



The perception of intention and emotion in non-cry pre-babbling infant vocalizations

*Lacheret-Dujour*¹, *A.*, *Bobin-Bègue*, *A.*², *Devouche*, *E.*³ & *Gratier*, *M.*²

¹ Modyco (UMR 7114), Department of Linguistics,
Paris Nanterre University, France

² Laboratory of Ethology, Cognition and Development (EA 3456), Department of Psychology,
Paris Nanterre University, France

³ Laboratoire de Psychopathologie et Processus de Santé, Department of Psychology,
Paris Descartes University, France.

anne.dujour.27@gmail.com, anne.bobin_begue@parisnanterre.fr, emmanuel.devouche@parisdescartes.fr,
mgratier@parisnanterre.fr

Abstract

From the sixth week of life infants start to produce non-distress vocalizations during social exchange, and adults interpret these as communicative expressions.

The prosodic qualities of these early vocalizations have not been extensively studied.

The aim of this study was to examine some of the factors which lead “naïve” adult listeners to perceive infant vocalization as intentional on the one hand and as expressing emotion on the other hand.

A sample of 24 non-distress vocalizations produced in the course of social interaction with a parent was selected for this study. Half of the vocalizations were produced by 6 week-old infants, and half by 18 week-old. The sample was also evenly distributed according to two other factors, duration (short vs. long) and continuity (continuous vs. discontinuous). Pairs of vocalizations were created by varying only one of the 3 factors (age, duration, continuity) at a time.

A hundred and ten participants heard a total of 48 pairs of vocalizations presented in counter-balanced trials and were asked to determine which vocalization was more communicative and, on a separate occasion, which expressed greater emotion. Binary response choice was recorded for each stimulus and for each participant.

Results show that listeners use different strategies in attributing communicative intent and emotional intensity. It also appeared that older infants’ vocalizations were generally perceived as more communicative whereas longer vocalizations were perceived to convey more emotion

Index Terms: infant vocalization, perception, emotion, communication

1. Introduction

Only a handful of studies have investigated how adults interpret infant non-cry vocalization in the pre-babbling stage of vocal development (see below). None of these studies deals with French infants. Yet the early stages of vocal development, covering the first six months of life, are well documented in the Anglo-Saxon literature. The earliest non-cry sounds infants produce are short, nasalized, have limited harmonic resonance and are considered to be involuntary [1] [2]

From 6 to 8 weeks of age infants produce cooing sounds that are more resonant and that elicit emotional responses from social partners [3]. These vocalizations are produced through articulatory movements at the back of the vocal tract, yet they lack the “concrete” phonological structure that permits phonetic transcription [1].

Around 4 months of age, infants gain increasing control of their vocal productions resulting in vocal productions with greater variation in acoustic features such as pitch, loudness and timbre.

From 5 to 10 months of age, they begin to produce fully resonant sounds with faster formant transitions, resulting in canonical syllables, which represent a language general unit of mature vocal production [4]. The repetition of these phonologically stabilized syllables is known as babbling.

It has been considered for a few decades now that pre-babbling vocalisations do not carry the imprint of the language spoken in the child's social environment. In fact, given the current literature, it is not easy to determine at what age infant vocalizations take on the specific phonological colouring of the emergent spoken language. However, it has been shown that the intonation contours of newborn infants' cries differ in France and Germany, in accordance with the predominant stress patterns of these two languages [5].

The rapid changes in the acoustic qualities of sounds produced by young infants cannot be dissociated from the social and emotional contexts within which they are produced. In spontaneous interaction, adults respond to infant non-cry vocalizations right from birth, often describing them verbally as expressions of agency [6] [7]. Indeed adults are sensitive to the acoustic qualities of infant vocalizations, preferring more resonant, fully voiced sounds [8] and responding more often to resonant sounds with an intonation contour [9]. These sounds are considered to be more intentionally communicative [9]. Acoustic analysis of infant vocalization indicates that adults rely on salient acoustic features in order to infer both communicative intent and emotionality [10].

Adults consistently categorize the sounds of 2-month-old infants on the basis of emotional content (happy-sad) [11]. In the babbling stage mothers are able to consistently categorize infant vocalizations, even those of unfamiliar infants [12] and parents can readily identify the first well-formed consonant-vowel syllables produced by their infant [13]. Furthermore, parents classify the babbling vocalizations of their infants as either emotive or communicative based on distinct acoustic features [10].

It is also of great interest to note that infants as young as 3 months tend to produce more mature sounds involving harmonic resonance and pitch contouring specifically in response to adult utterances [14]. It thus appears that infants' vocal development, associated with increasing resemblance of their vocal sounds to the sounds of intelligible speech, is embedded in bidirectional motivational processes. Indeed, if infants respond to affectionate well-timed adult speech with more controlled and harmonic sounds and if adults in turn prefer those same sounds and respond to them with greater emotionality, the whole process of learning to speak must be rooted in the mutual recognition of both intentionality and emotional expression.

The present study examines naive adult listeners' perception of pre-babbling infant vocalization. Its aim is to explore the effects of two characteristics, duration and continuity, on adults' choices in attribution of intentionality and emotionality to the vocalizations. Furthermore, we wished to ascertain the implicit effect of infant age on adults' perception of intentionality and emotionality. Because rapid alternation between sound and silence is a basic characteristic of all human speech, we expected that discontinuity would be considered characteristic of speech-like communicative vocal

behaviour in infants. Auditors should thus choose discontinuous vocalizations more often when asked which is addressed to someone. Emotional forms of speech such as sung speech or infant-directed speech tend to involve more continuous modulated sound. For this reason, we expected auditors to choose continuous vocalization when asked which is more emotional.

2. Method

2.1. Participants

The study involved 110 participants recruited at Nanterre university (mean age : 21.71 ; SD = 5.68). Participants all gave their written consent. They were 39 men (mean age: 21.96; SD = 5.93) and 71 women (mean age: 21.71 ; SD = 5.68). None of the participants had daily contact with an infant at the time of the experiment and none of them suffered from any form of hearing impairment.

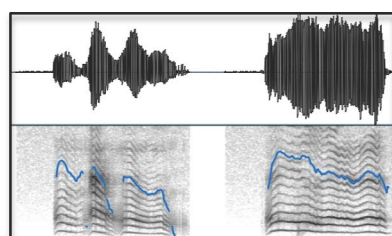
2.2. Stimuli

24 infant vocalizations were selected from naturalistic audio recordings of 4 infants made in the course of interacting with their mothers at 2 different ages. Each infant was recorded at 6 weeks of age and at 18 weeks of age. All of the selected vocalizations were non-cry and did not present any negative emotional valence or any concrete phonological structure.

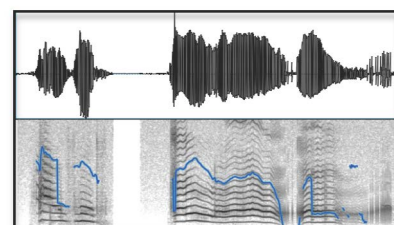
Each selected vocalization was characterized in 3 ways. First, it was produced by either a very young infant (a 6 week-old) or by a slightly older infant (a 18 week-old). Second, it was either of short duration, comprised between 200 ms and 900 ms, or of longer duration, comprised between 1400 ms and 2300 ms. Third, it was either continuous, with no interruption in sound, or discontinuous, including a silent pause lasting between 10 and 100 ms (See figures 1 and 2).

Thus eight categories of vocalizations were obtained, each comprising 6 different vocalizations.

24 pairs of vocalizations were made based on the aforementioned criteria. For each pair, only one feature varied (Age, Duration or Continuity). In each pair, the 2 vocalizations were separated by a silent pause of 500 ms. A silent pause was also added before the first vocalization so that each audio file had a total duration of 7000 ms.



Long, **discontinuous** vocalization, produced by a 5-week-old Long, **continuous** vocalization, produced by a 5-week-old



Short, discontinuous vocalization, produced by a 5-week-old **Long, discontinuous** vocalization, produced by a 5-week-old

Figures 1 and 2: Examples of pairs of vocalizations involving 2 variables: duration and continuity.

Table 1: Overview of the stimuli with 3 variables: age, duration and continuity (N=48 vocalizations)

Age	± continuity	± long	Number of stimuli
Young	Continuous (C)	Short (S)	6
		Long (L)	6
	Discontinuous (D)	Short (S)	6
		Long (L)	6
Old	Continuous (C)	Short (S)	6
		Long (L)	6
	Discontinuous (D)	Short (S)	6
		Long (L)	6

	Vocalization characteristics	Vocalization characteristics
Age	1 S C	2 S C
Age	1 L C	2 L C
Age	1 S D	2 S D
Age	1 L D	2 L D
Duration	1 S C	1 L C
Duration	1 S D	1 L D
Duration	2 S C	2 L C
Duration	2 S D	2 L D
Continuity	1 S C	1 S D
Continuity	1 L C	1 L D
Continuity	2 S C	2 S D
Continuity	2 L C	2 L D

Table 2. Construction of the 12 pairs of vocalizations with contrasting variables: Age (6 weeks vs. 18 weeks), Duration (Short vs. Long) and Continuity (Continuous vs. Discontinuous). Each pair is presented twice with alternating order of presentation so that participants heard 24 pairs of vocalizations per question.

2.3. Procedure

A two-alternative forced-choice experimental design was used. Participants were presented two series of 24 pairs of infant vocalizations and asked to choose, in the first experimental block, which one most clearly conveyed communicative intent, and in the second experimental block, which one expressed most emotion.

The experiment took place in a quiet room at the university. Participants were informed of the general procedure. They were seated in front of the computer and were given a set of headphones. A short training session was used in order to familiarize participants with the experimental set up.

The instructions for the experimental blocks appeared on the screen before each experimental block. After the instructions had been read, participants activated the audio files and had to make a choice within a 4 second time delay. The audio files were selected and played in a random order within each block. The question for the first experimental block was “which one of the 2 vocalizations you are about to hear seems to you to be most clearly addressed to someone”

and for the second “which one of the 2 vocalizations you are about to hear seems to you to most clearly express emotion?” Experimental blocks were counter-balanced across participants.

The experiment was conducted with the Perceval Software package (Perceval Package Manager 3.0.5.0) that recorded participants' responses for each of the 48 choices they made.

3. Results

Data were analyzed using a general linear model. Logistic regression was run on infant age, duration and continuity of vocalization, including gender, age, experience of participants and order as potential confounders.

All other things being equal, results show that adult listeners systematically attribute both greater communicative intent and greater emotionality to vocalizations produced by 18week-old infants than to ones produced by 6 week-olds. Overall, participants also responded less frequently to the emotionality question than to the intentionality question.

3.1. Communicative intentionality of infant vocalizations

Participants most often selected the vocalization of an older infant (18 week-old versus 6 week-old) when asked which expressed more communicative intent (58.9 % occurrences vs. 41.1% ; $p = .001$, CI 95% [.57, .61]). Adults did not rely primarily on durational aspects nor on continuity/discontinuity in their choice strategies for this question.

Results show however that adults attribute communicative intent to the older infant vocalizations more often when they are long (71.8%, $p < .0001$, OR= 2.89, CI 95% = [1.69, 2.48]), which is not the case when they choose a younger infant vocalization.

3.2. Emotional expression

Participants most often selected the vocalization of an older infant (18 week-old versus 6-week-old) when asked which expressed more emotion (59 % occurrences vs. 41% ; $p = .001$, CI 95% [.55, .59]).

Adults selected long vocalizations over short vocalizations when asked which of the two expressed more emotion (64.2% occurrences vs. 35.8%, CI 95% [61.9, 66.4]). Their choice strategy did not appear to be influenced by continuity-discontinuity of the vocalizations.

Furthermore, when participants chose an older infant vocalization they were not influenced by either its duration or continuity. But participants were more likely to select a long over a short vocalization if it was also continuous rather than discontinuous (74.4%, $p < .0001$, $OR = 2.31$, $CI\ 95\%$, [1.88, 2.84]).

4. Discussion

The aim of this study was to further investigate the strategies adults use in attributing intentionality of communication and emotion to infant non cry vocalizations produced in the pre-babbling stage of vocal development. Results show that adult listeners systematically attribute greater communicative intent and greater emotionality to vocalizations produced by 18 week-old infants. Overall, this study suggests that naïve adult listeners implicitly perceive a qualitative difference between the vocalizations of 6 week-old infants, who are just beginning to produce the more resonant sounds considered to be communicative, and those of 18 week-old infants who will soon begin to produce the sounds of well-formed babbling [4].

This study also reveals that adults do not rely on the same strategies in attributing intentionality of communication and emotion to infant vocalizations. The fact that the same set of vocalizations was interpreted differently according to the questions asked is of particular interest. Indeed, participants did not rely on duration or continuity of sound when asked to interpret communicative intent whereas they did rely on the durational aspect of vocalizations, but not on their continuity, when asked to interpret their emotionality. Thus, adults attributed greater emotionality to longer vocalizations than to shorter ones even when produced by very young infants but they did not consider 6 week-old infants' longer vocalizations to be more communicative than their shorter vocalizations.

The fact that despite the constancy of some central acoustic features (resonance, length, continuity or discontinuity), adults attribute greater intentionality to older infants' vocalizations suggests that their judgments are based on some other internal acoustical features that are still to be discovered.

Anatomically, by 20 weeks of age infants have greater maturity of the vocal tract and larger resonators which might well explain perceptible differences in the resonant quality of sounds produced. They have also gained greater motor control over the process of producing sound and it is not clear today how such prospective motor control [15] might translate to a measurable acoustic feature.

Furthermore, although many specialists consider that infraphonological competence (which precedes babbling) is not yet language specific [4] [16], it is worth investigating whether specific linguistic signatures might be woven into pre-babbling vocalization, thus facilitating its 'intelligibility' for adults.

An acoustic analysis of the infant vocalizations used in this study might help elucidate which features adult listeners recruit in formulating interpretations in terms of communicative intent on the one hand and emotion on the other hand. Features worth investigating include intonation contour, amplitude variation, harmonicity, and timbre. Insight into the role played by these features or by combinations of these features would provide support for generating stimuli in

further experiments on the perceptual strategies involved in interpreting pre-babbling non cry vocalization.

Finally, an important limitation of the study is the potential confusion between the two questions posed to auditors, i.e. the intentional versus emotional nature of vocalizations. Indeed, it is not clear whether emotional expression can be dissociated from a social context, that of being addressed to another person. Thus, in future studies attention should be paid to the manner in which naïve auditors are interrogated about their interpretation of pre-babbling infant vocalization.

5. References

- [1] D.K. Oller, "The emergence of the sounds of speech in infancy", in G.Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology, Vol.1: Production* (pp. 93–112). New York: Academic Press, 1980.
- [2] R.E. Stark, "Stages of speech development in the first year of life", in G.Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology, Vol.1: Production* (pp.73–90). New York: Academic Press, 1980.
- [3] K. Bloom, "Quality of adult vocalizations affects the quality of infant vocalizations". *Journal of child language*, 15(3), pp. 469–480, 1988.
- [4] D.K. Oller, *The emergence of the speech capacity*. Mahwah, N.J.: Lawrence Erlbaum Associates, 2000.
- [5] B. Mampe, A. D. Friederici, A. Christophe, and K. Wermke. "Newborns' cry melody is shaped by their native language." *Current biology* 19, no. 23, pp. 1994–1997 (2009).
- [6] S. Dominguez, E. Devouche, G. Apter, and M. Gratier, "The roots of turn-taking in the neonatal period". *Infant and Child Development*, 25(3), 240–255, 2016.
- [7] S. Dominguez, M. Gratier, M., K. Martel, A. Buil, G. Apter, & E. Devouche, "Le nouveau-né, un partenaire pour sa mère. Analyse du discours maternel". *Neuropsychiatrie de l'Enfance et de l'Adolescence*, 65(4), 201–210, 2017.
- [8] K. Bloom, and E. Lo, "Adult perceptions of vocalizing infants". *Infant Behavior and Development*, 13(2), 209–219, 1990.
- [9] K. Bloom, L. D'Odorico, & S. Beaumont, "Adult preferences for syllabic vocalizations: Generalizations to parity and native language". *Infant Behavior and Development*, 16(1), 109–120, 1993.
- [10] C. Papaeliou, G. Minadakis, & D. Cavouras, "Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech, Language, and Hearing Research*, 45(2), 311–317, 2002.
- [11] M. Papoušek, "Determinants of responsiveness to infant vocal expression of emotional state." *Infant behavior and development* 12, no. 4, pp. 507–524, 1989.
- [12] M.H. Goldstein, and M.J. West., "Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size." *Journal of Comparative Psychology* 113, no. 1, 52, 1999.
- [13] D.K. Oller, R.E. Eilers, and D. Basinger, "Intuitive identification of infant vocal sounds by parents." *Developmental Science* 4, no. 1, 49–60, 2001.
- [14] K. Bloom, A. Russell, and K. Wassenberg, "Turn taking affects the quality of infant vocalizations". *Journal of child language*, 14(2), 211–227, 1987.
- [15] N.A. Bernstein, *On the co-ordination and regulation of movements*. New York: Pergamon (A collection. of English-translated papers published previously in Russian and German journals and monographs), 1967
- [16] B.E. de Boysson-Bardies, "Ontogeny of language-specific syllabic productions", in B.E. de Boysson-Bardies, S.E. de Schonen, P.W. Jusczyk, P.E. McNeilage, and J.E. Morton (eds), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 353–363). Springer, Dordrecht, 1993.