations expressing emotions. Similarly, Papaeliou and Trevarthen (2006) found evidence that prelinguistic vocalizations can be a tool for both communicating and thinking. They observed four

English-speaking infants from 0;7 to 0;11 and classified their vocalizations as 'communicative' or 'investigative' according to concurrent non-vocal behaviors. They considered a vocalization to be investigative if the infant was holding an object, inspecting an object, or completing a task; they considered it communicative if the child was interacting with an adult, pointing, directing eye-gaze at the adult, and reaching or giving something. They observed that children displayed different prosodic patterns when vocalizations were classified as communicative relative to when they were classified as investigative: compared to investigative vocalizations, communicative vocalizations had a higher mean and maximum *f₀*, higher standard deviation of *f₀*, and shorter duration.

All in all, very few studies have investigated infants' use of prosodic contours to express distinct pragmatic functions when children are younger than 1;0. Even though it has been found that infants can produce adult-like prosodic patterns at the one-word stage, little is known about whether intentional differences influence the prosodic patterns of vocalizations at an earlier age. The purpose of the present study is to investigate whether children express intentionality by means of prosodic cues when they are still not able to produce words; and, if they do so, how they do it.

Thus, the goal of this study is twofold. First, it seeks to investigate whether Catalan-babbling children use prosodic cues such as pitch range or duration to distinguish between communicative and investigative (non-communicative) vocalizations during the second half of their first year, since it is during this period that children start communicating intentionally (Piaget, 1936; Trevarthen, 1977; 1979; 1982; and others). We analyzed a total of 2,701 naturalistic vocalizations recorded from four Catalan-speaking children at 0;7, 0;9, and 0;11. Following Papaeliou and Trevarthen (2006), our hypothesis was that children's investigative vocalizations would be produced with a narrower pitch range and longer duration than communicative utterances. If this hypothesis were corroborated, Papaeliou and Trevarthen's (2006) results would be strengthened with a language other than English and with a wider corpus, since that study tested only 193 vocalizations and the current study includes over 2,000 vocalizations. Second, our study aims at discovering whether Catalan-babbling infants are able to use such prosodic cues (pitch range and duration) consistently in order to express distinct pragmatic functions such as request, discontent, response, or statement. This second goal represents a step forward in the analysis of how the development of prosody is related to the emergence of communicative intention, given that previous studies on prosodic development of babbling children did not take into account pragmatic considerations. In general, we hypothesized that (1) babbling children will display a consistent use of prosodic cues to distinguish communicative from investigative vocalizations, based on results found in previous studies, and

that (2) when intending to communicate, babbling children will also select prosodic cues to convey specific pragmatic intentions. Previous studies found that prosody is used by babbling children to signal the communicative status of a vocalization. Therefore, the corroboration of the first hypothesis would confirm results from prior studies. However, to our knowledge, no studies have investigated whether babbling children use prosody to distinguish between specific intentions, even though the babbling period in language development is known to coincide with the children's development of intentionality. Verifying our second hypothesis, then, would suggest that prosody is a tool that children use during the babbling period to express communicative intentions.

METHOD

Participants

Four Catalan-learning infants participated in the study, two male (Bi and Ma) and two female (An and On). Infants were recorded weekly from 0;7 to 0;11. The present study analyzes children's vocalizations at ages 0;7, 0;9, and 0;11. If we take Piaget's four stages of cognitive development as a reference, the period of interest would be included in the late 3rd and the 4th sub-stages of the sensorimotor stage. It is during these sub-stages that intentionality and logic emerge, starting with intentional grasping of a desired object and differentiating between means and goals, and ending up with the coordination of schemes and intentionality, and planning steps to achieve an objective.

All parents of the four participants speak exclusively Catalan to their child and to each other. Parents were asked about their linguistic habits through a questionnaire, and results showed that all four mothers have Catalan-speaking parents, have lived in Catalonia all their lives, and have Catalan as their first language (L1). They use Catalan in all dealings with their family, work colleagues, and friends. As for fathers, three of them have Catalan-speaking parents, and have always lived in Catalonia. Catalan is their L1 as well as the vehicular language for family, work, and friends. An's father, however, has Spanish-speaking parents and uses Spanish as the primary language for communicating with his parents and work colleagues. However, he speaks and writes Catalan fluently, and uses it with his wife, daughter, and friends. The children come from four small towns in the same region of Catalonia, Alt Penedès, located 50 km to the south of Barcelona. According to the information available from the official statistics website of Catalonia (www.idescat.cat, Linguistic census from 2001), in three of these towns Catalan is spoken regularly by about 90 percent of the population, and in the fourth town Catalan is spoken by 80 percent of the population. Thus, it may be safely assumed (and also according to the

parents' reports) that there is very little Spanish influence in the children's linguistic input, since children are not exposed to Spanish at home and hear very little of it outside the home.

Data collection

All children were video-recorded at their homes during weekly 30-minute sessions between ages 0;7 and 0;11 using a SONY camera, model DCR-DVD202E PAL. Thus, they were all recorded three to five times per month, except for Bi at 0;9 and On at 0;11, who were recorded only twice during those months due to illness. Recordings were made by the first author of this study, who was previously acquainted with the families and children. Children were always recorded in the same room of their respective homes, typically their living-rooms, during free-play sessions. All children were recorded as they interacted with their mothers, except for one child, An, who was recorded while interacting with both her father and her mother in most of the sessions. A tripod was used, placed as close to the child as possible and positioned so that the camera was pointing toward the child's face.

In order to monitor vocabulary acquisition, the same set of toys was given to the child in all sessions. The first toy offered, a pyramid of four colored plastic stackable disks with animal heads, was common to all four infant subjects and available to them only during the recording sessions. When subjects lost interest in this toy (which tended to happen after about ten minutes), their parents offered them another toy from the child's own collection, usually the same toys from one recording session to the next.

From all the weekly sessions recorded during this six-month period, we selected for analysis vocalizations produced when the children were 0;7, 0;9, and 0;11. These ages were selected based on the hypothesis that these vocalizations would display the typical features of certain stages of development: before the onset of intentional communication, when intentionality starts, and when intentionality is already developed (Piaget, 1936; Trevarthen, 1977; 1979; 1982; and others).

Data analysis

The approximately 18 hours of recordings were segmented into 2,946 vocalizations. From these, 245 were excluded from the analysis because of the following circumstances: (1) child and parent overlapped when vocalizing, (2) ambient noise was too loud, (3) the child vocalized while having an object inside his/her mouth, or (4) the sound did not show a visible trace on the spectrogram. This yielded a corpus of 2,701 vocalizations.

Before segmenting the data, we established the unit of analysis of our study. Following Papaeliou and Trevarthen (2006), two utterances were

considered distinct vocalizations if they were separated by 50 ms or more. Additionally, when there were more than 50 ms between two vocalizations, but their prosodic contours were linked by a sustained fall at the end of the first vocalization followed by a second vocalization starting at that sustained *fo* level, they were not separated but considered the same vocalization.

Pragmatic analysis. All vocalizations were first annotated by one coder in terms of the communicative function they conveyed using the Phon software system (Rose *et al.*, 2006). Different authors have dealt with the classification of pragmatic functions of early vocalizations in different ways. As noted above, D'Odorico and Franco (1991) used the terms 'vocalizations during infant manipulation of a toy', 'vocalizations during shared experience' (manipulating a toy but looking at the adult), 'vocalizations during adult manipulation of a toy', and 'vocalizations during exchanges with the adult' (neither of them is manipulating the toy but they are both looking at each other). Blake and Boysson-Bardies (1992) classified their subjects' vocalizations using the following labels: fine object manipulation, gross object manipulation, upright movement, confined movement, request, comment, book-reading, demonstrative, response to adult's utterance, give and take, rejection-protest, or physical interaction. In addition, Sarriá (1991) and Karousou (2003) used these categories: request (object, help, or attention), rejection, protest, satisfaction, question (what, where, and how), statement, proto-conversation, narration, interactive game, imitation, non-social, or greeting.

Since the first aim of our study was to discover whether the vocalizations of Catalan-babbling children conveying communicative information are different from vocalizations that did not intend to communicate information, we first classified our data into one or the other, labeled respectively 'communicative' or 'investigative'. Following Papaeliou and Trevarthen (2006), a vocalization was considered to be investigative if the infant was holding an object, inspecting an object, or completing a task; a vocalization was considered to be communicative if the child was interacting with an adult, pointing, directing eye-gaze at the adult, and reaching or giving something. Thus, the distinction between communicative and investigative vocalizations relied mostly on gestural cues, as well as context and parental reactions before or after the vocalization.

Apart from the labels 'investigative' and 'communicative', an extra category was used to classify all those utterances that were difficult to label. Thus, 'not clear' was the label used when visual cues were not clear enough to decide whether a vocalization was communicative or not. For instance, when the child was vocalizing but her hand or face was not visible in the video (e.g. behind the sofa), it was included in the 'not clear' group. The presence of this third category enhances the reliability of the results, since no vocalization was forced to fit into one of the other two categories

described above. A total of 324 vocalizations were labeled as ‘not clear’ following this criterion. Thus, of a sum of 2,701 recorded vocalizations, our analysis yielded a total of 1,676 communicative vocalizations, 701 investigative vocalizations, and 324 vocalizations whose purpose was ‘not clear’.

In order to test the second hypothesis, i.e. whether children select certain prosodic cues to express distinct pragmatic functions, all communicative vocalizations were further classified into narrower categories depending on the specific pragmatic functions the child was judged to be performing. The pragmatic functions adopted were based on Sarriá (1991) and Karousou (2003). The specific intentions used were discontent (the child expressed ‘sadness’ actively), request (the child wanted the other person to do something), response (the child reacted to a stimulus, either a verbal stimulus uttered by an adult or an action performed by the adult), satisfaction (the child expressed happiness about the current situation), statement (the child vocalized simply because (s)he wanted the adult to know something), surprise (the child wished to express the idea that an unusual or unexpected event had occurred), and vocative calling (the child called somebody). Hence, the pragmatic analysis consisted not only of deciding whether a vocalization was communicative or investigative but also of deciding whether that vocalization bore a specific intentionality. In order to screen out the potential influence of prosodic cues in the audio material, this specific classification was performed only when the recording displayed clear contextual and non-vocal information. All those communicative vocalizations that were impossible to classify further into one specific pragmatic meaning were included in a category called ‘fuzzy intention’. Thus, when a vocalization was clearly communicative but too fuzzy to fit in any of these specific pragmatic categories, it was labeled as ‘fuzzy intention’. Such cases represented 745 out of the 1,676 communicative vocalizations. In sum, all vocalizations relevant for our study were first classified as ‘communicative’, ‘investigative’, or ‘not clear’. Next, the group of ‘communicative’ vocalizations was further subdivided into the specific pragmatic functions. These classifications were conducted on the basis of audio and visual cues in the recordings. Importantly, in order to minimize the potential influence of prosodic/acoustic cues in determining the communicative status and specific intention of vocalizations, the pragmatic and gestural analyses of all vocalizations (performed independently using Phon) were performed prior to the acoustic analysis (performed independently using Praat) (see the following sections).

To test the reliability of the pragmatic coding, an inter-transcriber reliability test was conducted with a subset of 20% of the total number of vocalizations in the target materials (which represented a total of 540 utterances), making sure that all children and ages were uniformly represented.

Three independent coders labeled a random selection of 20% of the data in terms of communicativeness and specific pragmatic intentions. The overall agreement was 82% when deciding whether the vocalization was communicative or not, and 74% when deciding on specific pragmatic intentions. The fact that the overall agreement was lower when rating specific intentionality than when rating the communicative status might be due to the fact that in the former case raters had to choose among a considerably higher number of categories or because some of the specific intentions were more difficult to categorize. For instance, raters sometimes found it difficult to distinguish between the categories 'discontent' and 'request' because in some cases a child might urge the adult to do something while expressing sadness. All in all, we think that these scores reveal a substantial agreement among raters and are comparable with other studies' scores (Chen & Kent, 2009; Papaïiou & Trevarthen, 2006). Chen and Kent (2009), for instance, achieved an overall agreement of 84% in their inter-transcriber reliability test.

Gesture analysis. The gestural analysis was performed in parallel with the pragmatic analysis described above. As is well known, children begin to gesture very soon in order to influence the mental state of others, i.e. because they want others to do, know, or feel something (Tomasello, Carpenter & Liszkowski, 2007). The first communicative gestures that typically developing children produce are deictics such as pointing, giving, showing, or requesting (Bates, Benigni, Bretherton, Camaioni & Volterra, 1979; Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005; Sansavini, Guarini & Stefanini, 2010; Tomasello *et al.*, 2007). Each vocalization was annotated in terms of the gestures displayed by children when vocalizing, using the Phon software system (Rose *et al.*, 2006). All vocalizations were labeled with gestural information regarding gaze direction, manual gestures, and facial gestures. A simplified version of Allwood, Cerrato, Jokinen, Navarretta and Paggio's (2007) categories was adopted in the present study for the annotation of infants' gestures: hand gestures were defined in terms of handedness (single hand, both hands), hand trajectory (up, down, sideways, etc.), and their semiotic and communicative value; facial gestures were defined in terms of general face, position of the eyebrows, eye position, gaze direction, form of the mouth, head position, and their semiotic and communicative value. This codification system was chosen because it enabled us to code gestures independently of their possible meaning or function, using the system's labels regarding the form of the gesture. Table 1 shows the gesture categories used in our study.

Acoustic analysis. The main aim of this study was to find out whether different prosodic patterns are at play when infants try to communicate or convey a set of pragmatic functions. In order to perform the acoustic analysis, we manually extracted all the audio files (in .wav format) from our

TABLE I. *Gesture categories used in the gesture analysis*

<i>Gaze direction</i>	absent gaze gaze at camera gaze at object gaze at parent
<i>Manual gestures</i>	clapping hands extending arms hugging parent manipulating object moving arms pointing at object moving hands shaking arms no specific manual gesture
<i>Facial gestures</i>	furrowing brows opening eyes closing eyes opening mouth closing mouth pouting shaking head smiling raising eyebrows no specific facial gesture

Phon corpus and analyzed them with the Praat software package (Boersma & Weenink, 2005). Also, no information on child, age, pragmatic intention, or gesture was at the coder's disposal when annotating the acoustic measures, in order to guarantee that there would be no influence of pragmatic coding on the determination of acoustic parameters.

Two prosodic features were manually labeled: duration and pitch range, i.e. start and end points of vocalizations, and pitch maximum and minimum points. The aim was to analyze the global pitch range of the contour and total duration, which are the features that are most commonly used in studies of the prosody of infants' vocalizations (Marcos, 1987; Papaeliou *et al.*, 2002; Papaeliou & Trevarthen, 2006; Scherer, 1986). As for pitch range, an overview of the data indicated that the best way to obtain this measure was to select three pitch points from the fundamental frequency contour. These three pitch points were distributed along the fundamental frequency line and included the lowest (f_0 min) and highest points (f_0 max). The first pitch point (p1) was selected at the onset of vocalization, since this point is usually referred to as the reference level of the speaker; the second pitch point (p2) was generally selected at a point in the middle of the f_0 contour; and finally, the third point (p3) was usually selected at the end of the vocalization. However, when the lowest or highest pitch values did not appear at the very beginning, at the very end, or right

in the middle of the vocalization, the points selected were moved according to our needs in order to make them coincide with the lowest and highest point.

In percentages, the lowest *f₀* point was mostly located at p₃ (50.47% of cases) or p₁ (40.02% of cases); the lowest *f₀* point was located at p₂ for just 9.51% of the vocalizations. The highest *f₀* point was located at p₂ in 72.75% of cases, and was less frequently located at p₁ (16.66%) or p₃ (10.59%). When these points were annotated, the pitch maximum and pitch minimum values were extracted using a Praat script, and the pitch range was calculated by subtracting the pitch minimum from the pitch maximum. In order to compare different pitch ranges across the four children, pitch values were extracted in semitones rather than in Hz.

Additional considerations for determining the *f₀* index measurements were as follows:

- When the vocalization had more than one peak point at the same level, the last point was selected.
- If the vocalization displayed no clear peak, a pitch point in the middle of the vocalization was selected.

In order to obtain the total duration of the vocalization, the first point (t₁) and last point (t₂) in the *f₀* line of the vocalization were selected. Following Papaeliou and Trevarthen's (2006) work, we considered two sounds to be distinct vocalizations if they were separated by at least 50 ms. When there was 50 ms between two vocalizations but they were prosodically linked, they were considered one vocalization.

Figure 1 illustrates how vocalizations were annotated in terms of pitch range and duration. Below the *f₀* contour, the first tier was used to annotate start and end time of the vocalization (t₁, t₂), and the second tier was used to annotate the three index pitch points (p₁, p₂, p₃) to later calculate pitch range values. The upper graph is an example of an investigative vocalization and the lower graph is a communicative vocalization.

RESULTS

This section includes two different parts. The first part presents the results of the analysis of the potential effects of the communicative status of the vocalization on prosodic cues (i.e. pitch range and duration). The second part presents the results regarding the potential effects of the pragmatic function on prosodic cues (i.e. pitch range and duration).

All statistical analyses in this article were performed by applying a linear mixed model (LMM; West, Welch & Galecki, 2007) using SPSS Statistics 15.0 (SPSS Inc., Chicago IL). West *et al.* (2007) state that LMMs are the appropriate model for analyzing unbalanced longitudinal data, since they

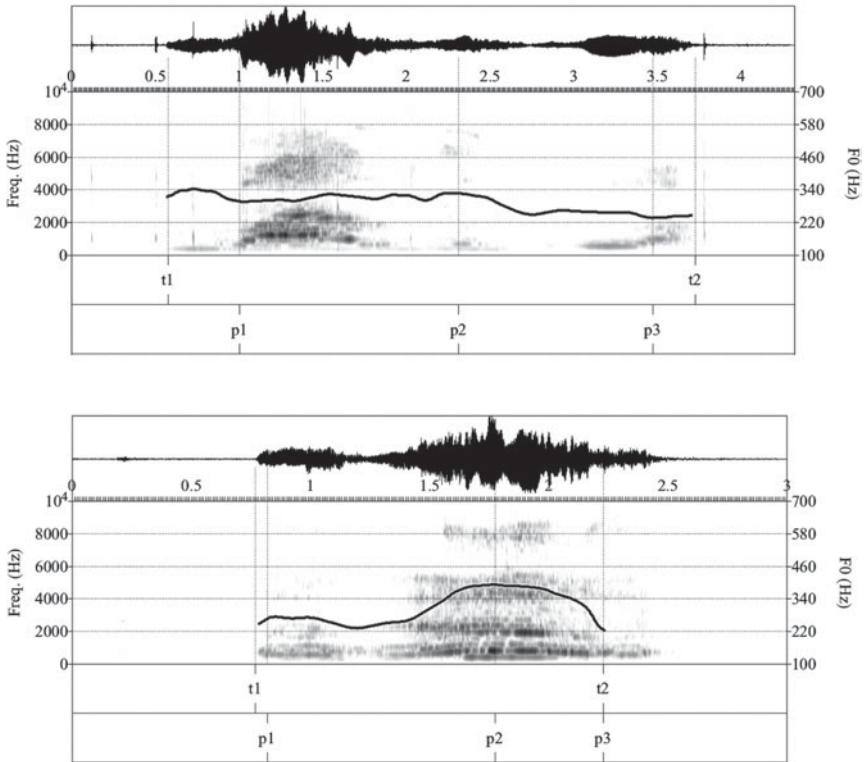


Fig. 1. Example of an annotated investigative vocalization (top) and a communicative vocalization (bottom) performed by Ma at 0;9.

allow for subjects with missing time points (i.e. unequal measurements over time for individuals), have the capacity to include all observations available or all individuals in the analysis, and cope with missing data at random. As West *et al.* (2007) point out, linear mixed models can accommodate all of the data that are available for a given subject, without dropping any of the data collected from that subject.

Prosodic cues and communicativeness

Table 2 and Figure 2 show a general overview of the data included in the analysis. Table 2 displays the number of vocalizations produced by each child at each age, and their classification according to the communicative status. Figure 2 shows the percentage of ‘communicative’, ‘investigative’, or ‘not clear’ vocalizations across the different ages. The results in both Table 2 and Figure 2 reveal that children produce more communicative

TABLE 2. *Number of vocalizations classified in terms of communicative status and age*

	0;7	0;9	0;11	Total
'communicative' vocalizations	384	465	827	1,676
'investigative' vocalizations	226	261	214	701
'not clear' vocalizations	85	102	137	324
TOTAL	695	828	1,178	2,701

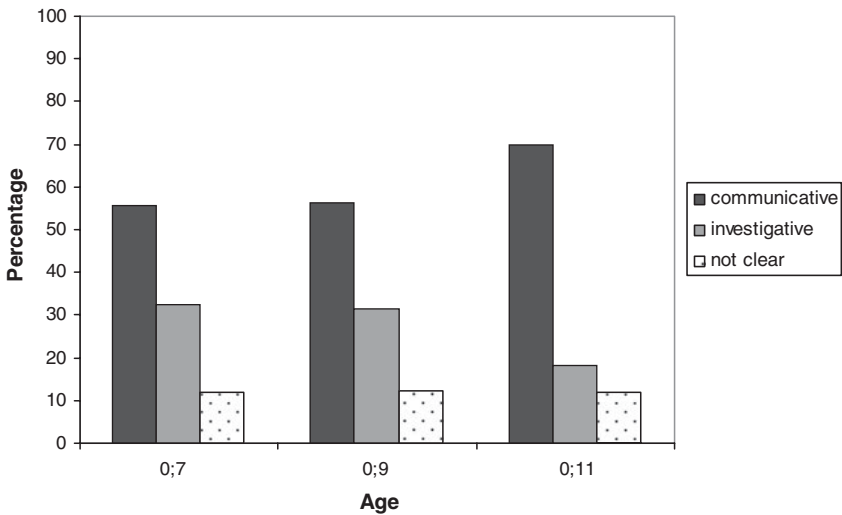


Fig. 2. Percentages of 'investigative', 'communicative', and 'not clear' vocalizations across the different age groups.

vocalizations than investigative vocalizations at all ages and that such expressions increased longitudinally: at 0;7 and 0;9 communicative vocalizations approximately double the number of the investigative ones, and at 0;11 the communicative vocalizations are four times more frequent than the investigative ones. They also show that 12% of the total number of vocalizations could not be identified as being either communicative or investigative. Chi-squared tests of independence were carried out in order to investigate whether the proportion of 'communicative' and 'investigative' vocalizations differed from each other and across ages. Results showed that the proportion of communicative and investigative vocalizations was statistically different at all ages ($\chi^2 (1, N=610)=41.512, p < 0.001$ at 0;7, $\chi^2 (1, N=726)=57.322, p < 0.001$ at 0;9, and $\chi^2 (1, N=1041)=35.308, p < 0.001$ at 0;11). As for the potential significant difference among

proportions of communicative and investigative vocalizations across ages, the chi-squared tests revealed that the proportions of communicative vocalizations differed significantly at all ages: from 0;7 to 0;9 ($\chi^2(1, N=859)=7.728, p=0.005$), from 0;7 to 0;11 ($\chi^2(1, N=1211)=160.861, p<0.001$), and from 0;9 to 0;11 ($\chi^2(1, N=1292)=100.465, p<0.001$). In contrast, the proportion of investigative vocalizations varied significantly only from 0;9 to 0;11 ($\chi^2(1, N=475)=4.651, p=0.031$), and not from 0;7 to 0;9 ($\chi^2(1, N=487)=2.667, p=0.102$), nor from 0;7 to 0;11 ($\chi^2(1, N=440)=0.276, p=0.600$).

In the following sections, we discuss the effect of the communicative status on pitch range and then we move on to its effects on duration. All statistical analyses were performed excluding outliers (13 in total) and vocalizations labeled as 'not clear' (324 in total).

Pitch range and the communicative status of vocalizations. The relationship between pitch range and the communicative status of vocalizations was analyzed using linear mixed model analysis (LMM). Pitch range (in semitones) was the dependent variable, and fixed factors were age (3 levels: 0;7, 0;9, and 0;11), communicative status (2 levels: communicative and investigative), and the interaction between age and communicative status. Child was classified as a random factor and not a fixed factor because the purpose of the study was not to investigate individual differences and also because previous analyses of the data revealed that the variable 'child' did not have a significant effect on the results. The analysis revealed a statistically significant effect of the communicative status of the vocalization on the pitch range ($F(1,2073)=12.690, p<0.001$). No significant effects of age were found on pitch range ($F(2,2047)=0.816, p=0.442$), and results on the interaction between communicativeness and age were also non-significant ($F(2,2073)=0.214, p=0.807$). Figure 3 shows the pitch range displayed by communicative and investigative vocalizations at the three ages analyzed.

Duration and the communicative status of vocalizations. The relationship between duration and the communicative status of the vocalization was tested using LMM analysis with duration (in milliseconds) as the dependent variable, and age (3 levels: 0;7, 0;9, and 0;11), communicative status (2 levels: communicative and investigative), and the interaction between age and communicative status as fixed factors. Again, child was classified as a random factor and not a fixed factor for the reasons stated above. The statistical analysis showed that duration was significantly affected by age ($F(2,2072)=22.602, p<0.001$) as well as the communicative status of the vocalization ($F(1,2072)=57.732, p<0.001$). The interaction between age and the communicative status, however, was not significant ($F(2,2072)=0.879, p=0.415$).

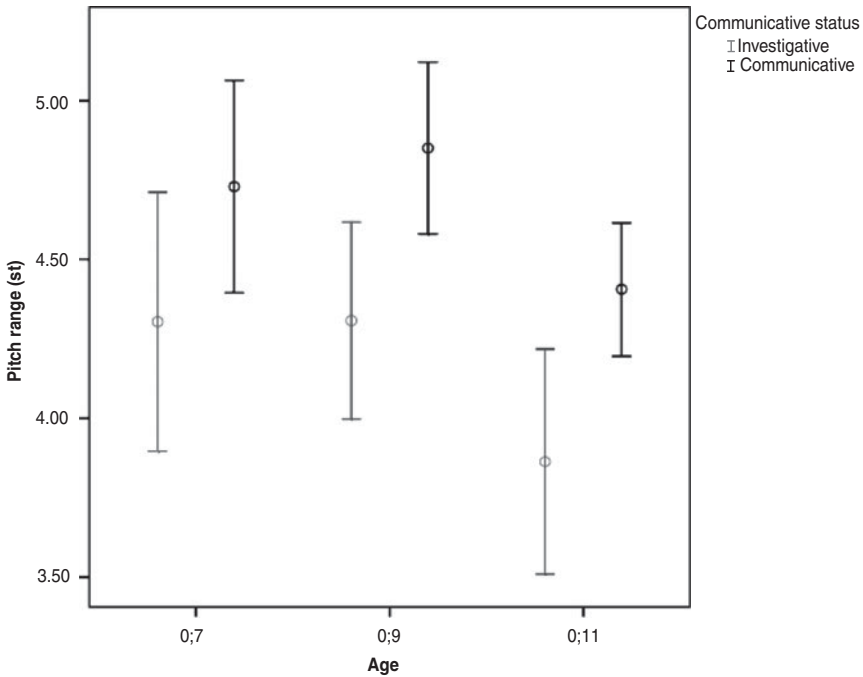


Fig. 3. Error bars of the pitch range of vocalizations (in semitones) as a function of communicative status and children's age.

Bonferroni-corrected pairwise comparisons revealed that the mean duration differed significantly from 0;7 to 0;11 ($p < 0.001$) and from 0;9 to 0;11 ($p < 0.001$) but not from 0;7 to 0;9 ($p = 0.062$). Thus, results for duration in relation to the communicative status of the vocalizations were more robust at 0;9 and 0;11 than at 0;7.

Figure 4 displays the error bars of the total duration of vocalizations (in milliseconds) as a function of communicative status. These results show that at all ages communicative vocalizations tended to be shorter than investigative vocalizations. It can also be observed that this difference is more prominent for some ages than others: at 0;7 the mean duration of a communicative vocalization is 890.30 ms ($SD = 631.668$) compared to 1090.94 ms ($SD = 770.798$) for an investigative vocalization; at 0;9 the mean duration of a communicative vocalization is 939.83 ms ($SD = 626.628$), compared to 1234.57 ms ($SD = 655.421$) for an investigative vocalization; and at 0;11, a communicative vocalization lasts a mean of 682.90 ms ($SD = 474.791$) compared to 881.16 ms ($SD = 679.777$) for an investigative vocalization.

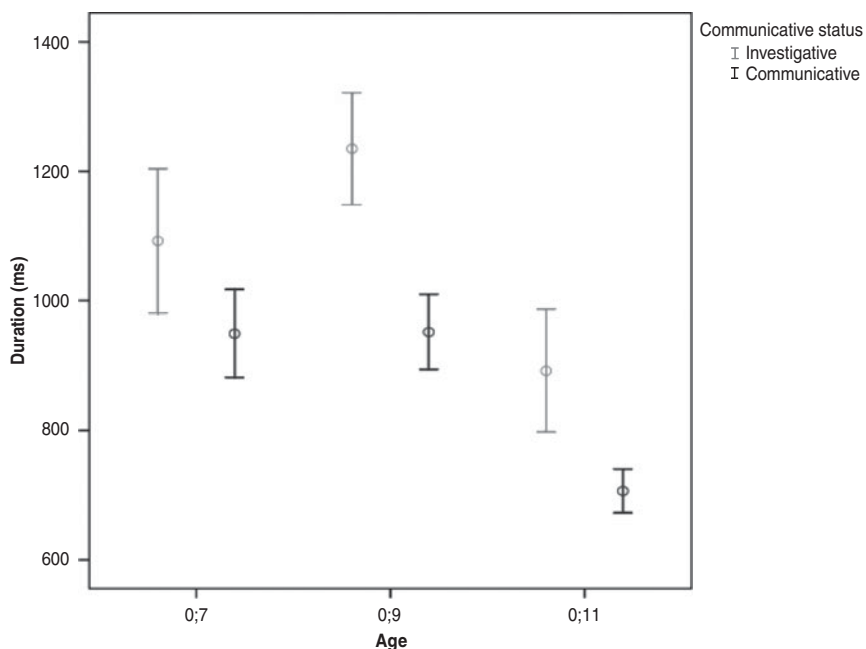


Fig. 4. Error bars of the duration of vocalizations (in milliseconds) as a function of communicative status and children's age.

In sum, statistical analyses of the data showed that pitch range and duration were both significantly affected by the communicative status of the vocalization. As for pitch range, vocalizations displayed a wider pitch range when children were communicating than when they were performing investigative vocalizations. In terms of duration, communicative vocalizations were shorter in general than investigative ones. Yet our results also seem to show that the duration cue was not controlled until children were 0;9. To clarify the picture, in the next section we will investigate whether the specific pragmatic meaning conveyed by the communicative vocalizations has an effect on the pitch range and duration patterns.

Prosodic cues and specific pragmatic functions

We investigated the prosodic cues within the communicative vocalization group by investigating how pitch range and duration patterns of the vocalization were influenced by the specific pragmatic function displayed. Table 3 shows the number of vocalizations analyzed classified in terms of age and specific intentional purpose. As the table shows, vocalizations expressing discontent and satisfaction are the most frequent in the corpus

TABLE 3. *Number of vocalizations classified in terms of pragmatic intention and age*

		0;7	0;9	0;11	TOTAL
<i>communicative vocalizations</i>	<i>discontent</i>	97	137	166	400
	<i>request</i>	24	30	43	97
	<i>satisfaction</i>	56	45	90	191
	<i>response</i>	4	17	57	78
	<i>statement</i>	5	9	129	143
	<i>surprise</i>	—	3	10	13
	<i>vocative</i>	—	—	9	9
	<i>fuzzy intention</i>	199	222	324	745
<i>TOTAL</i>	385	463	828	1,676	

(400 and 191, respectively), followed by statements (143 instances), requests (97 instances), and responses (78 instances). Interestingly, statements, responses, and requests are found more often in the corpus when children are 0;11 but not when they are younger, whereas expressions of discontent and satisfaction are regularly produced at the earliest stages analyzed. The fact that at 0;7 the children in our study expressed mainly discontent and satisfaction and that most of the pragmatic intentions did not appear until 0;11 is similar to what Snow and Balog (2002) and Snow (2006) found in their studies, namely that around 0;8 intonation is still influenced by emotional factors.

Specific intentions like ‘surprise’ and ‘vocative’ were seldom produced in comparison with other pragmatic functions like ‘discontent’ or ‘satisfaction’. The low frequency of occurrence of these two categories (see Table 3) meant that they could not be reliably compared with the other relatively abundant pragmatic functions and we therefore decided to exclude them from further analysis. The table also shows that the group including most vocalizations is the group labeled as ‘fuzzy intention’: the proportion of communicative vocalizations which did not have a clear intention was 51.69% at 0;7, 47.95% at 0;9, and 39.13% at 0;11. As noted above, this group included all those communicative vocalizations that could not be unambiguously identified as any specific pragmatic function.

Pitch range and pragmatic intentions

The relationship between pitch range and specific pragmatic intention displayed for communicative vocalizations was tested using LMM analysis, with pitch range (in semitones) as the dependent variable, and age (3 levels: 0;7, 0;9, and 0;11), pragmatic intention (5 levels: discontent, request, satisfaction, response, and statement), and the interaction between age and pragmatic intention as fixed factors. Again, child was classified as a random

TABLE 4. *Statistical p values of the pairwise comparisons of pitch range and duration between pragmatic intentions*

		Pitch range	Duration
Discontent	Request	1.000	0.000**
	Satisfaction	0.006*	0.000**
	Response	0.258	0.000**
	Statement	0.144	0.000**
Request	Satisfaction	1.000	0.002*
	Response	1.000	0.041*
	Statement	0.969	0.000**
Satisfaction	Response	1.000	1.000
	Statement	1.000	0.517
Response	Statement	1.000	1.000

NOTE: * $p < 0.01$, ** $p < 0.001$.

factor. Results revealed a significant effect of specific pragmatic intention on pitch range ($F(4,763) = 4.539, p = 0.001$). No effect of age was found for pitch range ($F(2,729) = 1.544, p = 0.214$), and there was no interaction of age or intention with pitch range ($F(8,784) = 1.356, p = 0.212$).

As Table 4 shows, Bonferroni-corrected pairwise comparisons revealed that there were no significant differences in pitch range across pragmatic intentions, except for expressions of discontent, which vary significantly from expressions of satisfaction ($p = 0.006$). When looking at mean pitch range values with all ages combined, distinct tendencies can be observed across pragmatic intentions: the mean pitch range for expressions of discontent was 5.37 st ($SD = 3.18$), 5.10 st ($SD = 2.85$) for requests, 4.46 st ($SD = 3.09$) for expressions of satisfaction, 3.82 st ($SD = 2.84$) for statements, and 3.73 st ($SD = 2.49$) for responses. Figure 5 shows the different tendencies across pragmatic intentions: expressions of discontent display wider pitch range, requests show a pitch range that is narrower than that of expressions of discontent but wider than that of the other intentions; expressions of satisfaction show a pitch range that is narrower than that of expressions of discontent and requests but wider than that of responses and statements; statements show a pitch range that is narrower than that of expressions of satisfaction but slightly wider than that of responses, and responses are the pragmatic intention that display the narrowest pitch range. Although the differences in mean pitch range are not statistically significant for the most part, they show clear tendencies across pragmatic intentions.

Duration and pragmatic intentions

The relation between duration and specific pragmatic intention displayed in the communicative vocalization was tested once more using LMM analysis.

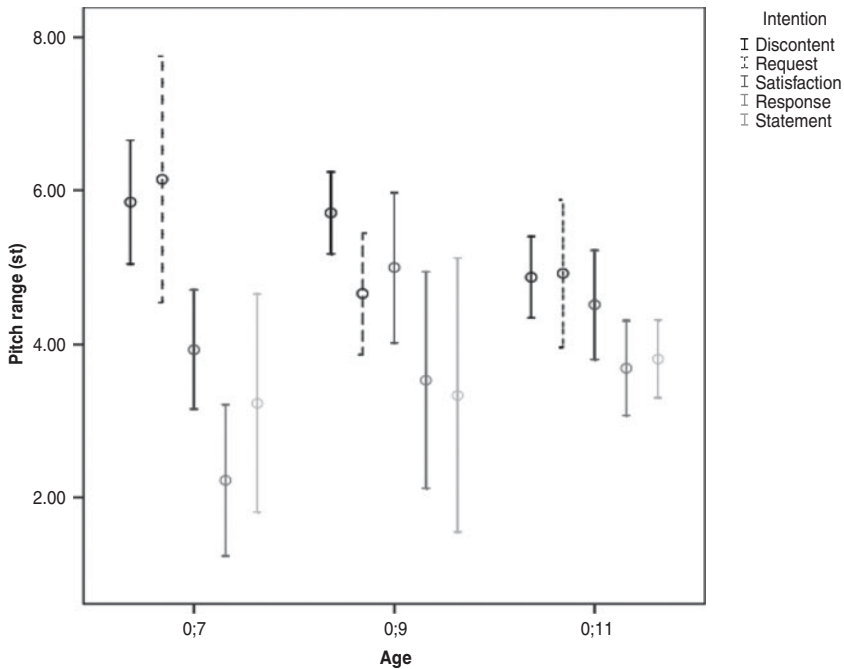


Fig. 5. Error bars of the pitch range of vocalizations (in semitones) as a function of the specific pragmatic intention and children's age.

The dependent variable was total duration (in milliseconds), and the fixed factors were age (3 levels: 0;7, 0;9, and 0;11), pragmatic intention (5 levels: discontent, request, satisfaction, response, and statement), and the interaction between age and pragmatic intention. Child was once again classified as a random factor. The results showed a significant effect of pragmatic intention on duration ($F(4,787) = 60.841, p < 0.001$). Neither age ($F(2,786) = 1.672, p = 0.189$) nor the interaction between age and intention ($F(8,787) = 1.015, p = 0.423$) had any significant effect on duration.

As Table 4 shows, Bonferroni-corrected pairwise comparisons revealed that some pragmatic intentions varied significantly from each other in terms of duration: vocalizations that express discontent or function as requests were significantly different compared to all other intentions; vocalizations expressing satisfaction had similar duration to responses and statements but differed from expressions of discontent or requests; and responses and statements differed from expressions of discontent and requests. Mean duration values across pragmatic intentions with all ages combined patterned in a similar way to the mean pitch range results reported in the previous section: expressions of discontent showed the longest duration

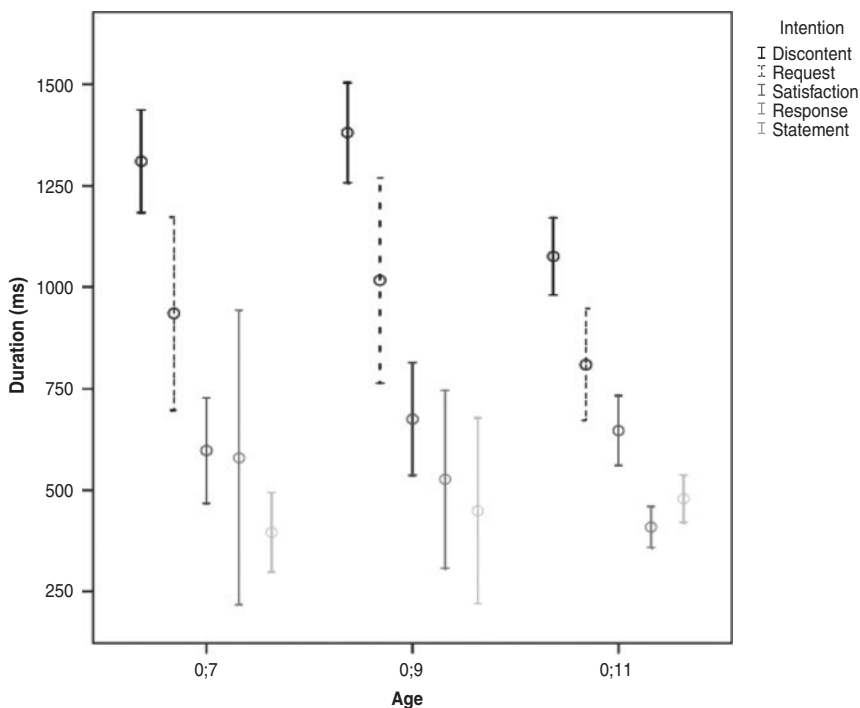


Fig. 6. Error bars of the duration of vocalizations (in milliseconds) as a function of the specific pragmatic intention and children's age.

(1241.83 ms, $SD=611.02$), followed by requests (899.91 ms, $SD=513.23$); expressions of satisfaction had a mean duration of 639.71 ms ($SD=429.16$), while statements lasted 479.59 ms ($SD=327.95$) on average. The pragmatic intention with the shortest duration (450.40 ms, $SD=276.89$) was responses. Figure 6 shows these tendencies with error bars. Note that results for the duration of responses and statements at 0;7 must be treated carefully, since only four vocalizations were classified as responses and only five as statements for that age.

Hence, the analysis of prosodic cues like pitch range and duration of early vocalizations showed that babbling children seem to control pitch range and duration as early as 0;7. In terms of pitch range, we observed that communicative vocalizations had a wider pitch range than investigative ones. Further analyses of communicative vocalizations revealed that depending on the pragmatic intention expressed, pitch range tended to be wider or narrower. Thus, expressions of discontent showed significantly wider pitch range than expressions of satisfaction. Further observation of mean pitch range values revealed that although it was not statistically

significant, expressions of discontent and requests had wider pitch ranges than responses and statements.

In terms of the duration of vocalizations, our results showed that it was also strongly affected by their communicative status. Investigative utterances were significantly longer than communicative ones. Our subsequent analysis of communicative vocalizations, whereby they were categorized into specific pragmatic intentions, showed that the patterns for the duration of the vocalizations were strongly influenced by their pragmatic function. Specifically, the shortest vocalizations were responses; statements were slightly longer than responses but still shorter than the other intentions; expressions of satisfaction were longer than responses and statements, but shorter than requests or expressions of discontent. Requests were longer than all the other intentions except for expressions of discontent, which were the longest.

DISCUSSION AND CONCLUSIONS

This study had two aims: first, to investigate whether infants use specific prosodic cues when attempting to be communicative with their parents; and second, to investigate whether these babbling infants are able to express specific pragmatic intentions by means of prosodic cues. The longitudinal analysis has revealed that between 0;9 and 0;11 children significantly increase their total number of communicative vocalizations. At 0;7 and 0;9 communicative vocalizations are double the number of investigative ones; however, at 0;11 communicative vocalizations are four times more frequent than investigative ones (see Figure 2). These results support previous studies stating that children develop intentional communication around 0;8–0;9 (Bates, Camaioni & Volterra, 1975; Piaget, 1936; Tomasello, 1993; Vygotsky, 1962).

With respect to our first goal, the prosodic analysis of the data revealed very consistent effects of the communicative status of the vocalizations on prosodic cues such as pitch range and duration. In terms of duration, communicative vocalizations are shorter than investigative ones. Even though this tendency was observed at the three ages recorded (namely at 0;7, 0;9, and 0;11), it was only statistically significant when children were 0;9 and 0;11. These results suggest that some children at 0;7 still do not control the use of duration as a prosodic cue to convey communicativeness, so it is not until children are 0;9 that this ability seems to be acquired. An analysis of a larger database is required to confirm the results on the interaction between duration and communicativeness at 0;7. As for pitch range, our data has shown that children produce vocalizations with a wider pitch range when seeking to communicate with their parents and vocalizations with a narrower pitch range when performing investigative

vocalizations. Children as young as 0;7 thus seem able to control their vocalizations' pitch range, displaying a wider pitch range when they attempt to communicate and a narrower pitch range when they do not. The patterns of results on pitch range and duration thus replicate Papaeliou and Trevarthen's (2006) conclusions that communicative vocalizations uttered by English-babbling children tend to have a wider pitch range and shorter duration than investigative vocalizations.

Our second goal was to test whether babbling infants were able to use prosodic cues selectively in order to express distinct pragmatic functions well before they produce their first words. First, as shown in Table 3, our data confirm that before producing words, children are able to communicate intentionally. At 0;7 and 0;9, children are able to communicate with their parents through expressions of discontent and satisfaction, and requests. As their communication skills develop, i.e. at 0;11, they intentionally produce a wide variety of pragmatic meanings such as expressions of discontent and satisfaction, requests, responses, and statements, apart from random instances of vocatives and vocalizations expressing surprise. These results are consistent with Bates *et al.* (1975), who state that before 0;10, children communicate through perlocutions, i.e. "communicative acts which have an effect on their listener, but which are not designed as conventions recognized by both speaker and listener"; after 0;10 children move on to the illocutionary stage, when the child "intentionally uses nonverbal signals to convey requests and to direct adult attention to objects and events". The fact that at 0;7 the children in our study expressed mainly discontent and satisfaction and that most of the pragmatic intentions did not appear until 0;11 is similar to what Snow and Balog (2002) and Snow (2006) found in their studies, namely that around 0;8 intonation is still influenced by emotional factors.

The results of the acoustic analysis revealed a consistent effect of the pragmatic intention of vocalizations on pitch range and duration patterns. Results of the statistical analyses revealed that utterances classified as discontent had significantly higher pitch range than expression of satisfaction. The observation of the mean pitch range values showed that utterances classified as expressions of discontent and requests have a wider pitch range and longer duration than utterances classified as responses and statements, which are shorter and have a narrower pitch range. Also, expressions of satisfaction lie in the middle ground, as they are shorter than requests and expressions of discontent but longer than responses and statements, and they have a narrower pitch range than expressions of discontent and requests but a wider one than responses or statements. Hence, before the first words are produced, children are able to select specific prosodic cues to express intentionality in their vocalizations. When children express discontent or make a request, they consistently use

prosodic features like expanded pitch range and longer duration; when they express satisfaction, they use wide pitch range but short duration; and when they produce responses or statements, they use narrow pitch range and short duration.

In sum, our study supports previous research on the prosodic features of prelinguistic vocalizations (D'Odorico & Franco, 1991; Papaeliou *et al.*, 2002; Papaeliou & Trevathen, 2006; Sachs, 1993) in the sense that infants select particular prosodic cues to express communicativeness. Our results corroborate the claim that prelinguistic infants produce longer vocalizations with a narrow pitch range when they are playing alone or with a toy and do not interact with their parents. In contrast, their utterances are shorter and show a wider pitch range when interacting with their parents. Yet our results go a step further and show that important prosodic differences are obtained when early vocalizations are related to intentional communication and specific pragmatic intentions. These results thus demonstrate the usefulness of investigating the development of early prosodic patterns at the babbling stage in relation to the development of intentional meaning.

We argued on the basis of our data that before children produce their first words, they are able to systematically use prosodic cues to express a set of distinct pragmatic meanings. Thus, children at 0;9 and 0;11 are able to distinguish expressions of discontent and requests from responses and statements by means of prosody. Recent findings also report the use of adult-like intonational contours to convey specific pragmatic functions in the one-word period (Frota & Vigário, 2008, for Portuguese; Marcos, 1987, for French; Prieto *et al.*, 2012, for Catalan and Spanish). Prieto *et al.* (2012), for instance, investigated the development of prosodic patterns in four Catalan children and two Spanish children and demonstrated that children at 1;1 and 1;3 are able to produce a set of adult-like intonation contours. Marcos (1987) analyzed the communicative functions of pitch range and pitch direction in French infants from 1;2 to 1;10, comparing the prosodic patterns of ten children when requesting, giving, showing, and labeling. In terms of pitch range, the highest pitch range was found in repeated requests, a somewhat lower range for initial requests, a still lower range for giving and showing, and the lowest range for labeling. For pitch direction, patterns were only clear with requests and labeling, since children used rising tones when requesting and falling tones when labeling.

Although our babbling data revealed a consistent use of target prosody by young infants, further research is needed to investigate the development of prosodic patterns from the early babbling period to the first-word period by taking into account the communicative uses of language, since it is during the babbling period that children start using language for communicative purposes. It might well be that the first signs of developmental language

impairment can be discernible in the early prosodic patterns that an infant uses when babbling.

REFERENCES

- Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. & Paggio, P. (2007). The MUMIN coding scheme for the annotation of feedback, turn management and sequencing. In J. C. Martin, P. Paggio, P. Kuehnlein, R. Stiefelbogen & F. Pianesi (eds), *Multimodal corpora for modeling human multimodal behaviour*. Special issue of *Language Resources and Evaluation* 41(3/4), 273–87. Heidelberg: Springer.
- Astruc, L., Prieto, P., Payne, E., Post, B. & Vanrell, M. M. (in press). Tonal targets in early child English, Spanish, and Catalan. *Language and Speech*.
- Balog, H. L. & Brentari, D. (2008). The relationship between early gesture and intonation. *First Language* 28, 141–63.
- Balog, H. L., Roberts, F. & Snow, D. (2009). Discourse and intonation development in the first-word period. *Enfance* 3, 293–304.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L. & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Bates, E., Camaioni, L. & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly* 21, 205–226.
- Blake, J. & Boysson-Bardies, B. de (1992). Patterns in babbling: A cross-linguistic study. *Journal of Child Language* 19, 51–74.
- Boersma, P. & Weenink, D. (2005). *Praat: Doing phonetics by computer (Version 4.3.01)*. University of Amsterdam 2005 [http://www.praat.org/].
- Chen, L. & Kent, R. (2009). Development of prosodic patterns in Mandarin-learning infants. *Journal of Child Language* 36, 73–95.
- D’Odorico, L. & Franco, F. (1991). Selective production of vocalization types in different communication contexts. *Journal of Child Language* 18, 475–99.
- Davis, B. L., MacNeilage, P. F., Matyear, C. L. & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustical study. *Child Development* 71, 1258–70.
- Engstrand, O., Williams, K. & Lacerda, F. (2003). Does babbling sound native? Listener responses to vocalizations produced by Swedish and American 12- and 18-month-olds. *Phonetica* 60, 17–44.
- Flax J., Lahey, M., Harris, K. & Boothroyd, A. (1991). Relations between prosodic variables and communication function. *Journal of Child Language* 18, 3–19.
- Frota, S. & Vigário, M. (2008). The intonation of one-word and first two-word utterances in European Portuguese. Paper presented at the XI International Conference for the Study of Child Language (IASCL).
- Furrow, D. (1984). Young children’s use of prosody. *Journal of Child Language* 11(1), 203–213.
- Furrow, D., Podrouzek, W. & Moore, C. (1990). The acoustical analysis of children’s use of prosody in assertive and directive contexts. *First Language* 10, 37–49.
- Galligan, R. (1987). Intonation with single words: Purposive and grammatical use. *Journal of Child Language* 14, 1–21.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. New York: Elsevier.
- Iverson, J. M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science* 16(5), 367–71.
- Karousou, A. (2003). Análisis de las vocalizaciones tempranas: su patrón evolutivo y su función determinante en la emergencia de la palabra. Unpublished doctoral dissertation, Universidad Complutense de Madrid.
- Kent, R. D. & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America* 72, 353–65.

- Levitt, A. & Utman, J. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language* **19**(1), 19–49.
- Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.
- Mampe, B., Friederici, A. D., Christophe, A. & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology* **19**(23), 1994–97.
- Marcos, H. (1987). Communicative functions of pitch range and pitch direction in infants. *Journal of Child Language* **14**, 255–68.
- Özçalışkan, S. & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition* **96**, 101–113.
- Papaeliou, C., Minadakis, G. & Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech, Language and Hearing Research* **45**(2), 311–17.
- Papaeliou, C. F. & Trevarthen, C. (2006). Prelinguistic pitch patterns expressing 'communication' and 'apprehension'. *Journal of Child Language* **33**, 163–78.
- Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Neuchâtel: Delachaux et Niestlé.
- Prieto, P., Estrella, A., Thorson, J. & Vanrell, M. M. (2012). Is prosodic development correlated with grammatical and lexical development? Evidence from emerging intonation in Catalan and Spanish. *Journal of Child Language* **39**(2), 258–83.
- Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P. & Warehem, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In D. Bamman, T. Magnitskaia & Colleen Zaller (eds), *Proceedings of the 30th Annual Boston University Conference on Language Development*, 489–500. Somerville, MA: Cascadilla Press.
- Sachs, J. (1993). The emergence of intentional communication. In J. Gleason (ed.), *The development of language*, 40–64. New York: Macmillan.
- Sansavini, B., Guarini, S. & Stefanini, C. (2010). Early development of gestures, object-related actions, word comprehension and word production, and their relationships in Italian infants. *Gesture* **10**(1), 52–85.
- Sarriá, E. (1991). Observación de la comunicación intencional preverbal: un sistema de codificación basado en el concepto de la categoría natural. *Psicotema* **3**, 359–80.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin* **99**, 143–65.
- Snow, D. (2006). Regression and reorganization of intonation between 6 and 23 months. *Child Development* **77**, 281–96.
- Snow, D. & Balog, H. L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua* **112**, 1025–58.
- Tomasello, M. (1993). On the interpersonal origins of self-concept. In U. Neisser (ed.), *The perceived self. Ecological and interpersonal knowledge of the self-knowledge*, 174–84. New York: Cambridge University Press.
- Tomasello, M., Carpenter, M. & Liszkowski, U. (2007). A new look at infant pointing. *Child Development* **78**, 705–722.
- Trevarthen, C. (1977). Descriptive analyses of infant communicative behaviour. In H. R. Schaffer (ed.), *Studies in mother–infant interaction*, 227–70. London: Academic Press.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (ed.), *Before speech*, 321–47. Cambridge: Cambridge University Press.
- Trevarthen, C. (1982). The primary motives for cooperative understanding. In G. Butterworth & P. Light (eds), *Social cognition: Studies of the development of understanding*, 77–109. Brighton: Harvester Press.
- Trevarthen, C. (1990). Signs before speech. In T. A. Sebeok & J. U. Sebeok (eds), *The semiotic web*, 689–755. Berlin: Mouton de Gruyter.
- Vihman, M. M. & DePaolis, R. A. (1998). Perception and production in early vocal development: Evidence from the acquisition of accent. In M. C. Gruber, D. Higgins, K. S. Olson & T. Wysocki (eds), *Chicago Linguistic Society* **34**, 373–86.

- Vihman, M. M., DePaolis, R. A. & Davis, B. L. (1998). Is there a 'trochaic bias' in early word learning? Evidence from infant production in English and French. *Child Development* **69**, 933-947.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.
- West, B., Welch, K. B. & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. New York: Chapman & Hall/CRC.
- Whalen, D. H., Levitt, A. G. & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language* **18**, 501-516.