J. Child Lang. **33** (2006), 163–178. © 2006 Cambridge University Press
doi:10.1017/S0305000905007300 Printed in the United Kingdom

Prelinguistic pitch patterns expressing 'communication' and 'apprehension'*

CHRISTINA F. PAPAELIOU

Department of the Preschool Education and of the Educational Planning, University of the Aegean, Rhodes, Greece

AND

COLWYN TREVARTHEN

Department of Psychology, The University of Edinburgh, Edinburgh, Scotland, UK

(*Received* 20 *September* 2004. *Revised* 7 *March* 2005)

ABSTRACT

This study examined whether pitch patterns of prelinguistic vocalizations could discriminate between social vocalizations, uttered apparently with the intention to communicate, and 'private' speech, related to solitary activities as an expression of 'thinking'. Four healthy ten month old English-speaking infants (2 boys and 2 girls) were simultaneously video- and audiorecorded in their homes under two conditions: (A) when the infant was playing with mother, and (B) when the infant was alone. One hundred and fifty-six vocalizations were classified as 'communicative', if accompanied by non-vocal communicative behaviour, or as 'investigative', if accompanied by explorative activities. For the acoustic analysis, an automatic PITCH PATTERN RECOGNITION (PPR) software system was developed. The PPR-system could distinguish 'communicative' from 'investigative' vocalizations with an overall accuracy of 91·67%. These findings confirm that prelinguistic vocalizations might serve both as means of purposeful communication and as a tool of thought. These are the functions later assumed by language.

[*] I am indebted to Professor D. Cavouras and his collaborator Mr. G. Minadakis for their invaluable help and great support in the acoustic analysis of the vocalizations. Address for correspondence: Dr. C. F. Papaeliou, Miltiadou St., 39, 175-63, P. Faliro, Athens, Greece. e-mail: papailiou@rhodes.aegean.gr

INTRODUCTION

Mature language mediates both intentional communication between persons and cognitive activity or 'thinking' of individuals. This two-fold function of language has been studied as it develops in young children (Vygotsky, 1934/1962; Diaz & Berk, 1992). There is evidence that these two psychological processes may be distinct prelinguistically, and that, around nine months, they undergo radical developmental changes, which are considered to be necessary preconditions for the emergence of language (Bates, Camaioni & Voltera, 1975; Trevarthen & Hubley, 1978; Tomasello, 1993).

Thus, one may hypothesize that in the pre-verbal stage, as in language, expressive vocalizations may mark as distinct two kinds of functions: (a) those that express the infant's participation in interpersonal exchanges, and (b) those that reflect internal perceptual and cognitive processing during solitary activity – a kind of 'thinking aloud' (Vygotsky, 1934/1962; Papousek & Papousek, 1981; Diaz & Berk, 1992).

Some studies have actually compared vocal expressions of infants uttered during communicative interactions and those emitted during solitary exploration of the inanimate environment. The prosodic features of these vocalizations have been rated subjectively. Halliday (1975), using his expertise as a phonetician, assessed the final movement of the pitch contour of his own son's vocalizations during the latter half of the first year and into the second year. As he notes, at 10 months it was possible to identify two kinds of vocal expressions which apparently served constantly distinct functions: the INTERACTIONAL which conveyed the motive for companionship and had a mid falling tone, and the PERSONAL which conveyed interest in the modifications of an object and had a low falling tone of narrower range. According to Halliday, Nigel's system developed further and at 1;4 the child systematically used a rising tone on utterances serving a social 'doing' function requiring a response from the person addressed ('pragmatic'), and a falling tone on utterances serving a 'learning' function requiring no response ('mathetic'). Similarly, in their study of four infants at 0;11 Dore and colleagues (Dore, Franklin, Miller & Ramer, 1976) found that communicative vocalizations, which accompany communicative actions (e.g. pointing or request gestures), and 'grouping' vocalizations, which reflect subjective states in relation to objects being attended to or manipulated by the infant, are characterized by different prosodic features. Infant prosodic patterns continue to distinguish between vocalizations uttered in intersubjective situations from vocalizations uttered during solitary activity from the prelinguistic period to the one-word stage (Furrow, 1984; Marcos, 1987; Flax, Lahey, Harris & Boothroyd, 1991). Such a subjective assessment, even one made by an expert phonetician, represents the hearer's perception, and may not accurately identify the utterer's intended vocal product and its purpose. Moreover, the hearer's perception may be affected, intentionally or unintentionally, by learned cultural and/or language norms, which function as a standard set of categories for classifying different prelinguistic vocalizations (Papaeliou, Minadakis & Cavouras, 2002). Infants, as uncultured agents, may have motivations for their utterances that are different from those expected by sophisticated adult observers, and the patterns of these utterances may have unexpected features.

Other studies (Delack & Fowlow, 1978; D'Odorico & Franco, 1991) have investigated longitudinally the role of single acoustic features in discriminating vocalizations uttered in contexts of infant–adult communication, or object inspection during the first year. However, it might be the case that while a single isolated acoustic feature may not discriminate vocalizations that have been attributed different functions, that same feature or dimension may contribute to the distinction of these vocalizations when it is combined with other features. For this reason, it is important to assess the discriminating power of feature combinations that may better represent the acoustic pattern of a vocalization (Papaeliou *et al.*, 2002). Moreover, in the most recent of these investigations D'Odorico & Franco (1991) extracted measurements of acoustic features by visual inspection of the spectrogram, which may not be sufficiently sensitive for important features.

Vocal expressions uttered in communicative situations have been particularly studied in relation to the gesture of pointing. Masataka (1995) observed that already at 3 months index-finger extension was significantly more often combined with syllabic sounds rather than vocalic sounds, while no other categories of nonvocal behaviours showed such a relationship. In other words, long before infants can actually talk, there are close connections between the pointing action and 'speechlike' prelinguistic vocalizations. Studies of children at the second year have demonstrated that the vast majority of pointing and reaching is associated with vocalizations which occur during the communicative gesture (Franco & Butterworth, 1996; Haynes, Zylla-Jones, Smith, Rodekohr, McEachern & Berry, 2004). Moreover, it has been shown that temporal coordination between pointing, gazing and vocalizations occurs at a significantly higher rate when 12-month-old infants initiated interaction with an adult compared to an object (Legerstee & Barillas, 2003). Nevertheless, these studies do not provide any information about the acoustic form of the vocalizations.

Although absolute pitch and pitch shift discrimination depends mainly on the acoustic feature of fundamental frequency, it is also strongly determined by duration. Thus, in order to examine pitch patterns in infant vocalizations more adequately, one needs also to investigate the role of rhythm and variations in timing. However, relatively few studies have focused on this issue. Lynch and his colleagues (1995) examined phrasing, a universal characteristic of human communication, in vocalizations produced by typically developing infants from age 2 months to age 12 months. Their findings showed that prelinguistic vocalizations were composed of a hierarchy of grouping structures with internal cohesiveness, as reflected by acoustic characteristics and in adult perception. The standard deviation of each of these rhythmic units matched the durations of units in mature hierarchical systems of language; i.e. the poetic line, the metrical feet and the syllable. Petitto and her colleagues (2001) have demonstrated that hearing babies born to profoundly deaf parents produced a class of rhythmically controlled hand movements that is distinct from other uses of their hands, which evidently have a clear communicative intention.

Rhythm is considered as a principal index of subjective and intentional states of mind. Current research suggests that all acts conveying meaning, vocalizations included, have their origin in a fundamental time-regulated brain activity described as Intrinsic Motive Pulse (IMP). The IMP constitutes the expressive form of a neuronal emotional system called the Intrinsic Motive Formation (IMF). The IMF includes the brain stem, basal ganglia and limbic structures, and operates to integrate the intentional cognitive systems of the neocortext which control attention, learning, planning and acting purposefully (Trevarthen, 1999). The IMP is modified to exhibit distinct rhythmic patterns that, with other features of the vocal act, may communicate the changing focus of awareness (Scherer, 1986; Trevarthen, 1999).

Taking into account the above considerations, the present study aimed to apply an objective and more comprehensive method to identify whether infants utter acoustically distinct vocal patterns in communicative exchanges and in explorative activities at the age of ten months. Selective production of different pitch patterns in different situations at this age may reflect that ability for connecting sound and meaning, on the basis of the fundamental distinction between self and other, develops steadily, at least since the beginning of intentional communication, despite any reorganization processes that might be taking place. Sound–meaning connection is considered as one of the fundamental criteria for the linguistic status of an utterance (Chafe, 1970).

Our study focused on children between the ages of 0;9 and 0;11 months for the following reasons: in order one to examine whether infants are capable of conveying information about communicative and self-directed purposes differently by means of their vocalizations in a typical environment, one needs to be assured that infants are able to conceive clear communicative purposes. Research has demonstrated that, in general, infants are clearly motivated to communicate intentions and to share initiatives in the use of objects only after nine months (SECONDARY INTERSUBJECTIVITY) (Trevarthen & Hubley, 1978). The expression of this level of communication requires the highly evolved social ability of taking interest in the others' feelings, attentions and intentions, establishing sympathetic 'mutual attention', as well as an awareness of how to coordinate in 'joint attention' a self-centred orientation to particular events in the outside world with a partner's other-centred orientation to the same events. This ability has been shown to develop at the last quarter of the first year (Bates, Camaioni & Voltera, 1975; Camaioni, 1993; Reddy, 2001).

In addition, between 0;9 and 0;11, the infant's vocal repertoire consists mainly of readily isolable units bounded by pauses. These units called PHONETICALLY CONSISTENT FORMS (PCFs) are particularly rich in prosody, unlike either the babbling, which appears earlier, or the meaningful words that will appear later (Dore *et al.*, 1976).

Vocalizations were sampled in two different conditions i.e. during mother–infant interaction and during solitary activity with an object. Subsequently each vocal expression was characterized either as 'communicative' or 'investigative'. This identification did not rely on the condition the vocalization was uttered, but on the functional analysis of the accompanying nonvocal behaviours from the video. For the discrimination of acoustic patterns, a PITCH PATTERN RECOGNITION software system (henceforth: PPR-system) was employed. The PPR-system not only calculates acoustic signal features, but it also automatically detects the arrangement of such features in patterns. These patterns are then used to classify vocalizations that have already been assigned in different functional categories by independent analyses of behaviours (Gonzalez & Woods, 1992; Papaeliou *et al.*, 2002).

METHOD

Participants

Four infants, two boys and two girls, participated in this study, with their mothers. All infants were healthy, full-term, first born, English-speaking Caucasians, living in Edinburgh, Scotland. Participants were recruited through announcements given to the mothers in the maternity hospital.

Data collection

The material of this study is part of material collected for a more extensive research on infant expressive behaviour in the second half of the first year, where infants were recorded at their home every fortnight from 0;7 to 0;11 months, under different conditions. For the present study we selected only the vocalizations produced at two sessions when mean age of the infants was 0;9·20 and 0;10·10 months old (*S.D.*: 1 week), and under

TABLE 1. *Actual duration of conditions A and B in every session for individual participants*

Conditions	Session (months)	Duration (minutes)				Total Duration (minutes)
		Tom*	Kathryn	Robin	Julie	
Playing with mother (A)	0;9·20	7	7	7	9	30
	0;10·10	8	7	7	7	29
Playing alone with objects (B)	0;9·20	2	7	8	4	21
	0;10·10	3	6	4	5	18
TOTALS		20	27	26	25	98

* Children's names are not real.

two conditions: (A) when the infant was playing with mother, and (B) when the infant was alone in the presence of various objects. The order of the observational conditions varied across infants and sessions, so as to counter possible effects of a circumstance resembling Strange Situation. In any case, the researcher had been visiting infants' homes for more than 2 months and the infants had become well acquainted with her.

Each condition was planned to last 7 minutes. In reality, it was rarely possible to obtain a satisfactory recording of this duration for condition B, 'the infant alone in the presence of various objects'. In agreement with findings of Bakeman & Adamson (1984), we observed that infants of this age prefer to spend time in coordinated joint play and do not like to be left alone to play. Even in their own home, infants often exhibited great resistance when the researcher asked the mother to leave the room, and showed persistent signs of frustration and loneliness (e.g. crying or trying to follow the mother out of the room). Table 1 summarizes the actual duration of each condition in every session for individual participants.

Each session was recorded simultaneously on video and audio. Video recording was made using a Panasonic HiFi camera, mounted on a tripod. In order to provide signals of good enough quality for the acoustic analysis, audio recording was made using a SONY Digital Audio Recorder with an external one-channel microphone. The microphone was also mounted on a tripod and placed next to the camera. The mother was asked to try to be sitting on the floor opposite to her infant, in such a position that camera and microphone would aim to the space between them, as close as possible so as not to distract their interaction.

The procedure described above yielded a total of 193 vocalizations (mean rate of vocal production: 1·97 vocalizations/minute). Two sounds were considered as distinct vocalizations when they were separated by an interruption of phonation lasting at least 50 msec, that is the period considered as the minimal possible pause. Vocalizations separated by pauses lasting more than 2 sec were considered as belonging to different sequences. This criteria were also used by Stark (1978) and D'Odorico and colleagues (1985). Cries, fussy expressions, laughters, and vegetative sounds (e.g. coughs, burps, grunts, or sighs) were excluded. Other vocalizations were discarded in the following circumstances: (a) when an infant's vocalization was overlapped with mother's voice or an external loud sound, (b) when the signal/noise ratio was unacceptable (e.g. vocalization was produced during manipulation of a very noisy toy), (c) when the vocalization was produced with the infant's hand or part of an object in the mouth, and (d) when the intensity of the vocalization was too low to produce a clearly visible trace on the spectrogram. The vast majority of vocalizations (83%) consisted of isolated vowels, 12% of sequences of vowels and 5% of CV syllables.

DATA ANALYSIS

Classification of vocalizations

Vocalizations were first assigned by inspection of the video in one of the two broad categories: 'communicative' and 'investigative'. Unlike the case in other studies (Delack & Fowlow, 1978; D'Odorico & Franco, 1991), in which the function of a vocalization has been identified merely by the presence or absence of other people or objects, in the present study it was felt that the functional classification of a vocalization should not be identified only from the context in which the vocalization was uttered. The infant may be vocalizing alone, but his or her gaze might be directed at the place where the mother had left the room. The conditions A and B were only introduced in order to encourage different kinds of spontaneous vocal performance, for different purposes. However, the functional attribution of each vocalization was not based on the circumstances in which it was uttered. The infant's quality of involvement in a particular context is reflected in coherent configurations of all expressive modalities, including gestures, facial expressions, orientations, and postures, as well as vocalizations (Halliday, 1975). For this reason, the characterization of each vocalization as 'communicative' or 'investigative' was based on a comprehensive analysis of the concurrent non-vocal behaviour. A non-vocal behaviour was considered as related to a particular vocalization, if it occurred in the temporal framework defined 3 seconds before the beginning and 3 seconds after the end of the vocalization. This timeframe was chosen because it has been demonstrated that human motoric and attentional function in communicative and non-communicative contexts, tends to be organized in coherent intentional groups or phrases lasting approximately 3 seconds (Pöppel & Wittmann, 1999).

A vocalization was classified as 'communicative' if it was related to at least one of the following categories of behaviour that may be reliably detectable from the video (Harding & Golinkoff, 1979; Franco & Butterworth, 1996):

LOOKS AT MOTHER: infant clearly gazes at mother's eyes, face or body;
TOUCHES MOTHER: The infant playfully touches mother's face, head or body;
SEEKS MOTHER: infant looks at the direction mother left in condition B;
SEARCHES MOTHER: infant moves towards the place mother left in condition B;
DIRECTS ATTENTION THROUGH GAZE: infant alternates gaze from an object of interest to mother at least once;
POINTS: infant's index finger is extended in direction of an object, while looking at the mother;
REACHES: the infant's arm is extended with the palm open in a grasp posture;
GIVES SPONTANEOUSLY: infant puts an object into mother's hand without this being requested;
SHOWS: infant holds out an object to mother, but does not permit her to take it;
OFFERS: infant holds out an object to mother and permits her to take it;
FOLLOWS GAZE: the infant looks at the same direction with the mother;
FOLLOWS POINT: the infant looks at the direction indicated by the mother's pointing or other directive action (i.e. giving, offering, showing);
TAKES: After the mother has offered an object the infant takes it;
COMPLETES ACTION: infant acts so as to complete a task with mother;
OBEYS MOTHER'S DIRECTIVE: infant acts according to the mother's directive immediately following the directive.

A vocalization was classified as 'investigative' if it was related to at least one of the following categories of behaviour (Dore *et al.*, 1976; Trevarthen & Hubley, 1978):

HOLDS OBJECT: infant has an object in hands and looks at it;
INSPECTS OBJECT: infant attempts to modify the orientation or configuration of an object by displacing it with the hand, while looking at it;
TASK COMPLETION: infant acts on an object so as to achieve a goal related to preceding behaviour (e.g. putting blocks in a box).

Two trained researchers independently assessed the video of each session, and classified each vocalization in one of the above two groups of behavioural categories. Cohen's kappa (κ) which measures interjudge agreement, was calculated in order to examine the reliability of measurements. The kappa was 0.89 for the 'communicative' group and 0.86 for the 'investigative' group. For the purposes of the present study vocalizations produced while the infant initiated an interaction and vocalizations produced as a response to a mother's communicative act were collapsed in one category.

The attribution of meaning in infant vocalizations that is identical to the meaning derived from the functional coding of accompanying gestures has been a widely applied method in studies of prelinguistic vocal expressions (Dore *et al.*, 1976; Harding & Golinkoff, 1979). However, following this method exclusively may lead to ambiguous data. Although infant vocal and gestural repertoires are combined in a coherent expressive system, it might be the case, especially in the first year that at a given time these two communicative modes convey different but yet not contradictory messages. For example, the infant may be inspecting a toy when the mother has left the room, while at the same time he/she is uttering a moan as a sign of loneliness. This inspection is further supported by research findings demonstrating that infants are particularly sensitive to the emotional regulation of interpersonal contact first and before any exploration of the inanimate environment (Trevarthen, 1999). Similarly, it has been shown that almost every instance of pointing is accompanied by a vocalization, which acts as an intensifier expression of emotion (Leung & Rheingold, 1981). A given psychological state of action does not consist exclusively either of emotional or of referential content; rather, these two aspects of the motive for action co-exist, and the relation between the emotion and the referring may be reflected in a particular expression. In order to avoid including such ambiguous data in the sample, two psychology students, blind to the purposes of the study, were asked to intuitively identify vocalizations conveying the primitive emotions of 'happiness' and 'sadness'. These vocalizations were also not included in the sample.

Acoustic analysis and pitch pattern recognition[1]

The procedures of selection of vocalizations from the collected data described above yielded a final corpus of 156 (81%) items, which were subjected to acoustic analysis. Eighty-one vocalizations were classified as 'communicative' and 75 as 'investigative'. Seventy-five (93%) of the 'communicative' vocalizations were produced in condition A, and 60 (80%) of the 'investigative' vocalizations were produced in condition B.

[1] Part of the procedure of the acoustic analysis and the design of the Pitch Pattern Recognition software system was also described in: 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.

Each vocalization was sampled at 44·1 kHz. An analysis window of about 25 ms was created using a 1024-point Fast Fourier Transform. The resulting spectrum was smoothed by means of 5-point averaging operation, and the first 512 points were employed to form the spectrogram of the vocalization. The number of analysis windows varied depending on the duration of the vocalization. The 'noisy' edges of the vocal signal which exhibit extraneous values that are not representative of the tonal movement of the vocalization, were identified from the spectrogram and were cut off. A Fast Fourier Transform (FFT) of the pruned signal was then computed using the parameters described above. At the second stage the fundamental frequency (F0) contour was calculated using a threshold based detection algorithm. This algorithm examines local amplitude differences in each amplitude spectrum and decides on the first maximum (i.e. F0) on the basis of a predetermined threshold value that may be altered interactively.

In order to examine whether vocalizations classified from the video as 'communicative' and 'investigative' can be discriminated on the basis of their pitch pattern, we developed the Pitch Pattern Recognition (PPR) software system. The design of the PPR-system was performed in two stages: (a) feature generation, and (b) classifier design. In the first stage 9 features were calculated from the spectrogram: the duration of the vocalization, the beginning, final, maximum, minimum, and mean fundamental frequency values, the range and standard deviation of fundamental frequency. These features are the ones most commonly used in studies of acoustic analysis of infant vocalizations (Scherer, 1986). The relative value of standard deviation of fundamental frequency depends on the duration of the vocalization. For this reason the ratio of standard deviation of fundamental frequency to duration of the vocalization was used in addition to the standard deviation.

In the second stage of the PPR-system design, the vocalizations were automatically classified on the basis of the 9 generated fundamental frequency features by employing a commonly used class-discriminating algorithm, the LEAST SQUARES MINIMUM DISTANCE (LSMD) classifier. The classifier maps each token from a feature space into a decision space wherein tokens are clustered around a pre-selected point. This transformation is chosen such that the overall mean-square mapping error is minimized. The 'leave-one-out' method was, then, applied in order to assess the reliability of the classifier. In particular, for the design of the classifier one token was left out to be employed as a test sample. After classification, the omitted token was reinserted into the data pool and the classifier was redesigned, this time leaving out another token for classification. This procedure was repeated until all tokens had been tried (Gonzalez & Woods, 1992). The leave-one-out method was applied for all possible feature combinations

TABLE 2. *Classification of vocalization by function and by pitch pattern*

	Pitch pattern			
	Communicative		Investigative	
Video analysis	f	%	f	%
Communicative	73	89·7	4	5·6
Investigative	8	10·3	71	94·4
TOTAL	81	100	75	100

TABLE 3. *Acoustic features of vocalizations*

	Communicative (N=81)		Investigative (N=75)	
Acoustic Features	Mean value	S.D.	Mean value	S.D.
Mean Fo (Hz)	396·18	56·54	356·63	55·86
Max Fo (Hz)	612·39	116·98	554·28	127·66
S.D. Fo (Hz)	65·30	30·06	53·27	24·21
Duration (sec.)	0·35	0·25	0·96	0·45
S.D. Fo/dur. (Hz/sec.)	264·10	174·71	73·38	58·83

(i.e. 2, 3, 4 etc. feature combinations) in order to determine the feature combination that achieved the highest classification accuracy.

RESULTS

The results obtained by the PPR-system demonstrate that vocalizations classified from the video analysis as 'communicative' and vocalizations classified as 'investigative' can be distinguished from each other on the basis of their pitch pattern as well as their duration with an overall accuracy of 91·67%. The discriminating accuracy for 'communicative' vocalizations was 89·7% and for 'investigative' vocalizations 94·4% (Table 2).

The leave-one-out method produced an optimum pitch pattern that distinguished between 'communicative' and 'investigative' vocalizations. This pattern consists of the following features: mean fundamental frequency (Mean Fo), maximum fundamental frequency (Max Fo), the standard deviation of the fundamental frequency (S.D. Fo), the duration of vocalization, and the ratio of standard deviation of fundamental frequency to the duration of vocalization (S.D. Fo/dur.). As shown in Table 3, vocalizations classified as 'communicative' exhibited higher mean and maximum fundamental frequency, higher standard deviation of fundamental frequency, higher ratio of standard deviation of fundamental frequency to duration of vocalization and shorter duration, compared to vocalizations classified as 'investigative'.

DISCUSSION

We have investigated the hypothesis that ten-month-old infants may produce acoustically distinctive vocal patterns that express their different intentions in communicative interactions and in solitary exploration of the inanimate environment. For the acoustic analysis of the vocalizations a Pitch Pattern Recognition (PPR) software system was designed, which automatically detects differences between classes of vocalizations on the basis of their combinatorial acoustic patterns, rather than single acoustic features. The results obtained by the PPR-system show that the fundamental frequency pattern of vocalizations identified by video analysis of non-vocal behaviours as 'communicative' could be distinguished from the fundamental frequency pattern of vocalizations identified as 'investigative' with an overall accuracy of 91·67%.

Specifically, it was shown that 'communicative' vocalizations display a higher pitch pattern compared to 'investigative' vocalizations. Other studies that have adopted different methods of vocal analysis (i.e. contrasting physical acoustic vs. subjective auditory analyses), have, similarly, shown that in the last quarter of the first year, vocalizations uttered in communicative contexts are likely to be marked by high-pitched melodic contours, while vocalizations uttered during object inspection are likely to be marked by low-pitched melodic contours (Dore *et al.*, 1976; Delack & Fowlow, 1978). These findings indicate that communicative vocalizations are more 'expressive' and appear to give information about more intense states of mind.

Moreover, in this study it was demonstrated that communicative vocalizations were shorter compared to 'investigative' vocalizations (approximately 0·35 vs. 1 second). The duration of 'communicative' vocalizations approximates the average duration of a syllable in the speech of adults (Lynch *et al.*, 1995). It is possible that the infants in the present study produce relatively short 'communicative' utterances in interactions with their mothers because their expressions are adapted to trigger an immediate response in an ongoing communicative exchange as well as to stimulate the communicative partner to take on her part. On the other hand, 'investigative' vocalizations that may be regarded as more 'thoughtful' do not require any response. These utterances of the infants were comparatively long, and they may have served in organizing the explorative activity and associated cognitive processes of the infant.

In contrast to our results, D'Odorico & Franco (1991) found that, after nine months, vocalizations produced in different contexts i.e. during communicative exchanges with an adult, and during inspection of a toy, do not differ in their a acoustic characteristics. In that study 'context' was defined on the basis of the infant's gaze direction; thus, a vocalization was classified as 'communicative' only if the infant was looking at the adult. Although gaze is an important social marker, it is not the only means of initiating or maintaining an interaction. Especially after nine months, gestures accompanying vocalizations are considered as one of the primary means infants have to express their communicative intentions in 'protolanguage' (Halliday, 1975). Gaze may sometimes not mark the communicative act. For example, the infant may deliberately communicate with the adult by giving her an object and vocalizing, while the infant's gaze remains directed towards the object. Thus, D'Odorico & Franco's failure to find acoustic differences may be related to their coding system, which did not take gestures into consideration.

Halliday (1975) reports that infants' ability to express a distinction between 'Pragmatic' and 'Mathetic' vocalizations, appears first at 1;4. The author proposes that it is only at this age that the child can mentally distinguish and, thus, vocally differentiate, 'external' acts addressed to significant others from 'internal' or 'self-monitoring' acts related to exploratory behaviours. Nonetheless, widely observed developments in the infants' communicative and cognitive abilities beginning at nine months are in agreement with our finding that infants of this age produce distinctly different vocal patterns in communicative and explorative contexts.

In communication, infants of this age begin to perceive a partner as a carrier of interests, intentions and feelings that may be related to their own. Infants begin to see people but not objects as psychological agents. Moreover, they can coordinate their attention to a communicative partner with their attention to an object of the surrounding environment. Thus, before the end of the first year infants are capable of directing the other's attention to an object of their own interest, as well as of following the other's attention to an object and they can participate purposefully in co-operative activities. This accomplishment may in turn sustain progress of joint attentional skills as infants begin to "tune in" to others and get others to "tune in" to them. These behaviours are not just intended to draw attention to self but they appear to be intended to share interesting aspects of the environment with others (Scaife & Bruner, 1975; Butterworth, 1991; Legerstee & Barillas, 2003). This ability, which is served by prelinguistic vocalizations combined with gestural behaviours, is considered as a prerequisite for the learning of language, since in language acquisition the learner does not learn FROM another but THROUGH AND WITH another (Tomasello, 1993).

Similarly, on the cognitive level, ten-month olds (who are in Piaget's sensorimotor stage IV) become capable of combining schemas of familiar means to achieve novel goals, and this is interpreted as the earliest form of 'problem solving'. The establishment of a 'means-end' relation is held to be a cognitive achievement necessary for the development of symbolic function (Piaget, 1954; Bates *et al.*, 1975). It is suggested that humans derive cognitive benefits from the exercise of vocal expression. Language certainly shapes cognition by mediating and integrating the basic psychological processes of attention, perception, memory, and emotion, while overt speech may direct motor actions. Neuromotor operations of speech production can keep track of the mental processes of analysis and synthesis (Vygotsky, 1934/1962; Diaz & Berk, 1992). The immature nervous system of infants may not allow representations of cognition to proceed in independence of the motor processes of articulation and manipulation, and even preschool children use language not only for social communication, but also to guide, plan, and monitor their own activity (Vygotsky, 1934/1962; Papousek & Papousek, 1981; Diaz & Berk, 1992). Indeed, adults may be doing the same in many occasions. Our results indicate that solitary activities with objects are often accompanied by an acoustically distinct set of vocalizations in infants around one year of age. It appears that even prelinguistic vocalizations can assume a self-guiding function.

Given the small number of participants, this study cannot pronounce on what may be found in a larger population. Individual differences in early vocal production have often been recorded. These differences are considered to be due to the various changes in brain function, as well as to effects of the variety of environmental inputs that might affect the infants' vocal behaviour (Flax *et al.*, 1991). However, although some behaviour patterns may be quite different in different infants, there may be also common features of form and function across individuals. Moreover, the PPR-system is designed to classify vocalizations on the basis of rather strict discriminating criteria. As previously mentioned, the LSMD is developed so as to minimize the overall mean-square mapping error. In addition, employing the leave-one-out method the classification is checked for each and every token for all possible feature combinations. Thus, any individual differences are controlled out and classification of vocal parameters is performed on the basis of common acoustic features across individuals.

Using this method we have demonstrated that prelinguistic vocalizations can serve a two-fold function; that is, they can be directed to perform intentional communication or exploration of the inanimate environment by the individual. Further research may investigate the functional differences between social and private acts of expression at prelinguistic and linguistic stages of development.

REFERENCES

Bakeman, R. & Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development* **55**, 1278–89.

Bates, E., Camaioni, L. & Voltera, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly* **21**, 205–26.

Butterworth, G. E. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (ed.), *Natural theories of mind*. Oxford: Blackwell.

Camaioni, L. (1993). The development of intentional communication: a re-analysis. In J. Nadel & L. Camaioni (eds), *New perspectives in early communicative development*. London: Routledge.

Chafe, W. (1970). *Meaning and the structure of language*. Chicago: University of Chicago Press.

Delack, J. B. & Fowlow, P. J. (1978). The ontogenesis of differential vocalization: development of prosodic contrastivity during the first year of life. In N. Waterson & C. Snow (eds), *The development of communication*. New York: Wiley.

Diaz, R. M. & Berk, L. E. (eds) (1992). *Private speech: from social interaction to self-regulation*. Hillsdale, NJ: Erlbaum.

D'Odorico, L. & Franco, F. (1991). Selective production of vocalization types in different communication contexts. *Journal of Child Language* **18**, 475–99.

D'Odorico, L., Franco, F. & Vidotto, G. (1985). Temporal characteristics in infant cry and non-cry vocalizations. *Language and Speech* **28**, 29–45.

Dore, J., Franklin, M. B., Miller, R. T. & Ramer, A. L. (1976). Transitional phenomena in early language acquisition. *Journal of Child Language* **3**, 13–28.

Flax, J., Lahey, M., Harris, K. & Boothroyd, A. (1991). Relations between prosodic variables and communicative functions. *Journal of Child Language* **18**, 3–19.

Franco, F. & Butterworth, G. (1996). Pointing and social awareness: declaring and requesting in the second year. *Journal of Child Language* **23**, 307–36.

Furrow, D. (1984). Young children's use of prosody. *Journal of Child Language* **11**, 203–13.

Gonzalez, R. C. & Woods, R. E. (1992). *Medical image processing*. New York: Wesley.

Halliday, M. A. K. (1975). *Learning how to mean: explorations in the development of language*. London: Edward Arnold.

Harding, C. G. & Golinkoff, R. M. (1979). The origins of intentional vocalizations in prelinguistic infants. *Child Development* **50**, 33–40.

Haynes, W. O., Zylla-Jones, E., Smith, J., Rodekohr, R., McEachern, D. & Berry, B. (2004). A longitudinal analysis of timing differences in visual checking and vocalization associated with declarative pointing in infancy. *Infant Behavior and Development* **27**, 230–45.

Legerstee, M. & Barillas, Y. (2003). Sharing attention and pointing to objects at 12 months: is the intentional stance implied? *Cognitive Development* **18**, 91–110.

Leung, E. H. & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology* **17**, 215–20.

Lynch, M. P., Oller, D. K., Steffens, M. L. & Buder, E. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology* **28**, 3–25.

Marcos, H. (1987). Communicative functions of pitch range and pitch direction in infants. *Journal of Child Language* **14**, 255–68.

Masataka, N. (1995). The relation between index-finger extension and the acoustic quality of cooing in three-month-old infants. *Journal of Child Language* **22**, 247–57.

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.

Papousek, M. & Papousek, H. (1981). Musical elements in the infant's vocalizations: their significance for communication, cognition and creativity. In L. P. Lipsitt (ed.), *Advances in infancy research*, Vol. 1. Norwood, NJ: Ablex.

Petitto, L. A., Holowka, S., Sergio, L. E. & Ostry, D. (2001). Language rhythms and baby hand movements. *Nature* **413**, 35–6.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
Pöppel, E. & Wittmann, M. (1999). Time in the mind. In R. Wilson & F. Keil (eds), *The MIT encyclopedia of cognitive sciences*. Cambridge, MA: MIT Press.
Reddy, V. (2001). Mind knowledge in infancy: understanding attention and intention in the first year. In G. J. Bremner & A. Fogel (eds), *Blackwell handbook of infancy research*. Oxford: Blackwell.
Scaife, M. & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature* **253**, 265–6.
Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin* **99**, 143–65.
Stark, R. (1978). Features of infants' sounds: the emergence of cooing. *Journal of Child Language* **5**, 379–90.
Tomasello, M. (1993). On the interpersonal origins of self-concept. In U. Neisser (ed.), *Ecological and interpersonal knowledge of the self*. New York: Cambridge University Press.
Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. In Rhythms, Musical Narrative, and the Origins of Human Communication. *Musicae Scientiae, Special Issue*, 1999–2000, Deliègel, ed. Liège, Belgium: European Society for the Cognitive Sciences of Music.
Trevarthen, C. & Hubley, P. (1978). Secondary intersubjectivity: confidence, confiding and acts of meaning in the first year. In A. Lock (ed.), *Action, gesture and symbol: the emergence of language*. London: Academic Press.
Vygotsky, L. S. (1934/1962). *Thought and language*. Cambridge, MA: MIT Press.