FORM AND FUNCTION OF NON-LINGUISTIC CALLS IN HUMAN INFANTS

Verena Kersken

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2012

Form and Function of Non-linguistic Calls in Human Infants

Verena Kersken

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy

at the

University of St Andrews

December 2011

1. Candidate's declarations:

I, Verena Kersken, hereby certify that this thesis, which is approximately 55,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as as a candidate for the degree of Doctor of Philosphy in Septmner 2008; the higher study for which this is a record was carried out in the University of St Andrews between 2008 and 2012.

Date …… signature of candidate ………

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Phylosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date …… signature of supervisor ………

3. Permission for electronic publication:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

(ii) Access to all of printed copy but embargo of all of electronic publication of thesis for a period of two years (*maximum five*) on the following ground(s):

publication would preclude future publication;

Date ……

signature of candidate ……

signature of supervisor ………

Abstract

Before infants speak their first word, they already produce a large variety of sounds. Whilst the developmental process that leads to speech production is very well documented, little attention is given to how non-linguistic sounds function in the child's everyday environment and whether they show acoustic consistencies similar to those found in the calls of non-human primates. This thesis investigated whether human infants between 11 and 18 months have "calls". The first study observed 22 infants in their everyday nursery environment in Scotland and identified a number of contexts in which infants produced vocal behaviour. Vocalisations in five of these contexts, giving, declarative pointing, food requests, protests and action requests, were then subjected to an acoustic analysis. Results of the discriminant analysis suggest that four categories of vocal behaviour can be distinguished on the basis of their acoustic properties alone. To investigate whether these calls are part of a universal human repertoire, we conducted a cross-cultural comparison of the acoustic properties of vocal behaviour showed that, despite a slightly higher level of variation; four categories of calls could still be discriminated above chance level. This suggests that human infants possess calls with rather fixed acoustic properties as part of their vocal repertoire in addition to other, more flexible vocal behaviours. In order to assess whether listeners can gain information from these calls, we conducted a playback study with parents, experienced and inexperienced participants. Results show that all participants can categorise all vocalisations above chance level. Parents were the only participants that showed significantly better scores in correctly classifying vocalisations recorded in Scotland over those recorded in Uganda. Overall, the studies demonstrated that infants, as part of their vocal repertoire, produce some classes of calls that have constant acoustic properties across infants from different cultures, and contain information about the infant's activities that can be picked up by a listener.

Die Zeit ist ein kostbares Geschenk, uns gegeben, damit wir klüger, besser, reifer, vollkommener werden. Thomas Mann

Many people gave their time to contribute to this thesis and for this I will always be grateful. I firstly want to thank my supervisors, Klaus Zuberbühler and Juan-Carlos Gomez. Klaus, thanks for your encouragement, patience and practical knowledge. I am deeply grateful to you for firmly believing in this project even when I as a bit shaky. Juan, your way of doing science and your depth and breadth of knowledge is amazing and I am lucky to work with you. Thank you for being such a patient and kind teacher – and for many, many capital letters.

This thesis would not have been possible without 50 little people, and you guys really deserve the biggest thank you. The same is true for those around you: staff in Acorn and WonderYears Nursery and all parents in Uganda and Scotland. You were invaluable and I cannot thank you enough for your hospitality and kindness. I am especially grateful to the Ugandan National Council of Science and Technology and the Budongo Conservation Field Station for allowing me to live and work in Uganda. God's blessing and asante sana to Magret, Santa, and Flora, who showed a clueless mzungu the ropes, were a source of endless encouragement and optimism, and taught me so much. Thanks also to the team at camp, particularly Anne, Tanja and Geoffrey – your advice and listening ears were greatly appreciated.

The Psychology Department here was always a good place to be. Thanks to all my fellow Wolfson teammates, particularly Katie, Cat, Marion, Cris, Emily, Lisa, Bess, Guill ('s powerchords), and Louis. And of course the general PhD support group, particularly Jessica, Stephane, Louise, Ashley, Joanne. Kate thanks for telling me that I'm rubbish (in the most constructive way possible) whenever I felt like a soufflé in a draft.

Katharina, thanks for numerous invitations to 'whine and dine', countless kilometres on the old running shoes, and generally for pulling me along Ali, Jo, Ceri and Bee – you deserve your own power-metal ballad. Thanks for never being far away. The Cosmos Thursday Group was the perfect place to get away. Thanks to you guys for reminding me that Abba and a cup of tea make everything better. Lucy, you deserve a huge thank you, especially for sending me cake in the post. Mariam, how can I thank you for hours of listening, all the tears and laughter, jumps through random rapeseed fields, visits to all kinds of places (including those with edges), and for sharing a home - other than by saying "I loaf you – dough much".

Martin, du oller Quallenknödel, könnte ich dir 'Postcards from Italy' schenken, ich würd's tun. Danke. Annette, deine Couch war immer wie Urlaub – vielen Dank für's da-sein, zuhören, Recht haben und sämtliche pädagogische Massnahmen.

Der grösste Dank geht an meine Mutter, Marion Kersken. Der olle Goethe sagt, man soll seinen Kindern Wurzeln und Flügel geben, und von dir hab ich nicht nur tiefe Wurzeln sondern auch besonders grosse Flügel bekommen.

There is one person, who really deserves a lot of gratitude but is unfortunately not around anymore. Oppa, I hope you would be proud of this.

Note to the reader:

Throughout this thesis I have used the pronoun 'we' instead of 'I'. The work here is my own in terms of hypotheses, analyses and conclusions, but is effectively the product of a ratchet effect (Tomasello 1999) of collaboration and discussion.

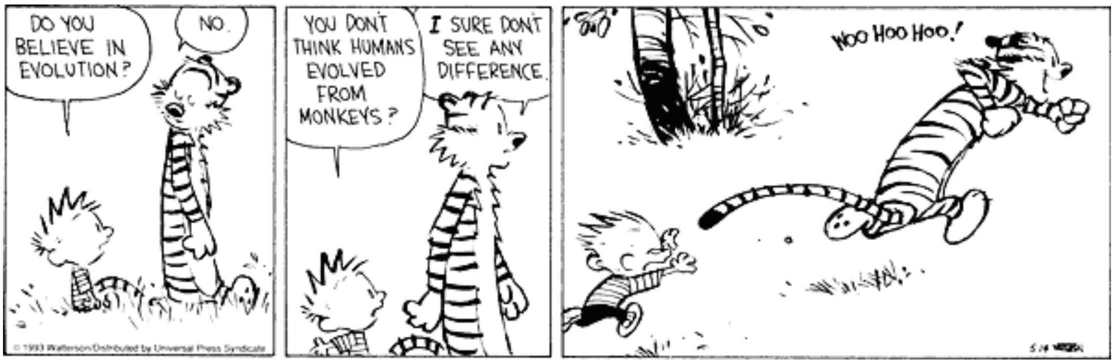

DO YOU BELIEVE IN EVOLUTION?
NO.
YOU DON'T THINK HUMANS EVOLVED FROM MONKEYS?
I SURE DON'T SEE ANY DIFFERENCE.
WOO HOO HOO!
© 1993 Watterson/Distributed by Universal Press Syndicate
5-14

Table of Contents

Chapter 3: Cross-cultural Comparison of Infant Calls from Scotland and Uganda

List of Figures

Chapter 2

Chapter 3

Chapter 4

"Look, they are one people, and they all have one language; and this is only the beginning of what they will do: nothing that they propose to do will now be impossible to them." Genesis 11:6

General Summary

Language is a powerful tool. Linguistic communication not only enables the construction of a tower that reaches the high heavens, but allows humans to share information and knowledge to a degree that likely surpasses the communication systems of any other animal. Language is amazing, and poses a number of puzzles and questions, and nothing short of centuries have been spend answering some of these. One particular puzzle is how human infants progress, within the first two years of life, from uttering screams and cries to producing words, maybe even sentences, and generally seem to be equipped for the task of sharing in a linguistic environment.

Initially scientists documented how their own children progressed from crying to the first words (examples include Darwin and Taine 1877), assuming that there must be certain developmental milestones that, following each other in a logical sequence, eventually enable the child to produce and comprehend words and sentences. The researchers started to transcribe the infant's utterances using the phonetic alphabet, again with the goal of identifying prelinguistic patterns and progression in vocal behaviour that were the basis for the production of language. Research progressed, aided by the advent of recording equipment and quantitative analytical methods, and today we have a very detailed picture of how an infant's vocal production changes and develops before she produces her first word (for example Oller 2000).

But language development is not limited to the acquisition of arbitrary sound patterns. One function of language is communication. At its most basic communication requires that a message of some kind is passed on between a sender and a receiver, independent of any intention to pass on this message (Franco 1997, Sperber and Wilson 1995). At its most complex, communication uses a common set of arbitrary symbols, which are combined by some rules and intentionally produced to transmit information, either by indicating the intention to communicate or by encoding some form of information, to a receiver - language. Most living organisms communicate in some form or another, and fall somewhere on this spectrum. Communication has therefore got a much older evolutionary history than language, and diverse communication systems have been identified in many animal species, for example red tails on sticklebacks (Janik and Slater 2000) or the functionally referential alarm calls of vervet monkeys (Struhsaker 1967, Cheney and Seyfarth 1980).

In this thesis we want to investigate whether human infants have some kind of sound-meaning system in place before and/or alongside the linguistic signals they learn to produce. Cries and grunts have been identified as signals that serve both a biological and social function, and caregivers readily recognize and assign meaning to them (Lester and Boukydis 1989, Papoušek 1992, McCune et al. 1996). Signals like these are often dismissed as trivial (Vihman, 1996). They are said to be linked to biological functions and do not appear to be too influential in language acquisition. Nevertheless, signals that have a reliable connection between sound and meaning can be valuable for the caregiver when it comes to inferring the infant's needs, wants and emotional states. Furthermore, it forms no contradiction to more complex forms of communication, for example those found in gestures (Liszkowski 2007) or the process of complex sound production as part of language acquisition (Oller 2000).

Here, we wanted to investigate whether prelinguistic infants between the ages of 11 and 18 months, an age commonly associated with the onset of intentional communication, possess 'calls', vocal behaviours whose acoustic properties vary systematically with the context in which they are produced. Secondly, we wanted to know whether this kind of vocal behaviour is influenced by the culture (and language) surrounding the infant. Lastly, we wanted to ask whether listeners could infer information about the production context from these signals alone.

The concept of calls, a vocal behaviour that is context-specific and functions to influence listener's behaviour, is commonly associated with the vocal behaviour of non-human primates. Here, we adopted a methodology that effectively treats human infants as an 'unknown primate species' (Gómez 2007) to avoid a linguistic bias and interpretation of results in light of later communicative developments. Leaning closely on studies of the vocal behaviour of non-human primates, our approach therefore consisted of observing the infants' vocal behaviour in everyday situations, identifying situations in which vocal behaviour was commonly observed, and using a quantitative acoustic analysis to compare vocalisations across situations.

Our first project was conducted with 18 infants in two day-care nurseries in Fife, Scotland. We recorded instances of vocal behaviour in everyday contexts such as meal times or play and identified a number of categories in which vocalisations were reliably observed across most infants. We aimed to describe these categories in behavioural terms and, where possible, recorded reactions they provoked in the listener. Five of these categories yielded enough good quality recordings to allow for a comparison based on their acoustic properties. We therefore measured 10 different acoustic variables that described infant vocalisations emitted in the categories of giving, declarative pointing, action requests, protests and food requests. Results from a discriminant analysis identified a number of basic algorithms that distinguish between vocalisations emitted in the different categories at a level above chance. This means that signals like these could theoretically transmit social information.

To further investigate the nature of these signals, and specifically whether they are part of a universal, basic sound-meaning system, we conducted a similar study with 22 infants from five villages in the Masindi district, Uganda. Again, we aimed to record instances of vocal behaviour in the infants' natural interactions. This time we were interested firstly whether Ugandan children would vocalise in the same situations as their Scottish counterparts, and, secondly, whether their vocalisations exhibited the same acoustic properties as those we recorded in

Scotland. We found that, despite general differences in vocal behaviour, Ugandan infants produced vocal behaviour in the same situations. Furthermore, after conducting the same analysis that we used for the Scottish sample, the Ugandan data also suggest that the acoustic properties of calls recorded in different contexts exhibit different acoustic properties. When we compared the Scottish and Ugandan data, however, there seemed to be some variation in the acoustic properties of the vocalisations.

In order to investigate whether differences found between the cultures are actually meaningful to a receiver, as well as to generally assess the information content of our recorded vocalisations, we designed a playback study that tested how well adult listeners (parents, experienced participants and non-parents) could infer the production context from the vocalisation alone. We found that all participant groups could match a vocalisation to its respective production contexts at a level above chance. Parents performed slightly better than the other two groups, but this effect was not significant. Participants listened to audio clips from both Ugandan and Scottish infants to investigate whether the infants' cultural background affected the listener's judgement of the calls. Neither experienced nor inexperienced participants showed a difference in performance between audio clips from either culture. Only parents displayed a significant difference: they were significantly better at categorising calls recorded in Scotland. Results from the playback study suggest that calls we identified have some inherent referential value that is informative to listeners.

Before we present the methods and results of our studies in detail, we will review previous studies and a body of literature from a variety of fields including phonology, pragmatics, gestural development and animal communication, to illustrate what is known about non-linguistic infant vocal behaviour, where the main emphasis in research lies and to identify the gaps that our research aims to address.

Chapter 1: Introduction

Before they begin to speak, human infants produce variable, flexible sounds that are the foundation of speech production (Vihman 1996). But speech production, despite receiving by far the most research interest, is not the only task these sounds accomplish (Locke 1993), nor are speech-related vocalisations the only sounds in the infant vocal repertoire. The first signals an infant produces are cries, screams and grunts, and these seemingly inflexible signals are used by caregivers as an indicator of the infant's needs and wellbeing (Papoušek 1992). Although the human infant quickly starts to produce novel signals related to speech acquisition, these earlier vocalisations remain part of the infant's repertoire at least until the end of their second year (Locke 1993, Franco 1997). There is, however, very little empirical research that focuses on these vocal signals and the role they play as a potential source of information about the infant's activities, emotional states and well-being for caregivers, particularly in infants who already produce more complex vocal behaviour.

In this thesis we wanted to address the question of whether infants at the onset of intentional communication, who already produce flexible vocal signals and are learning their first words, have calls – non-linguistic vocal behaviour with consistent acoustic properties that varies with the situations in which it is produced. In this first chapter we review studies from phonology, communicative development, and primatology to investigate what shape prelinguistic vocal behaviour takes, how the infant communicates before she can speak and how similar infant vocal communication is to that of our evolutionarily closest living relatives, non-human primates.

The research we will present in this review comes from many different areas, such as vocal and gestural development, phonology and animal communication, that do not necessarily focus on the same themes or overlap in a significant way. We will therefore treat each area of research as somewhat separate, whilst trying to answer the questions of what shapes do non-linguistic sounds take and what functions they serve, are there any correspondences between the sounds and their functions in human infants, what level of communicative competence do infants show in their second year of life and how do listeners react to the infant's non-linguistic communicative signals. Where possible, we will try and compare the human infant's communicative signals and skills to those of other primates.

Vocal development from birth to the first word

A lot of effort has been invested in describing the course of vocal development as the child moves from innate prelinguistic utterances, to the first vowels and syllables, to the first words and the first sentence. Many early observers, amongst them Darwin (1877) and Taine (1877), recorded the behaviour of their own children in great detail. Early methods focused on transcribing the infants' utterances using the International Phonetic Alphabet, but with the onset of technological advances such as audiotape and video recording and computerised acoustic analysis, the systematic study of a large sample of infants from different language backgrounds was realised (Bernhardt and Stemberger 1998).

We want to begin by trying to establish some sort of chronology of the development of vocal behaviour from birth to the second year. Here we want to present a common view of infant vocal development that focuses on what kinds of sounds the infant produces at what age and how she goes from simple, innate sounds to the more complex sounds that are observed in spoken language. In this section we want to introduce the vocal behaviours that have been documented in prelinguistic infants in order to question whether this description is sufficient, both in terms of the completeness of the described repertoire, and the possibility that a strong focus on phonological development ignores the functional aspect of prelinguistic communication.

Five stages of vocal development

Most current authors agree that preverbal vocal development consists of five stages (e.g. ,Vihman 1996, Oller 2000, Masataka 2003, Oller and Griebel 2008, Stark 1980). In the first stage (0-1 month) the infant produces sounds with no modulations, and there is no evidence that the infant is able to manipulate pitch, loudness or length of the vocalisations. These 'quasivowels' (Nathani and Oller 2001, Oller 2000, Bernhardt and Stemberger 1998) are produced spontaneously, both when the infant is alone and with a caregiver.

In the second stage (2-3 months) the infant now modulates the previously produced quasivowels by changing the length of the utterance, the loudness and the number of repetitions. New types of vocal behaviour enter the repertoire, such as laughing and cooing. During this time the infant also begins to use her freshly descended larynx to alter sounds – this is evident in the production of squeals and growls as well as the manipulation of pitch (Titze 1994, Oller and Griebel 2008). What is noteworthy is that the infant does not seem to use her vocal repertoire in a stereotyped way – that is, vocal behaviour is not limited to specific situations or to specific sound shapes, like it is in many animals (Snowdon 2008). Instead the infant vocalises when alone and in interaction with another (Papoušek 1992) and, as described before, varies the shape of her vocal behaviour.

The onset of vocal play (Stark 1980) marks the third stage (4-6 months). In this stage the child seems to discover how to produce systematic manipulations of sound properties, for example pitch patterns or intonation. She produces sequences of sounds with one particular property, for example high pitch, and is increasingly able to alternate them with sounds that have another property (Oller and Griebel 2008). The produced sounds also begin to fall into increasingly distinct categories with defined sound properties, like squeals, growls, yelling and blowing air. Recent studies (Warlaumont, Oller and Buder 2010) suggest that these sounds can be distinguished on the basis of their acoustic properties, for example frequencies, pitch and melodic contours, and neural networks can learn to classify them accurately

(Oller, Niyogi, Gray, Richards et al. 2010). This suggests that although the infant can flexibly combine and produce different aspects of a sound, she also produces some classes of vocal behaviour with constant acoustic properties. These are not associated with a particular production context or with contingencies between sound and function because they are not communicative in the way that crying or screaming is, but their primary function is assumed to be developing and tuning the vocal system for speech production (Oller and Griebel 2008). Although the production of vocal play is not associated with a specific communicative function, caregivers treat these vocalisations as meaningful and provoke the infant into producing them by making them part of a social turn-taking game that has been termed proto-dialogue (Papoušek 1992, Franco 1997).

In proto-dialogues, caregivers treat the infant's non-linguistic vocal behaviour as a meaningful contribution to a linguistic dialogue and comment on and assign meaning to the infant's utterance (Papoušek 1992, 2007). The infant is seemingly joining in; she times her vocalisation in a way that she makes sounds when the adult stops talking, and often matches the pitch and melody of the adult's part. Caregivers perceive these vocal exchanges as indicative of the infant's positive emotional state and willingness to communicate (Papoušek 1992, Masataka 2003).

In the fourth stage (7-10 months) the infant begins to produce regular syllables (Stark 1980, Oller 2000). These syllables are the same that are observed in spoken language and that make up words and sentences (Oller and Griebel 2008). Babbling, which is defined as a repetitive sequence of canonical syllables like /dadada/, also starts during this stage. Babbling is thought to be a uniquely human vocal behaviour because it seemingly serves no function other than practicing the manipulation of vocal productions that are increasingly speech-like (Griebel and Oller 2008, Vihman 1996, but see Elowson, Snowdon and Lazaro-Perea 1998 for a documentation of babbling in a non-human primate). Babbling is thought to illustrate the flexibility in human vocal behaviour because it shows that infants can produce sounds without a specific environmental trigger, such as pain or surprise, and can finely manipulate these sounds, for example alternating vowels or consonants, at will.

Contrasts between consonants or vowels observed in babbling sequences are the first evidence of the infant's native language in their prelinguistic productions. De Boysson-Bardies and Vihman (1991) analysed babbling sequences produced by 9-to-13-month-old infants with different native languages, and found that these sequences feature syllables that have a high frequency in the infant's respective native language. This shows that the infant is adapting her vocal productions to what she hears around her (Kuhl and Meltzoff 1996). This is evidence of vocal learning that requires flexible vocalisations to adapt to any given language and that is absent in many primate species (Janik and Slater 1997, 2000).

Parents perceive these babbling sounds as positive and rate them as more favourable the more speech-like they are (Bloom and Lo 1990). Papoušek (1992) observed that babbling often occurs when the infant is playing by herself or not engaged in a joint activity with a caregiver. In these instances, it also seems that more complex sounds are produced than when vocal behaviour is part of a communicative exchange (Papousek 1992).

During the fifth and final stage stage (11-12 months) the infant expands her babbling repertoire by including alternate syllables with different consonants and vocalic elements, for example '/dabida/'. The child is also able to match and imitate adult vocalisations (Kuhl and Meltzoff 1996). This 'successive approximation' can be observed during picture book reading or labelling with an adult, where the child is encouraged and rewarded for making sounds that correspond to the modelled word. At 12 months, the infant is motorically equipped to produce language – she has a high level of comprehension, she is able to produce fine manipulations of sound and able to imitate what she has heard (Kuhl and Meltzoff 1996).

Descriptions of vocal development often end when the child produces her first word, at around 12 months of age (Stark 1980). Whilst there are some cross-cultural studies on prelinguistic vocal behaviour, these centre on babbling, which shows evidence of the infant's native language. It is unclear whether the other stages of vocal development appear in the same sequence or at the same time in infants of other cultures, particular in those where infants receive less parental input (Keller 2007).

Summary: Stages of vocal development

From the material we have reviewed here we get the impression that infants quickly progress from acoustically simple sounds to skilfully produced sequences and combinations of sounds that reflect their native language. The infants' flexibility in sound production is often emphasized in descriptions. This indicates that, unlike many other animal species, human infants are not limited to an innate, fixed vocal repertoire: they are capable of vocal learning that adapts their own production to what they hear around them, and furthermore can produce sounds in the absence of a specific environmental or emotional trigger. Only the very early productions of the human infant have been considered to reflect an innate repertoire and thought to be an indicator of the infant's emotional state (Papoušek 1992, Lester and Boukydis 1989, Scheiner, Hammerschmidt, Jurgens and Zwirner 2002).

When considering the developmental progression of vocal development, one might get the impression that the vocal behaviour of infants quickly progresses from simple stereotyped sounds to the complexities and flexibility of emerging speech. However, it is important to remember that the vocalisations that are present in the repertoire of younger infants do not disappear and are not replaced by more complex forms of vocal behaviour. Vocal behaviour like crying or screams make up a large part of the infant's repertoire in the first few years of life (Blake 2000, Locke 1997) in addition to more complex vocal behaviour such as babbling.

The descriptions we reviewed here tend to see spoken language as the outcome of vocal development. Infant vocalisations that are not thought to contribute to language learning and exhibit few linguistic characteristics, such as screams or cries, are neglected in favour of those that show language-like characteristics. In addition to neglecting some parts of the vocal output, the study of vocal behaviour also shows a bias in favouring the structure of vocalisations over their communicative functions. Infant vocalisations are described and assessed as to how well they match different units of spoken language, for example syllables or vowels. The emphasis is on the form of the infants' utterances rather than the situations in which they are produced or what communicative functions they serve.

Communicative development

In this section we want to complement our initial phonetic account of language development by giving a brief outline of the human infant's communicative development focusing on the period from birth to the end of the second year. The problem we face is that, to the best of our knowledge, there is not one account of communicative development that combines phonology, cognition and gestures. Therefore our task is to pull together a body of research from a very wide spectrum in order to present a coherent picture of the skills and tools that infants possess in their non-linguistic communication.

The majority of studies in communicative development are conducted in the gestural domain. Whilst vocal signals have mainly been linked to the acquisition of language sounds, gestures are thought to be the road to understanding communication (Tomasello 2008). This has led to different emphases. The vocal domain is mainly investigated with regard to speech building whereas gestural studies emphasize communicative functions. Moreover, the infant's gestures have often been the subject of extensive comparisons to other primates, mainly the great apes.

We will now introduce Bates, Benigni, Bretherton, Camaioni and Volterra's (1979) model of communicative development and apply this to both vocal and gestural infant communication. We aim to give an indication of the infant's communicative skills and how these change from about 9 months of age. We will briefly introduce the field of gestural development with a particular focus on what kinds of gestures there are, how these are studied, what level of reference they display and lastly how they compare to those of other primates.

Bates et al.'s (1979) model of communicative development

As a model for communicative development we would like to use the stages proposed by Bates et al. (1979). In this view the infant progresses through three distinct stages of communicative complexity. Initially the child's communication is *perlocutionary* - the infant has a systematic effect on listeners without any intentional control or awareness of it. For example, the infant screams because she is in pain, and although a listener will hear and interpret the scream, and try to help her, this was not her explicit intention when she produced the scream. The second stage is the *illocutionary* stage in which the child uses non-linguistic signals instrumentally to direct a caregiver's attention and formulate requests. In this stage the child shows increasing awareness of the demands of communication. She is aware that communicative signals change the behaviour of a recipient and that these signals have to be adapted to the recipient's needs, for example clarified or repeated. The commonly used examples here are requests in which the infant uses her gestures and vocalisations to enlist the help of a partner. The last stage is the *locutionary* stage in which the infant begins to produce conventional, arbitrary signals and use them as a means to inform a recipient of something. These signals include both the words of the infant's native language and conventional gestures such as 'thumbs up' to mean a positive evaluation of something.

Communicative competence in prelinguistic vocal communication

In the first three months of life, the infant's communication mainly revolves around her primary needs – nutrition, hygiene, warmth and reduction of pain (Franco 1997). Whilst the infant produces signals that are likely nothing more than a reflection of her emotional state (Oller 2000, Franco 1997, Papoušek 1992, Vihman 1996), it is the caregiver's task to interpret these cries and assign meaning to them. This meets the criteria of Bates et al.'s (1979) definition of the perlocutionary phase. It is important to note that the notion of vocal behaviour as expression of moods or an affective state is somewhat circular: as concepts such as anger or pain are not measurable per se, affective states are usually recognized by vocal behaviour or facial expressions associated with the emotion, or the context in which they are observed and which contains the likely cause of the vocal behaviour.

In the first few months of life, the caregiver's task is made easier by the infant's limited activities as she cannot yet engage in complex behaviour, cannot move much and her needs and desires are possibly limited to warmth, food, comfort, and affection (Blake 1999). When interpreting the infant's vocalisations caregivers therefore engage in a form of "suitably constrained guesswork" (Sperber and Wilson 1995)

Although the vocal repertoire expands significantly between the ages of six and nine months, communicativelythere is little progress. The infant increases the quantity of cry sounds she produces (D'Odorico 1984), and some researchers speculate that she now produces cries or screams simply to obtain a caregiver's attention (Lock 1980). There are not enough studies to confirm impressions that previously reflexive vocal behaviour, produced in reaction to an immediate event, is now produced intentionally with the aim to communicate and/or obtain a caregiver's attention, even in the absence of the original cause of the vocal behaviour. A popular example that rests mainly on anecdotal evidence is the production of fake cries that are thought to emerge as early as six months of age (Wolff 1969, Leung and Rheingold 1981). In fake cries the infant uses a signal that is normally an honest indicator of her emotional state and reproduces it to serve some other means – often to get the immediate attention of an adult (Papoušek 1992).

The production of biological signals, that is signals that are normally triggered by a certain affective state, such as anger, or an external stimulus such as pain, and are part of the infant's innate repertoire where they act as honest indicators of these states, as a means to an end is probably the best indicator that the infant has some understanding of the effect these signals can have on a listener. For example, when the infant produces a vocalisation that is normally observed when she experiences pain, even in the absence of an event causing her pain, solely in order to obtain the caregiver's attention. This shows that the infant on some level understands, possibly through simple associative learning, that vocal signals can be used to obtain and redirect the caregiver's attention – that could then be further manipulated. This shows that there is a decoupling between producing vocal behaviour and their initial cause, that is the intentional production of vocal signals in the absence of environmental or emotional triggers. Furthermore, she shows that she can produce certain types of vocal behaviours even in absence of the environmental or emotional trigger that normally underlies this behaviour and purely as a means to influence a listener's behaviour. For example, the infant can produce cry sounds if she wants attention even in absence of pain, hunger or discomfort. Wolff (1969) stated that fake cries have different acoustic properties from real cries, a possible sign that the acoustic markers of urgency are absent, but does not offer an analysis of his results.

At around twelve months of age the infant will begin to produce locutionary signals, that is, words and eventually sentences in her native language and conventional gestures.

As the relative paucity of studies in this section shows, the transition from perlocutionary to illocutionary signals is not well documented for the vocal signals the infant produces before speech. The infant's communicative competence is mainly illustrated in studies on gestural behaviour.

Communicative competence in prelinguistic gestural communication

Gestures have been proposed to be the cradle of true, human communication (Tomasello 2008). They are thought to be intentional, to use conventional, arbitrary signals and to be adapted to the receiver's needs. Consequently most studies on the transition from perlocutionary to illocutionary communication have focused on infant gestures. Here we would like to present a selection of studies that show how the infant uses gestural signals and how these change through the course of development.

How are gestures studied?

For a gesture to become more than a simple body movement, it needs to be embedded in a communicative context. This firstly requires a receiver, who is often motivated to interpret the infant's communicative signals. Secondly, there are seemingly several contextual constraints that allow the correct interpretation of a gesture. A reaching gesture will almost always be directed at a visible target; this allows the receiver to combine the gesture 'request' with 'that object'. Thereby a very specific meaning is communicated by a gesture with a rather broad functional frame.

Ochs and Schieffelin (1979) argue that gestures require an ecological description that takes into account the shape of the gesture, its proposed function from the infant's point, its meaning as observed from the recipient's reaction, and environmental constraints that help to specify the meaning of a gesture. Descriptions of gestures have a very pragmatic approach, focusing on the goal of the infant's gestural signal, the steps necessary to achieve it and the reaction from a recipient (Blake et al. 1992). This is true for both observational studies that aim to catalogue the infant's gestural repertoire and how it is used in her everyday activities and in experimental studies that test how the infant uses and adapts her gestures in different situations.

In contrast to prelinguistic vocal behaviour, researchers do not strongly associate infant gestures with language acquisition (with the possible exception of declarative pointing). This is evident in the description and interpretation of the results. Gestures are thought to be a sort of interim state of communication and mainly serve the purpose of learning what communication is about, how it is achieved and what effect it has on recipients. Some researchers see gestures, particularly pointing as 'the royal road to language', mainly because a child learns to express communicative intent through gestures (Bruner 1983, Butterworth 2003). These authors propose that, in contrast to vocal behaviour, gestures are not strongly coupled with emotions, and are first and foremost tools of communication for the infant. For example, when an infant fails to reach an object, but tries to grasp it, and a caregiver consistently responds to this failed attempt and gives it to her, eventually she will perform the hand movement not as attempt to reach the object, but as tool to get the adult to help her in obtaining her goal (Bruner 1983). Pointing is even more relevant, as it is likely not originating in a failed action, but is immediately used as a way of singling out objects in the environment – and to show them to or request them from a communicative partner. Other groups, mainly the one around Mike Tomasello (1999) see gestures as a more or less specific communication system that is in place before the infant uses linguistic means that does not necessarily contribute to language. In this view, it seems that gestures are thought to constitute a more primitive or phylogenetically older form of communication. Indeed, many parallels have been drawn between infant gestures and those of other primate species, most notably the great apes.

Two types of gesture

Gestures have traditionally been divided into two types, proto-imperative gestures that mainly include requests for actions or objects and proto-declarative gestures, amongst them showing, giving and pointing, that aim to direct a recipient's attention to a specific aspect of the environment (Bates et al. 1979).

Declarative and requestive gestures have been proposed to have different origins and highlight different aspects of the infant's communicative competence. Some authors propose that requests originate in failed attempts at reaching and are later recruited to serve a solely communicative purpose (Schieffelin and Ochs, 1979). Non-human primates are also observed to use gestures to request things, mostly food (Call and Tomasello,2007), but also objects and actions (Gómez, 2007). Declaratives on the other hand might reflect a uniquely human motivation to communicate about aspects of the environment and share one's attitude towards it (Gómez, Sarriá and Tamarit 1993, Tomasello et al. 2007, Tomasello 1999). Therefore, it seems that, whilst requests are a pragmatic means to an end, declaratives are mainly about the communicative act of sharing one's impression of the environment.

Proto-imperative gestures

The most primitive form of gestures can possibly be termed 'failed actions', for example an infant tries to reach an object, stretches her body and performs grasping actions (Lock 1980). It is debatable whether actions like these are truly communicative, that is whether they are performed to serve as a communicative act and are goal-directed this way, rather than a goal-directed action (Bates et al. 1979). Nevertheless, recipients would probably find it easy to interpret them and help the infant achieve her goal. In a way these gestures are comparable to the crying behaviour that we described earlier. They are in reaction to an immediate event and probably display little or no awareness of their communicative value to the partner. There is, however, to the best of our knowledge no study that assesses whether infants alternate gaze or make sure they have the attention of a caregiver during these failed reaching attempts.

The infant's understanding of the requirements of communication is displayed in how she adjusts her signals to the needs of a listener. For example, Golinkoff (1983, 1986) conducted a study that investigated how 13-month-old infants react to misunderstandings. In this quasi-experimental study infants were placed in a highchair opposite a number of interesting and uninteresting objects that were out of their reach. Parents sat next to the children and were asked to respond differently to the infants' requests by either complying with their request and giving the desired object to the child, partially understanding the request and giving the wrong object, or not showing any reaction. Golinkoff (1986) described the following response patterns for the conditions: When infants were given the desired objects, they stopped performing request gestures, partially understood infants kept repeating their gestures and tried to direct the adult to the correct object, misunderstood infants kept gesturing and displayed a wider range of gestures before eventually giving up and showing signs of frustration.

In terms of communicative competence, the study clearly shows that infants use gestures in a goal-directed manner. When their goal to obtain the interesting object was reached, they stopped. When their goal was not yet met but they saw a chance that this might happen, they continued and tried to provide the partner with the missing information – not only did they want an object, but a specific one. And when their partner ignored them they gave up (presumably not without making their displeasure known). Golinkoff (1983, 1986) concluded that the infants obviously had some idea about the effectiveness of their communication and could adjust their own behaviour as a reaction to that of their partner. What the study is lacking, is a systematic assessment of the infant's communicative intent, as measured by attention monitoring or gaze alteration, rather than a pure focus on the instrumental act of trying to obtain an object, and associated behaviours that need not be intentionally communicative. Grosse, Behne, Carpenter and Tomasello (2010) offer an interesting follow-up study that dissociates being understood by the communicative partner from simply obtaining a goal. Participating infants in their study continue to communicate with gestures and vocal behaviour even after they have obtained their goal if they have been misunderstood (Grosse et al 2010).

The partner plays a vital role in requestive gestures. Bruner (1983) proposes that, initially, the infant does not understand communication in how it affects the attention and actions of a listener, but simply observes a person consistently and willingly respond to her actions. The partner, and recipient, of the infant's gestures is therefore initially understood as a sort of tool – a means to reach a certain goal (Bates Camaioni and Volterra 1979). Only when the infant is well into her second year does she have a fuller understanding of the communicative partner and how his attention should be manipulated to make communication more effective.

The descriptions of requestive gestures in general are interesting for a number of reasons. Firstly, they include the context in which gestures are produced as well as their consequences. In fact, the success of a request can only be assessed by considering the infant's reaction to the recipient's reply. If the infant keeps signalling, her request is not yet fulfilled, if she stops signalling, her goal has been met. This is a marked contrast to how infant vocal behaviour is described – here the focus is on form of the utterance, not function. And, as a later section will show, also a contrast to how declarative gestures are investigated, where a lot of emphasis is put on underlying psychological states and motivations. Maybe not surprisingly, requestive gestures in human infants are thought to be similar to those of non-human primates.

Proto-declarative gestures

Proto-declarative gestures are argued to be a uniquely human expression of the desire to share some aspects of the environment and the signaller's attitude about this object or event (Tomasello 1999). These gestures include showing, giving and declarative pointing (Bates et al. 1979). The infant performs these gestures from the end of her first year, and their main goal seems to be to elicit a communicative response in the recipient, for example commenting on what has been given to him, orienting towards the target of pointing or voicing surprise (Liszkowski 2008).

Declarative pointing gestures have received by far the most attention in research on gestural development. The pointing gesture is thought to single out an object or aspect of the environment (Kita 2003). Pointing is argued to have many 'linguistic' attributes - it is very specific, referential, occurs in interactions with adults and the environment and, as the child develops, shows a growing awareness of the receiver's attentional state (Tomasello 2008, Tomasello et al. 2007, Liszkowski, Carpenter, Henning, Striano and Tomasello 2004). Recently, researchers go as far as claiming that pointing gestures show displacement, referring to objects that are not immediately present or to past or future events (Liszkowski, Schaefer, Carpenter and Tomasello, 2009). Index finger pointing has been reported in most human cultures and is possibly unique to humans (Tomasello et al. 2007, Kita 2003, Callaghan, Moll, Rakoczy, Warneken, Liszkowski, Behne, and Tomasello 2010).

A large number of studies investigate how the infant coordinates her gestures with a growing knowledge of various communicative demands. Results suggest that by 12 months of age the infant uses the pointing gesture for a number of things: imperatively to request things, declaratively to point out certain aspects of the environment, and informatively to help a receiver (Behne, Liszkowski, Carpenter and Tomasello, 2011, Liszkowski 2006). Furthermore, pointing has also been observed in private, non-communicative situations where the infant presumably points to direct her own attention (Delgado, Gómez and Sarriá 2011, Masur ,1980), and when the infant wants to gather information about something, for example the name or function of an object (Southgate, Senju and Csibra 2007).

In terms of communicative competence, by 12 months the infant is able to take into account for a recipient's visual attention and will point more when someone is looking at her (Liszkowski et al 2008). By 18 months of age, the infant is increasingly able to obtain a partner's visual attention prior to her pointing gesture (Liszkowski 2006), and alternates her gaze between the target of her gesture and the partner (Franco and Butterworth 1996). It has also been suggested that infants take into account past interactions when pointing for another person (Liszkowski et al 2004), for example they will point at an object that their partner has not seen before, but with which they are familiar. This has been suggested as further evidence that infants tailor their gestures to their partner's attention and knowledge state. These abilities only become evident from about 12 months of age, suggesting that between the onset of pointing and 18 months, the infant must undergo some developmental transition.

Interestingly, the pointing gesture seems to be culturally universal. Children from many cultures have been reported to point towards interesting objects and events using the same gesture (Liszkowski, Brown, Callaghan, Takada and De Vos in press, Callaghan et al. 2010). Research has not yet established whether pointing gestures serve the same functions of requesting, informing and declaring cross-culturally (Callaghan et al. 2010, Blake, Vitale, Osborne and, Olshansky 2007). Little is known about the development of culturally specific deictic gestures, for example lip-pointing or using other fingers for the gesture (see Kita 2003) and how well joint attention behaviour such as gaze alternation between object and partner and ensuring a partner's visual attention prior to gesturing are present in infants from other cultures.

The wealth of contexts in which pointing is observed shows that the gesture can serve a number of functions. It also poses the question that if the gesture is the same in all of these contexts, what further information is available to a listener to distinguish between them. For example, Gómez (2007) cites an informal observation in which a child points to the mailbox with letters in it. The pointing gesture offered multiple interpretations: it could indicate the child's desire to pick up the letters, inform the partner that there is mail, or request to open the mail box and empty out the letters. This example again makes it clear that although pointing is thought to be referential, it is so in a rather broad sense, and further specification is needed to understand the infant's goal.

When we look at descriptions of the infant's gestural repertoire (e.g., see the detailed example in Blake, McConnell, Horton and Johnson 1992), it becomes apparent that many gestures cover broad referential categories, for example comment, protest, requestive or emotive (Blake 2000). A number of specific movements are associated with any of these gestures, but can also be observed in multiple contexts, as our previously cited example of ambiguous pointing has shown. Therefore at least some gestures contain broad categories of information, and combined with additional cues, such as other gestures, vocalisations or environmental constraints, serve specific communicative functions (Blake et al. 1992, Blake and Dolgoy 1993).

There is an obvious contrast between how gestures and vocalisations are described. Whereas gestures are embedded in a communicative context and evaluated with regard to the functions they serve and what effect they have on the infant's environment, descriptions of vocal behaviour mainly focus on the form of the utterance. Similarly, infant gestures are often compared to those of other primates, whereas studies on vocal behaviour strongly emphasize the differences between humans and other primates.

Similarities and differences with gestures in non-human primates

In contrast to vocal behaviour, which is thought to be fixed and limited to a few, innate sounds, primates are thought to have a flexible repertoire of gestures (Call and Tomasello 2007, Tomasello and Zuberbühler 2002). Many studies have been conducted that directly compared the gestural behaviour to that of infants. For example Cartmill and Byrne (2007) replicated Golinkoff's failed request study with an orang-utan and found similar results. And indeed, descriptions of the gestural repertoire in both human infants (Blake et al. 1992) and apes (Hobaiter and Byrne 2011) are rather similar.

The conclusion is that primates are very good at performing goal-directed gestures – they take into account their audience and their past experience with the audience (Call and Tomasello 2007), they use their gestures flexibly and tailor them to the recipient (Cartmill and Byrne 2007), and show some awareness of the receiver's attentional states (Hare, Call, and Tomasello 2001), and any constraints they might have in fulfilling the signaller's request (Liebal et al 2007). Chimpanzees have been reported to have a common gestural repertoire as well as some idiosyncratic gestures that have only been observed to be used with only one particular partner (Goodall 1986, Hobaiter and Byrne 2011). They are sensitive to the recipient's visual attention, particularly in social gestures that are used to initiate play or grooming (Kaminski 2011). Great apes have even been shown to perform an open-hand pointing gesture to request food from humans (Leavens and Hopkins 1999).

Ape gestures have been observed to serve a number of functions, for example play, travelling, begging, or aggression (Hobaiter and Byrne 2011,Pika, Liebal, Call and Tomasello 2007). In a longitudinal observation of gesturing in wild chimpanzees, Hobaiter and Byrne (2011) resported that multiple gestures are associated with any functional category and they are often observed to be combined. For example, play can be initiated through an arm shake or through a play face, and both of these gestures can be observed in one sequence of play initiation directed at one individual.

What ape gestures seem to lack almost entirely, is the class of declarative gestures (Tomasello et al 2007). Apes do not seem to have anything in their repertoire analogous to the pointing gesture, which human infants use to direct another's attention, share their attitude about some event or object, or even provide information to a recipient (Tomasello et al 2007).

Summary – Gestures

Between the ages of 9 and 18 months human infants begins to use several types of gestures (imperative, declarative, informative, interrogative, etc.), and increasingly meets the demands of complex communication as shown by an increasing awareness of the receiver and his needs. Traditionally gestures have been divided into two types: Proto-imperative gestures are goal-directed requests that can be used flexibly; infants can clarify their requests and use a number of gestures to express them. Similar gestures are also observed in the great apes; in fact, descriptions of both gestural repertoire and skill are remarkably similar across the species. The human infant also performs proto-declarative gestures; here the goal seems to be to share attitudes, attention and information for the benefit of a communicative partner. Skills such as joint attention, attention monitoring and strategically combining signals to reach a goal have been observed in these gestures (Carpenter et al 1998). Proto-declarative gestures are not observed in other primate species. This suggests that the infant's gestural repertoire represents both phylogenetic continuity with other primates and uniquely human forms of communication. In contrast to vocal behaviour, gestures are studied and described with regard to the function they serve and the effect they have on the child's environment. Descriptions usually include the form the gestures take as well as what responses they provoke in listeners. Gestures are broadly referential and are made more specific through additional information provided through environmental cues or other communicative signals.

The interplay of vocal and gestural behaviour

Communication is rarely unimodal but often consists of signals from different domains, for example gestures and vocalisations. In this section we aim to review evidence for the co-ocurrence of vocal behaviour with gestures in prelingusitic infants. We are particularly interested in whether vocal behaviour can serve additional communicative functions that either complement gestures or perform different, independent functions to the gesture. Furthermore, we want to focus on whether there are any acoustic indicators, for example pitch, melody or intonation, that vary systematically between these functions, to gain further insights about the pragmatic functions of vocal behaviour.

From the start, infant communication and signalling involve body movements and vocalisations. In the period where speech is not yet available to the infant, this multi-modality can be important in two ways. Firstly, the combinations of vocal and gestural signals can expand the infant's repertoire (Papoušek and Papoušek 1989) and, secondly, it can help to refine signals that are otherwise very general, for example a scream that normally indicates excitement can be refined when it is combined with a pointing gesture directed at, for example, a dog. Papaeliou and Trevarthen (2006) suggest that gestures and vocalisations could convey different, but not contradictory messages, and thereby serve to make a signal more specific – just like in the above example of how a vocal signal can disambiguate a pointing gesture by effectively adding another function to it.

Studies on the infant's gestural repertoire often mention concurrent vocal behaviour, but rarely comment on what form this takes or if and what additional functions it could serve. The earliest examples are observations of physical resistance during cries, reaching out for food when the child is hungry, and calling and raised arms when the child is crying and wants to be picked up (Blake 1999, Masataka 2003, Papoušek 1992).

Most systematic research into the functions vocal signals fulfil in addition to gestures has been conducted on vocal behaviour that occurs in combination with pointing gestures. A number of sources suggest that in 12-month old infants the vast majority of pointing gestures are accompanied by vocalisations (87% - Blaket et al. 1992, 70% - Franco and Butterworth 1996). There are at least two possible ways in which vocal behaviour can contribute to pointing gestures. Firstly, it can serve to attract a recipient's visual attention by making them orient towards the source of the sound. For this to be effective, the infant needs to vocalise prior to, or during gesturing in order to ensure attention to the gesture – or possibly at the same time as performing the gesture. Secondly, acoustically different vocalisations can allow a distinction of different types of pointing, for example distinguish between declarative or imperative gestures.

Liszkowski et al. (2008) conducted a study that tested whether vocalisations serve to attract a recipient's attention to the gesture. They observed pointing gestures of 12-and-18-month olds whilst manipulating the recipient's attentional state, for example facing the infant or not facing the infant, and recorded the frequency and timing of vocal behaviour. Results suggest that although both age groups combine vocalisations and gestures, only the older group can use vocal behaviour strategically to direct an adult's attention to their pointing gesture. That is, they produce vocal behaviour prior to the gesture and only perform the gesture when they have their recipient's visual attention. Unfortunately, Liszkowski et al. (2008) only code the presence or absence and timing of the vocalisations. It would have been interesting to investigate whether there is a difference in the vocalisations at different times during the pointing gesture, for example between a vocalisation that acts to get the recipient's attention and one that expresses the child's attitude towards the target of the gesture.

The second question is whether the acoustic make-up of vocalisations that accompany the pointing gesture differs between different types of pointing, for example declarative or imperative pointing. To the best of our knowledge there is only one completed study that has investigated acoustic differences in vocalisations that occur with pointing. Leroy, Mathiot and Morgenstern (2009) followed two French-speaking children from 8 to 23 months of age, collecting data in monthly video sessions in the infant's home environment. They recorded any vocal behaviour that co-occurred with the infant's pointing gestures and analysed it with regard to the timing of the vocal behaviour and its acoustic features, in this case mainly rising or falling intonation. The authors found that whilst vocalisations during points were the rule from about 12 months onwards, there were no specific prosodic cues associated with either imperative or declarative pointing. Rising intonations were found in both cases and therefore a simple distinction between rising intonations for requests and falling intonations for assertions was rejected. Instead Leroy et al. (2009) propose a finer distinction. They observed that rising intonations were associated with episodes in which the infants tried to obtain the attention of a caregiver, regardless of whether they wanted to request something or point out something in the environment. A further distinction could be made in the data – assertions were associated with falling intonations, whereas requests almost exclusively contained rising intonations. Vocalisations with the pointing gesture therefore serve two functions: Firstly to obtain the recipient's attention, in a similar way to what has been reported by Liszkowski et al. (2008), and secondly to distinguish between declaring and requesting.

Leroy et al.'s (2009) study also provided an interesting addition to Liszkowski et al.'s (2008) study. Liszkowski et al. (2008) reported that the child alternated her gaze between the target of her gesture and the recipient. This observation of joint attentional behaviour during the pointing gesture has been frequently made during laboratory studies, to the extent that it has become a defining feature of the gesture (for some examples see Tomasello et al. 2007). Leroy et al. (2009), observing pointing under natural conditions, found that "instances of pointing gestures where the child gazed exclusively at the target of the pointing without any visual alternation were predominant over the whole period of filming". Therefore, the fine-tuning to a recipient the infant exhibits in laboratory setting are apparently not very frequent in a natural setting. This might either be because the child is already reassured of the caregiver's attention in activities they do together, or that the child cannot display their receiver awareness in a setting where multiple things happen at the same time.

Leroy et al.'s (2009) study makesimportant contributions in two ways. Firstly, it has demonstrated that systematic differences in vocalisations emitted during pointing gestures, could be an additional cue to narrow down the gesture's meaning. Secondly, findings, particularly with regard to gaze alternation, suggest that pointing gestures in the child's everyday environment and interactions might be different from findings obtained in laboratory settings. This highlights the importance of field studies in infant research. The obvious problem with their study is the small sample size of only two infants. Furthermore, the authors provide few descriptions of pointing episodes that go beyond a simple classification into declarative and imperative. Lastly, the authors simply determined whether an utterance had a rising or falling intonation. This ignores the possibility that other acoustic variables contribute to finer distinctions between different classes of pointing.

Summary: Interplay of vocal and gestural behaviour

Whilst the presence of vocal behaviour is often mentioned in gestural development studies, its function mostly remains guesswork. Pointing studies have mentioned vocal behaviour as a potential means to obtain a recipient's attention (Liszkowski et al. 2006), or as an indicator of communicative intent (Blake et al. 1992). It is however conceivable that vocal behaviour carries more fine-grained information in its acoustic make-up. Studies like Leroy et al.'s (2009) suggest that meaning can lie in the acoustic properties of non-linguistic utterances. We will now turn to review studies that have explicitly looked at non-linguistic vocal behaviour in this way.

Sound-meaning correspondences in prelinguistic vocal behaviour

We now want to review studies that investigate whether vocalisations can serve a similar function to gestures or calls of other primates, that is, whether they can transmit meaning in a broadly referential way and thereby function communicatively in that recipients can gain information about the infant's emotional state and activities from them. We are particularly interested in whether changes in acoustic variables, such as intonation, melodic contours or pitch, are associated with specific emotions or communicative functions.

Systematic relationships between sounds and meaning have been documented in a number of non-human primate species. Some primate calls have been found to be functionally referential, that is, calls with consistent acoustic properties are observed in similar situations, and it would be interesting to investigate whether human infants produce similar vocal behaviour, despite having a flexible repertoire. In the gestural domain we have already observed categories that are shared with primates and those that differ from the start. We now want to ask whether the same is true for at least some forms of vocal behaviour.

Crying

The first vocalisation the infant produces is a cry. Crying is the beginning of human vocal communication and the initial mode through which the infant's needs are expressed (Lester and Boukydis 1989). In the neonate crying is a reactive, automatic response to unpleasant states such as hunger, pain or discomfort (Lock 1993). Cries emerge in the absence of auditory experience and are presumably produced without communicative intent (Lester and Boukydis 1989). Despite being an involuntary behaviour of an immature vocal tract, crying still transmits information to listeners (Papoušek 1992).

From the above description, it is relatively easy to make the link from crying to the vocal behaviour of other primates. Their vocal behaviours are thought to fall mainly into the same categories – they are innate, emerge in the absence of auditory experience, are motivational and provoke systematic responses in the listener (Oller and Griebel 2008, Snowdon 2008). Judging from this description, crying seems to be very animal-like and displays few of the characteristics of human language. Motivational animal signals show strong links between external events and the acoustic characters of the vocalisations. Similar observations have been made in cries produced by infants less than three months of age (Masataka 1999).

Lynip (1951) was one of the first to differentiate newborn's cries based on visual spectrographic analysis and identified different acoustic characteristics of cries recorded under different circumstances. He identified two types of cries, one for pain and one for hunger. Papoušek (1989, 1992) and Scheiner et al. (2002) confirmed these results in a more extensive study of infant vocal behaviour. Both studies looked at the entire vocal repertoire of infants from birth to the third month of life and used an analysis that measured different aspects of the cry for example melody contours, pitch, intensity and fundamental frequency. In addition to pain and hunger cries, Papoušek's study also found "acoustic correlates of state information" for comfort, discomfort, joy and neutral emotional states. Therefore, in her first three months the infant seems to have different vocal expressions that correspond to different positive and negative emotional states.

Testing a similar hypothesis, Scheiner et al. (2002) asked parents to record and describe the vocal behaviour of their infants in 12 distinct, environmentally defined categories ranging from hunger, to dirty nappies, to comfortable after a meal. The team then used quantitative acoustic analyses, a method adapted from Hammerschmidt's studies on primate vocal behaviour that measured a large number of acoustic variables in the vocal behaviours (further discussed in Schrader and Hammerschmidt 1997). Results from the study suggest that infant vocalisations can roughly be divided into positive and negative emotional states, which are associated with different acoustic patterns.

That different acoustic patterns observed in cries are meaningful to a receiver was confirmed by Papoušek (1989) in a playback study that presented different types of cries to participants and asked them what emotional state they reflected. They tested six different subject groups, ranging from parents to inexperienced adults to 8-year-old children. Participants listened to 50 infant sounds, divided into comfort, discomfort, cry and joy sounds. All participant groups could distinguish between positive and negative sounds, but how well they did varied as a function of age and experience (also confirmed by Lester, Garcia-Coll and Valcarel 1989).

As mentioned in the description of vocal development, research on the acoustic properties of crying all but stops in infants older than 3 months despite studies reporting that crying makes up a large part of the infant's repertoire at least until the end of her second year. The question is now whether similar sound-meaning correspondences exist in older infants, who display goal-directed, intentional communication in their gestures, and whether these acoustic consistencies that correspond to certain emotional states or functions are present in vocal behaviour other than crying.

In sum, cries seem to be an inflexible, stereotyped vocal behaviour that shares many properties with animal signals and is consequently distinctly unlike language. Acoustic analyses of cries emitted in different situations revealed different acoustic patterns in each production context. These patterns could also be identified by listeners, suggesting that the acoustic differences correspond to changes in the perception of the meaning of the cry. The finding that crying has this presence in the infant's vocal repertoire is interesting as it suggests that primitive, animal-like vocal behaviours are still present as the infant uses her flexible vocal repertoire. The question arises whether crying is the only example of such behaviours or whether there are more vocal behaviours whose acoustic properties systematically change with the function they serve or the meaning they transmit.

Sound-meaning correspondences other than crying

From about six months of age the infant is able to manipulate certain acoustic variables in their vocalisations, for example pitch, intonation or timing. The differences in these acoustic parameters in vocal behaviour could solely be related to practising the sounds of speech and enlarging the infant's repertoire. But, as the aforementioned example of crying suggests, some variations in the acoustics of the call could also transmit specific kinds of information. The existence of such categories could potentially have an adaptive value in that it informs caregivers of the infant's emotional states or needs when she is still unable to speak but not as dependent as a neonate.

Halliday (1975) observed and transcribed the vocal development of his son, Nigel, from 9 to 18 months of age. At the initial stage Halliday proposed a number of functions for Nigel's utterances that covered the child's physical and material needs, the interactional frame with a caregiver, his awareness of the environment and personal utterances that are unique to one child. Halliday proposed that there are early 'content-expression pairs' – this means that certain vocal behaviour reliably indicates a certain wider ‘content’, for example ‘I want that’. Unfortunately Halliday does not elaborate on what these expressions sounded like.

Instead he focuses on ‘protowords’: utterances that emerge prior to language and refer to a specific entity and are consistently observed in the child’s vocal behaviour. His prominent example is that of Nigel’s toy bird, consistently referred to as “/bø/”. In his observations Halliday wanted to show that protowords are present before the onset of conventional words and that the infant understands the relationship with objects and how sounds refer to them. He does, however, not offer a way of showing that protowords are not just simply inaccurate pronounciations of conventional words.

Halliday (1975) did not explicitly investigate the role of acoustic features in Nigel’s utterances. His account can possibly best be seen as anecdotal evidence that is shared by many parents, but confirms the suspicion that prelinguistic vocalisations fulfil communicative functions and carry meaningful distinctions in their acoustic make-up. It is important to note the (not always obvious) disctinction Halliday makes in the vocal behaviour he describes – there are broadly functional vocalisations akin to gestures and behaviours that form a sort of ‘proto-language’ and display many linguistic characteristics.

Indication that distinct sound patterns occur in similar situations to gestural behaviour has initially been reported by Lewis (1936) and Leopold (1939), who both observed the vocal behaviour of their own children and found that there were similar sounding vocal behaviours that accompany anger, pain or requests. Dore, Franklin, Miller and Ramer (1976) later elaborated on this finding by observing whether the infant produces consistent vocalisations in specific communicative situations.

Dore et al. (1976) observed 3 children over a span of 8 months, starting when the infants were 11 months old. Video data were collected every month in play sessions at the children’s home. The video data identified five functional categories in which vocalisation occurred. The first were expressions of affect, which included joy, anger, satisfaction and protest. These were strongly correlated with facial expressions (Ekman and Friesen 1979), posture and body movement, for example kicking and clenched fists when a child is angry. Secondly there were instrumental expressions, which were basically requests directed at an adult. Thirdly indicating expressions, the most prominent being vocalisations that occur during the pointing gesture. Lastly, Dore et al. (1976) proposed grouping expressions, utterance forms that notated the same emotional response for a group of objects, for example excitement upon seeing animals.

Dore et al. (1976) termed these expressions phonetically consistent forms (PCFs) and defined them as readily isolable units that occurred repeatedly in a child's repertoire and could be correlated with specific, recurring conditions such as the experience of different emotions, requests or indicative expressions. As Dore et al. (1976) did not provide any further analysis or transcript of the vocal behaviour, it is reasonable to assume that the classification of the sounds was made on the subjective impressions of the authors which might have missed or indeed over-emphasized regularities in the vocalisations. Despite this, Dore et al. (1976) concluded that "different children use different forms for different functions", which suggests that there is consistency in the form and function of vocal behaviour within but not between the three individuals. Because of a lack of supportive evidence, we consider this conclusion as being mainly speculative in nature.

In a later publication, Dore (1983) elaborated more on the properties of PCFs and now called them "indexical expression"- they are indicators of the infant's affective state or attitude towards something. The main advancement was that Dore proposed that PCFs basically consist of phoneme-like sounds with affective qualities, and that these affective qualities changed according to context. More detailed description of the sound patterns of individual children is provided in the article; however Dore does not address the question whether these patterns are similar across infants.

What Dore's study lacked in quantitative analysis is provided in a study by D'Odorico and Franco (1991), albeit investigating a much smaller part of the infant repertoire. They wanted to investigate the hypothesis that "vocalisations are selectively uttered in relationship to their production context". In order to test this, they recorded the vocalisations of 5 infants from 4 to 11 months of age in an adult-infant-toy interaction. Their aim was to investigate how vocal behaviour varied in different contexts and whether there were similarities across infants.

The infants' vocal behaviour was recorded during play sessions with the mother in four different contexts: when the infant was manipulating an interesting toy, when she looked at an adult, when the adult showed the infant something with the toy and finally when no one did anything to the toy. They then measured a number of acoustic parameters of these vocalisations and compared these using a Discriminant Function Analysis. Results suggest that under the age of 9 months, children produced acoustically distinct vocalisations in each of the contexts. Intonation contours were found to be the greatest contributor to the discrimination between vocalisation types. However, this 'sound-meaning connection' seemed to disappear once the infants are older than nine months, where an increased variation in the vocal behaviour was observed. Furthermore, this system seemed to be increasingly idiosyncratic, and thereby confirms Dore et al.'s (1976) hypothesis.

Whilst we applaud the methods used in D'Odorico and Franco's study as well as their hypothesis, we think that a major point of criticism concerns the contexts they compared. Firstly, toy interaction is only a very limited part of the infant's behaviour. Secondly, looking at different stages of an interaction with an adult and a toy does not necessarily lead to different functional or communicative categories. For example, the infant could make request sounds when the adult was manipulating the toy as well as when no one manipulated the toy. It might be necessary to derive functional categories on the basis of different classification of the infant's actions rather than just on the timing of an action. More specifically, a distinction could be based on the different acts involved in the play sessions, for example requesting, giving or pointing. In that way, an acoustic analysis would compare vocal behaviour that might actually serve different functions and therefore more directly address their hypothesis. Ideally, this would be supplemented with an independent measure of the intentionality of the vocalisations – if this could be achieved, vocalisations could truly be analysed with regard to their function and intended goal, rather than having to rely on environmental descriptions. Although environmental descriptions can provide some measure of intent as indicated by reactions of communicative partners or caregivers.

Summary: Sound-meaning correspondence is prelingusitic vocal behaviour

In sum, the studies we have just reviewed suggest that there is some link between the acoustic properties of non-linguistic infant sounds and the contexts in which they are produced or the function they serve. The evidence is fragmentary at best, and relies heavily on anecdotal evidence. The study by D'Odorico and Franco (1991) proved a notable exception but also had its flaws. What is notable is that the contexts suggested for sound-meaning correspondences in vocal behaviour are very similar to those in which gestures have been documented (compare Dore et al. 1976, and Blake et al. 1992). Vocal behaviours seem to contain similarly broad meanings than gestures, whilst not necessarily being related to language acquisition or linguistic concepts. In the next section we will review research that suggests that adult-like patterns of prosody are already present in prelingusitic vocal behaviour that would give further evidence of the connection between certain aspect of a sound and the function it serves.

Prelinguistic 'Prosody'

In this section we will review studies that document links between certain aspects of infant vocal behaviours, such as melody or pitch, and the contexts in which they are observed or the functions they serve. The studies we present here use the term 'prosody' for these variations. Prosody is defined as the rhythmic and intonational aspects of language that carries additional information in speech, for example about the speaker's emotional state or makes different functional categories such as questions or statements (Bernhardt and Stemberger 1998, Papaeliou and Trevarthen 2006).

From about six months onwards, the infant starts to manipulate acoustic variables, such as pitch or intonation, in her vocal play (Oller 2000, Vihman 1996). These changes in the acoustic structures of an utterance could not only be the by-product of the infant's playful sound manipulation, but could serve different communicative functions. Just like in adult speech, systematic differences in these acoustic features could form meaningful distinctions, for example between requests and indicative sounds (Papaeliou and Trevarthen 2006).

Harding (1982) recorded the vocal behaviour of children between the age of 6 and 11 months. She found that from 8 months, children would use different intonation patterns for indicative and declarative utterances in a play session with mother and a toy. Furthermore, Harding noted that children would use the same sounds in the different contexts, e.g. "/da/" but would modify the intonation contour. Indicative utterances were marked by a falling and declaratives by a rising intonation contour. This pattern apparently also persists after the child learns her first words (Galligan 1987) and is still observed when she has a vocabulary of 50 words and produces two-word combinations (Flax et al. 1991 – but see Leroy et al 2009 for contrary evidence).

Papaeliou, Minadakis and Cavouras (2002) investigated the differences in acoustic patterns between vocal behaviour that was either communicative or emotional. According to the author's definition, communicative behaviour requires a reply from a partner – for example requests, enquiries, or vocal exchanges with a turn-taking pattern. Emotional expressions are the reactive broadcasting of an emotional state, for example screaming in pain. The team got mothers to classify their children's utterances as either communicative or emotional. They then conducted a computer analysis of a number of acoustic variables to offer a quantitative description of the recorded sounds. Results suggest that emotional and communicative vocalisations have differing acoustic properties and that these are sufficient to classify vocalisations into the correct context.

Critical Comments

Prosody usually denotes "rhythmic and intonational aspects of language" (Oxford English dictionary), this includes both the rhythm of speech itself and any additional meaning that is conveyed through variables such as intonation, melody contours or pitch. According to this definition, it makes little sense to talk about prosody as a concept separate from language. Acoustic variation, for example in intonation patterns, inflection, intensity or melodic contours, is present in infants before speech begins. But they do not function in the same way as prosody does in spoken language. Instead of providing additional meaning these acoustic variables are the main, or only, source of 'meaning' in prelingusitic infants (Marler, Evans and Hauser 1989).

The question is, whether researchers simply use the term prosody as a grouping expression for the acoustic variables found to play a role in changing the structure of nonlinguistic vocal behaviour, for example systematic changes in pitch, melodic patterns or length of utterances that are associated with vocal behaviour produced in certain cintexts, or whether they imply that the infant uses linguistic classes of vocal behaviour prior to speech, for example a rising intonation to indicate a question or falling intonation to mark a statement. The studies reviewed here provide evidence for the former – regularities in the acoustic make-up of prelinguistic vocalisations vary with the contexts in which they are produced or the function they serve. The variables that show these variations are the same that are observed in linguistic prosody – melody, pitch and intensity, but obviously occur in the absence of speech.

We think that this is yet another case in which a prelinguistic vocal behaviour is seen in light of later developments – prelinguistic prosody is seen as precursor to prosody in speech, although speech is not yet present and consequently prosody cannot function in the same way. In prelinguistic infants prosody is the only way in which acoustic variance can transmit meaning, as conventional signals that encode semantic messages are still unavailable. Acoustic variation in prelinguistic infant behaviour is therefore by definition more like animal signals, where systematic variation in acoustic parameters alone serves broad functions.

Summary: Prelinguistic prosody

Studies on prelinguistic 'prosody' provide further evidence that certain acoustic parameters in infant vocalisations vary with their production context in a systematic way across a number of infants. Although infant vocal behaviour contains the same prosodic features as speech, the term might be misapplied as variables such as intonation, melodic contours or intensity are not an additional source of 'meaning' but the only one.

Conclusion– Form and Function of Prelinguistic Vocal Behaviour

Evidence for functional categories of vocal behaviour

We reviewed a number of studies that suggest that the acoustic make-up of some classes of vocal behaviour observed in particular functional contexts varied between these contexts, for example declaratives, requests, or comments. The acoustic markers at the base of this variation are possibly the same as those that carry meaning in animal signals. The instances in which infants have been observed to produce these vocal behaviours are similar to those in which gestures have been observed to function communicatively. However, the evidence for this, as reviewed here, is rather fragmented and tentative, with a strong tendency to explain patterns found in prelingusitic behaviour in relation to later emerging linguistic concepts. There are, to date, few studies that explicitly test the hypothesis that infants produce acoustically similar vocalisations in the same situations that presumably serve some communicative function.

The studies that report acoustic similarities in vocalisations produced by infants in the same categories do not provide sufficient evidence that these vocal behaviours can serve communicative functions. In the next section we will review studies that investigated how caregivers perceive and react to prelingusitic vocal behaviour. Whilst descriptions of gestures generally include descriptions of how recipients react to the infant's signals, this is only seldom the case in studies on prelinguistic vocal behaviour. This lack of description again highlights that prelingusitic vocal behaviour are not well understood in terms of their communicative function.

The other end: Receivers

Parents and infants are an unusual match when it comes to communication. Initially the parent accomplishes the lion's share of any exchange. Many authors report that parents treat their infant's productions as intentional and as a genuine contribution to a dialogue. This signals the parent's readiness to try to 'understand' their infant's signals (Papoušek 2007, Bruner 1983).

Bruner (1983) advocated that the parent's willingness to understand their child and attribute intention to any gestures and utterances the child produces are the engine for the development of intentional communication in the infant. By observing that gestures and utterances have an effect on people around her, the infant begins to understand these signals as tools to change her environment (Bruner 1983, 1975).

A large body of research has concentrated on how parental 'scaffolding' can aid the child in language acquisition (Papoušek 1992, Papoušek 2007). Infant directed speech, in particular, has been the focus of research and much effort has gone into offering analyses of its both form and content (for example Fernald and Kuhl 1987, Fernald 1989, Masataka 2003). Although this is an important part of how infants and parents interact, here we want to forego an in-depth analysis of the interactional frame between caregiver and infant, and instead concentrate specifically on the information content of prelinguistic vocal behaviour. Detailed accounts of infant directed speech and parental scaffolding can be found elsewhere (Papoušek and Bornstein 1989, Papoušek 1992, Kuhl 2004 for a review).

When parents assign meaning to their infant's vocal or gestural productions, what clues do they use as the basis of their judgement of the infant's state? In terms of vocal behaviour, one possibility is that systematic acoustic variation in the infant's utterances, of the kind presented in the previous sections, can provide the listener with information about the infant's emotional state and activities.

It is conceivable that vocalisations function communicatively in a similar way to gestures – they fall into broad categories, such as requests or comments whose meaning is further specified through additional signals or contextual cues, as well as a constrained frame of reference (Sperber and Wilson 1995, Bruner 1983). Listeners can then exploit all these sources of information in order to infer the meaning of the infant'sutterance. Until now, there are no studies on the acoustic properties of vocalisations in these contexts, but some studies investigated how listeners react to them, which can give some idea about their function.

One way of testing whether listeners can gain information from vocal behaviour is to use playback designs. This method is best known from studies on non-human primate vocal behaviour (for example Cheney and Seyfarth 1980, Zuberbühler , Noe and Seyfarth 1997, Slocombe and Zuberbühler 2006). Studies usually involve playing previously recorded sounds back to listeners of the same species and recording their reactions. For example, the alarm calls of Diana monkeys were recorded when either aerial or ground predators were present. An acoustic analysis revealed differences between the calls emitted in each condition. To investigate whether these differences are meaningful to a listener, calls are played back to the monkeys and their reaction is recorded. In the case of Diana monkeys, playing back alarm calls emitted to different predators led to different flight responses in recipients (Zuberbuhler et al 1997).

There is, to the best of our knowledge, only one study that combines both a quantitative analysis of vocal behaviour and a playback study that tested whether these differences make sense to listeners. We already described Papoušek's (1989, 1992) study that found acoustic differences in the vocalisations of 3-month-old infants in the contexts of pain, discomfort, hunger or joy. The same researchers then conducted a series of studies that used a playback paradigm to demonstrate that parents can differentiate sounds produced with different motivations based on the acoustic information alone (Papoušek 1989). Participants were asked to listen to 50 infant vocalisations and classify them as either sounds of comfort, discomfort, joy or cries. It was found that crying, comfort and discomfort vocalisations could be differentiated reliably, whereas vocalisations recorded as an expression of joy seemed to transmit ambiguous information. Participants ability to correctly classify vocal behaviour varied significantly as a function of age and previous experience with infants.

A further experiment tested whether infant cries contained information about the child's identity. Interestingly, Papoušek (1989) found that only about half of the participating mothers could correctly identify vocal behaviour from their own child as opposed to that of other children. All vocal behaviour was recorded from infants less than three months of age. It seems that at this age, there is little or no information about the signaller's identity contained in the vocal behaviour. What is interesting is that, in contrast to identity, parents could reliably discern information about the production context from cries. This suggest that there are commonalities in infant vocal behaviour and that, at least in the first few months of life, there might be vocalisations that are not very flexible. The question of whether similar vocalisations are also present in the repertoire of older infants remains unanswered.

Goldstein and West (1999) conducted the only study that tested how accurate parents are in drawing information from vocal behaviour of infants in their second year of life. They collected video episodes of three infants aged 9, 11.5, and 19 months in play sessions with their parents. They recorded the infant in four different behavioural categories: Giving/Showing, Naming, Showing Concern and No Response. Participants were shown a video of an infant playing and were then presented with a vocalisation in any of the four aforementioned contexts. Despite judging playback episodes from unfamiliar infants, participants performed well above chance in assigning the correct category of vocal behaviour to the episodes of vocal behaviour. Because the category of vocal behaviour did not always match the video episodes parents saw, the researchers measured what source of information, video or audio, corresponded to the parent's choice. Results suggest that parents mainly used the audio clips to make judgements about the infant's behaviour.

The study has some limitations – firstly it did not explicitly assess the information content of vocal behaviour. Vocalisations were always presented with a video episode and were not considered separately to assess the relative contribution of either source to the parent's choices. Nevertheless, participants mainly relied on the audio information when making judgements about the infants' activities. Secondly, audio samples came from only three infants; therefore we do not know whether the vocalisations are similar across infants or unique to individuals. In sum, the study provides tentative evidence that listeners can use infant vocalisations to gain information about their activities but does not investigate whether acoustic variation in the vocal behaviour is at the source of this.

Summary - Receivers

It is well-known that parents attribute meaning to their infant's pre-linguistic utterances. Whilst there is a lot of research documenting how parental interaction aids the child to learn the value of communicative symbols, less is known about what sources of information parents use to make accurate judgements about the infant's communicative goal. Studies using playback paradigms to assess listeners' reactions to vocal behaviour are common in studies on animal communication, but rare in human infants. We presented two studies that used this methodology and showed that listeners can use the information contained in prelinguistic vocal behaviour to make inferences about the infant's emotional state or activities. This provides further evidence that infants possess vocal behaviours that show acoustic regularities that are meaningful to listeners.

Overview: The functions of non-linguistic vocal behaviour

In the first 18 months of life the infant moves from producing inarticulated, reflexive cries over vocal play that reflects her native language and is voluntarily controlled to 'proto-words' that use unconventional but constant vocal constructs to refer to the same objects in the infant's environment. Infants do, however, produce sounds other than cries and vocal play. The question is if crying is on one end of the spectrum and protowords are on the other, where do the other vocalisations fall and what exactly are they?

The studies we reviewed here suggest that infants produce at least some sounds that might have relatively constant acoustic properties that vary with the situations in which they are produced. Those situations might be rather similar to the ones in which prelinguistic gestures have been observed – for example requests, comments or declaratives. A number of studies report that vocal behaviour consistently accompanies gestures in such situations, and could possibly be an important source of information for a listener trying to understand the infant's goals or needs in everyday situations. Studies explicitly investigating whether there is a relationship between sounds produced in certain situations and their acoustic variation, although rather fragmented and tentative, suggest that there might be a relationship between sound and meaning for some infant vocalisations. Furthermore, it has been suggested that infants use 'prosody' in their vocal behaviour as a marker of affective state or function – this suggests that, like the communicative signals of other primates, infant vocal behaviour shows systematic correspondences between variables such as pitch, melodic contour and inflection, and the meaning or function of the vocal behaviour. The hypothesis that prelinguistic sounds transmit information about the infant's emotional state or activities gains some support from a couple of playback studies, where listeners showed consistent responses to the vocal behaviour of infants and were able to infer information about the production context from the vocal behaviour.

All these studies suggest that there might be sound-meaning correspondences in categories of non-linguistic infant vocal behaviour that has been largely neglected by studies of infant communication, for example screams or grunts. These might not necessarily be related to language acquisition or the practice of language sounds but might instead serve the pragmatic functions of informing caregivers about the infant's well-being, her mood, attitude towards things and maybe even activities and thereby contribute to the modulation of meaning in the whole communicative system of infants together with gestures and linguistically related vocalisations.

In this thesis we will explore the structure and functions of these non-linguistic vocalisations produced by human infants. Our specific hypothesis is that the sound-meaning correspondences of those vocalisations might be more like signals in animal communication, where variations in the acoustic property of a vocal behaviour are associated with different functions and contexts. We propose that some forms of early human vocal behaviour resemble those of our closest living ancestors and reflect the evolutionary trajectory of communicative signals. In accounts of infant vocal development, there is a constant emphasis on the unique flexibility of human infant vocal productions and their relation to speech and language acquisition. This does not exclude the existence in humans of phylogenetically older, less flexible vocal signals that have an adaptive value in informing caregivers of the infant's needs and goals.

The questions of whether there are sound-meaning correspondences in infant non-linguistic vocalisations, and whether these are related to specific production contexts and serve specific functions cannot be conclusively answered from the existing literature. As this is the aim of this thesis, we now want to introduce our methodological approach.

Studying human infants from a primate communication perspective

We would like to propose a methodological approach that is borrowed from the study of animal communication, and more specifically the study of the vocal communication of non-human primates. In this section we want to offer a brief overview of vocal communication in non-human primates, and ask whether this is in any way similar to non-linguistic vocal behaviour in human infants. We want to introduce a comparative methodological approach that can investigate whether human infants have similar sound-function relationships as other primates. Throughout, we want to offer some thoughts on how this methodology can possibly add to our understanding of prelingusitic vocal behaviour.

Vocal Communication in Non-Human Primates

We firstly want to offer a brief overview of vocal communication in non-human primates. Extensive reviews and descriptions of form, function and underlying cognitive capacities can be found elsewhere (Cheney and Seyfarth 1996, Tomasello and Call 1997, Zuberbühler 2006, Tomasello and Zuberbühler 2002, Rendall, Owren and Ryan 2009, Fischer forthcoming).

Primates produce vocalisations to avoid predators, travel together, discover food and maintain social relations (Tomasello and Zuberbühler 2002). In general, the vocal signals of primates show little flexibility: they are part of a fixed repertoire, emerge in absence of auditory experience, individuals of one species use the same repertoire and there is no evidence of new signals being invented (Fitch 2000). Vocal behaviour does display some signs of functional learning and flexibility, however, examples of learning include the correct usage of calls, for example performing an alarm call only to a potentially dangerous aerial and not to every flying object (Seyfarth and Cheney 1986). Flexibility is observed primarily when vocal behaviour is adjusted to the composition of an audience, for example during food calls in Capuchin monkeys (Pollick, Gouzoules and de Waal 2005) or travel grunts in chimpanzees (Mitani and Nishida 1993).

Whilst primate vocal production shows little flexibility, receivers are still able to extract a lot of information from it (Tomasello and Zuberbühler 2002, Cheney and Seyfarth 1996). Individuals perform specific flight responses upon hearing different alarm calls even when they have not themselves seen the predator (Seyfarth, Cheney and Marler 1980, Zuberbühler et al. 1997) or can discern dominance hierarchies and social encounters from hearing others vocalise (Kitchen, Cheney and Seyfarth 2005, Gouzoules and Gouzoules 1999). Listeners can therefore extract a wealth of information from hearing a conspecific vocalise, for example the caller’s identity, environmental events associated with the call and the activity of others – all in absence of actually seeing what event caused the vocalisation.

Vocal behaviours serve a variety of functions. Many monkey species have distinct alarm calls that often show systematic acoustic variations for different predators (e.g. Seyfarth, Cheney and Marler 1980, Zuberbühler et al. 1997, Zuberbühler 2000), and most primate species have signals that facilitate social behaviours such as dominance or aggression (Cheney and Seyfarth 1995, Slocombe and Zuberbühler 2005, Clay and Zuberbühler 2009). Vocal behaviour is also observed in the context of feeding where food calls can either inform others of the presence of food or its quality (Slocombe and Zuberbühler 2005, Pollick et al. 2005).

Although some studies suggest that ape vocalisations show an adjustment to the audience (Townsend et al 2008, Clay et al 2011), there is little evidence for intentional production on behalf of the signaller; instead vocal behaviour is thought to emerge as a reaction to an immediate event or as a reflection of the signaller's emotion (Cheney and Seyfarth 1996).

A basic distinction can be made between monkey calls that fall into discrete acoustic categories with a fixed meaning, and those of the great apes, whose repertoire consists of graded signals. Whereas monkey calls can be seen as inflexible units with a specific meaning or function, ape vocalisations are graded and therefore fall on a spectrum between distinct classes of vocalisations, for example grunts, barks or hoots (Snowdon 2008, Slocombe and Zuberbühler 2010).

The level of reference contained in calls of non-human primates is still debated. Whilst most researchers agree that primate calls do not exhibit the same level of reference as words, the exact information that is contained in a call is subject to speculation. Some researchers propose that primate calls are almost entirely affect-driven (Owren and Rendall 2001) and contain information to external events only by associations made by the listeners. Others suggest that primate calls are based on affect as well as containing information about some external event, for example the presence of a predator and how dangerous the situation is for the monkey (Cheney and Seyfarth 1996, Fischer, forthcoming).

In the graded vocalisations of apes, referentiality has been treated slightly differently. Ape species do not show anything like the alarm calls of monkeys that have a very strong sound-meaning relationship. Instead vocal behaviour of certain kinds is associated with certain contexts, for example aggression, reconciliation or travel, but is generally more flexible than that of monkeys. For example chimpanzee hoots can be part of a contact call but are also observed in the context of travel or aggression (Slocombe and Zuberbühler 2010). Because of the graded nature of the system, there is more diversity in vocal behaviour emitted by apes in the same contexts. In all cases, receivers are very apt at gaining information from the call, even when they do not have visual access to the situation in which the call has been produced (for a fuller review see Cheney and Seyfarth 1996) and make use of additional contextual cues in the environment to supplement the information contained in a vocalisation (Leger 1993).

To summarise, non-human primates produce a variety of either stereotyped or graded vocalisations in a number of contexts, for example social relations, food or predation. These calls contain information that allows receivers to draw inferences about the signaller's affective state and corresponding events in the environment. Monkeys mainly produce distinct types of vocal behaviour whereas the repertoire of apes seems to be graded. Whilst calls might be largely reactionary and display little intention to influence the listener or change their behaviour, listeners can gain a wealth of information from them and adjust their behaviour accordingly.

Are there possible parallels in infant vocal behaviour?

Primate vocal behaviour shows a strong relationship between the acoustic structure of a sound and the function it serves or the information it transmits. Whilst form and function relationships seem to be very fixed in monkeys, apes show more flexible sound-meaning correspondences in their graded signals.

Infant vocal behaviour has been traditionally seen as very different. Throughout our review we highlighted the flexibility in usage, the ability to produce sounds without any environmental trigger, and the variety found in the vocal behaviour of human infants. But we also found evidence that not all infant vocalisations display these features. A large number of studies suggest that human infants, similar to other primates, show correspondences between the sounds they produce and the situations in which the sounds are emitted, but this needs to be supported by solid, systematic evidence. Previous research has mainly concentrated on proto-words, sound-meaning correspondences that are similar to words but not yet using a conventional form (Halliday 1975, Dore et al. 1976, 1983). Furthermore, listeners seem to be able to gain information through these vocal behaviours much in the sense that other primates do.

Throughout the review we presented studies that compared some aspects of prelinguistic communication to those of later emerging language. Whilst it would be expected to find linguistic precursors in the infant's vocal production, it is also possible that there are parts of prelinguistic communication that are not immediately related to later language development. It is therefore necessary to investigate this behaviour with a more balanced view that avoids overly hasty, and sometimes inappropriate comparisons to linguistic concepts but does not revert to simple stimulus-response descriptions (Gómez 2007).

We think that research investigating the vocal behaviour of non-human primates offers exactly this balanced approach. A communication system is studied in its own right and in the species' environment to gain more information about form and function. Whilst the investigation of relationships between form and function are in the foreground, a comparison to linguistic concepts is secondary. Consequently, we want to apply this approach, albeit with some modifications, to study nonlinguistic infant vocalisations.

Methodological Approach

We now want to offer a sketch of how primate vocalisations are studied (for a 'manual' see Wittig and Zuberbühler 2011). For this, we want to use the example of vervet monkeys, a species well known for its alarm calls. Struhsaker (1967) was among the first to report on the vocal repertoire of the vervet monkey. His aim was to "present and describe… the audible behaviour of the vervet monkey, (…) catalogue … sounds" and provide "information on the … conditions evoking the sound and the communicative function".

Struhsaker (1967) recorded the vervet monkey's sounds in the field and gathered detailed data on the contexts in which vocal behaviour was produced. Descriptions not only included information on the sounds the monkeys produced, but also on how sounds affected conspecifics, the timing of vocalisations and any environmental events correlated with sound production, e.g., presence of a predator. From this information Struhsaker extrapolated the presumed function of the monkeys' vocalisation. The most famous example is the alarm calls; he observed that monkeys produced acoustically different calls for different predators, e.g. eagles or leopards, and that conspecifics react differently upon hearing each type of alarm call.

More specific investigations followed this initial finding; most of them lead by Cheney and Seyfarth. Using quantitative acoustic analyses, they found that alarm calls emitted in the presence of different predators show different acoustic structures and could therefore serve as referential signals to a receiver (Seyfarth, Cheney and Marler 1980). They tested this with systematic playback studies in which different alarm call were played to conspecifics and their reaction was observed. Conspecifics reacted differently to each type of alarm call, they ran up trees for an alarm call recorded in the presence of ground predators and hid in shrubs or remained still for alarm calls recorded in the presence of aerial predators (Cheney and Seyfarth 1980).

All these results together suggest that vervet monkeys produce acoustically distinct alarm calls that reliably produce the same responses in listeners and these responses differ with the context in which the alarm call was given. These alarm calls have been termed 'functionally referential'; they might be the result of affect and are fixed in the monkeys' repertoire but show a strong correlation with specific events in the environment.

This methodological approach, as it is commonly applied to the study of primate vocal behaviour, consists of three elements: The observation of situations in which sounds are produced reliably by a number of individuals, the acoustic analysis of these sounds to investigate whether there are communalities across individuals, and systematic differences between sounds produced in different situations, and lastly systematic playback studies that investigate whether possible variations found in the sounds lead to different responses in the receiver.

Research Questions

We employed the previously illustrated methodological approach to investigate whether prelinguistic human infants show consistent relations between sound properties of non-linguistic vocalisations, the contexts in which they occur, and the functions they serve. Previous researchers explicitly looked for precursors to words, that is sound-function relationships that were arbitrary and often targeted one particular object. In contrast we were interested in whether infants display calls: non-linguistic vocal behaviours with consistent sound patterns that vary with the contexts in which they are produced, and contain broadly referential information that can be used by a listener to gain information about the infant's mood, attitude or activities. In other words, we wanted to investigate whether human infants have vocal behaviours in their repertoire that are similar to the calls that have been documented in many non-human primates.

The aim of our first study was to observe the vocal behaviour of prelinguistic infants in their everyday environment and identify situations in which all or many of these children produced vocal behaviour. Rather than providing a complete description of the infant's prelingusitic vocal repertoire (which has been attempted elsewhere, e.g. Halliday 1975, Oller 2000), we wanted to investigate whether there are situations in which infants consistently produce vocal behaviour, and whether this behaviour served some communicative function as indicated by the reaction of the recipient(s). Secondly, using the results from the observation, the aim was to identify possible similarities between sounds produced in the same situations by a number of infants. We therefore conducted an acoustic analysis of the observed sounds in one situation and compared them to those recorded in other contexts. This study aimed to identify possible constant sound-function relationships in the infants' natural interactions.

Our second study aimed to identify and assess the possible influence of culture and native language on the previously described vocal behaviours. The question here was whether prelingusitic sound-function correspondences are part of a primate-like, innate repertoire that is not influenced by culture or whether culture and native language, which already shape other classes of vocalisations, were already evident in non-linguistic vocal behaviours like these. To this end, we aimed to conduct a cross-cultural comparison with infants growing up in Scotland versus rural Uganda. We firstly wanted to investigate whether infants produce vocal behaviour in the same or similar situations, and secondly to compare the acoustic properties of vocalisations produced by infants of each culture.

Lastly, we wanted to investigate whether infant vocal behaviour alone was sufficient to allow listeners to draw inferences about the infant's emotional state or activity. To investigate this hypothesis we conducted a playback study that asked listeners to rate prelinguistic vocalisations with regard to their information content. This could provide further clues to whether and to what extent prelinguistic vocal behaviour contains information and, if this is the case, how listeners use this information.

Chapter 2: Non-linguistic calls in human infants

Summary

Vocal behaviour in non-human primates is usually investigated by describing the morphology of the signals, their contextual use, and their effects on conspecifics. Using the same approach, we observed the vocal behaviour of 22 pre-linguistic 11-to-18-month-old infants in their everyday nursery environment to investigate whether infants produce acoustically similar categories of non-speech vocalisations in the same behaviourally defined categories. From video episodes, we first described a number of contexts in which many of the infants regularly produced vocal behaviour. Of these, five categories of vocal behaviour (protests, food requests, action requests, declarative pointing and object sharing), yielded enough good quality samples for further analysis. We conducted acoustic analyses on all vocalisations produced in these situations, followed by a cross-validated discriminant function analysis. Results showed that the acoustic properties of vocalisations in four of the five presented categories of infant vocalisations varied significantly with the associated context and therefore can potentially transmit referential information to listeners. We conclude that these observed acoustic regularities in infant vocal behaviour are an important part of communicative acts that could function as communicative signals in a similar way to non-human primate calls.

Introduction

Vocal communication in primates has long been a focal point in comparative psychology. A standard methodological procedure is to collect recordings of calls in different situations, provide a detailed acoustic description of these calls and conduct statistical comparisons between them in relation to their context of emission. Such studies have been conducted with a large variety of species, for example squirrel monkeys (Winter, Ploog and Latta, 1966), vervet monkeys (Struhsaker 1967), Barbary macaques (Fischer and Hammerschmidt 2002), capuchin monkeys (Oppenheimer 1973), lemurs (Macedonia 1993), or bonobos (Bermejo and Omedes 1999).

A thorough description of a species' repertoire is often the basis for further experimental or observational research, for example concerning the function and meaning of calls (Seyfarth, Cheney and Marler 1980; Zuberbühler et al. 1997), the role of call combinations (Crockford and Boesch 2005, Arnold and Zuberbühler 2006) or the development of the adult repertoire (Seyfarth and Cheney 1986).

For example, Struhsaker (1967) provided a catalogue of the vocal behaviour of the vervet monkey, listing the circumstances in which vocal behaviour occurred, what effect this had on listeners and his impressions on what the vocalisations sounded like. These initial studies were the basis for subsequent acoustic analyses that identified functionally referential calls in many monkey species (Cheney, Seyfarth and Marler 1980, Zuberbühler 1997). In the vervet monkey, acoustically distinct alarm calls were identified that corresponded to the predator type (aerial or ground predator) and reliably provoked different flight responses in listeners (Cheney and Seyfarth 1980).

The aim of most communication research with nonhuman primates is to investigate the function and meaning of the various species-specific signals. Some studies have attempted to relate function and meaning in the vocalisations of non-human primates to concepts from human language, such as reference, syntax, grammar or signal flexibility. The general conclusion from this research is that nonhuman primates, particularly monkeys, have a genetically fixed vocal repertoire that is applied to different situations in more or less intentional ways. In some instances there is evidence that some call types serve a referential function in that they contain information about some external event, and others understand their meaning and how to react to them (for a review see Snowdon 2008, Zuberbühler 2003).

The vocal repertoire of one particular primate has been investigated very differently, namely that of the human infant. The main concern of many phonological studies has been to highlight the transition of sound production from unmodified, innate calls, like crying or screaming, to the flexible signals observed in spoken language, such as well articulated vowels or syllables and the combination of syllabic sounds (Owren and Goldstein 2008, Masataka 2003). The main focus of phonological research are the technical aspects of language; how an infant controls and manipulates her vocal production, how she learns the sounds of its native language and how she comes to combine these elements (Oller 2000, Vihman, Ferguson and Elbert 1986). The other main strand of investigation into early communicative development has been the analysis of gestural signals, which are thought to be more communicatively complex than vocalisations (e.g. Liszkowski 2008, Liszkowski et al 2008, Capirci, Contaldo, Caselli and Volterra 2007, Iverson and Goldin-Meadow 2005, Bates and Dick 2002), most notably pointing.

In this study, we were concerned with a question that is central to the study of primate behaviour: what is the natural communicative function of human infants' non-linguistic vocal behaviour in their everyday environment? In the first few months of life human infants are already capable of producing sound of considerable acoustic flexibility, beyond what is normally observed in primate vocal systems (Snowdon 2008, Oller 2000). This ability is thought to serve speech acquisition in that it enables the child to match and produce the sounds heard around her (Oller and Griebel 2008). There is, however, good evidence that the function of early vocal behaviour is more than simple practise in order to acquire spoken language, and that some vocalisations have important communicative functions of their own. Although rehearsing sound production may represent a biological function in its own right, there is also good evidence that, in addition, some vocalisations are expressions of simple messages. For example, in newborns vocalisations and associated facial expressions have been shown to convey information to caretakers about the infant's physical health and emotional state (Papoušek 1992). Furthermore, Scheiner et al. (2002) recorded vocal behaviour from 4-month-old infants in twelve different contexts, ranging from feeding to nappy changing. An acoustic analysis suggested that there were no measurable differences between the sounds recorded in the different contexts, but some differences were found between vocal behaviour in 'positive' (feeding, time with caregiver) and 'negative' circumstances (hunger, pain). Papoušek and Papoušek (1989) made a similar observation when they played back recordings of infant sounds to parents and let them decide whether the baby was in a positive or negative emotional state. As these examples suggest, vocalisations may play a key functional role in influencing caregiver or peer behaviour (Owren and Goldstein 2008, Papoušek 1992, Lock 1980) or modulating the meaning of accompanying gestures (Gómez 2007, Leroy et al. 2009).

Pre-linguistic human infants produce a large number of non-linguistic sounds from birth and continue to do so even after language competence is achieved, well into adult life. Nevertheless, the main aim of most studies on the pragmatics and function of non-linguistic communication has typically been to illustrate the transition to speech. For example, in one early study, Halliday (1975) reported on a 12-month-old infant's production of acoustically distinct calls when referring to a toy or when requesting food, suggesting these were transitions to words. A similar longitudinal observation by Lock (1980) suggested that ritualised patterns of vocalisations and gestures in the child's interaction with a caregiver are precursors to linguistic referential communication. Goldstein and West (1999) presented parents with sound clips of 9 to 16 month old infants during different activities. They found that parents were able to make judgements of the child's activity at a lvel above chance from the sound clips alone, when asked to choose from a set list of behaviours e.g. infant playing alone, infant playing with mother or infant hungry. In newborns, vocalisations and associated signals, such as facial expressions, have been shown to convey information for caretakers about physical health and emotional state (Papoušek 1992, Scheiner et al. 2002). Although these studies all point to a communicative role of pre-linguistic vocal signals, they do not provide direct empirical evidence for a systematic relation between acoustic structure and external events.

Although human infants have a flexible vocal repertoire that increasingly expands through vocal play and babbling (Griebel and Oller 2008), there might be some classes of calls with fairly constant acoustic properties. D'Odorico and Franco (1991) provided one of the few studies that tested whether infants between 4 and 11 months of age produced vocal behaviour that was acoustically related to different contexts. They recorded the vocalisations of four infants when obtaining a toy from an adult in three successive stages, joining attention, giving the object, and examining it. In a control condition, the adult showed the object, but refused to give it to the infant. By analysing spectrograms of the vocal behaviour that infants produced, the authors found context-specific vocal behaviour that was acoustically related to the different stages, particularly from eight months onwards. Another interesting finding was that calls were acoustically consistent across the four individuals up until the age of nine months, but showed more individual differences after that.

The goal of our study was to expand on these first results of D'Odorico and Franco (1991) by investigating the vocal behaviour produced during socially relevant situations by infants between the ages of 11 and 18 months, an age commonly associated with the onset of intentional communication and the entry into linguistic communication (Tomasello 2008, Bates et al. 1979).

In our initial study, we firstly wanted to observe and catalogue behavioural contexts in which all or most infants produce vocal behaviour. These descriptions are the foundation of our second aim: to explore whether young infants produced non-linguistic vocal behaviour that is acoustically similar in similar situations. Gómez (2007) suggested that one way of investigating human infants' non-linguistic communication is by treating the infants methodologically as an "unknown primate species" with a different sets of cognitive and vocal skills than other primates, but without the interpretative bias of later communicative and linguistic developments (see also Liszkowski 2008).

A key finding in many non-human primate studies has been that individuals produce context-specific call types that can be meaningful to others (e.g. Zuberbühler 2006). Whether or not similar sound-function correspondences can be allocated to some of the vocalisations produced by prelinguistic infants as part of their otherwise very flexible vocal repertoire is currently unclear. For example, work on infant cries has emphasised its graded nature, both in terms of production and perception, and its alleged function in expressing degrees of arousal (e.g. Protopapas & Eimas 1997; Owren, Rendall and Bachorowski 2005, Owren and Goldstein 2008) with no evidence for context-specificity. Similar points are usually made in relation to babbling, a communicative behaviour not normally seen in non-human primates (but see Elowson et al. 1998). Babbling is associated with a positive emotional state in the infant but is not associated with any particular production context or set acoustic variations (Oller and Griebel 2008). It is certainly true that rehearsing sound production and expressing inner states are important biological functions of human infant vocal communication but this does not exclude the existence of simpler sound-meaning associations that guide caregiver behaviour or attention in a period when the child's communicative expressions are limited.

Our main question was therefore whether there are reliable associations between some of the acoustic structures produced by infants and the events that usually trigger vocalisations or are associated with them. In order to investigate this question, we investigated the spontaneous occurrence of vocalisations in a naturalistic environment of many pre-school Western children, the nursery.

We are not the first ones to advocate this ethological approach. A number of studies in the 1970s and 1980s adopted the same stance, arguing that in order to get a realistic picture of the infant's communicative skills, these have to be studied in their natural environment (Bruner 1983, Locke 1993). By employing this method, we also hoped to gain further insights about the functional significance and evolutionary relevance of human vocal signals during a time span when language production is either not available yet or limited to a small number of words.

Method

Study Sites and Subjects

Data were collected from two groups of infants in two daytime nurseries in Fife, Scotland, between January and November 2009. The first group ("Acorn Nursery") consisted of 11 normally developing children between the ages of 11 and 18 months (see Table 2), who caregivers reported to have vocabulary of < 5 words.

Observations were collected while the infants moved around freely in one room (approximately 12 x 4 m). The room was furnished with cots and buggies used for naps, shelves, a table, a small climbing frame with a slide, a toy tent, a playpen with various toys and a number of highchairs arranged in a semicircle for mealtimes. There were various toys available at all times, such as a toy kitchen, building blocks and picture books. Additional toys were available on a changing schedule, such as musical instruments, play-dough, stuffed toys, plastic animal models and various kinds of buggies to move around the room. All infants had meals together; the younger ones in highchairs, the older ones at a table. Weather permitting, the infants spent time in an adjacent playground with a climbing frame, swing and slide and access to nearby farm animals, namely rabbits and cows. During data collection it snowed heavily so yard visits occurred rarely. Food was available as a mid-morning snack (e.g. toast, cereal, fruit), lunch (cooked meal with dessert) and afternoon snack (e.g. crackers, scones, cheese or bread sticks). Water was available ad libitum during meal times.

The caregiver-child ratio was 1:3 with at least three caregivers present at all times. Caregivers took care of basic needs, such as nappy changing, comforting, enforcing nap-time and feeding. Additionally they offered activities, for example singing songs and nursery rhymes, arts and crafts or baking.

The second group (“WonderYears Nursery”) consisted of 11 normally developing children between the ages of 11 and 18 months (see Table 2) and a reported vocabulary of <5 words. The group used two adjacent rooms (approximately 8 x 8 m each), one for play activities, the other for meals and naps. The children were allowed to move freely around the playroom, which was furnished with two tables and chairs, a seating area, a toy kitchen, a small ball-pool, a trampoline and a small sand pit. The second room contained tables and chairs and a small kitchen unit in one half and beds and buggies in the other. In the playroom toys, such as cars, stuffed animals, plastic animals, musical instruments or picture books, were available at all times. The caregivers provided one or two activities per day, such as painting, arts and crafts, and often provided additional toys, such as magnets or play tunnels. Weather permitting, infants had access to an adjacent playground with a climbing frame, various structures to walk over or crawl under and toys such as seesaws or tricycles. Food was available three times a day, in similar composition and schedule to the “Acorn Nursery.” After lunch nearly all children went for a nap. The caregiver-child ratio was 1:3. As the infants in this group were slightly older, greater emphasis was placed on encouraging independence, for example by practicing walking and standing, and independent spoon-feeding with associated praise. A typical nursery day is illustrated in Table 1. The structured routine in both nurseries allowed us to observe different situations frequently and with multiple children. Recruitment of participants was through the nurseries. Information sheets and consent forms were given to the parents, which they returned directly to the nursery or to the experimenter in case they had further questions. In terms of race and ethnicity, all infants were Caucasian/white. Several infants had a multi-language background with one or both parents being non-native speakers of English (WonderYears: N=4; Acorn: N=0). Overall, there were N= 14 male and N=8 female infant participants. In terms of socio-economic status, the nursery staff provided us with the information that all came from middle-class families, typically with both parents working.

Time	Activities
8.30 – 9.30	Arrival at nursery Playtime: toy cars, building blocks, animal models, picture books, toy kitchen, musical toys - available at all playtimes Varying structured activity, e.g. Play-dough is offered
9.30– 10.00	Morning snack
10.00 – 11.15	Playtime Varying structured activity, e.g. caregivers offer watercolours
11.15 – 11.30	Songs
11.30 – 12.00	Clean-up and lunch
12.00 - 13.15	Putting infants to bed and nap-time
13.15 – 15.00	Playtime Going for a walk, visit playground (weather dependent) Varying structured activity, e.g. caregivers set up tunnels for the children to crawl through
15.00 – 15.15	Snack time
15.15 – 17.00	Playtime Infants are being picked up from 15.30

Table 1: Daily Routine in two nursery groups

ID	AGE start	AGE end	Sex	Group
AL	1;5	1;6	m	AC
AR	1;2	1;6	m	WY
BE	0;10	1;1	m	WY
CY	1;1	1;3	f	WY
EM	0;11	1;2	f	WY
ET	1;3	1;5	m	AC
GA	1;5	1;6	m	AC
HA	1;3	1;5	m	AC
HU	1;0	1;3	m	WY
IG	1;5	1;6	m	WY
JO	1;2	1;6	m	WY
KA	1;3	1;5	f	AC
KV	0;10	1;1	m	AC
LA	0;11	1;1	f	AC
LX	1;3	1;6	f	WY
LI	1;0	1;2	m	AC
MC	1;4	1;6	f	WY
ME	0;10	1;0	f	AC
MI	1;5	1;6	m	AC
SP	0;11	1;2	m	AC
ST	1;2	1;6	m	WY
ZO	1;4	1;6	F	WY

Table 2: Age (start and end of data collection) and sex of participating infants. Nursery group is notated in the last column (AC= Acorn Nursery, WY= WonderYears Nursery)

Data Collection

Observation usually began midmorning and lasted until the afternoon snack. The observer participated in the normal nursery routine, and occasionally engaged with the infants in some activities, e.g. play or book reading. This was important because it habituated the infants to the presence of the observer and recording equipment and thereby ensured natural behaviour during data collection. We aimed to record any instances associated with the production of infant vocalisations. All events were document on video, using all-occurrence sampling (Altmann 1974), i.e. filming whenever episodes of vocal behaviour were observed. Filming was sometimes continuous, for example during lunch preparation or play sessions, to anticipate episodes of vocal behaviour.

Materials

Episodes of vocal behaviour were recorded using a Sony Handycam DCR-HC19E camcorder and analysed with the 'iMovie' software package on an Apple MacBook Pro computer. Sound was extracted from the videos using the 'QuickTime' software package. Audio recordings were transferred digitally onto the computer using 'Praat 5.1.03' (www.Praat.org). Images were created using Raven Pro (www.cornell.edu/birds)

Variables

Functional Categories

Following a traditional ethological approach, we firstly tried to characterize the contexts in which the recorded vocal behaviour occurred. We tried to classified these into distinct categories. Our general approach was to allocate calls to specific social situations, that is, the context in which they occurred, including the reaction provoked in listeners, who could hear the infants vocalise. Hence, we first created a list of situations in which vocalisations were systematically recorded (Table 3). In this list we present more categories than were included in the subsequent analysis (Giving, Declarative Pointing, Protests, Food Requests and Action Requests) to illustrate when the infants produced vocal behaviour and what shape this could take. We are, however, aware that these categories do not represent the infant's full repertoire. For our purposes, we wanted to concentrate on the categories of vocal behaviour that were observed across a great number of infants.

Call Classification

On the basis of these descriptions, we identified specific markers that determined the classification of a call and when the call began and ended. Part of these descriptions were the action that the infant was performing or was trying to perform, for example stacking bricks, reaching for an object, or events in the environment, such as meal preparation or getting ready to go out. Similarly, the ending of a call was determined on the basis of the recipient's reaction, for example making comments or complying with requests, or the infant's response, for example being satisfied and moving on to another activity or calming down. Calls were assigned to specific categories on the basis of these criteria. Naïve coders were then asked to categorise 20% of the calls in our dataset using the aforementioned criteria. Furthermore, two naïve coders were presented with a set of video samples (40 samples) of vocal behaviour that were included and excluded in the final analysis to control that the exclusion criteria used here were consistently applied. Coders were given a sheet with descriptions of the contexts in which vocal behaviour occurred and an additional box in case they thought the video did not fit any of the descriptions, and asked to indicate to which category the video sample belonged.

Excluded Vocalisations

The majority of vocal behaviour in our samples consisted of quasi-vowels, unmodulated sounds such as screams or grunts and seldom consonant-vowel pairs. In the acoustic analysis we excluded vocal behaviour that was fully linguistic (i.e. words like car, tractor, mama or duck), or proto-linguistic in character (word approximations and word attempts such as /'ook/ for book or /aed-y'/ for teddy).

We also excluded as vocal behaviour that was not readily classifiable into the proposed categories once they were derived. This was a large proportion of the sample but was in line with our aim to identify categories in which many infants produced vocal behaviour, rather than concentrate on a full description of the infant repertoire or individual differences.

Furthermore, we excluded samples, which were overridden by background noises or other voices, as these would have produced unreliable measurements. The nature of the nursery environment makes it a difficult environment to obtain pristine sound recording. It is very common that multiple children vocalise alongside caregivers or each other, and consequently a large proportion of the data was excluded because it would not have led to reliable measurements. If Praat was unable to give a clear visualisation of different call parameters, these measurements were also excluded. Furthermore, we excluded laughs, cries or screams from the acoustic analysis to avoid false differences between the groups.

Combining Multiple Contributions of Individuals

In order to minimise the impact of individual contributions on the vocalisations in the acoustic analysis, we randomly picked two calls out of multiple contributions by one individual and randomly assigned them to either the construction or the test data set. To ensure randomisation, we used the random numbers generator in PASW 18.0 and assigned these to an individual's set of calls. We chose calls assigned the highest and second highest random number and assigned the first to the construction set and the latter to the test set. This greatly reduced the overall sample size in favour of a more conservative measure.

Acoustic Analysis

We measured the acoustic structure of vocalisations recorded in the different contexts using Praat 5.1.03 (praat.com) with the following settings: pitch range 0 – 2000 Hz, spectrogram view range 0 – 25 kHz to determine the number of harmonics and 0 – 10 kHz to measure fundamental frequency. Intensity measures were taken using the program's algorithm to obtain the median intensity for the entire call and the maximum intensity was measured using the option to view a list of intensity measured for the entire call and identifying the highest value. The following spectral measurements were taken: (1) 'mean F0' = fundamental frequency across the entire call (Hz), (2) 'early F0' = fundamental frequency at call onset (Hz), (3) 'mid F0' = fundamental frequency at middle of the call (Hz), (4) 'end F0' = fundamental frequency at call offset (Hz), (5) 'max F0' = maximum fundamental frequency (Hz), (6) 'min F0' = minimum fundamental frequency (Hz), (7) median intensity (dB), (8) peak intensity (dB), highest intensity across the entire call (dB), (9) 'N harmonics' = number of visible harmonic bands, (10) 'N units' = number of consecutive call units that formed a call segment (not separated by more than 3s), and (11) 'duration' = total length of the call segment (s).

Statistical Analysis

In order to explore whether the calls recorded in different contexts differed significantly with regard to the 11 measured acoustic parameters we conducted a Discriminant Function Analysis (DFA) using PASW 18.0. The DFA compares the vocalisations in each context category with regard to the acoustic parameters and explores whether there are differences between eachcategory. The DFA then classifies data points using the parameters it derived from the original data. This method is commonly used in the assessment and classification of primate calls (e.g., Clay and Zuberbühler 2009, Slocombe and Zuberbühler 2005, Crockford and Boesch 2003) and compares whether a simple statistical algorithm can generate a higher rate of correct classifications than chance. To cross-validate the results from this analysis, we split our data into two sets. We used one set, the construction set; to derive a number of funtions that would discriminate between different call categories. This model was then applied to the other set of data, the test set, to investigate whether the discriminant functions derived from the construction set could classify new data points at a level above chance.

Terminology

We describe the observed episodes of vocal behaviour as 'calls', although this term is adopted from research with nonhuman primates, we emphasize that we do not imply or compare the vocal behaviour of human infants to the more fixed categories found in apes and monkeys. We use the term calls, because we believe that a comparative method also warrants comparative terms and because it is largely free of a linguistic interpretative bias.

Results

Behavioural contexts and qualitative descriptions

Overall, we recorded 624 instances of infant vocal behaviour. Our general approach was to allocate calls to specific social situations, that is, the environmental and behavioural context in which they occurred. In some cases, distinctions were also based on the apparent function, inferred from the reaction they provoked in listeners. Where possible, we will provide a description of some of the features of the sounds that the infants produced. Our goal was not to provide a detailed phonetic transcript of infant vocal behaviour as these can be found elsewhere (Oller 2000,Bernhardt and Stemberger 1998).

Vocalisations recorded in the	Description
Object request	Request to obtain an object, mainly directed to caregivers
Private	Solitary activity, no searching or scanning for partner, not addressed to anyone.
Social Routine	Singing, reciting songs or nursery rhymes; speech-based dialogues ('hello', 'bye-bye', 'sorry', 'thank you') often accompanied by gestures
Aggression	Attacking or fighting with another infant, e.g. pushing or hitting, gestures observed were clenching of fists, arm waving, angry facial expressions
Food context	Requesting or attaining food during food preparation and distribution, frequently accompanied by reaching gestures
Object sharing	Giving, receiving or game-like exchanges of objects, often accompanied by extended arm and open palms, 'begging'
Protest	Reaction to unpleasant events by another or the environment, co-occurs with averse body orientation, clenched fists, arm shaking and waving, kicking legs
Declaratives	Reaction to a new or exciting external event; sometimes with pointing gestures
Action request	Request help with an action that infant is unable to perform, mainly directed to caregivers.

Table 3: Overview of categories in which vocal behaviour was observed

(a) Private/non-directed vocal behaviour:

Infants sometimes produced vocalisations that were not obviously directed at anyone. In these episodes the infant was generally involved in a solitary activity, for example examining a toy, riding a tricycle or playing with bricks. Generally, there was no searching, calling for or scanning the surrounding area for a partner. For example, the infant was riding a tricycle across the playground with no observable goal while vocalising. Vocalisations were audible, but generally quiet. Babbling and variegated babbling occurred occasionally. Often sounds were produced either continuously or in longer units, sometimes laughter was observed.

(b) Social routines:

Infants commonly vocalised during social routines, such as singing or reciting of familiar songs or nursery rhymes. In the typical case, caregivers led the infants in singing songs and most infants contributed with vocal behaviour and frequently gestures, for example waving their arms or executing set routines. The other notable contexts were acts like saying "hello" or waving "bye-bye", saying "sorry" or "thank you". These were strongly encouraged, often modelled by caregivers and verbally rewarded. For example, the infant was picked by his mother, who turned him to face the rest of the children, and said "Say bye-bye everybody". The infant waved and vocalised, often over a caregiver's speech. With increasing age, there was a tendency to vocally imitate the mother's speech sounds.

(c) Aggression:

Overt aggressive behaviour with vocalisations was almost exclusively addressed to peers, usually in competition over toys. Vocal behaviour was often accompanied by pushing or hitting the opponent, as well as angry facial expressions.

For example, two infants wanted to grab a tennis racket and got one end each. One attempted to hit the other, vocalised and pulled the racket towards him. A carer approached and took him away. Calls emitted during aggressive behaviours were very loud, and often contained high intensity shrieks on behalf of the aggressor.

(d) Food context:

Vocal behaviour was very common when infants sat down for a meal and food was prepared or handed out and already visible. Vocalisations were often accompanied by reaching gestures and orientation towards the food. For example, at snack-time a caregiver handed out biscuits. One infant stretched out in his chair, oriented himself towards the food, vocalised, extended his hands towards the caregiver and repeatedly performed grasping motions. The calls were characterised by high intensity, un-modulated frequency contours and repeated emissions. Our subjective impression was that calls in response to high-preference foods, such as chocolate cake, were of higher amplitude and acoustically different from other calls.

(e) Object exchange:

Vocal behaviour was observed when infants gave or received an object from a peer or caregiver, or as part of a game-like interaction where objects are passed to and fro between infant and recipient (see Ratner and Bruner 1977). For example, an infant sat on the floor playing with a toy car. Another infant approached, handed him an interesting toy, vocalised, turned away and continued playing. Vocal behaviour that accompanied the reception of objects was less frequent than when giving objects. In all instances vocal behaviour was characterised by short utterances, with falling intonation and few harmonics.

(f) Protests:

Infants often vocalised in response to unpleasant events, either undesirable actions performed on them or unpleasant environmental events. For example, infants often vocalised when a caregiver performed a necessary action on them, which was almost always accompanied by physical resistance, including kicking, arm shaking, throwing or pushing objects or refusing to move. For example, an infant playing outside was asked by the caregiver to go inside. Because the infant did not comply the caretaker attempted to pick the infant up, but he held on to the tricycle, vocalised and jammed his feet into the ground. The caretaker lifted the infant who vocalised more and cried, oriented towards the tricycle and tried to push away from the caretaker. Protest calls were long and had very high amplitude with a large number of harmonics often combined with crying and screams.

We differentiated protest vocalisations from crying, innate vocal behaviour associated with the production of tears when the infant is unhappy or hurt, as these behaviours are acoustically quite different from voiced sounds (Bernhard and Stemberger 1996). Crying sometimes occurred after protest vocalisations.

(g) Declaratives:

In the classic case, declarative behaviour is the production of the pointing gesture to direct another's attention towards a particular object or event (Bates et al. 1979). However, we often observed declarative behaviour without pointing in a diverse range of situations, such as spotting a favourite toy, spotting the mother who has arrived to pick the infant up, or showing a caregiver an interesting object. Vocalisations were regularly produced during such declarative episodes, as if the infants wished to 'comment' on a specific aspect of the environment, either something immediately relevant for the current activity or something new and exciting. For example, an infant and caregiver were both looking out the window towards the sea. A boat approached the harbour. The infant pointed and vocalised. In another example, the infant was playing with stuffed animals, saw a favourite caregiver walk past the door to which he picked up the animal and ran towards the door while vocalising. Vocal behaviour in these situations was very diverse, perhaps reflecting the different classes of declaratives. Calls accompanying declarative points were usually short, with rising intonation and high intensity, mostly produced as single units.

(h) Requests for actions and objects:

Infants vocalised when they were either unable to perform an action or wanted an object that was out of reach or otherwise not available. These calls were very co-ordinated with eye contact with the addressee and sometimes involved pointing, reaching gestures, iconic motions of the action needed or leading a person to the problem or object. Addressees were almost exclusively caregivers with vocalisations functioning to alert the addressee to the problem. For example, an infant and adult were building a tower with Lego bricks. The infant was unable to put the bricks together alone and seemingly wanted the caregiver to do this for him. He looked at the caregiver, vocalised and pointed to the bricks and eventually picked up a brick to pass it to the adult. Calls consisted of short, mid-amplitude units with audible exhalation that were repeated quickly.

Classifying video episodes of vocal behaviour

We divided episodes of vocal behaviour by using specific environmental and behavioural markers that helped us determine the start and end point of an episode (Table 4). These were also used to classify the video clips into one of the categories.

Context	Start of Call	End of Call
Private	- Infant is by herself - Infant starts vocalising - Does not look around for others - Others are not engaging in an activity with the infant	- Infant stops vocalising - Infant stops current activity, gets up and does something else - Other person interrupts the infant's activity and initiates exchange
Social Routines	- Nursery rhymes or songs being sung - Verbal cues from caregivers such as 'Say bye-bye' - Routine gestures such as waving and hand actions associated with certain songs and rhymes	- End of song or rhyme - Praise or comments from caregiver after vocal behaviour has occurred
Aggression	- Infant is or is attempting to hit or kick another - Infants throws object in direction of other, attempting to hit them - Infant attempts to violently take object from another	- Caregiver interrupts and removes the aggressor or victim - Conflict is resolved, both victim and aggressor are peaceful again
Food Context	- Food is present but not yet available for the infant - Meal preparation	- Food has been consumed and is no longer available - The infant no longer shows interest in eating

	begins, e.g. sitting down in high chairs, cutlery and plates being distributed - Additional or different food is made available	the food present - Meal times have finished and food is removed
Giving	- Infant is in the presence of objects and potential communicative partner - Infant stretches out arm and offers the object to someone - Infant holds an object, moves towards someone else and hands the object over - Playful object exchange	- Partner receives object - Caregiver comments upon receiving object - Infant turns away from partner and object, engages in another activity
Receiving	- Someone offers an object to an infant - Infant is waiting for a response after requesting an object	- Infant takes object and leaves the partner - Infant returns the object to partner
Protests	- Someone performs an action on the infant - Infant physically resists this action by kicking, pushing, winding away - Infant tries to avoid action - Someone takes a desired object or toy away from the infant	- Infant calms down and engages in another activity - Caregivers offer soothing actions, infant grows quiet - Displacement activity is offered, infant engages in this activity

	- Someone takes the infant away from an object or place	
Declaratives	- Interesting event has just happened, for example an animal or boat or familiar person went past - Presence of moving objects such as balloons or bubbles - Reading a picture book and being asked a question by a caregiver that requires finding and object - Seeing a familiar or well-liked object or person, for example older sibling playing in the yard - Something surprising spontaneously occurs for example infant discovers that a mat plays music when walking on it	- Infant stops pointing to event or object - Infant received a response from caregiver and either continues with activity or moves on - Event ends - Infant no longer shows interest in activity, object or event
Action Requests	- Infant tries to reach an out-of reach object - Infant needs help operating a toy - Infant can't perform an	- Caregiver responded to infant's request and infant is satisfied with the answer - Problem is solved, infant

	action for example putting Lego bricks together or tying her shoes - Caregiver is present and in a position to help the infant - Infant and caregiver do an activity together where each of them contributes a different part - Infant approaches caregiver in order to seek help - Infant brings object to caregiver that needs fixing, putting together or operating	goes away - Caregiver does her part of the activity

Table 4: Markers for start and end of calls

Interrater reliability

In order to validate our method for classifying video episodes we asked two naïve observers to classify the context of 20% of our video samples entered into the acoustic analysis using the descriptions we presented earlier (declarative pointing, food, giving, action request and protests). The raters' classification of video episodes corresponded 100% with ours.

Two coders were asked to classify video episodes using the above coding sheet and descriptions of vocal behaviour. They were asked to watch 60 video samples of infant vocal behaviour that were randomly chosen from the entire body of data (624 calls). This sample also contained videos that were excluded in the acoustic analysis. Coder's rating of videos corresponded 95.12 % with ours.

Excluded video samples

Out of the 624 video episodes of vocal behaviour we recorded, 295 samples were excluded from the subsequent analysis. Reasons for exclusions were poor sound quality caused by multiple individuals vocalising at the same time (109/295), large amount of background noise (83/295), large distance to recorded individual (13/295) , the individual moving out of the microphone's range whilst vocalising (4/295) or failure of recording equipment (2/295). Furthermore, we excluded samples of vocal behaviour that did not match any of the categories we described earlier (84/295), these behaviours were, for example, vocal play or babbling, ritualised vocal exchanges with caregivers or other vocal behaviour that was not easily classifiable. After this initial exclusion, 329 samples of vocal behaviour were used for the analysis.

Vocal activity in different contexts

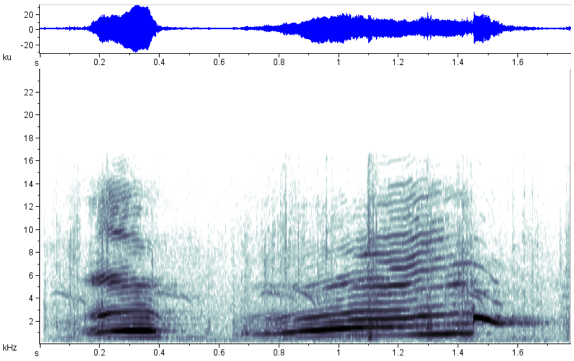
The most frequently observed calls were in protest conditions. They were present from 11 months of age and accounted for 22.2% of all recorded calls (73/329). Nearly all children in the sample vocalised in protest situations. Another frequently observed category was food calls, accounting for 21.6% of recordings (71/329), possibly due to the fact that the feeding context was frequent, occurring three times a day. We did not observe any food calls in the absence of food or in anticipation (e.g. during food preparations). Calling in declarative situations, often accompanied by the pointing gestures, was less frequently observed (31/329, 9.4%). Vocally supported, game-like object exchanges and action requests were mainly observed in older infants, perhaps a function of increasing joint attention skills. One of the less frequently observed calls was action requests, accounting for 8.8% of recordings (29/329). Acoustically, they were comparatively complex, addressed to one specific recipient, repeated or modified and often involved a combination of vocal behaviour and gestures, until the recipient met the request. Vocal behaviour during aggression was the least frequently observed category (4/329, 1.2%), perhaps due to caregivers' quick interventions. Finally we observed a reasonably high number of private and social routine episodes (46/329, 13.9%). The contribution of each individual infant is shown in Table 5.

ID	PT	FRQ	ARQ	DP	GV	RC	PR	IP	SOC	AG
AL	1	4	4	3	3	1	1	3	3	0
AR	1	2	0	0	0	0	0	0	0	0
BE	8	2	1	1	0	0	0	0	1	2
CY	5	3	0	1	0	0	1	1	1	0
EM	1	1	0	0	0	0	0	0	0	0
ET	1	1	0	0	0	1	2	0	1	0
GA	5	1	2	1	0	0	3	0	1	1
HA	0	1	0	1	0	0	1	1	0	0
HU	13	18	2	3	1	1	2	1	1	0
IG	6	4	1	0	3	0	2	0	2	0
JO	5	9	14	7	6	3	9	1	4	0
KA	2	0	2	0	0	0	0	0	0	0
KV	8	6	0	1	3	1	3	0	0	0
LA	0	1	1	1	0	0	1	1	0	0
LX	1	0	0	1	2	1	6	0	2	0
LI	2	0	0	1	1	0	0	0	0	0
MC	6	4	0	4	1	2	6	1	11	0
ME	1	1	0	0	0	0	0	0	0	0
MI	0	11	0	2	0	0	4	1	1	0
SP	5	0	1	2	2	3	4	0	1	1
ST	0	1	1	0	0	0	1	0	1	0
ZO	2	1	0	2	0	0	0	0	0	0
Tot	**73**	**71**	**29**	**31**	**22**	**13**	**46**	**10**	**30**	**4**

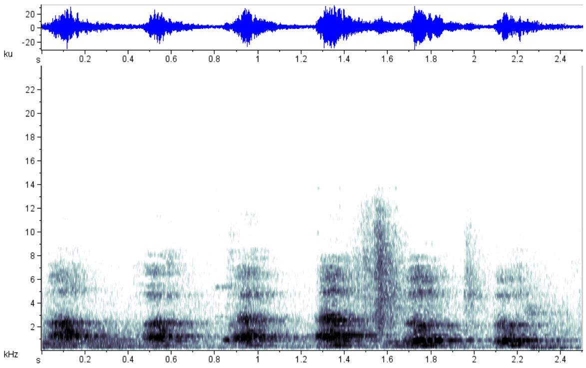
Table 5: Individual's contribution to dataset (PT- Protests, FRQ- Food Requests, ARQ – Action Requests, DP – Declarative Pointing, GV – Giving, RC – Receiving, PR – Private Vocal Behaviour, IP – Imperative Pointing, SOC – Social Routines, AG – Aggression)

Call structure in different contexts

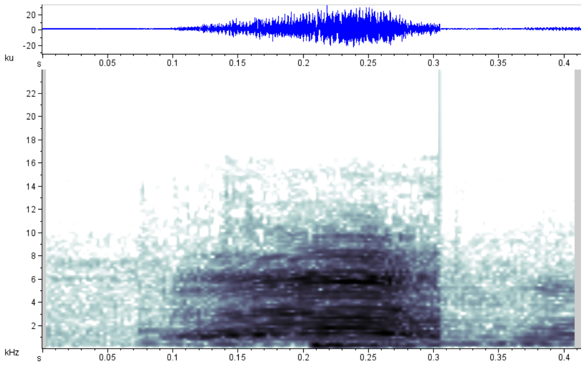
Five social conditions yielded a large enough sample size of good quality recording (n= 168) for quantitative acoustic analyses: (a) food requests (n=50), (b) action requests (n= 29), (c) protests (n=57), (d) declarative pointing (n= 16), and (e) giving (n= 16). Figure 1 depicts spectrographic illustrations of typical vocalisations produced in these situations.



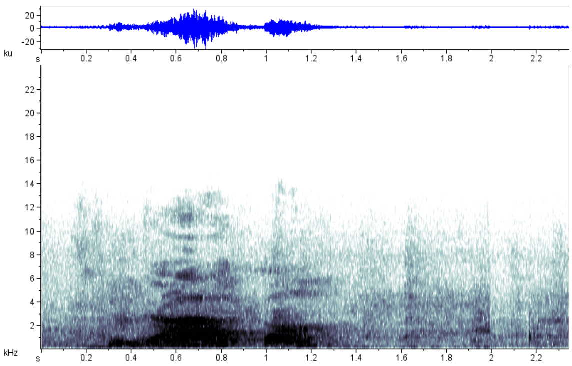
a) *Food call by MC (1;5), requesting more of a favourite food*



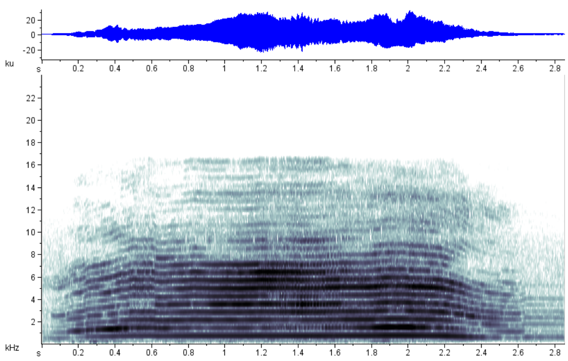
b) *action request by JO (1;4), asking adult to connect two Lego Bricks*



c) *vocalisation accompanying declarative pointing gesture by AL(1;4)*



d) *EM (1;1) giving a toy to a peer*



e) *protest call by HU (1;2)*

Figure 1: Examples of time-frequency spectrograms illustrating the five different call types that were compared in the acoustic analysis

We first checked for co-linearity by regressing all parameters and removing variables with a variance inflation factor > 10 (i.e. mean fundamental frequency). We entered all other nine uncorrelated acoustic parameters for further analysis.

The next step was to create a model based on a set of data. For this we randomly picked one data point from each individual. In cases where individuals only contributed one data point, we used this in the construction set to calculate the model. This controlled for individual contribution but greatly reduced the sample size (from n= 168 to n= 58 in the test set as follows: food requests: n=15 ; action request n= 11; declarative pointing n= 9; action requests n= 10; protests n= 16). The mean values and standard deviations for each of the acoustic variables in each call type are provided in Table 6.

Variables	Food Requests	Action Requests	Declarative Pointing	Giving	Protests
f0 onset (kHz)	386.35 (97.67)	366.66 (105.22)	385.05 (89.00)	366.11 (120.92)	561.64 (245.14)
f0 middle (kHz)	451.92 (113.11)	446.77 (82.29)	409.33 (104.05)	384.33 (60.62)	610.79 (455.13)
f0 offset (kHz)	709.54 (402.31)	404.40 (124.62)	330.85 (32.44)	562.60 (392.21)	727.41 (558.49)
f0 max (kHz)	764.30 (408.70)	459.75 (84.01)	442.82 (94.64)	668.31 (372.99)	963.73 (673.07)
f0 min (kHz)	331.30 (104.53)	334.12 (114.02)	261.62 (97.25)	284.72 (82.27)	416.81 (227.22)
med Intensity (dB)	76.85 (3.28)	72.25 (5.09)	70.73 (4.86)	73.17 (3.89)	76.25 (3.88)
peak Intensity (dB)	81.35 (1.97)	77.31 (4.30)	75.95 (3.95)	78.13 (3.69)	81.22 (2.64)
harmonics	22.60 (4.45)	15.73 (1.27)	18.17 (5.04)	16.30 (5.31)	21.44 (4.75)
call units	1.93 (1.22)	3.55 (1.37)	1.33 (0.82)	1.20 (0.63)	2.63 (2.09)
duration (sec)	1.11 (0.59)	0.48 (0.14)	0.68 (0.29)	0.54 (0.71)	1.54 (0.98)

Table 6: Means and standard deviation for each acoustic variable measured for each call category

From these data we conducted a Discriminant Function Analysis (DFA) to construct a model to classify the other datapoints. Out of the four functions used in the DFA, two functions discriminated significantly between the five call types and explained a significant amount of the variation in the acoustic structure in the call types (Wilk's Lambda = 0.152, χ^2= 93.316, p< 0.001 – Table 7).

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.362[a]	50.9	50.9	.759
2	.860[a]	32.2	83.1	.680
3	.295[a]	11.0	94.1	.477
4	.158[a]	5.9	100.0	.370

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 4	.152	93.316	40	.000
2 through 4	.359	50.769	27	.004
3 through 4	.667	20.046	16	.218
4	.863	7.268	7	.401

Table 7: Eigenvalues and Wilk's Lambda for the four discriminant functions used in the analysis.

Basic analysis

The model successfully classified 79.3% of all calls according to context. The success rate of call classifications was highest for the category action requests (90.9%), followed by protests (81.3%), food request (80.0%), giving (70.0%) and declarative pointing (66.7%). The greatest source of confusion was in the declarative pointing category, where calls in this category were often classified as either food requests (20.0%) or instances of giving (16.7%).

As they were five categories in which a given call could be classified, we would have expected a random model to be successful in 20% of the cases. Consequently, the significance level for the model was a correct classification rate of 20% or above. Therefore all categories of calls were correctly classified at a level above chance.

Cross-validated analysis

In order to cross-validate the model, we created a test set from our raw data. This test set consisted of data points from infants who contributed multiple calls. From these we randomly chose one call for each individual (n= 30, food requests n= 11, action requests, n= 4, giving n= 3, declarative pointing n= 5, protests n=8). We then used the functions derived from the previous DFA to classify the test. Overall, the model correctly classified 60% of the test cases. Classification rate was highest for action requests (75%) followed by food requests (71.4%), protests (70.0%) and giving (66.7%). Declarative pointing was only correctly classified 16.7% of the time and was often misclassified as either food request (50%) or giving (33%). Figure 2 illustrates the distribution of data points along the two discriminant functions identified by the analysis.

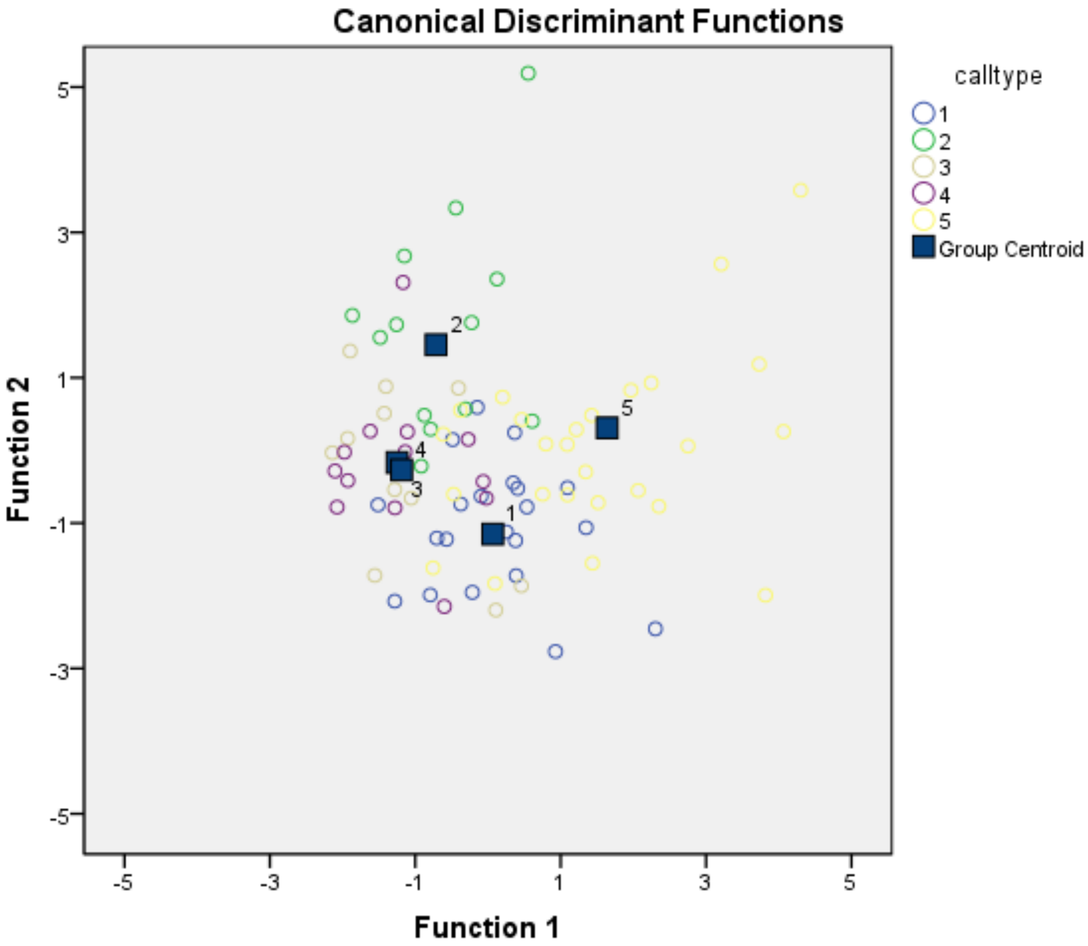


Figure 2: Diagram illustrating distribution of discriminant scores along the two canonical discriminant functions shown to discriminate between the different call types (Eigenvalues: Function 1= 1.362, Function 2= 0.860). Squares represent the group centroids, 1) food requests, 2) action requests, 3) declarative pointing, 4) giving, 5) protests

Post-hoc test

In order to test whether the rate of correct classification obtained by the DFA was significantly above chance, we conducted a binomial test. We analysed the number of correct and incorrect classifications for each call category in the cross-validated analysis. As there were five possible responses, the chance of correct classification was 20%. The categories of protests, action requests and food requests and giving all had a correct classification rate significantly above chance ($p>0.05$).

Calls in the category declarative pointing did not display a significant level of correct classification.

Control for infants from a multi-lingual background

Four infants in our sample came from multi-lingual background (one language other than English spoken at home). As explained earlier, this could possibly distort the data as language is thought to be evident in the infant's nonlinguistic sound from about six months of age (Vihman 1996, Stark 1980). We therefore re-ran the analysis excluding any data points produced by multi-lingual infants. The construction set now correctly classified 76.1% of the cases. We cross-validated the model excluding contribution by individuals from a multi-language background by entering a test set of data and found that the model correctly classified 61.9% of the data.

Discussion

We documented the natural non-linguistic vocal behaviour of human infants in one of their natural habitats, a nursery environment. Infants produced a high variety of vocalisations in a number of situations. We classified these call episodes using behaviourally and situationally defined categories. Five categories provided large enough samples for quantitative acoustic analyses. Results suggested that human infants produced four classes of vocal behaviour that varied systematically with their production context: Giving an object to a peer or caregiver, requesting an action, protesting and requesting food. Context-specific vocal behaviour could conceivably contain enough information to guide caregivers' behaviour towards the infant even in absence of additional visual information. For example, the caregiver can use the infant's vocalisations to decide whether she is safe and happy or whether it is necessary to check on her and find out what she is doing. Our results thus confirmed earlier reports on context-specific acoustic differences in infant vocal behaviour (e.g. Dore et al. 1976, Locke 1993, Bruner 1983; Papoušek 1992).

In contrast to many primate studies, we were unable to provide a full catalogue of the infants' repertoire or any further analysis with regard to frequency and proportion of the observed calls compared to other vocal behaviour. Such data would be an important addition to our analysis but was not realisable in the current set-up as it would require at least one other researcher to monitor the infant's activity whilst the other is recording. Furthermore, it is important to consider that the nursery environment is a rather difficult environment for audio recordings. It is often loud, many individuals vocalise at the same time and many activities overlap. Given these constraints, it is not surprising that a large proportion of our recordings had to be rejected for the acoustic analysis. Nevertheless, we gathered a suitably large body of data that could be analysed.

Although our study was methodologically similar to field studies with nonhuman primates, we are not implying that the documented vocal behaviour can be easily or directly compared to nonhuman primate vocalisations. For example, it has been suggested that infant vocalisations can be influenced by the infant's native language (Mampe, Friderici, Christophe and Wermke 2009), by ritualisation (Lock 1980), or by maturational processes (Oller 2000). Nevertheless, we were able to demonstrate that young infants produced non-linguistic vocalisations with stable acoustic core features and used these calls to interact with their everyday environment and the people around them.

A significant majority of the measured calls could be classified based on their acoustic structure alone, suggesting possible intrinsic informational content. The fact that classification was not perfect is not a major concern as non-human primates and humans tend to perceive sounds categorically, typically by relying on a restricted number of acoustic features (Oller and Griebel 2008, Bernhardt and Stemberger 1998). For example, human listeners might discriminate two calls with similar intensities but different durations, a pattern that will generate only a weak statistical difference with a DFA. Computer modelling and neural network-based approaches are increasingly used in the analysis of acoustic data and could potentially provide a more precise analysis (Pozzi, Gamba and Giacoma 2010; Warlaumont, Oller and Buder 2010).

In the only non-significant call category, the DFA seemed to systematically misclassify calls in the category of declarative pointing as either food requests or giving, suggesting that they were acoustically similar and perhaps based on related underlying psychological experiences, for example arousal or attitude. Bates et al. (1979) suggested that declarative pointing and giving an object to an adult so that he may comment on it or initiate an activity are both part of the broader category of indicating expressions. Giving in order to indicate or show an object to a caregiver is argued to precede declarative pointing in development and to be a good predictor for the emergence of declarative pointing (Bates et al. 1979). In case of ambiguous calls like these, contextual information helps to disentangle possible interpretations and interpret the calls correctly. For example, if a long, high intensity call was produced from a child as a caregiver distributed food to others, and the child also produced reaching gestures, the caregiver often reassured them that they would get their share or sometimes even handed them a portion of food. In contrast, when an acoustically similar call was produced whilst a child was pointing at a distant target, for example a boat or animal, the caregiver looked over to the child, identified the target and often commented on it or labelled it. Both calls might reflect high arousal level or the underlying function of obtaining the caregiver's attention to either provide food or turn their attention to a specific event. The reaction of caregivers and their verbal explanations imply an understanding of the child's vocal behaviour – this understanding could, however, also be based on other communicative signals accompanying the vocal behaviour, or be solely based on an interpretation of the context in which the vocalisation was observed. Although we could speculate that this might help the child to form sound-meaning associations, which are crucial for an understanding of how language works (Bruner 1983), further experimental evidence is needed to assess the value of the vocal signals we documented here. A first step will be provided in the forthcoming chapters.

An alternative explanation would be that because of caregivers' consistent reactions to the child's behaviour, the child produces ritualised vocalisations to serve certain functions (Locke 1980, Halliday 1975). The acoustic consistencies found in our sample would therefore be the product of ritualised vocal gestures that have been practised many times with a caregiver. This also implies that the produced vocal behaviour is intentional if the child produces it consistently as a means to an end. We think that this explanation cannot account for all of our findings, firstly because we collected data in two different day care facilities and it is unlikely that they would share the same ritualised communication patterns. Secondly, ritualisation downplays the role affect plays in the shaping of vocalisations (Owren and Rendall 2001, Owren and Rendall and Bachorowski 2005), for example we can imagine that protest calls are mainly motivated by affect as they are the infant's immediate reaction to an aversive stimulus. The call patterns we presented might also have differing levels of affect as underlying motives. For example action request often occurred with close monitoring of the listener's attention and reaction and might therefore be less emotionally motivated than protest calls. More discussions on the motivation underlying infant vocal behaviour can be found elsewhere (Oller 2000, Bruner 1983,Locke 1980, Bates et al. 1979).

In request calls vocalisations rarely occurred alone – they were nearly always combined with gestures, such as reaching, showing, pointing or leading another by pulling them towards the desired goal. In this particular case it seemed that the vocalisations indicated a broad functional category, which was then specified through the help of a gesture or action that identified the referent. For example, an infant sat in a pile of connectable plastic shapes and failed to connect them. She picked up two pieces and turned around to face a caregiver sitting near her. She then uttered a request call, showed the pieces to the adult and moved them together and apart. The adult commented that she understood and put the pieces together. During the following play sequence this request was observed multiple times, becoming less elaborate as play moved on and the rules of the game were established.

It is conceivable that one call served both functions in this instance - obtaining the recipient's attention and telling him something about the reason for the call, i.e. request the listener's help. Other authors have also reported the close association of vocalisations and gestures. For example, Golinkoff (1983, 1986) investigated the request-behaviour of infants between 12 and 14 months by placing them in a highchair, facing them with an object they would want and manipulating the adult's response. Although it was not their major finding, they reported that in most cases requests, repetitions and corrections consisted of gestures and vocal behaviour. Similarly, in a study on the co-ordination of pointing gestures with attention-getting behaviour, Liszkowski et al. (2008) found that 12-and-18-month-old infants reliably combined calls and points, and that these combinations were present even before the infants used the vocal channel to ensure an adult's attention to their pointing and consequently the goal of their gesture. Bates et al. (1979) reported that as much as 85% of points were accompanied by vocal behaviour, and Gómez (2007) cites personal observations suggesting infants systematically combine pointing with vocalisations.

The aforementioned examples highlight that, in addition to acoustic information contained in infant calls, there are other important channels in the whole communicative act, namely gestures and context. These can help to specify and determine the meaning of an infant call or, alternatively, the infant call may help determine and specify the meaning of the gestures. Non-linguistic infant communication is accomplished using various channels that work together to serve specific communicative functions or, as Gómez (2007) puts it, children communication emerges out of an intersection of different lines of communicative behaviour.

Taking all this into account, the fact that a very basic model of acoustic analysis can classify the most common classes of naturally produced infant vocalisations, independently of any contextual information and with an accuracy of 60.0%, is good evidence that non-linguistic vocal signals play a more important role in human infants than typically assumed by contributing an important share in the communicative acts for humans who are not yet competent language users.

Anecdotal evidence suggests that listeners were able to make basic inferences about contextual situations experienced by the caller. Additional information simultaneously available, such as gestures, specific behaviours, or external situations, might help to narrow the content of such broad messages, such as the cause of a protest call or the caller's momentary state in a more complex event, such as during feeding. Protest calls, for instance, were contextually very unspecific and only obtained their precise meaning within an on-going event (referring to undesirable object, such as when trying to pull off a bib or to undesirable action, such when being pulled away from a toy). The perception of general call types might help listeners decide if they require further contextual information, for instance by visually inspecting the scene.

Our results are consistent with the more general hypothesis that vocal communication of human infants contains at least two types of signals: (a) context-specific call types that serve as pragmatic tools and could inform the caregiver of the caller's needs and (b) non-referential unspecific calls, such as produced during vocal play or babbling, which may serve the child as a training ground for speech production (Owren and Goldstein 2008).

Chapter 3: Cross-cultural Comparison of Infant Calls from Scotland and Uganda

Summary

In the previous chapter, we provided evidence that 11-to-18-month-old human infants raised in Scotland produce acoustically distinct non-speech vocalisations to the following behaviourally defined categories; protesting, requesting food, requesting actions, and giving objects to a peer or caregiver. The acoustic properties of the infant calls varied significantly with the associated context and therefore could function referentially to listeners. Here, we replicated this study in a remote rural area of Western Uganda with infants growing up in non-English speaking households. We found that, despite distinct cultural differences, infants produced calls in the same key behavioural contexts, and the acoustic structure of the calls showed similar systematic variations between the contexts they were produced in. In a subsequent cross-cultural acoustic analysis, we found significant acoustic commonalities in four of the five categories of vocal behaviour, despite a slightly higher level of variation in the various call types. The results suggest that despite some cultural variation, the basic call types are largely similar and could therefore be part of a non-linguistic vocal system that consists of simple sound-meaning correspondences and can inform caregivers about the infant's mood, attitude or activities.

Introduction

Human infants possess a highly flexible ability for vocal production that enables children to learn any given language well before they begin to speak (Oller and Griebel 2008, Oller 2000, Vihman and De Boysson-Bardies 1994). The question arises whether this flexibility is restricted to speech or also evident in non-linguistic vocal productions. Some authors claim that the vast majority of infant vocal behaviour is flexible (Oller and Griebel 2008, Vihman 1996), whereas others propose that a basic distinction must be made between flexible speech-related vocal behaviour, such as babbling and vocal play, and stereotyped, inflexible non-linguistic vocalisations, such as grunts or crying (Papoušek 1992, Scheiner et al. 2002, Lester and Boukydis 1989). Whilst the former do not only vary between different cultures but also between individuals, this might not be true for the latter. Instead, these non-linguistic vocalisations might be part of some sort of species-specific signalling system with relatively stereotyped sound-meaning correspondences. Previous work has shown that the acoustic properties of cries produced by young babies vary with their production contexts, for example hunger vs. pain, but not between individuals (Lester and Boukydis 1989, Papoušek and Papoušek 1989). Similarly, infants between 4 and 8 months produce similar vocal behaviours during different stages of object interaction with an adult (D'Odorico and Franco 1991), suggesting that some forms of non-linguistic vocal behaviour are generally acoustically inflexible and thus unrelated to the vocal behaviour underlying to language acquisition.

Very little is known about the non-linguistic vocal behaviour of older infants, despite occasional acknowledgement that this still makes up a large part of the communicative repertoire until at least the second year of life (Tomasello 2008, Vihman 1996, Bruner 1983). McCune et al. (1996), for example, showed how grunts that were initially uttered as a sign of effort eventually came to function as an attention-marker around the age of 14 months, in addition to their original function. In the previous chapter, we recorded the vocal behaviour of 22 infants between the ages of 11 and 18 months in their everyday nursery environment in Scotland. We found that five categories of social behaviour (giving, declarative pointing, food requests, action requests and protests) were commonly associated with vocal behaviour and, more importantly, that the acoustic properties of vocalisations emitted in four of these situations showed significant correspondences between the acoustic make-up of the vocalisation and the contexts in which they were produced. We suggested that this vocal behaviour might be part of an early, basic, non-linguistic signalling system, tied to specific social situations, that is comparable to the vocal systems of modern non-human primates (e.g. Cheney & Seyfarth 1996; Zuberbühler 2003). Despite showing limited flexibility in their production, non-human primate calls are capable of transmitting information about the caller's affective state and/or external events experienced (Zuberbühler 2006, Cheney and Seyfarth 1996).

The idea of sound-meaning correspondences that inform a caregiver of the infant's emotional states and attitudes has been suggested by several researchers (e.g. Lester and Boukydis 1989, Dore et al. 1976, Papoušek 1992, Franco 1997), but the hypothesis has never been formally tested, especially for older infants from 11 months of age, i.e. at the onset of intentional communication (Bates et al. 1979), and alongside early language acquisition. Acoustic signals like these are most likely not arbitrary but linked to basic biological functions, such as obtaining attention or expressing anger (Owren and Rendall 2001, Marler, Evans and Hauser 1989).

If such a primitive vocal signalling system exists in human infants alongside the vocal system related to speech acquisition, and if this were related to the vocal systems of non-human primates, we would expect non-linguistic calls to show less flexibility and relatively little impact of external influences, such as cultural background or native language. One way of testing the degree of flexibility in non-linguistic sound meaning correspondences is to conduct a comparison between the sounds produced by infants growing up in two very different cultures. Therefore, in order to test whether the sound-meaning correspondences observed in Scottish infants are also present in infants from a different cultural background, we observed the vocal behaviour of infants growing up in rural Uganda in a non-English speaking environment. We analysed the data with the same methodological tools as the vocal productions of the Scottish sample from our former study, which allowed us to carry out direct comparisons.

Results will not only provide a picture of the non-linguistic vocal behaviour of children from a non-Western cultural background, but also identify how much non-linguistic vocal behaviour is influenced by the infants’ linguistic and cultural environment.

Method

Study sites and subjects

Data were collected in five villages in the Masindi District, Uganda, between September and December 2010. Participants were 28 normally developing infants between the ages of 11 and 18 months. Observations were collected in the infants' family homes and compounds. Homesteads generally consisted of several mud houses used as bedrooms, kitchen, and storage units. Few families lived in simple brick houses. Farm animals, such as chickens, goats or ducks, were housed on all compounds. Houses did not have access to electricity or running water. Immediate and extended family members, older siblings and other children were nearly always present during data collection. The infants' mothers were sometimes present but often occupied with household tasks or working in the fields. Infants rarely had one-to-one interactions with primary caregivers but were sometimes on their mother's arm, lap or back whilst she did other tasks. During the day, infants were often left at the compound as mothers went to work on the fields. Older siblings or grandparents were then in charge of the infant. Mothers returned regularly to breastfeed and rest, especially after lunchtime.

Infants were free to move around their home compound. Everyday objects, like jerry cans, plastic containers, farming tools, baskets, plates, cups and spoons, were available to the infants to explore and play with. Hardly any children had access to toys. Infants were often given aforementioned objects to play with, as well as natural objects, like sticks or stones. Older siblings sometimes practiced walking with the infant, and dancing and drumming were some of the favourite activities.

All but three children (ANG, BAG, DOR) were breastfed. Additionally, they received solid foods, such as sweet potato, cassava or fruit during family meals or as snacks. Water was offered regularly during the day, and in the mornings most children had some tea. Infants had between one and two unscheduled naps a day when the child was tired.

Recruitment of participants was through our local collaborators. Information sheets and consent forms were translated to the parents, which they filled in and returned to the research team. Participants were encouraged to ask questions. Most families supported themselves through small-scale agricultural activities, which covered their basic needs. Some fathers were in full- or part-time employment as carpenters, motorbike taxi-drivers, hairdressers, or field assistants of a local research station. The education level was generally low. Some mothers had never gone to school, most had only a few years of formal schooling, and only a minority had completed primary school education. Fathers were often somewhat better educated. Most families spoke a local dialect of Swahili as their primary language; some additionally spoke Nilotic languages, such as Alur or Acholi. Although English is the official language in Uganda, most mothers spoke little or no English. Infants occasionally heard spoken English, for example in songs or announcements in the radio, when siblings sung English songs or dod their homework, or when speakers of other languages visited their homes. In general, their exposure to English was rare, and almost never directed towards them. Overall, there were N= 17 male and N= 11 female participating infants (Table 1).

ID	AGE start	AGE end	SEX
ANG	1;5	1;6	m
BAG	1;2	1;5	m
BAK	1;1	1;4	f
BRI	1;1	1;4	m
DEO	0;11	0;11	m
DOR	1;3	1;6	f
EZR	0;11	1;1	m
FIL	0;11	1;1	m
FRE	1;2	1;5	m
GLO	0;11	1;2	f
JUL	1;1	1;2	f
LNG	0;11	1;0	m
LNY	0;11	1;1	m
MAN	1;4	1;6	m
MIL	1;2	1;3	m
NAN	1;2	1;5	f
ODO	1;4	1;6	m
OLI	0;11	1;1	f
OPI	1;4	1;6	m
PAI	0;11	1;2	m
PLY	1;2	1;5	m
PNG	1;5	1;6	m
RIT	0;11	1;1	f
SAF	1;2	1;5	f
SUN	0;11	1;0	m
TAN	1;5	1;6	f
TED	1;2	1;4	f

Table 1: Sex and age of participating infants at beginning and end of recording

Data Collection

Observation took place either in a morning or afternoon session and lasted between 45 minutes and 2.5 hours. The observer and a field assistant visited families at their home compounds and occasionally engaged with the infants in little games like 'giving and taking', or pointing to animals. Infants usually took a while to get used to the presence of a Caucasian observer, and interaction in play helped them to habituate to the observer and the recording equipment. This habituation also ensured natural behaviour during recording sessions. Communicative episodes were recorded on video using all-occurrence sampling (Altman 1974), by filming whenever episodes of vocal behaviour occurred or could be anticipated, for example when food was distributed or an interesting object was taken away from the infant.

Materials

Episodes of vocal behaviour were recorded using a Canon digital camcorder in conjunction with a Sennheiser K6P/ME64 directional microphone. Data were analysed with the 'iMovie' software package on an Apple MacBook Pro computer. Sound was extracted from the videos using the 'QuickTime' software package. Audio recordings were transferred digitally onto the computer using 'Praat 5.1.03'. Images were created using Raven pro 1.3 (www.birds.cornell.edu/raven).

Variables

Functional categories

Similar to the previous study conducted in Scotland, we initially aimed to characterise and categorise the contexts in which vocal behaviour occurred. As described in chapter 2, our general approach was to allocate calls to specific social situations, that is, the context in which they occurred, as well as the reaction provoked in listeners. One of the aims of our study is a comparison between non-linguistic vocal behaviour of the Ugandan infants to those growing up in Scotland. We therefore investigated whether the categories of vocal behaviour observed in Scotland also applied to some or the entire Ugandan sample. We did not specifically target vocal behaviour that matched the descriptions from the Scottish sample but aimed to identify and describe contexts in which the majority of Ugandan infants vocalised. These categories are not a full description of the infants' social behaviour and vocal repertoire but they were regularly observed in the majority of infants.

Call classification

On the basis of the qualitative descriptions, we identified specific behavioural and environmental markers that determined the classification of a call as well as when the call began and ended.

Excluded vocalisations

We excluded vocalisations that were fully linguistic, (e.g. words such as maj, mbuzi or mtoto and holophrases such as kuja, guapi or letta), or proto-linguistic in character (e.g. word approximations such as /'mj/ for maj or /'dodo/ for mtoto). Native speakers of the infant's languages assisted in identifying these vocal behaviours in the video material.

We excluded vocal behaviour that did not fit easily into the proposed classes. This was a small proportion of our sample, but was in line with the aims of a cultural comparison and the identification of contexts in which many infants vocalised. We excluded samples in which the infants' vocal behaviour was overridden by background noise, for example other people's voices or animal noises, as these would have produced unreliable measurements. Most of our recordings happened outside in the presence of other people and animals. It was therefore difficult to obtain pristine audio recordings. If Praat was unable to give a clear visualisation of different call parameters, these measurements were also excluded. Furthermore, we excluded laughs, cries or screams from the acoustic analysis to avoid false differences between the groups.

Acoustic analysis

We measured the acoustic structure of vocalisations recorded in the different contexts using Praat 5.1.03 (praat.com) with the following settings: pitch range 0 – 2000 Hz, spectrogram view range 0 – 25 kHz to determine the number of harmonics and 0 – 10 kHz to measure fundamental frequency. Intensity measures were taken using the program's algorithm to obtain the median intensity for the entire call and the maximum intensity was measured using the option to view a list of intensity measured for the entire call and identifying the highest value. The following spectral measurements were taken: (1) 'mean F0' = fundamental frequency across the entire call (Hz), (2) 'early F0' = fundamental frequency at call onset (Hz), (3) 'mid F0' = fundamental frequency at middle of the call (Hz), (4) 'end F0' = fundamental frequency at call offset (Hz), (5) 'max F0' = maximum fundamental frequency (Hz), (6) 'min F0' = minimum fundamental frequency (Hz), (7) median intensity (dB), (8) peak intensity (dB), highest intensity across the entire call (dB), (9) 'N harmonics' = number of visible harmonic bands, (10) 'N units' = number of consecutive call units that formed a call segment (not separated by more than 3s), and (11) 'duration' = total length of the call segment(s).

Statistical Analysis

In order to investigate whether the calls recorded in the different categories differed firstly within the Ugandan sample and, secondly, between the Ugandan and Scottish samples, we conducted a Discriminant Function Analysis (DFA). The DFA compares the vocalisations with regard to their acoustic parameters, and derives a number of functions that discriminate between calls in each category. To cross-validate the analysis, the discriminant functions derived from one dataset are applied to another dataset to investigate whether the discriminant functions can classify these data at a level above chance. Results are expressed as a percentage value of correct classifications that can then be compared to the percentage of correct classification at chance level.

The DFA consists of two steps, one basic analysis and one cross-validated analysis, that each used one separate dataset. We conducted two analyses, one with the Ugandan data only and one cross-cultural comparison between calls recorded in Scotland and Uganda.

Datasets

Two datasets were used in the DFA, one construction set and one test set. The construction set consisted of one call per individual in each category. The calls were randomly chosen from the body of raw data. The construction set was used in the basic analysis. The second dataset was the test set. This consisted of another randomly chosen call per individual in each category, in those cases where individuals made multiple contributions per call category.

Multiple contributions of individuals

If individual children contributed multiple data points to one category of vocal behaviour, we randomly picked two of these and assigned one to the construction set and one to the test set. We used the random number generator in PASW 18.0 and assigned these numbers to each call of an individual. We then picked the two highest values and assigned the first to the construction set and the second to the test set. This procedure greatly reduced the overall sample size, but ensured thatoriginal data points were used (rather than values derived through an averaging process), and avoided that one individual's contribution biases the dataset.

Basic analysis

The construction set was entered into the DFA to derive a number of discriminant functions that can classify the data based on the acoustic variables. The resulting discriminant functions are then applied to the original body of data to see how well the functions can classify the data that is whether the classifications derived through the functions match the original recording contexts. Results are expressed as a percentage value of what proportion of classifications derived from the analysis matches the original recording context.

Cross-validated analysis

In order to cross-validate the analysis, we applied the discriminant functions derived from the basic analysis to another data set, the test set. This allows us to investigate whether the discriminant functions can correctly classify data that was not used in the construction of the model. As before, results are expressed as a percentage value that expresses how well the derived classifications match the original recording context.

Analysis Uganda

To investigate whether Ugandan infants have calls, we conducted a DFA using data collected in Uganda for the construction and the test set.

Analysis cross-cultural comparison

In order to compare the acoustic structure of calls recorded in Scotland and Uganda, we conducted a DFA using a construction set derived from Ugandan data which was then applied to a test set consisting of data produced by Scottish infants, and vice versa applied a construction set derived from the Scottish data to a Ugandan test set.

Results

Qualitative Descriptions

Overall we collected 468 episodes of infant vocal behaviour, most of which in the following categories:

(a) *Food Context*

Infants produced vocal behaviour when food was given out or when another family member had desirable food, for example jak fruit, mango or biscuits. Vocal behaviour was often accompanied by moving, often running or walking, towards the person that held the food, and performing reaching gestures and grasping motions. Demands to be breastfed were usually not associated with vocal behaviour. Instead the child climbed onto the mother's lap and tugged on her shirt repeatedly. For example, the father returned with a small pack of biscuits. The infant quickly moved towards him, held on to his leg and, with the other hand reached upwards whilst vocalising. Vocal behaviour was characterised by unmodulated, high intensity call units that were sometimes repeated.

(b) *Action Requests*

Infants vocalised when they were unable to accomplish something by themselves, needed help in obtaining an object or wanted an adult to perform a certain action, e.g. pick them up or play with them. Calls were often coordinated with getting the attention of the recipient either through touching them, leading them by the hand or eye contact. Vocal behaviour was almost exclusively addressed to adults or older children. For example, the grandmother fetched a ball and showed it to the child. She then dropped it and made it bounce. The infant watched this, then gave the ball to the grandmother and vocalised, bouncing her hands on the ball. She stopped when the grandmother repeated the action. Vocal behaviour was characterised by short-mid-intensity units that were repeated quickly after one another.

(c) *Declarative Pointing*:

Infants produced vocal behaviour when they pointed to a distal target in the environment, for example a familiar person or animal. Parents often verbally encouraged pointing by directing the child's attention to a specific target and were often observed pointing for the child. For example, child and mother are sitting outside; the mother is occupied with a household task. One of their goat walks by. The child points to the goat and vocalises, the mother looks up and comments. Vocal behaviours accompanying pointing were often short, single units with a rising intonation and high intensity.

(d) *Giving*

When giving an object to a peer or adult, infants commonly produced vocal behaviour. For example, an infant was playing with object, then turned around and gave it to the adult, vocalising whilst putting it into his hand. Vocal behaviour in this category was characterised by audible single units with falling intonation.

(e) *Protests*

Infants protested when something unpleasant or undesirable was done to them. For example, an infant was sitting next to a big bowl of dried maize, running her finger through it and playing with it. The mother approached, told the infant off, picked her up and put her down somewhere else. From the moment the infant was picked up, she vocalised in protest, tried to wriggle out of her mother's arms and eventually started crying. Protest calls were long vocalisations with high intensity and little modification. They often co-occurred with crying and physical resistance.

(f) *Social Routines*

This category of vocal behaviour was observed in familiar and repeated exchanges between infant and caregiver that followed a specific pattern. One routine we often observed was a pretend hitting/crying exchange between infant and caregiver. For example, the infant would hit the caregiver and she would bury her face in her hands and pretend to cry. The infant would then hit again, expecting the same reaction. Religious songs and associated rhythmic clapping were frequently observed. During these the infant produced vocal behaviour that approximated the song's melody and also clapped. Boys were often encouraged to kick objects as if to play football. Caregivers and older sibling then shouted 'goal' and the infant often vocalised at the same time. Furthermore, infants were encouraged to greet older members of the family and adults visiting the compound. Mothers then pushed the infant towards the adult, and encouraged them to repeat their greetings. Adults then shook the infant's hand and greeted them in turn.

(g) *Movement*

Infants often vocalised when they were moving around by themselves, for example walking around the compound, playfully running with other children or dancing alone. For example, the infant was walking in big circles round the compound, vocalising as she moved. Caregivers were engaged in other activities and the infant made no attempts to obtain their attention. The vocalisations emitted during these dances were often quite melodic and could be an approximation of tunes heard during songs. During running, vocalisations were often one long, unmodulated sound, shaped by the infant's breathing rhythm. The vocal behaviour did not seem to be addressed to anyone in particular and caregiver's or peers did usually not react to them. In rare cases, an older sibling joined the infant in dancing and sang for her.

Classifying Video Episodes of Vocal Behaviour

Episodes of vocal behaviour were divided by using specific environmental and behavioural markers that helped us determine the start and end point of an episode. These were also used to classify the video clips into one of the categories described in the previous section (Table 2).

Category	Markers for Start of Call	Markers for End of Call
Food Context	-Food such as jak fruit or papaya is prepared for consumption -Other children have food that they might be willing to share - Mother is present and settling down, breastfeeding is therefore possible	-Infant received food -Infant starts breastfeeding -Food is removed from infant, for example siblings eat their share and show empty hands -Food is no longer avilable

Action Requests	- Infant needs help obtaining an object - Infant wants caregiver to manipulate an object for them, for example play a ringtone on their phone - Infant wants caregiver to do a certain action for or with them, for example drum on a jerry can	- Caregiver performed the action the infant desired and she seems satisfied - Infant stops signalling and engages in other activity
Declarative Pointing	- Pointing gesture directed at interesting object or event - Animal or familiar person walks past - Cars or trucks drive past - Caregiver verbally encourages pointing or points for the infant - Surprising or unfamiliar event happens, for example infant encounters a squeaky toy	- Caregiver comments on infant's gesture - Receivers are no longer present - Infant moves on to another activity Infant examines object
Giving	- Infant is in possession of an object and gives it to a caregiver - Infant is taking part in a giving-taking exchange - Caregiver asks infant to give or bring an object	- Caregiver receives object and infant engages in another activity - Object leaves infants hands

Protests	- Aversive action performed on infant such as washing face, putting on clothes - Object is taken away from infant - Infant is taken away from object or people - Infant is put down - Familiar person leaves or no longer engages with the infant in an activity	- Infant calms down and engages in another activity - Caregiver offers soothing, infant calms down - Infant is distracted with object or activity

Table 2: Behavioural and environmental markers that started and ended a call

Interrater reliability

In order to validate our method for classifying video episodes we asked two naïve observers to categorize 15% of our video samples of vocal behaviour using the descriptions we presented earlier (declarative pointing, food, giving, action request and protests). The raters' classification of video episodes corresponded 100% with ours. A further two naïve raters were given 50 video clips randomly selected from the entire dataset, including video episodes that were excluded from any further analysis, and asked to classify them using the criteria described above. Raters' coding matched ours in 95.12% of the cases.

Excluded vocalisations

Out of the 468 episodes of vocal behaviour we recorded, 142 samples were excluded from the further analysis after an initial assessment of the video data. Reasons for exclusion were poor sound quality of recordings caused by multiple individuals vocalising at the same time (60/142), high levels of background noise (35/142), or the individual moving out of the microphone's range (2/142). Furthermore, some samples of vocal behaviour were not readily classifiable into the proposed categories (45/142). These included instances of vocal play.

Vocal Activity in Different Categories

We firstly recorded the frequency with which we observed vocalisations in each category, while excluding words or word approximations. The most frequently observed category of calls was protests (121/326, 37.11%) observed in all but one infant, followed by action requests (67/326, 20.55%), declarative pointing (59/326, 18.09%), giving (44/326, 13.50%), and food requests (35/326, 10.74%). The individual contribution from each infant is illustrated in Table 3 and the overall distribution of data in Table 4.

ID	FRQ	ARQ	PING	GIVE	PROT
ANG	1	2	4	0	5
BAG	2	7	6	4	13
BAK	2	2	13	6	4
BRI	0	1	1	0	6
DEO	0	0	0	0	1
DOR	4	5	4	3	5
EZR	2	3	5	4	4
FIL	0	8	1	3	8
FRE	1	0	2	3	8
GLO	2	3	0	0	8
JUL	1	5	0	0	2
LNG	0	0	1	1	3
LNY	1	1	1	1	0
MAN	1	0	1	0	2
MIL	0	0	0	0	1
NAN	0	1	0	0	1
ODO	0	6	0	6	2
OLI	0	0	0	0	4
OPI	1	6	3	1	9
PAI	4	1	1	1	2
PLY	0	0	0	0	5
PNG	1	1	0	0	3
RIT	3	2	3	1	6
SAF	2	1	3	2	5
SUN	2	5	6	2	5
TAN	0	1	0	1	1
TED	5	6	4	5	8
Total	35	67	59	44	121

Table 3: Individual infants' contributions to each category of vocal behaviour (FRQ – food requests, ARQ – action requests, PING – declarative pointing, GIVE – giving, PROT – protests)

Category	Frequency	Percentage
FRQ	35	10.74
ARQ	67	20.55
PING	59	18.09
GIVE	44	13.5
PROT	121	37.11

Table 4: Frequency and percentage of each call category in the Ugandan sample

Call structures in different contexts

Five social contexts yielded large enough samples of good quality recordings suitable for acoustic analysis (food requests: n=35, action requests: n=67, declarative pointing: n=59, giving: n=44, protests; n=121; Table 4). We regressed all parameters to check for co-linearity and removed any variables with a variance inflation factor >10 (mean fundamental frequency). The remaining ten acoustic parameters were entered into the DFA.

We carried out two types of analysis, a basic analysis to derive discriminant functions and a cross validated analysis to test whether the functions can correctly classify a new set of data.

Analysis – Uganda

We compiled a construction set by randomly choosing, from the Ugandan data set, one call from each individual in each category. As we only chose one contribution per individual, this reduced our overall sample size to n= 82 in the construction set as follows: food requests n=14, action requests n=19, declarative pointing n= 12, giving n= 11 and protests n=26). Because not all infants contributed data to all categories of vocalisations the numbers of contributions in each category are not equal.

We then constructed a test dataset that was made up of data points by individuals that provided more than one contribution to each category. From these we randomly selected one call per individual in every class (n= 43; food requests n=4, action requests n=8, declarative pointing n=10, giving n=7 and protests n=14). As not all individuals contributed multiple calls per category, this dataset is smaller than the construction test.

Basic analysis Uganda

We conducted a basic analysis on the test set to derive functions that could discriminate between the five different categories of vocal behaviour. Four functions were identified, two of which explained a significant amount of the variation in the data (Wilks' Lambda = 4.14, χ^2= 213.1, p< 0.001 – Table 5).

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	5.513[a]	91.0	91.0	.920
2	.313[a]	5.2	96.2	.488
3	.153[a]	2.5	98.7	.364
4	.080[a]	1.3	100.0	.272

a. First 4 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 4	.094	192.729	40	.000
2 through 4	.612	40.011	27	.051
3 through 4	.804	17.830	16	.334
4	.926	6.262	7	.509

Table 5: Eigenvalues and Wilk's Lambda for the four discriminant functions used in the analysis

The model correctly classified 82.9% of the test set data into the five categories. The rate of successful classification was highest for action requests (100%), followed by protests (96.2%), giving (72.7%), food requests (64.3%) and declarative pointing (58.3%).

Cross-validated analysis Uganda

In the cross-validated analysis, we used the test set to investigate whether the functions derived in the basic analysis can correctly classify other data points from our sample. The model correctly classified 58.1% of the original data. Correct classification was highest for protests (96.2%), followed by food requests (75.0%), giving (57.1%, misclassified as pointing in 42.9%) and action requests (50.0%, misclassified as food requests in 50.0%). Declarative pointing displayed the lowest percentage of correct classification (40.0%, frequently misclassified as food requests (30%) or giving (20%)). Figure 1 illustrates the results of classifications.

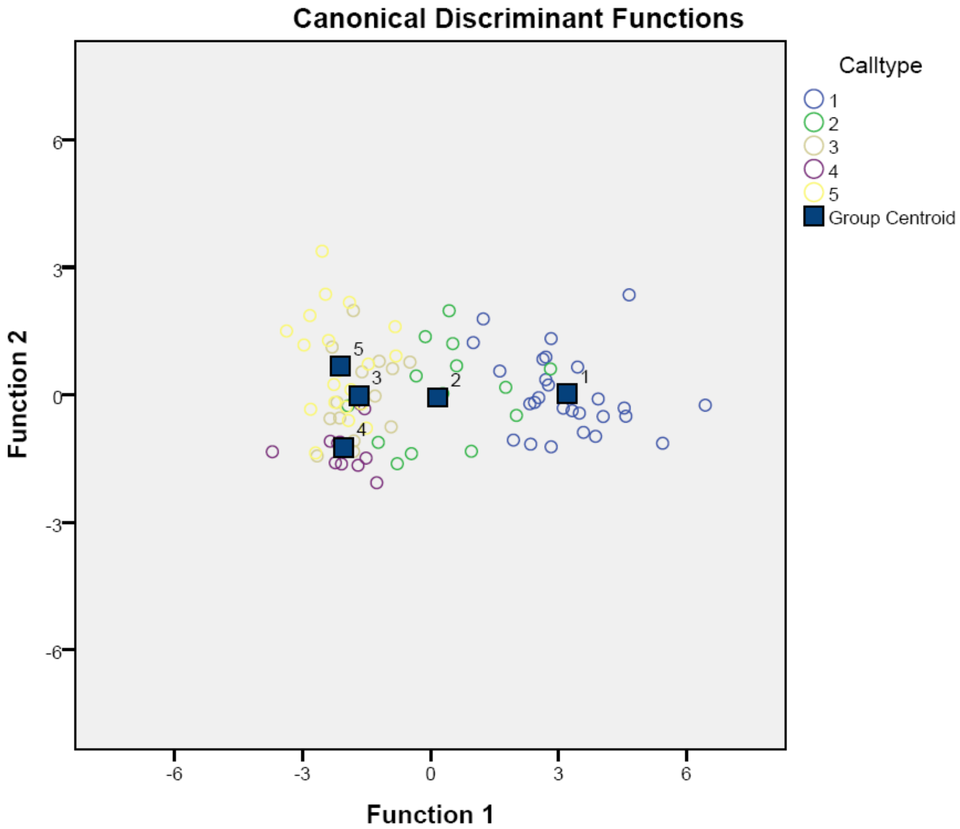


Figure 1: Distribution of calls along the two discriminant functions. 1) food requests, 2) action requests, 3) declarative pointing, 4) giving, 5) protests. Boxes mark the group centroids

Post-Hoc test Uganda

A binomial test was used to investigate whether the proportions of correct classification calculated by the cross-validated DFA were significantly higher than chance. As there were five possible categories of vocal behaviour, the level of a correct response by chance was 1/5. Results of the binomial test were significant in all five call categories of vocal behaviour ($p < 0.05$).

Cross-cultural Comparison

In order to conduct a cross-cultural comparison we used the data from Scottish infants described in the previous chapter and data collected with Ugandan infants described earlier. We firstly compared the frequencies of calls produced in the five categories of vocal behaviour (Table 6).

Call Type	Scotland		Uganda	
	Frequency	Percentage	Frequency	Percentage
Protests	79	33	128	38
Giving	27	11	42	13
Action Requests	35	15	67	20
Declarative Pointing	27	11	61	19
Food Requests	72	30	34	10

Table 6: Comparison of frequencies and proportion of calls in each category in each culture

To investigate whether infants produced acoustically similar vocal behaviour across the two cultures within the five social categories, we conducted a further DFA that applied the discriminant functions derived from the Ugandan construction set to a Scottish test set. We used a set of 82 Ugandan calls (food requests n=14, action requests n=19, declarative pointing n= 12, giving n= 11 and protests n=26) to test 58 calls from same age Scottish children (food requests: n=15; action request n= 11; declarative pointing n= 9; action requests n= 10; protests n= 16). The overall rate of correct classification was 48.3%. Action requests had the highest rate of successful classification at 72.7%. The second highest rate of correctly classified calls were food requests (60.0%). The correct rate of classification for giving was 50%; however, 20% of the cases were misclassified as declarative pointing. Calls in the category of declarative pointing were correctly classified 33.3% of the time. Another third of the cases were misclassified as giving (33.3%). The lowest rate of correct classification was observed in the protest category (18.8%), and this category also observed the highest rate of misclassification, where half of the calls were classified as food requests (50.0%).

We conducted an additional analysis, this time using the Scottish dataset as construction set and the Ugandan dataset as test set. The proportion of correct classifications was the same as with the reversed test and construction set.

The level of correct classification for the cross-validated DFA within each culture is very similar, 60% for the Scottish sample and 58.1% for the Ugandan data respectively. Results of the analysis are summarised in Table 7.

Call Category	Scotland	Uganda	Cross-Culture
Action Requests	75.0*	50.0*	72.7*
Protests	70.0*	96.2*	18.8
Declarative Pointing	16.7	40.0*	33.3*
Giving	66.7*	57.1*	50.0*
Food Requests	71.4*	75.0*	60.0*
Overall	60.0*	58.1*	48.3*

*Table 7: Percentages of correct classification from the Discriminant Function Analysis. * denotes statistically significant values.*

Post-Hoc test Cross-cultural comparison

We used a binomial test to investigate whether the rates of correct classifications obtained in the cross-validated DFA are significantly above chance. The level of obtaining a correct response by chance was 1/5. Calls in the categories of action requests, declarative pointing, giving and food requests were correctly classified at a level significantly above chance ($p < 0.05$). The rate of correct classification in the category of protests was not significant.

Discussion

Our study shows that infants growing up in rural Uganda display a variety of vocalisations whose acoustic properties change systematically with the context in which they are produced. These non-linguistic calls have also been demonstrated in Scottish infants (see previous chapter), and are possibly part of any infant's developing vocal repertoire in addition to other and better-documented signals, such as vocal play, babbling or crying. Non-linguistic vocalisations like the ones we presented here might not be directly related to language acquisition but instead might form a subsystem of vocal behaviour that could transmit information about the infant's emotional state and activities to the caregivers on the basis of systematic variations in the acoustic make-up of the vocalisations. These could be comparable to calls found in non-human primate species that have been proposed to be functionally referential and transmitting information about the listener's emotional state (Fischer, forthcoming, Winter et al. 1989, Cheney and Seyfarth 1996, Zuberbühler 2006).

The discriminant analysis compared non-linguistic vocalisation of Ugandan infants that were produced in five different contexts: giving, food requests, action requests, protests and declarative pointing. Results from the acoustic analysis suggest that calls recorded in these contexts can be discriminated on the basis of their acoustic properties alone. Although all of the five call types could be classified correctly at a level significantly above chance within each culture, there are some variations in the level of correct classification between the call types. Protests exhibited the highest rate of successful classification, suggesting that these calls have very distinct acoustic features that exhibit little variability across infants. Action requests, giving and food requests showed a slightly lower rate of successful classification, but the acoustic features still showed systematic variations between recording contexts. The lowest, yet still significant, rate of correct classification was observed in the context of giving. This suggests that there is more variability in the vocalisations recorded in this category compared to the others. The variation found in this class was, however, not random. Giving was often misclassified as declarative pointing; suggesting that vocal behaviour in these two categories was acoustically similar. A possible explanation for this similarity is that both vocalisations have similar psychological underpinnings. Bates et al. (1979) suggest that giving on behalf of the infant is a protodeclarative gesture, similar to declarative pointing in that it shows an object to a communicative partner, and the infant expects them to comment or express their attitude towards this object.

The rates of discriminability in the Ugandan sample are comparable to similar studies that found sound-meaning correspondences in chimpanzees (Crockford and Boesch 2003, Slocombe and Zuberbühler 2010) and monkeys with a graded system of vocalisations (Hammerschmidt and Fischer 1998). For example, a discriminant analysis of chimpanzee vocal behaviour recorded in six contexts showed a 56% rate of successful classification. In fact, the degree of similarity between calls produced by different infants is surprising, given that infant vocal behaviour is thought to be very flexible and therefore exhibit a large degree of individual variation – or certainly more variation than any other primate species (Oller and Griebel 2008).

Although vocalisations with relatively consistent acoustic properties might exist within each culture, they might be different between them and therefore reflect a variability that is similar to that observed in human languages. We tested whether these calls are universal across infants from both Scotland and Uganda by conducting a cross-cultural comparison. The first important thing to note is that children from both cultural backgrounds produced vocal behaviour in the same situations despite the fact that we did not explicitly look for the same contexts or even bring those about artificially. Infants in both cultures mainly produced protest calls, these made up more than a third of the observed calls in each group.

Furthermore, we observed similar frequencies in the categories of giving and action requests. The largest difference was found in the category of food requests, which made up nearly a third of the observed vocalisations in the Scottish group but only accounted for about 10% of the Ugandan sample. We think that the reason for this difference in frequency is related to cultural differences in how infants are fed.

Most of the Ugandan children were still breast-fed and requested feeding by climbing on their mother's lap and tugging at her shirt. Solid food is often dropped on the child's lap, mostly without previous requests from the infant. When someone prepares food such as fruit, the child is often given her share very quickly. Therefore, the child is hardly ever in the presence of prepared food that she is unable to obtain, a situation, which might motivate her to perform food requests. In contrast, meal times in the Scottish nurseries involved a great deal of preparation and waiting. Infants firstly had to be strapped into high chairs, and then they had to wait for their food to be prepared, for example, portioned or cooled, and wait until a caregiver had time to help them feed. Therefore, the Scottish infants were often in the presence of food that was unobtainable to them, and this could have been the motivation for their vocal behaviour.

Further differences in the frequency of observed categories of vocal behaviour were found in the domain of declarative pointing, which made up nearly 20% of the Ugandan sample and only 11% of the Scottish data. In Ugandan families we often observed that declarative pointing was part of a social game – the mother would point to an object for the child or would verbally direct the child's attention to this object, for example an animal or truck driving past, and the child would then point at the same object for her. When the child was pointing, the mother would frequently name the object the child was pointing at and also perform pointing gestures with the infant. This social game seems to be similar to the book-reading format described by Ratner and Bruner (1977) in which the child points to pictures in order to get the adult to name the target object of the pointing gesture.

Our study suggests that both Ugandan and Scottish infants have a number of discriminable non-linguistic calls whose acoustic properties vary with the context in which they are produced. The question is whether these vocalisations are actually the same and thereby part of a possibly universal infant repertoire or whether they differ between the cultures. In order to answer this question we conducted an acoustic analysis that applied the discriminant functions derived from the Ugandan dataset to the Scottish sample. Results from the acoustic analysis suggest that for four of the five call types (action requests, giving, declarative pointing and food requests), the rate of correct classification is significantly above chance, suggesting that the acoustic properties of the calls are constant in both Ugandan and Scottish infants. Protests are the only category of vocal behaviours that display random variations in their acoustic make-up. This class does, however, show a systematic misclassification – the Scottish protests are frequently classified as food calls based on the model derived from Ugandan data. Similarly, a number of calls in the class of declarative pointing, the category with the second lowest rate of correct classification, were classified as giving. We already mentioned that this tendency was also seen in both mono-cultural samples, and argued that one possible explanation could be that both belong to the same class of protodeclaratives (Bates et al. 1979). The finding that errors are not random but belong to particular categories suggests that acoustically similar calls might be observed in the different contexts and could highlight a similar underlying motive for the calls, for example attitude or mood.

Overall results support the hypothesis of species-typical calls in human infants. Despite some degree of cultural variation, calls generally exhibit a larger degree of commonalities across cultures, which suggest that they have stable, core acoustic properties that also exhibit some variations.

Low-level acoustic variations can be caused by a number of factors: differences between the senders (body size, voice pitch, gender – Snowdon 2008, Titze 1994), or differences in the recording environment (distance to caller, background noise; Wittig and Zuberbühler 2011). Although we expected these variables to be distributed randomly between each of the cultures, they could account for some of the slight differences we found. A further possible explanation is that the infant's native language is already reflected in their non-linguistic vocal productions (Vihman and De Boysson-Bardies 1994). Some prelinguistic utterances, mainly those that have been argued to be the main contributors to speech development, already display features of an infant's native language. For example the syllables in babbling sequences reflect those that are frequent in the language the infant's caregiver speaks (De Boysson-Bardies and Vihman 1991). It is conceivable that the non-linguistic vocal behaviour of human infants that we described here displays similar influences of a native language. The variation we found between the cultures could be the influence of a flexible vocal apparatus that is increasingly specialised to produce language sounds on vocalisations with relatively constant acoustic properties. Despite these possible sources of variation, our results demonstrate that enough common acoustic structure remains in four types of vocalisation for it to be classified contextually above chance, which supports the notion of at least some classes of relatively stereotyped calls, which show a small degree of cross-cultural variation.

The question that arises is whether the variations we found between the cultures are relevant to a listener in that they affect or change the perceived meaning of a call. This raises the issue of the general functional value of the vocal behaviours we documented here. If there are systematic differences in the acoustic properties of vocalisations, can listeners not only tell the calls apart but also gain information from the call alone? For example, can a mother, who is not visually attending to her infant, infer whether her infant wants a toy that she is unable to reach or is protesting because her brother took an object away from her?

Our results suggest that in each cultural sample the acoustic structure of calls did convey context-specific information that could potentially allow listeners to correctly identify the context provoking the call and that despite some variations within each call type, the underlying acoustic structures might be sufficient for a listener to gain specific information from a call even without taking into account other available information. That is, calls could work as part of a wider system that consists of gestures and contextual information in the environment to gain specific meanings, but, on their own, they might also act as primary indicators of urgency, affect, and the general nature of the event or situation, helping the caregiver decide whether it is necessary to gather further information about the infant's on-going activity.

That calls are part of a wider, multi-modal communicative system was evident in observations of caregivers' and parents' interactions with the infants: Parents seemed to use the infant's vocal behaviour to adjust their behaviour towards them, or to gauge whether they have to gather more information about the infant's activities. This is particularly obvious in verbal comments that are made to the child. Both Ugandan and Scottish caregivers have been observed to offer verbal reassurance or explanations to protesting infants, often making reference to the wider context in which the call occurred and comforting the infant, a reaction to her emotional state. For example, a Ugandan infant was drumming with a stick on a jerry can, her sister took the stick from her and she started protesting. An older sibling went over to the infant, offered reassuring comments, hugged the infant and tried to distract her by drumming on the jerry can with her hands instead. Scottish caregivers often verbally 'translated' the infant's utterances, for example when the child voices an action requests and performed grasping actions to reach pencils on a shelf, the caregiver expressed the child's request in language and added a verbal answer whilst fetching the object for the infant.

A further question is concerned with whether the cultural variations between the infants' vocalisations make a difference to a receiver. For example, can an English-speaking listener classify the Scottish calls correctly and gain information about their production context from them, but can gain only little information from the calls of Ugandan children? This question is central when we consider the nature of these calls- are they part of a universal repertoire that is the same across all or most typically-developing children or a display of flexible vocal behaviour that is influenced and adapted to a particular culture. It is, however, important to remember that despite a lower rate of discriminability between cultures, the rate of correct classification by the model was still at a level significantly above chance, indicating that the acoustic structure of the calls may contain universal context-specific components of biological origin, consistent with the idea that human infants possess natural calls like any other primate species.

In order to both find out more about how receivers perceive these calls and whether they are able to gain information from them, and investigate the relative impact of the variations and commonalities in acoustic structure across cultures, we conducted a playback study that is presented in the next chapter. Systematically testing what information listeners can gain from these calls helps to investigate what function they might play in adult-infant interaction. It can also shed further light on the referential value of these calls and whether they are part of a universal system or already under specific cultural influence.

Chapter 4: Playback experiments with non-linguistic infant calls

Summary

The previous chapters demonstrated that 11-to-18-month-old infants from two different cultures each produce nonlinguistic vocal behaviours whose acoustic properties vary with the context in which they are produced. These systematic acoustic differences could potentially be a source of information about the infant's activities or emotional state to a listener. In this chapter we want to investigate this hypothesis. To this end we designed a playback study that tested whether Scottish parents and listeners who are either experienced or inexperienced with young infants can listen to an infant call and can gain information about the call's production contexts from the acoustic information alone. Furthermore, audio samples in the playback study contained vocal behaviour produced by both, Ugandan and Scottish infants to investigate whether the higher degree of variability found in the discriminant analysis between calls collected in each culture is also reflected in the listener's ability to match the calls to their respective production context. Results firstly showed that all listeners could correctly classify the calls to their respective production context at a level above chance, regardless of their level of experience with young infants. Secondly, only parents showed a significant difference in performance with audio samples from different cultures. They correctly classified more audio samples produced by Scottish infants, who share their cultural background, than those produced by Ugandan infants. The results confirmed that listeners can indeed gain information from infant calls, and that the variation found between calls from Scotland and Uganda only affected the judgement of the most experienced listeners. Results are consistent with the hypothesis of a primitive, universal call system in infants.

Introduction

Parents readily assign meaning to their infant's prelinguistic utterances – this tendency can often be observed in 'proto-dialogues' between caregiver and infant where parents 'translate' the infant's vocalisations into linguistic utterances, comment on the sounds the infant produces, and make references to the infant's current activities, needs and emotions (Papoušek 1989, Papoušek 1992). This behaviour is commonly observed in parents from Western cultures and seems to be based on intuition (Papoušek 1992, Keller 2007).

The question arises whether parents assign meaning to these utterances based on acoustic cues contained in the infant's vocal behaviour that can form the basis of accurate inferences about the infant's emotional state, well-being, attitudes or activities. A number of studies have reported that the acoustic properties of some of infants' prelinguistic productions changed systematically with the situations in which they are produced or the functions they serve. For example, new-borns display different cry patterns when they are in pain or when they are hungry (Lester and Boukydis 1989, Papoušek 1989) and there seems to be a general distinction between the acoustic properties of vocalisations emitted in a positive or negative emotional state (Scheiner et al. 2002, Papoušek 1992). Similar consistencies between the shape of a sound and its function are also observed in older infants. For example, D'Odorico and Franco (1991) found that 4-to-8-month-olds produce acoustically similar behaviour during different phases of toy interaction. Dore et al. (1976) suggested that infants at the onset of language produced phonetically consistent forms that served functions such as the expression of affect, indicating an event or object, or requesting a caregiver's help. Our previous studies provided empirical evidence for Dore et al.'s (1976) assumptions – we found that Scottish infants between the ages of 11 and 18 months produce certain classes of vocal behaviour whose properties vary systematically with their production context (chapter 2). Our previous studies identified five contexts in which infants reliably produced vocal behaviour: Requesting food or requesting actions, protests, giving, and declarative pointing. Four of these categories displayed vocal behaviours that were readily discriminable by a simple model. Declarative pointing was often misclassified as giving, suggesting that these two categories could be classified together as indicating expressions (Bates et al. 1979). We also found these phonetically consistent classes of vocal behaviour in infants growing up in rural Uganda, suggesting that they might form a subset of vocal behaviour that displays relatively little variability across individuals and cultural backgrounds (chapter 3).

The question is whether these systematic acoustic variations in infant vocal behaviour across the aforementioned categories are meaningful to listeners, allowing them to gain information about the infant's current emotional state and activities in absence of additional contextual information. To the best of our knowledge there are only two studies that systematically investigated what kind of information listeners can gain from the vocal behaviour of prelinguistic infants. Papoušek (1989) designed a playback study that investigated how parents rate audio clips of two-month-olds' vocalisations that were associated with comfort, discomfort, joy or crying. Participants listened to an audio clip and indicated what they thought the infant's emotional states might be, using a specially designed infant state barometer. Papoušek (1989) tested different groups of participants to assess the influence of experience – parents of infants that were of the same age as those who produced the audio clips, parents of new-borns, eight-year-old children, and scientists who regularly work with this kind of data. Results suggested that sounds "effectively transmit both, discrete information pertaining to the categorical distinction between comfort and discomfort, and graded information pertaining to the relative intensity of affective arousal" (Papoušek 1989). Furthermore, experienced listeners were seemingly better at picking up finer nuances in the infant's vocalisations and were generally better at classifying the different types of vocal behaviour.

Papoušek's (1989) study demonstrated that listeners could make correct inferences about infants' emotional states from their vocal behaviour alone. This study raises the question of whether listeners can also gain other types of information from infant vocalisations and whether similar information can be gained from the vocal behaviour of infants older than three months. Goldstein and West

(1999) provide a partial answer to these questions with their investigation of the functional effect of non-cry vocalisations produced by three infants between the ages of 9 and 19 months. Fourty mothers were shown videos of an infant playing and were then played audio clips of vocalisations recorded in different circumstances. Mothers were then asked to choose one of several response options that they thought best fit what they had just seen and heard, for example 'baby wants something', 'baby is upset' or 'baby is commenting'. The authors were interested in the consensus that mothers showed when categorising the audio clips, rather than whether the choices they made actually corresponded to the context in which the sound was recorded. The authors found that mothers made differential responses based on changes in the infant's vocal behaviour and concluded that "prelinguistic infant behaviour […] contained sufficient information to guide playback mother's consensus as to communicative content" (Goldstein and West 1999) – that is, mothers agreed between themselves in their ratings of the playback clips and presumably based their judgement on the audio rather than visual stimuli, but it is unclear whether the mothers' classification matched the actual recording context of the audio material . Furthermore, Goldstein and West (1999) did not investigate whether their audio clips showed systematic variations in acoustic morphology that could have formed the basis of the mothers' judgement.

Our earlier studies found precisely these variations in the vocal behaviour of infants from two very different cultural backgrounds. Here, we wanted to introduce a playback study to test whether the acoustic variations in infant vocal behaviour could be used by a listener to gain information about the context in which the infant produced the sound. Furthermore, the cultural comparison we conducted (chapter 3) with infants from Scotland and Uganda showed both similarities and differences in the acoustic structure of certain categories of vocal behaviour. Testing a listener's ability to categorise infant calls from both cultures could help determine whether the differences we found at the acoustic morphology level affect a listener's assessment of the calls, and crucially whether listeners from other cultures and other languages could still discern the information contained in the call. Results, therefore, will be important for addressing the more general problem of whether prelinguistic calls in humans are influenced by the infant's cultural background or are part of a more structurally rigid universal repertoire.

A second relevant point is whether listeners' experience with young infants affects their ability to categorise infant calls and whether this is mediated by the infant's cultural background. Results will be relevant for the on-going debate on whether non-linguistic calls with specific acoustic properties have communicative value in that listeners can gain information from them.

Method

Participants

Overall 61 adults volunteered to participate in this study; 21 of these were parents of children older than 18 months of age, 20 reported to have experience with children up to two years of age, and 20 reported to have very little or no experience with children in this age range. Information about participants' gender and native language is supplied in Table 1, information about participants' age in Table 2. Participants were recruited through posters, word of mouth and the St Andrews School of Psychology's research participation system. Most participants were students or staff at the university. Participants' experience with toddlers was established through a self-report questionnaire where participants were asked whether they had any children and how old they were, or alternatively the open question of whether they had any experience with children under the age of two.

Group	Female	Male	Native English	Non-native/ Bilingual
Parents	17	4	21	0
Experienced	17	3	16	4
Inexperienced	16	4	12	8

Table 1: Participants' gender and native language

Group	18-20y	21-25y	26-30y	31-35y	36-40y	41+y
Parents	0	0	0	5	7	9
Experienced	7	6	3	1	0	0
Inexperienced	10	6	3	0	0	0

Table 2: Participants' age (years)

Playback Stimuli

Playback stimuli were extracted from video episodes of infant vocal behaviour that were collected as part of our previous research projects conducted in Scotland and Uganda (see chapters 2 and 3). Audio clips were extracted from video episodes of infant vocal behaviour that were recorded in the infants' natural interactions with their caregivers and environment in five different categories. A description of these categories is provided in Table 3. All infants were between 11 and 18 months of age. Infants from the Scottish nurseries were either raised with English as their first language or as one of two native languages. For the Ugandan children, Swahili, Acholi, or Alur were predominantly spoken at home, often in a mixture.

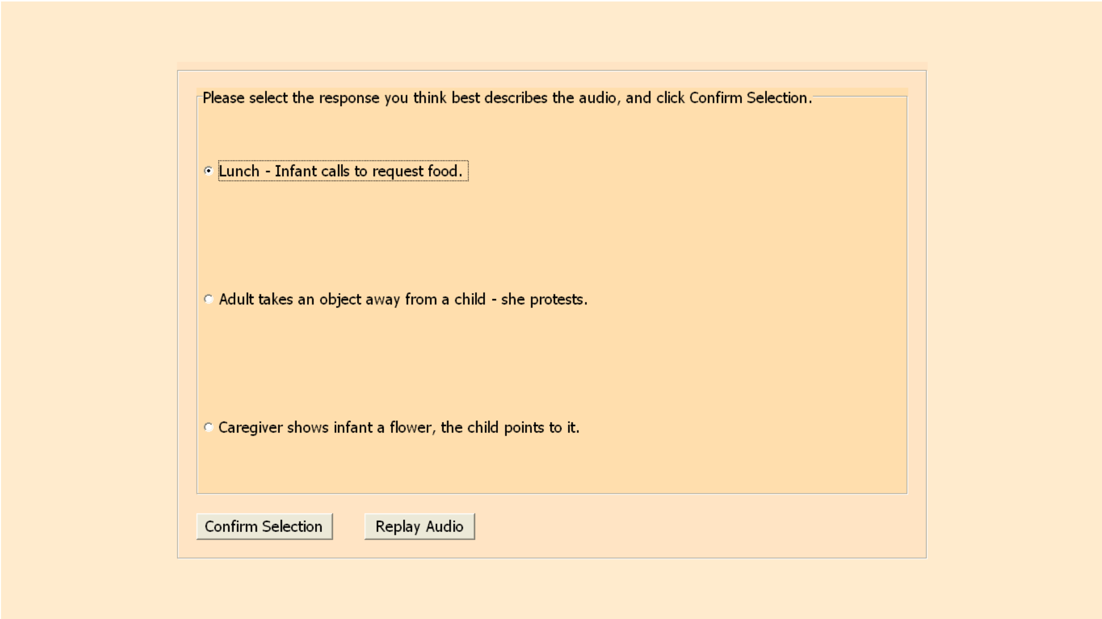
We randomly selected eight audio clips from each of the five categories, four produced by Scottish infants and the other four produced by Ugandan infants, by assignin random numbers to each of the clips in the datasets and selecting the four highest ones. Audio clips were between 2 and 15 seconds long. Using Adobe Audition, we removed any background noise from the clips that could provide additional clues to the infant's activity (for example hearing cutlery during food preparation). On some clips the stimulus amplitude was enhanced to match those of other clips and ensure that participants could hear the stimuli well. Otherwise the clips were not changed to ensure that the vocal behaviour was natural.

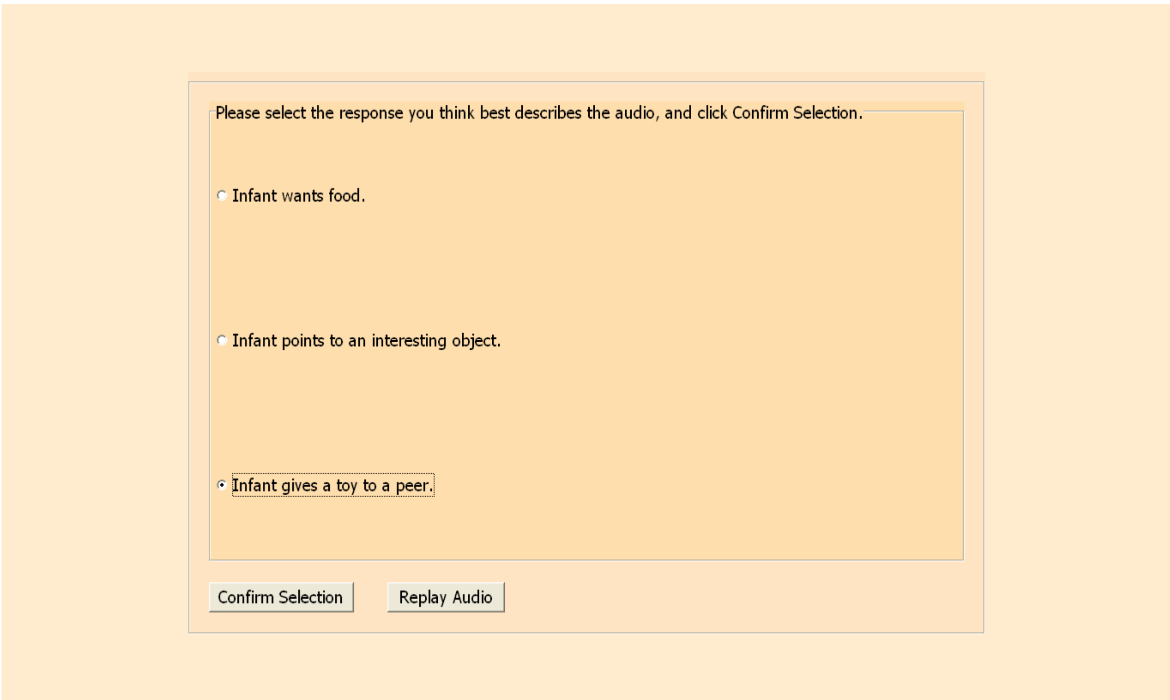
Category	**Description**
Giving	Infant gives object to peer or caregiver
Action request	Infant requests an action or object from a caregiver
Protests	Infant vocalises in reaction to an unpleasant event or action
Declarative Pointing	Infant points at an interesting object or event and vocalises
Food contexts	Infant is in the presence of food and requests some

Table 3: Categories of vocal behaviour presented in audio stimuli

Experimental Set-up

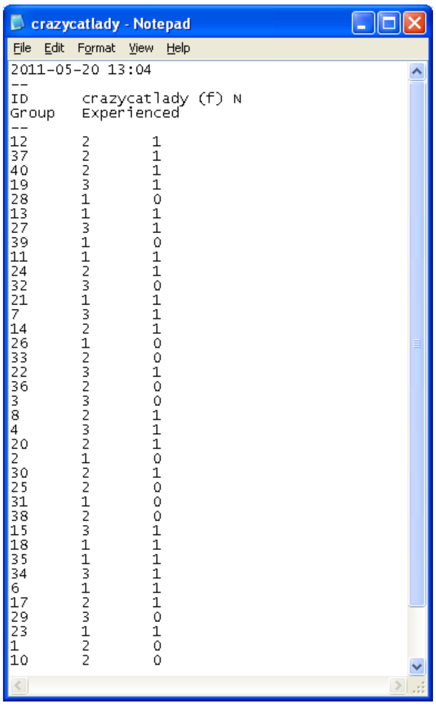
In the experiment, participants were presented with 40 different audio stimuli. In each trial, participants were asked to select one of three descriptions of infant vocal behaviour that they thought would best fit the audio clip they just heard. Each audio clip was paired with three descriptions. All descriptions were taken from the transcripts of the original video episode that contained the infant call sample. Some examples are provided in Picture 1. Each of the options offered brief descriptions of infant behaviour, without assigning meaning to the infant's vocal behaviour. In each trial, participants were first presented with an empty screen and one of the various audio clips. They were then shown the three possible response options that they could choose. Participants could replay the sound sample as often as they would like by operating a replay icon provided at the bottom of the page. Participants were unaware that the audio samples were produced by infants from two different cultural backgrounds. Any cues in the original descriptions of infant behaviour that would have revealed the infant's cultural background were removed or neutralised (e.g. descriptions of certain types of food or particular playthings). Audio clips were presented to participants in random order to avoid any effects due to presentation order. The programme recorded participants' responses and whether they were correct, i.e. whether the chosen behavioural description matched the audio clip, or incorrect. There were two practice clips presented at the start of the experiment to familiarize participants with the experimental procedure.





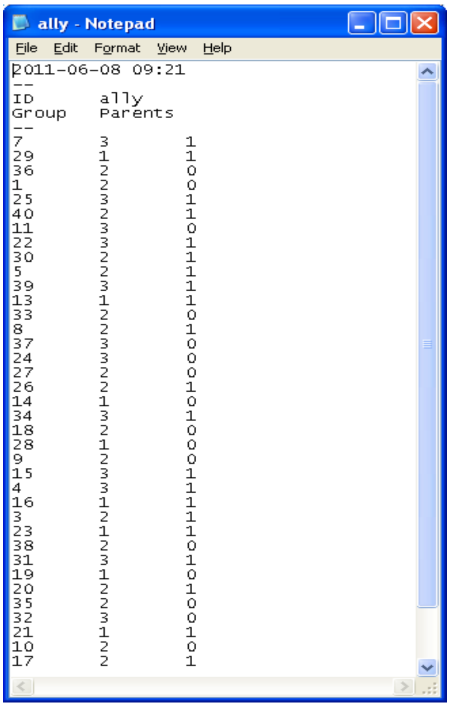
Picture 1: Screen shot of playback experiment. Initially the audio clip was presented with an empty screen, and then the answer screens appeared.

Participants' answers were recorded in a .txt file. Responses were coded on a binary scale, where 1 denoted a match between the particpant's response and the original recording context and 0 a mismatch (Picture 2). Additionally the programme recorded which of the three answers the particpants chose as a number from 1-3 (Picture 1).


crazycatlady - Notepad

File Edit Format View Help

```
2011-05-20 13:04
--
ID       crazycatlady (f) N
Group    Experienced
--
12       2       1
37       2       1
40       2       1
19       3       1
28       1       0
13       1       1
27       3       1
39       1       0
11       1       1
24       2       1
32       3       0
21       1       1
7        3       1
14       2       1
26       1       0
33       2       0
22       3       1
36       2       0
3        3       0
8        2       1
4        3       1
20       2       1
2        1       0
30       2       1
25       2       0
31       1       0
38       2       0
15       3       1
18       1       1
35       1       1
34       3       1
6        1       1
17       2       1
29       3       0
23       1       1
1        2       0
10       2       0
```


ally - Notepad

File Edit Format View Help

```
2011-06-08 09:21
--
ID       ally
Group    Parents
--
7        3       1
29       1       1
36       2       0
1        2       0
25       3       1
40       2       1
11       3       0
22       3       1
30       2       1
5        2       1
39       3       1
13       1       1
33       2       0
8        2       1
37       3       0
24       3       0
27       2       0
26       2       1
14       1       0
34       3       1
18       2       0
28       1       0
9        2       0
15       3       1
4        3       1
16       1       1
3        2       1
23       1       1
38       2       0
31       3       1
19       1       0
20       2       1
35       2       0
32       3       0
21       1       1
10       2       0
17       2       1
```

Picture 2: Example of participant's result files

Procedure

Participants who consented to take part in the study firstly completed a questionnaire collecting information about their age, gender and previous experience with young infants. All participants were from Western cultures. Participants then classified themselves as one of the three experimental groups – parents, experienced, or inexperienced. Audio stimuli and the possible answers were then presented on a computer screen, either on a desktop pc or laptop. Before starting the experiment participants received instructions on how to work the technical equipment and what the experiment required of them. Participants were then given headphones and started the study by completing two practice trials. After completing these, participants began the 40 experimental trials. There was no time limit for the completion of the study and participants could replay every audio stimulus as often as they liked. Participants were asked to confirm their choice and were then presented the next trial (Picture 1).

Statistical analysis

Each stimulus was coded as to whether the participant's answer matched the original description of the recording context of the audio clip. For each trial there was one correct response and two incorrect responses; therefore, the level to correct response by chance was 1/3. Distracters were chosen randomly from a list of descriptions of contexts in which vocal behaviour was observed. Distracters were chosen from contexts other than those in which the audio sample was recorded.

We first investigated whether participants performed better than chance when fitting their responses to the perceived audio clips. By conducting binomial tests within each participant group (inexperienced, experienced, and parents) for each of the five categories of vocal behaviour (giving, declarative pointing, action requests, food requests, and protests) we investigated whether participants' ratings of the audio clips were consistently above chance-level. We then compared the number of correct responses between each of the three participant groups by generating mean scores for each category of vocal behaviour in each participant group. In order to investigate whether there is a difference between the mean scores we conducted a one-way Analysis of Variance (ANOVA).

Our final enquiry concerned the question whether participants showed differences in their correct judgement between audio stimuli of infant calls recorded in Uganda and those recorded in Scotland. We therefore established how participants' correct scores were split with regard to the culture the audio stimulus was recorded in. We calculated mean score and standard deviation for each participant group and culture. We then conducted uncorrelated one-way ANOVAS to establish whether there was a statistically significant difference between the number of correct scores for Ugandan and Scottish sound samples.

Results

We evaluated the self-report questionnaires that participants filled in with particular regard to the group reported as having previous experience with children under the age of two. The experience participants reported was very diverse; examples include babysitting, offering sporting activities to toddlers, living with smaller siblings or half-, as well as nephews and nieces, and previous jobs in nurseries or as an au-pair. Two participants did not supply sufficient information to form a judgement on their experience with young children and were therefore excluded from the analysis, leaving a total of 18 participants in this group. No participant was removed from the other groups.

Overall Scores

We coded participants' correct and incorrect answers for each audio sample. The mean scores for correct responses for each participant group are shown alongside each category of vocal behaviour in Table 4 and Figures 1 and 2.

	Declarative Pointing	Giving	Protests	Action Requests	Food Requests
Inexperienced	4.05	4.35	6.25	4.50	3.40
Experienced	4.94	4.94	6.88	4.71	3.71
Parents	4.48	4.71	6.52	5.29	4.33

Table 4: Mean scores of correct answers by all participant groups for each category of vocal behaviour (maximum = 8; chance =2.67)

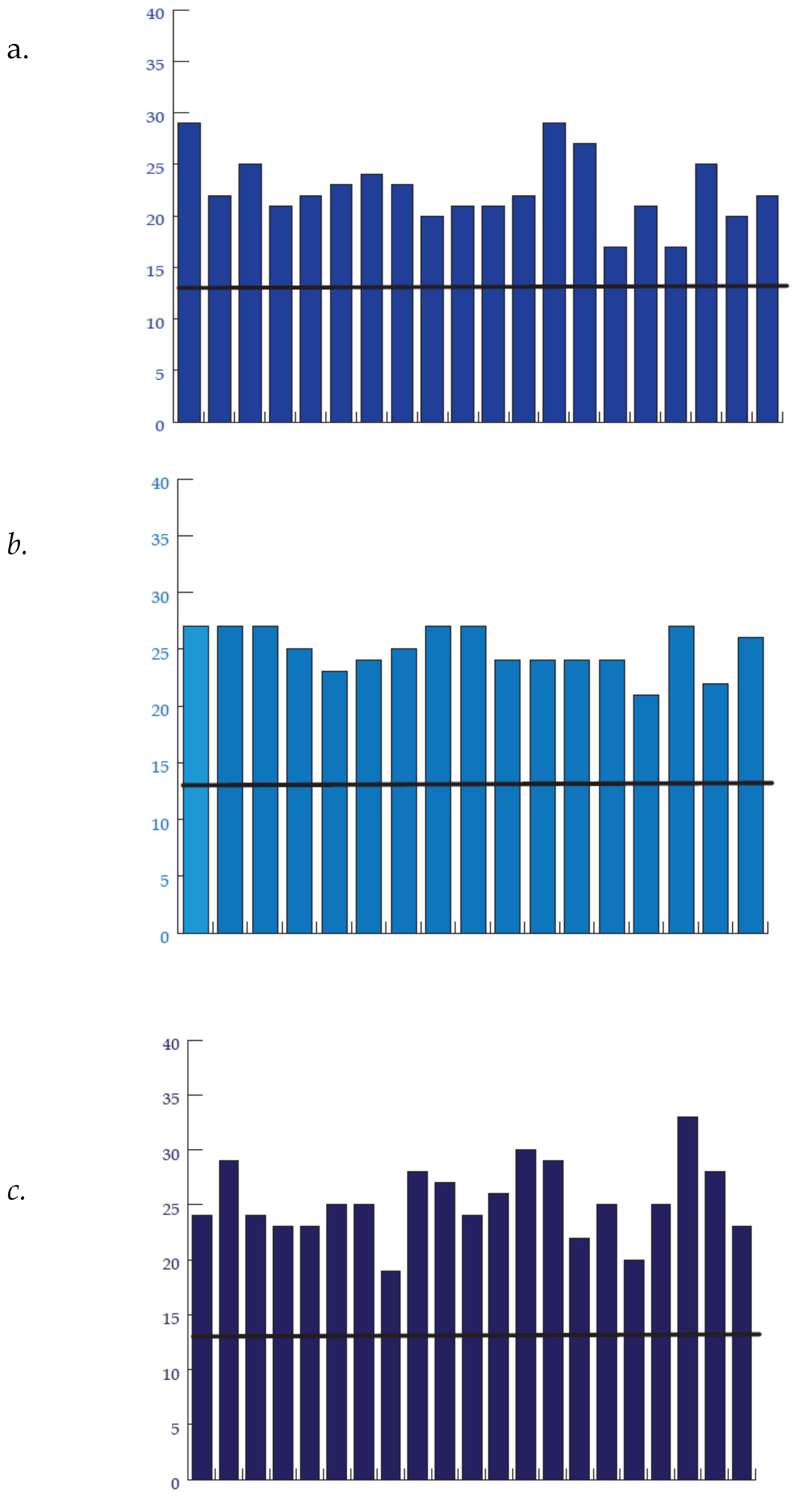


Figure 1: Overall scores of each participant group, a) inexperienced, b) experienced and c) parents. Black horizontal lines mark chance level (40/3).

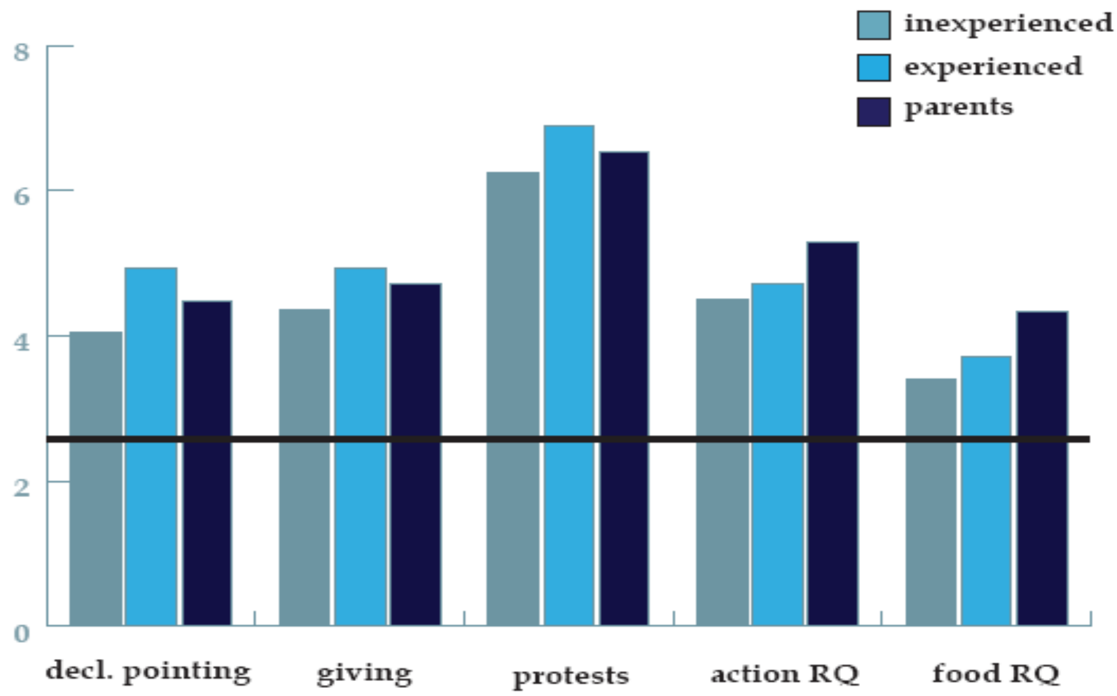


Figure 2: Raw scores of each participant group in each call category. Black horizontal line marks chance level (8/3).

In order to test whether participants picked the correct response at a level above chance (33.3%), we conducted binomial tests in PASW 18. We analysed each participant's score in each participant group for all categories of vocal behaviour. Table 5 shows the overall number of correct and incorrect scores, mean proportion of correct scores and standard deviations. Results from the binomial tests are significant across all participant groups and categories of vocal behaviour. Therefore all participants in the study chose the correct description of the audio clip at a level significantly above chance.

Participant Group	Category	Correct	Incorrect	Mean proportion	STDV	Binomial
Inexperienced	Pointing	81	79	0.51	0.50	*
	Giving	87	73	0.54	0.50	*
	Protest	125	35	0.78	0.41	*
	ARQ	90	70	0.56	0.49	*
	FRQ	92	68	0.46	0.49	*
Experienced	Pointing	84	52	0.62	0.49	*
	Giving	84	52	0.62	0.49	*
	Protest	117	19	0.86	0.35	*
	ARQ	80	56	0.59	0.49	*
	FRQ	73	63	0.46	0.50	*
Parents	Pointing	94	74	0.56	0.49	*
	Giving	99	69	0.58	0.49	*
	Protest	137	31	0.82	0.39	*
	ARQ	111	57	0.66	0.48	*
	FRQ	91	77	0.54	0.50	*

Table 5: Frequency of correct and incorrect scores, mean proportion of correct responses and standard deviations for each call category and each participant group (indicate that the binomial test is significant at the $p<0.01$ level).*

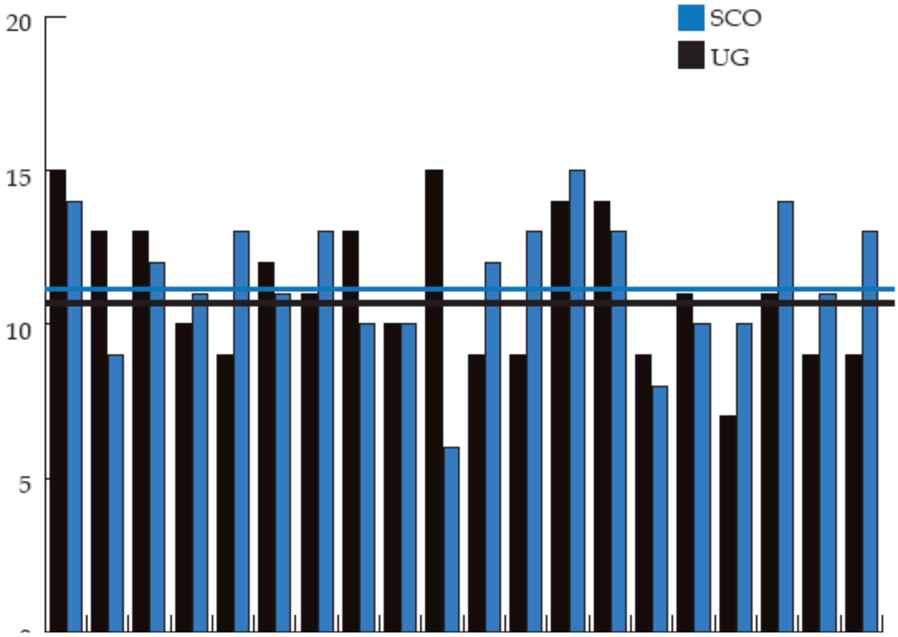
Differences between Participant Groups

In order to establish whether the mean scores of each group of participants (Figure 2) differed, we conducted a one-way ANOVA, which revealed no significant difference between the mean correct scores of each participant group ($F_{2,12}=0.452$, $p= 0.647$).

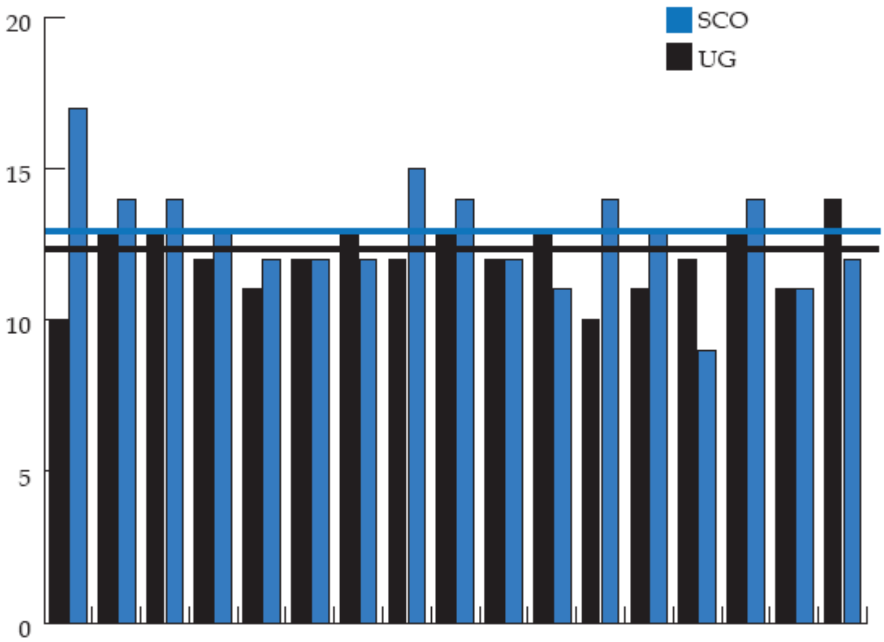
Participants' responses to stimuli from Uganda and Scotland

In order to establish whether there was a performance difference classifying audio stimuli produced by either Ugandan or Scottish infants we first calculated the proportion of correct responses participants scored for audio stimuli produced by either Scottish or Ugandan infants followed by a one-way uncorrelated ANOVAS. (Figure 3).

a.



b.



c.

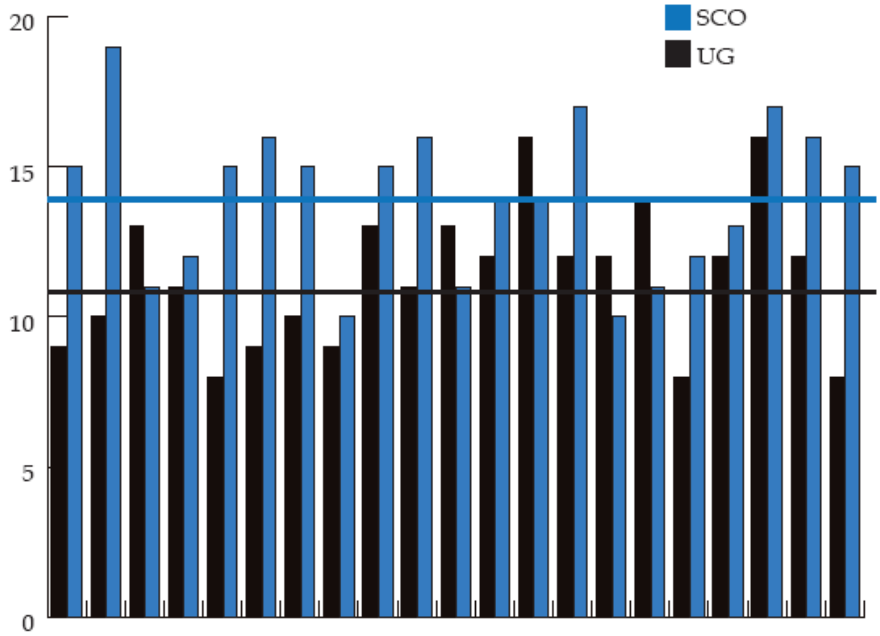


Figure 3: Proportion of correct responses for participant groups for stimuli from each culture (a. inexperienced, b. experienced, c. parents). Horizontal lines mark the mean scores of each culture for each group, blue line for calls produced by Scottish infants, black for Ugandan infants

For the inexperienced group all calls were correctly classified above chance regardless of the caller's cultural background and there was no significant difference between correct scores for the Ugandan and Scottish stimuli (F= 0.119, p=0.723). Similarly, scores for the experienced group were classified above chance in each culture group and showed a trend towards a higher proportion of correct scores for Scottish audio samples, though this trend was not statistically significant ($F_{1,32}$=2.871, p=0.100). Parents, finally, also classified calls above chance level regardless of culture but showed a significant difference between the proportions of correct scores from each culture (F= 13.022, p< 0.001). Scottish parents thus correctly classified a significantly higher proportion of calls produced by Scottish infants than those produced by Ugandan infants.

Additional Analysis: Protest Calls

Results from the statistical analysis show that participants' scores in all groups are most accurate for vocalisations recorded in protest contexts. It could therefore be the case that the results were mainly driven by the high rate of correct responses in this class of vocalisations in that participants could only distinguish protests from other calls but not make finer distinction between the other categories of vocal behaviour. For example, participants might be able to readily identify protest calls and able to distinguish which stimuli are not protests. This would mean that participants are not making fine distinctions between five categories of vocal behaviour, but that their responses are primarily driven by their ability to identify vocalisations as either protest or not protest. As a consequence this changes the probability of correct classifications in trials where descriptions of protest episodes are present as a distracter (16/32 stimuli). For example, if participants can reliable discern that a vocalisation is not a protest behaviour, this leaves them with a 50% chance of choosing the correct description of vocal behaviour from the remaining two. Therefore, the chance level of correctly matching a vocalisations and a behavioural description would be 50% rather than 33%.

In order to investigate whether participants' rate of correct responses is mainly driven by correctly identifying protest calls, we conducted two further analyses that excluded Protests from participants' responses.

The first analysis concerned only those experimental stimuli that did not contain protest vocal behaviour as either the correct match or distracter (17/40 stimuli). We conducted a Wilcoxon signed rank test to investigate whether participant's performance in matching the audio stimuli to their respective recording contexts in four different categories of vocal behaviour (giving, declarative pointing, action requests, and food requests) was significantly above chance. If participants correctly matched samples of vocal behaviour to their respective recording contexts, their performance should exceed 33% of correct responses. We entered participants' scores for all trials that did not contain a protest answer (17 overall) into the analysis, and compared the proportion of correct responses to chance performance (33% correct responses). Results suggest that participants could correctly match vocalisations to their respective production contexts at a level significantly above chance (for the inexperienced group: z= 3.93 p< 0.01, experienced group: z= 3.75, parents: p< 0.01, z= 4.02, p< 0.01). Therefore, participants can correctly classify four categories of vocal behaviour even when protest vocalisations are not included in the analysis.

The second analysis was conducted on experimental stimuli in which descriptions of descriptions of protests served as a distracter. In order to exclude the possibility that participants correctly matched audio samples and descriptions in these trials because they knew that the vocalisation was not recorded in the contexts or protests, we conducted an additional Wilcoxon signed ranks test. For this analysis we calculated the proportion of participant's correct and incorrect responses in trials in which protests were present as a distracter (15 overall), and changed the level of classifying vocalisations correctly by chance from 33% with three possible options to 50% between two possible, non-protest options. We excluded any trials in which particpants incorrectly chose the protest option. Participant's performance was then expressed as a proportion of correct responses in those trials in which participants chose one of the two non-protest answers. With exclusion of the protest answer, we wanted to investigate whether participants still correctly classified the audio stimuli at a level significantly above chance, where this level is defined as 50%. Results from the analysis indicate that all three participant groups correctly match samples of vocalisations to the contexts in which they were recorded at a level significantly above chance even when protests are excluded as a possible answer (for the inexperienced group: z= 3.26, p<0.01, for the experienced group: z= 3.47, p<0.01, and for parents: z= 3,67, p<0.01).

Discussion

It has long been assumed, by parents and researchers alike, that infant vocalisations contain acoustic properties that provide information about the infant's wellbeing and activities (Goldstein and West 1999, Oller 2000, Bates et al. 1979, Golinkoff 1986, Papoušek 1992). In previous studies we found empirical evidence for this in several different categories of vocal behaviour (action requests, protests, food requests, declarative pointing, and giving) produced by infants between the ages of 11 and 18 months in Scotland and Uganda. The playback study here aimed to investigate whether these systematic differences in the acoustic make-up of calls recorded under different circumstances could be used by listeners to gain information about the contexts in which the vocalisation was produced. To this end, we presented 40 audio stimuli produced by Scottish and Ugandan infants to three groups of Western participants, parents and adults with or without experience in interacting with infants, and asked whether they could use the auditory information contained in a call to match it to its respective production context.

The first notable finding was that all participant groups performed above chance level when asked to match audio samples of infant vocal behaviour in the five categories to descriptions of a situation in which vocal behaviour was observed. Therefore, the vocal behaviour produced by the infant contained valuable information that was a sufficient basis for a listener to make accurate judgements about the infant's general activities, attitudes and goals. Furthermore, these scores were not fundamentally influenced by the participants' previous experience with young infants. Although parents and experienced participants exhibited a slightly higher proportion of correct classification for the overall audio stimuli and descriptions, this was not statistically significant. This suggests that correctly identifying infant vocal behaviour is not necessarily a result of experience or practice, but could instead be based on a universal perceptual mechanism that associates certain sounds with a certain emotional state or function. Similarly, the acoustic properties of infant calls could serve an adaptive function by ensuring that a listener will pay attention to the infant, even though he is unable to see her (for initial research into this topic see Chang and Thompson 2011). Additional statistical analysis suggests that the effects we reported here are not driven by one or two categories of vocal behaviour, but that participants can match all five contexts above chance level.

Participants did not attain perfect scores when they matched audio samples to contextual descriptions. This does not distract from the potential referential value of these calls, for there are a number of additional clues that listeners can use to make inferences about the function or origin of infant vocal behaviour in everyday situations. These additional cues could be provided by the environment, the circumstances and activities the child is engaged in, and by gestures that co-occur with vocal behaviour and could specify the meaning of a call (Franco and Butterworth 1996, Bruner 1983). For example, when hearing a loud call that could indicate either a protest or a request for food, seeing the infant in a high chair while reaching out to a bowl of food and performing grasping motions, would immediately dissolve any ambiguities. Calls can therefore be treated as containing broad categories of information and gain more specific meaning through additional communicative signals or contextual cues.

Whilst some calls might heavily rely on additional information, some calls seem to be more readily understood than others. All participants scored highest for calls given in protest situations – suggesting that acoustic properties of these call-types are easily distinguishable from others. All other classes displayed a similar rate of correct classification, suggesting that these calls are maybe more ambiguous than protests but nevertheless can be reliably matched to a context at a rate significantly above chance.

In order to establish whether the rate of correct classification observed in participants' scores was mainly driven by the ability to correctly identify protests, we conducted further statistical tests. Results indicate that even when protest calls are excluded from the analysis, all three groups of participants are still able to correctly classify the four remaining categories of vocal behaviour at a level significantly above chance. This suggests that when matching the audio samples to the description, participants might make fine distinctions between each of the contexts and the actions associated with them, rather than just a general distinction between vocalisations associated with positive or negative emotional states, as previous research suggests (Papousek 1989, Scheiner et al 2002).

In our previous studies we compared the acoustic structure of infant calls recorded in either Uganda or Scotland in the same contexts to determine whether there are commonalities in the acoustic structure of the calls. The analysis revealed that functions derived from the data collected in Uganda could successfully classify calls produced by Scottish infants. And despite some cultural variations, calls were still correctly classified at a level above chance. This low level of variation between the two cultures identified in the cross-cultural Discriminant Function Analysis also seems to be reflected in the results of the present study. Members of all three participant groups showed slightly higher mean scores for the Scottish samples with the only significant difference seen in the parents' scores. This suggests that experience can possibly sensitize a listener to subtle differences between cultures that less experienced participant groups overlook.

The fact that participants can match audio stimuli to contextual descriptions at a level above chance, regardless of the caller's cultural background, is further evidence suggesting that part of a human infant's vocal repertoire might consist of vocalisations exhibiting systematic variations between acoustic form and content, such that this content can be perceived by naïve and experienced listeners. In addition, it seems that acoustic variations between calls recorded in either culture have little or no effect on a receiver's judgement. This is particularly interesting as it has been found that infants from the age of six months onwards already show influences of their native language in other manifestations of their vocal production, most notably in babbling (Boysson-Bardies and Vihman 1991, Vihman 1996, Mampe et al. 2009). Our study seems to suggest that these variations might well exist, but they have little influence on the function of infant calls in the categories we presented here. Furthermore, it is possible that sounds that are related to language acquisition, such as those studied by Vihman and De Boysson-Bardies (1994) or Oller and Griebel (2008) show more acoustic variation than those we hypothesized as belonging to a less flexible primitive vocal system, at least in relation to the core features of the vocalisations that transmit their broad meaning.

Significant differences between the proportions of correct scores between each culture are only found in parents. Parents are seemingly better at classifying audio samples that have been produced by children who are from the same cultural background as themselves and their own children. It is possible that parents have the most practice with these kinds of vocal behaviour through interaction with their own children. This could explain this sensitivity to subtle cultural variation in acoustic make-ups of the calls that is not present to the same degree in the other participant groups. Additionally, it is reasonable to assume that if parents had more exposure to these calls, so their better performance could be due to practice effects that these apply to infants of one culture – the one in which their own children are raised. In order to confirm this hypothesis, it is necessary to also test Ugandan parents and investigate whether they exhibit higher scores with audio samples produced by children from their native culture, the topic of forthcoming research.

Previous playback experiments (Papoušek 1989) found that parents can use infant vocal behaviour to decide whether their infant is in a positive or negative emotional state when the infant is younger than three months. Here, we show that parents and people with relatively less experience with infants can draw inferences about the infant's activities and communicative goals, based solely on acoustic information contained in non-linguistic vocalisations. Therefore, it seems that calls transmit more information than simply the infant's affective state. Alternatively, it is possible that the scale of affective information contained in prelinguistic vocalisations is, firstly, more complex than simple references to positive and negative states of arousal and, secondly, also provide listeners with clues about the infant's behavioural activities. For example, declarative pointing is likely to be coupled with a certain degree of excitement about discovering and wanting to share attention to an interesting object or event in the environment, but the gesture and call could also contain information about the referent of the gesture (Tomasello 2008). Acoustic variations in the makeup of the call could be correlated to different emotional states that are in turn related to the infant's activities, for example anger with situations in which the infant experiences something she does not like or excitement with pointing gestures.

Our study showed that the systematic variation in the acoustic properties of non-linguistic infant calls from two cultural backgrounds can be used by experienced and inexperienced listeners alike to make accurate judgements about the contexts in which they were produced. Cross-cultural variation found in the calls and documented in the previous chapter did not seem to greatly influence the listener's ability to classify the calls. Parents alone showed a significantly better rate of correct classification for calls produced by infants from their own culture. It is possible that the calls we presented here do incorporate some flexible or idiosyncratic elements beyond the core features; these elements could further specify a call's meaning – but always in addition to a fixed, pre-existing core present in infant across cultures and that can be understood by listeners regardless of their experience.

Chapter 5: General Discussion

In this thesis we wanted to address the question of whether human infants at the onset of speech production produce non-linguistic calls in functional ways - that is, whether they produce vocal behaviour whose acoustic properties vary systematically with the context in which they are produced and are meaningful to a receiver. With the results of our empirical studies, we now to offer a preliminary answer to this question and explore whether context-specific calls exist in 11 – 18 month old infants, what functions they serve, their similarity to the vocal behaviour of other primates and, lastly, how they fit into and contribute to our current view of infant vocal behaviour. Prior to answering these questions, we begin with a critical evaluation of our chosen methodology.

Methodological Comments

The human infant as 'Unknown Primate'

The methodology we applied here to study infant vocal behaviour is directly comparable to studies investigating the vocal behaviour of non-human primate species (for example Clay and Zuberbühler 2008, Fischer and Hammerschmidt 1998, Zuberbühler et al. 1997). The question is whether this methodology is appropriate, particularly as general opinion considers the vocal behaviour of humans to be different than that of other primate species in many important respects (Oller and Griebel 2008).

One criticism of this approach may be that it suffers from oversimplification. It has been argued that the human vocal system differs anatomically from other primates and animals in general (Fitch 2000, Liebermann 1991). Human vocal capabilities include producing diverse sounds, vocal learning, and readily combining different sound elements according to a defined set of rules. The human infant already displays precursors of these skills in their non-linguistic production.

With such complexity evident in the vocalisations of the prelinguistic infant, is it suitable to apply a methodology developed for simpler types of vocal signals? Or are we, by doing so, forcing infant vocal behaviour into categories that are not relevant for development because of eventual engagement in linguistic communication?

We think this criticism misses a key point. The acoustic analyses we present here are somewhat simpler than those commonly applied in studies on infant phonology (for example, Warlaumont et al. 2010 use neural networks that make distinguish between sound classes on the basis of over 100 acoustic variables). We were precisely interested in whether the acoustic variables successfully used to analyse non-human primate vocalisations can be appropriately applied when studying infant vocalisations.

Our data were collected under field conditions. This means we worked in the infant's everyday environment to ensure realistic behaviour from infants and caregivers. This ecological validity did, however, come at the cost of reduced quality of sound recordings. The nursery and family environments were noisy places, and because of this large amounts of data were not suitable for acoustic analysis. But despite minor setbacks we still obtained a suitably large dataset for acoustic comparison that offers a realistic picture of how infant vocalisations work in their everyday environment.

It is unlikely that every sound an infant produces is a precursor to language development. That some sounds do not directly relate to language learning is widely accepted in the case of crying or vegetative sounds, but might include other sounds as well. Treating the infant as an unknown primate species allowed us an 'uncontaminated view' (Gómez 2007) on early communication that opened opportunities for comparison with both human language and animal communication. This approach provides us with a realistic view of how infant communication works in the first two years of life, before speech becomes available. Methodologically comparable studies of different species' vocal behaviour validates a comparative approach by viewing the human infant as having her own communicative system, therefore clarifying whether human development can be explained more broadly in terms of general trends in vocal development.

The infants in our sample were aged between 11 and 18 months, an age range commonly associated with the onset of intentional communication and the transition from perlocutionary to illocutionary communication that we described in the introduction (Bates et al. 1979, Bruner 1983). During this age range, the infant also accomplishes major steps in vocal development – most children produce their first words, and some their first sentence, thereby showing a general grasp of linguistic communication. Despite entering the initial stages of language learning, infant communication remains primarily non-linguistic. Vocal behaviour during this age range is therefore very interesting for the infant shows growing communicative competence in gestures (Bates et al. 1979, Tomasello 1999), and begins to engage in linguistic communication. By asking whether simpler, more primitive vocal behaviours are still present in the infants' vocal behaviour during this stage, we reflect upon possible continuities in the infant's phylogenetic history.

We now would like to turn towards an evaluation of the results of our empirical studies and how they answer the question of whether infants have calls. To do so, we examined to what extent this notion is supported by our empirical work, what functions these calls serve, whether they are comparable to calls documented in non-human primates and lastly, how our results contribute to the current literature.

Do Infants Have Calls?

As seen in our data, it seems infants do have calls. That is, infants possess at least some non-linguistic vocalisations whose acoustic properties vary systematically with the contexts in which they are produced and contain enough information for listeners to make accurate judgements about the contexts in which they were produced. Our first two studies observed the natural vocal behaviour of 50 children in their everyday interaction in either their nursery environment in Scotland or their homes in Uganda. These studies identified a number of contexts in which infants from both cultures reliably produced vocal behaviour, for example when protesting against an aversive action or pointing to an interesting event. We specifically identified five situations (protesting, requesting an action or an object, requesting food, declarative pointing, and giving an object to another person) that are associated with vocal behaviour that can be distinguished on the basis of a few, simple acoustic parameters. Moreover, these distinctions were not only apparent in the statistical analysis of the acoustic structure that we conducted, but could also be made by human listeners, the typical recipients of these infant calls. In our playback study, participants from a Western background could classify calls they heard into their respective production contexts at a rate significantly above chance – and at a rate even better than the acoustic analysis. Results from this study showed that listeners could gain information about the production context from infant calls even in the absence of additional gestural signals or direct contextual information and that these calls could therefore serve a referential function.

These results support what a few researchers had previously only assumed (Dore et al. 1975, D'Odorico and Franco 1991, Halliday 1975): that human infants produce non-linguistic vocal behaviour that systematically varies with its production contexts and reliably conveys information to the listener.

Is the rate of successful classification high enough?

Results showed that the cross-validated rate of correct classification by the discriminant analyses was around 60% in both the Scottish and the Ugandan monocultural analysis, and is therefore far from a perfect classification score of 100%. This suggested that the calls, despite apparent commonalities in their acoustic structure, retain some limited variability, which might reflect the infant's otherwise more flexible vocal behaviour. Nevertheless, calls might have an acoustically stable core that contains broad functional messages that can be picked up by the receiver and also allow some flexibility in other acoustic parameters. Furthermore, calls still have the potential to transmit more specific information as they are typically interpreted in light of other contextual information. The rates of correct classification of the discriminant functions in the analysis of the Scottish and Ugandan sample were, however, comparable to those obtained in similar studies with apes and monkeys who have graded vocal repertoires.

For example, Crockford and Boesch (2003) analysed chimpanzee barks and co-occurring vocalisations in six different contexts (neighbouring group present, travel, aggression, contact, hunt, and presence of a snake). Results from a cross-validated DFA showed a correct classification of 56%, with the contexts of hunting and presence of a snake showing a correct rate of classification of over 80%. Other examples with chimpanzees that used a similar form of analysis showed 79% of correct classifications for two different food calls (Slocombe and Zuberbühler 2005) and 76% for victim's screams under different level of severity in aggressiveness (Slocombe and Zuberbühler 2007). Hammerschmidt and Fischer (1998) analysed the vocal repertoire of Barbary macaques and identified vocal behaviour that varied between eight different contexts. The subsequent DFA correctly classified vocal behaviour in five of these contexts at a level between 30% and 60%, and two at a level above 60%. Vervet monkey alarm calls to different predators show a successful classification rate between 75 and 95%, suggesting that there is very little variability in these calls between individuals (Cheney and Seyfarth 1980).

The slightly higher classification rates in some studies with non-human primates, compared to the results obtained in our analyses, could be due to the comparison of fewer contexts and a less conservative test to cross-validate the DFA (leave-one-out method rather than model and test set that we used). This suggests that the variation observed in infant vocalisations is similar to that of other primates who have a system of graded vocalisations such as chimpanzees, bonobos or Barbary macaques, as opposed to more fixed calls observed in many monkey species such as vervet or Diana monkeys.

Overall, authors of the aforementioned studies on non-human primates emphasize that additional information available to a listener is required to assign specific meaning to a call, therefore suggesting that even if that supplementary information from other sources can act to clarify the meaning of a call that shows some acoustic variability. For example, chimpanzees bark in the context of travelling, hunting, or aggression, but these are associated with additional contextual cues such as presence of other individuals or additional vocal and gestural signals (Crockford and Boesch 2003, Slocombe and Zuberbühler 2010). We and other infant researchers (Bruner 1983, Golinkoff 1986, Blake 2000) agree that contextual cues and additional communicative signals of either modality help in deciphering the meaning of human infants' vocal signals.

Differences between calls

Results from the Ugandan and Scottish monocultural analyses suggested that the calls we recorded in four (SCO) or five (UG) behaviourally and environmentally defined contexts can be distinguished at a level above chance, but the correct classification rates differ between the categories. In the Ugandan and Scottish infants, action requests, protests, and food requests showed the highest rate of successful classification (70.0 – 96.2%) by the discriminant functions in the monocultural analysis, suggesting that these calls possess rather unique acoustic categories that distinguish them from others.

Misclassifications within the analysis seemed to be systematic rather than random. In the Scottish and Ugandan sample, calls emitted in the contexts of declarative pointing and giving were often been misclassified with each other, suggesting that they are acoustically similar.

Bates et al. (1979) documented correlations between the behaviours of showing, giving and pointing. The emergence of one of these behaviours reliably predicted the emergence of the others. This suggests that there are underlying commonalities in the motivation and functions of these gestures. Bates et al. (1979) further elaborated on this by suggesting that all three are part of a broader category of indicating expression, uttered by an infant to direct a listener's attention to an event in the environment (pointing) or object they encountered (giving). In both cases, the infant might want to share their interest in an object and event with the expectation that the caregiver comments upon their interest. Though Bates et al. (1979) would argue that declarative pointing takes over from giving as the sole gesture of protodeclarative expressions, with giving being a transitory phase, we would instead argue that both gestures co-exist beyond the onset of pointing and share similar functions.

The systematic misclassification of calls in the category of declarative pointing and giving supports Bates et al.'s (1979) notion that both are expressions of the same protodeclarative function or at least have a common underlying motive or emotional correlate. It does, however, also indicate that pointing and giving are present in the infant's repertoire at the same time and serve similar functions. Interestingly, we often observed that caregivers in both cultures reacted similarly to giving or declarative pointing. They often named objects or events, commented on them or expressed excitement. Our findings support Bates et al.'s (1979) notion of a broader class of indicating expressions that consist of different gestures, for example giving or pointing, that are used in tandem with acoustically similar vocal behaviour to attract a partner's visual attention and express the infant's attitude towards the object or event they encountered. Alternatively, a proportion of vocalisations in the categories of giving and declaratives were classified correctly, suggesting that it is maybe necessary to have more nuanced classification that make more fine-grained distinctions, for example giving in order to show something to an adult.

Results from the playback study show, on average, a higher successful rate of classification than the rates obtained in the mono- and cross-cultural analyses. This finding suggests that some of the variability we observed in the quantitative analyses does not seem to affect the perceiver's judgement of the calls and might instead reflect acoustic variation that is irrelevant to the listener when extracting information from a call. For example, in the Scottish analysis, the DFA found a correct classification of vocalisations produced during declarative pointing at a level of 16.7% and a level of 40% in the Ugandan sample. Yet in the playback study, all participant groups correctly classified vocalisations in this group at a level above 50%. This suggests that although the acoustic features of vocalisations produced in the context of declarative pointing are quite variable, they can nevertheless reliably be recognized by receivers. Protests obtained the highest rate of classification in the results from the monocultural DFAs (SCO – 70%, UG – 96.2%). This was reflected in participants' correct rate of classification in this category ranging from 78.1 – 86.0%. As the cross-cultural DFA did not show a significant rate of correct classification for protest calls, participants' ratings show information can reliably be extracted from the call despite a large degree of acoustic variation between the cultures.

It is important to note that all categories of vocal behaviour were correctly classified above chance level by all participants of the playback study. This strongly suggests that some of the variation observed in the call types and identified by the DFA did not affect how the call was perceived or the information it contained, emphasizing the acoustic commonalities that are present in the calls, and exhibiting strong links to specific functions or meaning perceived by listeners. Based on these data, we speculate that there are core acoustic features in a call that can be reliably identified by a listener despite some acoustic variation in the call.

These core acoustic characteristics, present in calls produced by infants from two very different cultures, and the finding that participants could gain information from them are strong arguments in favour of describing the vocalisations we documented here as calls.

Cultural Influences

One source of acoustic variability is related to the infant's cultural background. When we combined the data collected in Scotland and Uganda in a cross-cultural analysis, the rate of correct classification by the discriminant functions decreased – suggesting that infants growing up in the different cultures produced vocal behaviour with slightly different acoustic properties. Obviously, the differences in acoustic make-up of the calls could be explained as a reflection of the infant's native language, similar to that observed in babbling (De Boysson-Bardies and Vihman 1991). The contribution of native language, although significant (because it lowers the rate of correct classification) must be relatively small, as the cross-cultural variability was a mere 10% difference as compared to the monocultural samples. Even if all of this variability were attributable to the native language of infants, ignoring other possible factors such as differences in background noise or body size, this would still be a negligible part of the vocal production: Listeners from other cultures were still capable of identifying the broader context of the vocalisation at a rate significantly above chance. This suggests that we might indeed be dealing with a basic, more primitive system of calls in human infants that is largely unaffected by cultural differences in vocal behaviour.

Indeed the commonalities between calls in the different groups are greater than their differences, with the notable exception of declarative pointing, where the variability was so great that the rate of correct classification in the Scottish sample was random. A possible explanation would be that declarative pointing is associated with different underlying motivations. For example the infant might want to obtain the listener's attention and direct it to their gesture. Alternatively, vocal behaviour during the pointing gesture could be an expression of the infant's attitude towards the target of the gesture, for example happiness, excitement or even fear. Nevertheless, even if some classes of infant vocalisations do reflect the flexibility advocated by many researchers (Oller and Griebel 2008, Vihman 1996), at least four of them do not and can properly considered to be calls. These findings go against the common line of opinion in infant vocal research (represented by Oller and

Griebel 2008, Vihman 1996), advocating the absence of any fixed or stereotyped vocalisations with the possible exception of crying in humans. It is interesting to not that in the Scottish the sample, the category of vocal behaviour that showed the largest variability was declarative pointing, as declaratives are a class of communicative behaviour not observed in other primates.

One source of variability, the cultural background of the caller only impacted on the classification on one of the participants groups, the parents. For the other two groups of listeners, it did not make a difference, suggesting that the differences in the acoustic make-up of the infant calls are rather fine grained and can maybe only be identified by those who have had a lot of practice with infant vocal behaviour. We did, however, not explicitly ask participants to make judgements about the culture in which the call was recorded to test whether they can explicitly make this distinction.

Some of the calls we presented here, namely food requests, action requests, giving and declarative pointing are comparatively invariable across cultures with regard to their acoustic make-up. When considering results from the playback study, it seems that all call categories show little variation within and across cultures.

A good comparison to the calls could be made with the pointing gesture that seems to be universal in humans. Although humans in many cultures point with their index finger from about 10 months onwards, there seem to be cultural differences with regard to accompanying joint attention behaviour (Kaller and Slocombe, forthcoming, Callaghan 2010), the conditions that elicit pointing (Callaghan et al. 2010) and, as a later development, even the shape of the gesture, for example lip pointing or using other fingers (Kita 2003). Despite these relatively secondary variations, the fundamental aspects of the pointing gesture remain the same and, although to the best of our knowledge, the appropriate empirical studies remain to be done, all these forms of pointing would be understood cross-culturally most of the time.

Summary – Do infants have calls?

Our findings provided empirical evidence that human infants produce non-linguistic calls, vocalisations whose acoustic properties differ systematically with their production contexts in four situations. These calls have been identified in infants from two very different cultural backgrounds and an acoustic analysis showed very similar rates of successfully classifying calls based on their acoustic properties. When compared cross-culturally, this rate decreased slightly, suggesting additional variability caused by cultural differences in the acoustic make-up of vocalisation. Despite this increased variability, four calls in four categories could still be classified at a level significantly above chance. The suggestion that commonalities outweigh the variability found in the cross-cultural analysis is confirmed by the results of the playback study. All participants could match calls to their production contexts at a level above chance and only parents from the sample showed a significantly better performance for the calls produced by Scottish infants over those produced by Ugandan infants. This suggests firstly that listeners are able to extract information about the production context from the call alone, and secondly, that cultural variation only affects listeners with a very high level of experience with infant vocalisation.

These calls are another type of vocal behaviour present in the repertoire of the developing infant, one that could reflect our primate heritage. They are akin to universally human signals of communication such as the pointing gestures (Kita 2003), or facial displays of emotion (Ekman and Friesen 1979).

What functions do these calls serve?

Results from the playback study show that listeners can functionally interpret calls based on the differential audio information contained in them. This suggests that the non-linguistic calls produced by the infants have inherent communicative value, that is, the acoustic structure contains information that allows listeners to make inferences about the infant's activities or affective states. The type of information the calls contain remains unclear– do they refer to external referents, or are they largely a reflection of the infant's mood or attitude about some particular object or event?

Motivational

Most of the call categories we presented here fit the latter hypothesis, an affective or mood interpretation of the informational value. These calls mainly broadcast the infant's attitude or affect towards an action, object or event. For example, protests are likely a reflection of the infant's negative attitude towards some action. Similarly, the underlying motivation for food requests likely stems from the physical state of hunger, the resulting excitement of seeing and anticipating food, or the frustration of not obtaining the food. We postulate that all call categories reflect the infant's mood, for example excitement upon seeing an unusual event, or frustration when action requests are not met.

Some categories of calls are not traditionally associated with one specific emotion, such as declarative pointing or giving. It is nonetheless possible that they contain information about the infant's attitude towards a specific event or object that finds expression in the vocal behaviour. For example, a child pointing towards a dog walking by could express her excitement and positive attitude towards the dog through her vocalisations. We can similarly imagine a scenario in which the infant points to a dog and vocalises anxiously to express that uncertainty or even fear of this animal. Diverse emotions that accompany protodeclarative gestures could thus explain the variety found in the acoustic make-up of the calls, and why listeners generally perform with slightly less success in these categories in the playback study.

Referential

Non-linguistic calls might also contain information about external referents in addition to information about the infant's mood. It is possible that this information is not contained in the call itself, but in a combination of all available sources of information such as gestures, the environment, gaze direction, or even the fact that some calls are consistently produced in the same situational context. This does not mean that calls are a lesser type of communication; instead, it is possible that they are part of a system in which calls help the listener determine the need for gathering more precise information about the underlying reason for the vocal behaviour. To give a more concrete example: through an action request call, the listener might be informed that the child requests something (through past experience of when this vocalisation usually occurs), and that she is repeating the request with increasing urgency (through acoustic correlates of frustration or anger in the child's voice). Lastly, visual inspection can identify the external referent of the call, that she needs help unzipping her jacket (after visual checking following the vocal behaviour).

Therefore, non-linguistic infant calls could contain information about the infant's affective states, her mood and attitudes towards a referent and, at the same time, also refer to an external event or referent. Vocalisations might refer to external event in a very broad sense, for example implying that the child needs help with something but not actually giving information about the particular problem. Calls are thereby acting as a reason for the listener to visually check on the infant to gather further information about the reason for the call.

How does this compare to non-human primate vocal behaviour?

Signals produced by different species of non-human primates fall on different places of a spectrum between motivational and referential signals. The debate about what exact information primate calls contain and whether they are motivational, referential or both is still ongoing (for example Owren et al. 2005, Fischer, forthcoming, Zuberbühler 2006, Cheney and Seyfarth 1996). Alarm calls in monkeys, for example vervet (Cheney and Seyfarth 1980) or Diana monkeys (Zuberbühler et al. 1997), probably lean very far towards the referential end of the spectrum, as they refer to the presence of a specific type of predator in the immediate environment. They do, however, also contain information about the levels of arousal and threat experienced by the signaller. For example, the acoustic make-up of the call changes with the distance between monkeys and predator, which reflects the monkeys' arousal level to different level of threat. Cheney and Seyfarth (1996) consequently suggest that calls contain both referential and affective information.

Vocal behaviour of apes, on the other hand, might be situated more towards the motivational end of the spectrum. Their vocalisations are graded and do not display the same, clear category boundaries that are found in those of monkeys (Slocombe and Zuberbühler 2010, Tomasello and Zuberbühler 2002). For example, Slocombe and Zuberbühler (2005) showed that barks and screams during aggressive episodes vary with the severity of threat experienced by the caller. This suggest that ape vocalisations are context specific in that certain types are produced in defined contexts, but additionally show systematic acoustic variations reflecting the callers' arousal level. Even signals that are very motivational, or mainly thought to contain affective information, must contain a minimal degree of signal specificity (Marler et al. 1989).

For example, different types of chimpanzee vocal behaviour might vary acoustically as a function of the signaller's affective state, but are also linked to- and contain information about- specific environmental or situational contexts such as hunting, alarm calls, travel, or aggression (Crockford and Boesch 2003). Therefore if barks are only observed as part of hunting or agonistic events, and although the acoustic properties of a 'bark' might differ as a function of the signaller's arousal, these signals contain some information about the production context (Crockford and Boesch 2003, Marler et al. 1989). In order to gain specific information to disambiguate the call, the listener has to rely on additional contextual information, for example visually checking for aggression in other group members or whether the group is travellling.

Similarly, Slocombe and Zuberbühler (2003) reported that victim screams during aggression contexts change with the severity of aggression experienced as well as the composition of the audience. A listening chimp might hear the scream and gain the information that there is a severe aggressive episode taking place, but will still have to investigate the cause of the episode and the identity of the aggressor.

In displaying a combination of affect and information about an external event, infant calls might be very similar to the vocal behaviour of other primates. Although both human infant and other primates' calls contain information about the signaller's affective state and links to an external event, attributing meaning to vocal behaviours is largely the listener's task and involves integrating information from a variety of sources such as acoustic information contained in the call, any additional signals that are produced, and further environmental clues (Cheney and Seyfarth 1996, Sperber and Wilson 1997, Leger 1993).

The listener's reaction to a call, whether in human infants or other primates, is therefore central to the question of how calls function and what information they contain. In the next section we want to discuss how listeners react to infant calls and how this gives further information of what functions infant calls serve.

Reactions from listeners

In the natural, everyday situations where we collected data for our acoustic analysis, we frequently observed that caregivers from either cultural background commented on the infant's vocalisations. For example, when an infant requested an object a caregiver would say: "You want this? I will get it for you. Here you go". Other reactions included picking the infant up and providing soothing comments, complying with a request, or providing verbal explanations of why something is happening (For example that an infant is not allowed to have a favourite toy to play with now, because they were getting ready to go out). These were not just reactions to the vocal behaviour alone, but comments that clearly referred to the whole behaviour and the infant's situational context.

We observed similar responses to infant vocal behaviour in Ugandan and Scottish caregivers – most importantly, caregivers in both cultures did not just comment on the vocal behaviour alone but included information gained from contexts, including specific objects, events or activities. Although Ugandan caregivers were often less inclined to talk to their infant, they nevertheless offered similar responses in similar situations. When the infant was protesting, caregivers in both cultures tried to calm the infant down by distracting from the discomfort, offering alternative activities, or explanations as to why things are happening. Infant giving was often followed by some display of interest in the object. Similarly, in pointing contexts the caregiver expressed an interest in the target by naming it, pointing to something else (often something similar), or acknowledging the infant's excitement. Food requests were usually met by speeding up the action needed to feed the child, giving some food to the child or explaining why she will not get any. When the child vocalised in order to request an object or an action, she was often obliged.

Caregivers in both cultures were willing to engage in joint activities with the child, with some differences; where Ugandan parents often reacted to what the child did, Scottish caregivers tended to engage the child into specific activities. For example, when a Ugandan infant requested an object, she was given it and the caregiver returned to their previous activity. Scottish caregivers would instead provide the object and engage in an activity with it and the child, for example giving the child a toy car and encouraging the child to push it to them (similar observations have been reported in Keller, 2007).

As mentioned in the introduction, it has frequently been reported that parents comment on the infant's state and activities in response to their vocalisations (Papoušek 1992, Locke 1993, Bruner 1983, Franco 1997, Goldstein and West 1999). While these observations show that infant vocal behaviour can influence a listener's behaviour, we suggest here that some of these reactions to infant calls might be guided by information contained in the acoustic structure of these calls.

In our playback study we explicitly tested how well listeners can match vocal behaviour to its production context, even in the absence of additional cues. Participants were able to do this at a level well above chance, suggesting that the audio clips alone contained enough information to match calls and descriptions of contexts in which they were produced. The results suggest that some calls contain more information than the infant's attitude or mood, for example with calls such as declarative pointing or giving where there is no immediate association with any particular emotion. Interestingly, a number of participants, mainly parents, reported that the way they solved the task of matching vocalisations to contextual descriptions was by visualising the descriptions and seeing whether the vocal behaviour they heard would match the scene or not. Again, this implies that vocalisations have a strong relationship with additional information available in the infant's environment while still containing enough acoustic information for listeners to make judgements about the infant's activities.

Summary – What functions do calls serve?

Until now, the study of pre-linguistic infant vocalisations has been limited to the acquisition of speech sounds and marginalized in the study of infant communication. Our studies show that among pre-linguistic vocalisations there seems to exist a sub-set of non-linguistic calls that are an important part of the communicative repertoire of one- to two-year-old infants: these calls often accompany gestures and goal-directed actions, and results demonstrate that they convey enough information to allow human listeners to identify the contexts in which they occur on the basis of acoustic information alone. The level of information might differ between call types, but likely contains indicators of the infant's mood as well as broad information about the external referent involved. Vocalisations might serve an important function in providing a listener who is not visually monitoring the infant with initial information about the infant's mood and activities, informing the decision of whether it is necessary to gather further information through visual checking. Furthermore, they can also provide listeners who are monitoring the infant with information about how strongly the child feels about a request (urgency), problem, or event.

Are Infant Calls Comparable To Primate Calls?

Comparing any vocal behaviour of humans to that of other primate species is not without its pitfalls. Perhaps the biggest criticism is that, in principle, human infant vocal behaviour is thought to be intrinsically different to that of other primates. It is more flexible, shows more variation and is produced by a very different vocal apparatus (Fitch 2000). As mentioned in the introduction, precursors to elements that are characteristic of language, such as reference, displacement, and cultural variation, are already present in the prelinguistic productions of the human infants. However, the fact that infants at the end of their first year are already able to produce rather complex sounds, for example the syllables observed in babbling (Oller 2000), word approximations, or sound matching (Kuhl and Meltzoff 1996), does not exclude the presence of simpler vocalisations that reflect our primate heritage.

One of our research questions was whether infants produce vocal behaviours that share properties with those produced by non-human primates. Primate calls are thought to be innate, for they emerge in the absence of auditory experience, with few changes to the acoustic form as the individual matures (Snowdon 2008). What seems to happen as the primate develops is that the individual learns to apply the calls to the appropriate situations, for example to an eagle flying over the monkeys rather than just anything in the air (Cheney and Seyfarth 1986) or, in the case of chimpanzees, producing pant-grunts to the higher ranked individual (Laporte and Zuberbühler 2010).

Primate calls might not always be produced intentionally on behalf of the signaller, but nevertheless provide the listener with quite specific information about the signaller, such as identity, affective state, and about events in the environment, for example presence and type of a predator or presence of a dominant individual (Tomasello and Zuberbühler 2002). Evidence for the information content of primate calls mainly comes from playback studies. For example, monkeys who hear an artificially played back call associated with an aerial predator hide in the canopy, whereas they climb high into trees upon hearing a call for a ground predator, even in the absence of any real predator (Marler, Seyfarth and Cheney 1984). Therefore calls function in a referential way; that is, they carry some meaning that influences the listener's behaviour in relation to an external event. Additional contextual clues, such as presence of other individuals, the actions they are engaged in, and the state of the environment, each act to specify the meaning of a primate call (Leger 1993).

Many of the above features apply to the infant vocal behaviours presented here. In our studies we found that infant calls share acoustic properties across a number of individuals, and transmit information about the production context to a listener and provoke consistent responses in people around the infant.

Of course, the existence of non-linguistic calls does not reduce the infant's vocal repertoire to limited, stereotyped signals. It merely suggests that in addition to the highly flexible vocal behaviour that plays a prominent role in language acquisition and as well as the flexible and intentional gestures that have been documented in early communicative development, human infants also possess relatively stereotypical, non-linguistic calls that have an adaptive value in informing the caregiver of the infant's needs and wants before she is able to speak.

In light of our findings, it seems that infants produce calls with some acoustically stable core characteristics that show little flexibility and are readily perceived by receivers regardless of their previous experience with this kind of vocal behaviour. At the same time, results from both mono-cultural and cross-cultural analysis show that there is still a level of acoustic flexibility in infant calls, suggesting that they are not as rigid as alarm calls of vervet or Diana monkeys but retain stereotyped acoustic core characteristics.

Stereotyped, inflexible vocal behaviour is not absent from the human repertoire; cries have been identified as sharing many of the same properties as animal vocalisations (Lester and Boukydis 1989, Scheiner et al. 2002). Even in adult humans, vocal behaviour like laughter, grunts, and cries are present and fall into the same category of inflexible, relatively stereotyped vocal behaviour (Locke 1993). Prosodic variables such as pitch, or certain melody contours, are often associated with the speaker's emotional state and are important signals that supplement the meaning of an utterance (Papaeliou and Trevarthen 2006).

The infant calls we presented here are relatively stereotyped in that they show significant acoustic similarities across different individuals and are even largely similar across cultures, with only a small degree of variation. Furthermore, we observed them in infants who are already capable of producing flexible vocal behaviour and beginning to use linguistic signals – this division suggests that some parts of the infant's repertoire are less flexible than others, just like the vocal repertoire of adult humans, which also contains highly flexible vocal behaviours, such as speech sounds, and more stereotyped sounds such as laughter or cries (Marler et al. 1989).

The suggestion that human infants produce "calls" could prove contentious (indeed an anonymous reviewer suggested that only nonhumans have calls, and that infants vocalisations are vastly more elaborate communicatively than those of any primate 'call'). However, our studies provide empirical evidence of the existence of such systematic non-linguistic vocalisations (not mere prosodic inflexions of words or protowords, but genuine non-linguistic vocalisations) capable of conveying contextual information through consistent relationships between sound properties and production contexts. We suggest that the term "call" is the most appropriate way to characterize them, highlighting the non-linguistic nature of these vocalisations.

The non-linguistic character of these vocalisations finds further evidence in the results of the cross-cultural comparison. When the categories of calls were compared between Scotland and Uganda, four of them showed little variation between the cultures, suggesting they are part of a universal human repertoire rather than part of flexible vocal behaviours that are influenced by the signaler's native language. Moreover, the cultural background of the signaler only affects the most experienced listeners, again providing evidence for the universality of these calls. The small degree of variation we found indicates that even relatively stereotyped vocal behavior we present here is not immune to cultural variations such as those primarily observed in other, more flexible vocal behaviours, babbling being one example.

Summary – Are infant calls comparable to primate calls?

Infants produce non-linguistic vocalisations that can be properly characterised as calls and that share many properties with primate calls. They seem to be context specific, and have acoustic properties that are consistent across a number of individuals in the same contexts. Further, listeners can gain information about the caller's affective state or environment from them, even in the absence of additional visual information. With similarities to facial expressions of emotions (Ekman and Friesen 1969), and some gestures such as pointing, this vocal system might be partially species-specific with similar functions across cultures and languages. In tandem with the more flexible, linguistic vocal and wider gestural repertoires that have been described in human infants, this system is part of the infant's prelinguistic vocal repertoire.

How do these results contribute to the current literature?

Most studies on infant vocal behaviour are focused on the transition from prelingusitic vocalisations to speech, instead of the role vocalisations play in infant communication. In contrast, non-linguistic communication has almost exclusively been studied in the gestural domain, with only a few studies hinting at the possibility of non-linguistic vocalisations serving communicative functions (D'Odorico and Franco 1991, Dore et al. 1975, Papoušek 1992, Leroy et al. 2009, Locke 1993). Despite providing some evidence for the relationship between sound and function in prelinguistic infants, we show in our review in Chapter 1 that these studies provided fragmented and rather asystematic evidence of the communicative value of infant vocal behaviour.

Our studies, using ethological methods adapted from studies of non-human primates, provide a systematic investigation of infant calls that encompass a range of vocal behaviour observed in environmentally and behaviourally defined production contexts. Our study provides one of the first pieces of empirical evidence that acoustic variables, such as utterance length, loudness, and melody contours, contain information about the infant's state and activities that can be picked up by a listener; in other words, that infants produce 'phonetically consistent forms' (Dore et al. 1976). In contrast to what other researchers propose (Dore et al. 1976, Papaeliou and Trevarthen 2006), we do not think these consistent forms are precursors to words or word-like, but are instead non-linguistic and largely uninfluenced by the flexibility observed in other areas of infant vocal productions.

Non-linguistic vocal behaviour is not only present in infants as they begin to produce speech, but our studies show that it can fulfil important functions in that it contains specific information about an infant's affective state and activities that can be decoded by listeners. As our studies show, this kind of behaviour is not limited to cries, grunts, and screams, as suggested by previous literature, but encompasses a wider range of vocal behaviours that are associated with diverse functions. This suggests that vocal behaviour is communicative before the infant learns to speak and that it could fall into similar functional categories than gestures.

Our studies are significant in that they provide a systematic cultural comparison of the vocal behaviour of prelinguistic infants, a type of study that is extremely rare, as most cross-cultural studies centre on the differences in child-raising across different cultures (for example, Keller 2007). Considerable research effort has been devoted to documenting elements of the infant's native language in their prelinguistic production, most notably babbling (Vihman 1996 for a summary of her studies, Boisson-Bardies and Vihman 1991, Boisson-Bardies and Vihman 1994). There is a growing body of studies that investigates whether the standard sequences and model of communicative development, in both the vocal and gestural domain, are actually the same across infants growing up in different cultures (Callaghan et al. 2010). Here, we showed that there are similarities in the non-linguistic vocal behaviour of human infants, suggesting that calls like those we identified might be part of a universal repertoire in developing humans that is not hugely affected by cultural variability. This provides an interesting contrast to studies that emphasize (and mainly look for) cross-cultural differences.

Rather than emphasizing cultural differences in vocal behaviour we want to highlight the commonalities we found in calls produced by Ugandan and Scottish infants. These indicate that there are at least some classes of universal, acoustically stereotyped vocal behaviour in human infants. We speculate that these could be part of an innate, species-specific vocal repertoire that shows relatively little variation between infants and is similar to calls documented in other primate species.

Most importantly, our playback study is the first to compare the listeners' interpretations of infant vocalisations from different cultures and thereby offers a unique contribution to the field. Contrary to Papoušek's study (1989), the study we presented here did not find that the listener's level of experience with infants is significantly reflected in their performance at classifying vocalisations. Our results found that experience plays only a limited role when comparing listeners' performance with calls produced by Ugandan and Scottish infants. Parents were better at categorising vocal behaviour produced by infants who share their cultural background compared to those growing up in a different cultural environment– possibly an effect of practice with these vocal behaviours.

Clearly, experience can only have an effect on participant's' judgement of the calls from one culture if there indeed are cultural differences between the calls. The acoustic analysis showed that when calls were compared cross-culturally, there was more acoustic variation in the rate of correct classifications between cultures than in the mono-cultural samples. Although this variation was rather small (the rate of correct classification dropped from ~60% within cultures to 50% between cultures), it might nevertheless leave subtle traces in the infant's vocal behaviour that can be perceived by experienced listeners and affect their interpretation of the call. Infants from about six months onwards show language specific elements in their vocal play (Vihman 1996). It would therefore be unsurprising that these variations also effect other forms of vocal behaviour, albeit to a lesser degree. To further assess the effect of the signaller's cultural background on listeners, it is also necessary to test whether Ugandan parents are also better at classifying vocalisations produced by infants of their own culture. We are currently in the process of planning this study.

The methodology we presented here is directly comparable to studies investigating vocal behaviour in non-human primates. This allows for a direct comparison of the form and function of vocal behaviour and could contribute to a better understanding of the phylogenetic continuity of vocal behaviour between humans and other primates.

Summary – How do results contribute to the current literature?

Results from our empirical studies contribute to the current literature in a number of ways. They provide systematic evidence for what other studies have only assumed – that infants' non-linguistic vocalisations vary systematically with the situations in which they are produced, and can thereby act as source of information about the infant's attitudes and activities to caregivers. However, contrary to the opinion of previous researchers (Dore et al. 1975, Papaeliou and Trevarthen 2006, Halliday 1975), we emphasize the non-linguistic character of these vocalisations and the commonalities they share with the vocal behaviour of other primates. This is supported by the cross-cultural analysis that showed more commonalities than differences in vocal behaviour produced by infants from Uganda. This same analysis demonstrated that the cultural difference had little influence on how well listeners could classify the calls. Again, this supplements the current view in the literature that emphasizes the emergence of cultural differences in prelingusitic vocal production in flexible vocal behaviours (Boisson-Bardies and Vihman 1992). Results from the playback study are particularly important as this paradigm has rarely been used to assess the function and effectiveness of infant vocal communication.

Concluding Remarks

Our study suggests that, in addition to the rich developing repertoire of vocalisations that are part of the process of speech acquisition, human infants display vocalisations that are better characterised as calls and that might be comparable to the vocal systems of other primates. Infant calls seemingly have stable acoustic core characteristics that show little variation across individuals or between cultures, and contain information about the infant's affective state, attitude or ongoing activities. This alludes to the evolutionary history of vocal behaviour, particularly in light of the vast differences in vocal communication between humans and other primates that have been reported so far. Our studies suggest that human infants still have some traces of phylogenetically older vocal behaviour that is comparable to vocalisations documented in other primate species.

Non-linguistic calls seem to have an adaptive value in that they can inform caregivers of the infant's activities, and help determine the any need for attention, even when the infant is not visually monitored. This is particularly useful as it allows caregivers to engage in other activities and could stem from a need to do tasks crucial for survival, such as gathering food, whilst still looking after their offspring. This allows the infant adequate care through broadcasting signals that vary with her needs. Infant calls can also contribute to the overall meaning of the infant's non-linguistic communicative system by interacting with gestures, contextual information and other types of vocalisation more related to the ongoing tasks of language acquisition.

Infant calls show that prelinguistic vocal behaviour is not limited to practicing the sounds of speech or to cries that broadcast the infant's affective state, but also includes classes of vocal behaviour with relatively fixed sound-meaning correspondences. Our studies show that these vocal behaviours are still present in infants who are beginning to engage in linguistic communication, and use gestures with communicative intent.

All in all, the findings we reported here contribute to better understanding the continuity with other primate species in the development of human vocal communication.

References

Altman, J. (1974). "Observational study of behaviour: Sampling methods". *Behaviour*. Vol. 49, p. 227.

Bates, E., and Dick, F. (2002). "Language, gesture, and the developing brain". *Developmental Psychobiology*, Vol. 40, pp. 293-310

Bates E., Benigni L., Bretherton I., Camaioni L., and Volterra V. (1979). *"The Emergence of Symbols – Cognition and Communication in Infancy"*. London, UK: Academic Press

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2011). "Twelve-month-olds' comprehension and production of pointing". *British Journal of Developmental Psychology*. Advance online publication. doi:10.1111/j.2044-835X.2011.02043.

Bernhardt, B.H. and Stemberger, J.P. (1998). *"Handbook of Phonological Development – From the Perspective of Constrained-Based Nonlinear Phonology"*. Academic Press, London

Bermejo, M. and Omedes, A. (1999). "Preliminary vocal repertoire and vocal communication of wild bonobos (Pan paniscus) at Lilungu (Democratic Republic of Congo)". *Folia Primatologica*, Vol. 70 (6), pp. 328-357

Blake, J., Vitale, G., Osborne, P. and Olshanksy, E. (2007). "A cross-cultural comparison of communicative gestures in human infants during the transition to language", in K. Liebal, C. Mueller and S. Pika (eds), *Gestural Communication in Nonhuman and Human Primates*, Amsterdam, NL: John Benjamins

Blake J. (2000): *"Routes to Child Language – Evolutionary and Developmental Precursors"*, Cambridge, UK: Cambridge University Press

Blake J. and Dolgoy S.J. (1993): “Gestural Development and its Relation to Cognition During the Transition to Language”, *Journal of Nonverbal Behaviour, Vol.* 17 (2), pp. 87 – 100

Blake J., McConnell S., Horton G., Benson N. (1992): “The Gestural Repertoire and its Evolution Over the Second Year”, *Early Development and Parenting*, Vol.1 (3), pp. 127 – 136

Bloom, K. and Lo, E. (1990). “Adult perception of vocalizing infants”. *Infant Behaviour and Development,* Vol. 13 (2), pp. 209-219

Bruner, J. (1975) “From Communication to Language – A psychological perspective”. *Cognition*, Vol. 3 (3), pp. 255-287

Bruner, J. (1983). *“Child’s Talk – Learning to Use Language”* Oxford, UK: Oxford University Press

Bruner J. (1983): “The Acquisition of Pragmatic Commitments”, in Golinkoff R.M. , *The Transition from Prelinguistic to Linguistic Communication*, London, UK: Lawrence Earlbaum Associates

Butterworth,G. (2003). “Pointing Is the Royal Road to Language for Babies” in S. Kita (2003). *“Pointing : Where Language, Culture and Cognition Meet”*. London, UK: Lawrence Earlbaum Associates

Call, J. and Tomasello, M. (2007). *“The Gestural Communication of Apes and Monkeys”*, New York, NY: Lawrence Earlbaum Associates

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T. and Tomasello, M. (2011): “Early Social Cognition in Three Cultural Contexts”, *Monographs of the Society of Research in Child Development*, Vol. 76 (2)

Capirci, O., Contaldo, A., Caselli, M.C. and Volterra, V. (2005). "From Action to language through gesture: a longitudinal perspective". *Gesture*, Vol. 5 (1/2), pp. 155-177

Carpenter, M., Nagell, K., & Tomasello, M. (1998)." Social cognition, joint attention, and communicative competence from 9 to 15 months of age". *Monographs of the Society for Research in Child Development,* 63 (4, Serial No. 255)

Cartmill E. and Byrne R. (2007): "Orangutans Modify Their Gestural Signalling According to Their Audience's Comprehension", *Current Biolog , Vol.* 17, pp. 1345 – 1348

Chang, R.S. and Thompson, N.S. (2011). "Whines, Cries and Motherese: Their Relative Power to Distract". *Journal of Social, Evolutionary and Cultural Psychology*, Vol. 5 (2), pp. 10-20

Cheney, D.L., Seyfarth, R.M. & Silk, J.B. (1995). "The role of grunts in reconciling opponents and facilitating interactions among adult female baboons". *Animal Behaviour,* Vol. 50, pp. 249-257

Cheney, D.L. and Seyfarth, R.M. (1996). "Function and intention in the calls of nonhuman primates", *Proceeding of the British Academy,* Vol. 88, pp. 59 – 76

Cheney, D.L. and Seyfarth, R.M. (1980). "Vocal Recognition in free Ranging Vervet Monkeys", *Animal Behaviour*, Vol. 28, pp. 362 – 367

Clay, Z. and Zuberbühler, K. (2009). "Food-associated Calling sequences in Bonobos". *Animal Behaviour,* Vol. 77(6), pp. 1387-1396

Crockford, C. and Boesch, C. (2005). "Call Combinations in Wild Chimpanzees". Behaviour, Vol. 142, pp. 397 – 421

Crockford, C. and Boesch, C. (2003). "Context-specific calls in wild chipmanzees, *pan troglodytes verus*: an analysis of barks". *Animal Behaviour*, Vol. 66, pp. 115-125

Darwin, C. (1877). "A Biographical Sketch of an Infant". *Mind*, Vol. 2 (7), pp. 285 – 294

De Boysson-Bardies, B. and Vihman, M.M (1991). "Adaptation to Language: Evidence from Babbling and First Words in Four Languages". *Lanuguage,* Vol. 62 (2), pp. 297 – 319

Delgado, B. , Gómez ,J.-C., and Sarriá, E. (2011). "Pointing Gestures as a cognitive tool in young children: Experimental evidence". *Journal of Experimental Child Psychology,* Vol. 110, pp. 299 – 312

D'Odorico, L. and Franco, F. (1991). Selective production of vocalisation types in different communication contexts. *Journal of Child Language*, 18, 475–499

D'Odorico, L. (1984). "Non-segmental features in prelinguistic communication: an analysis of some types of infant cry and non-cry vocalisations". *Journal of Child Language*, Vol. 11, pp. 17 – 27

Dore J. (1983): "Feeling, Form and Intention in the Baby's Transition to Language" in R.M. Golinkoff (ed.), *The Transition from Prelinguistic to Linguistic Communication*, London, UK: Lawrence Earlbaum Associates

Dore J., Franklin, M.B., Miller, R.T. and Ramer, A.L.H. (1976): "Transitional Phenomena in Early Language Acquisition", *Journal of Child Language,* Vol.3 (1), 13 – 28

Ekman P. and Friesen W. (1969): "The Repertoire of Nonverbal Behaviour: Categories, Origins, Usage and Coding", *Semiotica* , Vol. 1,pp. 49 – 97

Elowson, M.A., Snowdon, C.T. and Lazaro-Perea, C. (1998). "Infant 'Babbling' in a Non-Human Primate". *Behaviour*, Vol. 135 (5), pp. 643- 664

Fernald, A. (1989). "Intonation and Communicative Intent in Mother's Speech to Infants: Is the Melody the Message". *Child Development,* Vol. 60, pp. 1497 – 1510

Fernald, A. and Kuhl, P. (1987). "Acoustic Determinants of Infant Preference for Motherese Speech". *Infant Behaviour And Development*, Vol. 10, pp. 279-293

Fischer, J. (forthcoming). "Information, Inference and Meaning in Primate Vocal Behaviour". To appear in U. Stegmann (Ed.) *Animal Communication Theory: Information and Influence.* Cambridge, UK: Cambridge University Press

Fischer, J. and Hammerschmidt, K., (2002). "An Overview of the Barbary Macaque, *Macaca sylvanus*, Vocal Repertoire". *Folia Primatologica*, Vol. 73(1), pp. 32 – 45

Fitch, T.W. (2000). "The evolution of speech: a comparative review", *Trends in Cognitive Sciences,* Vol.4 (7), pp. 258 -268

Flax J., Lahey M., Harris K., Boothroyd A. (1991): "Relations between prosodic variables and communicative functions", *Journal of Child Language*, Vol.18.(1), pp. 3 – 19

Franco F. and Butterworth G. (1996): "Pointing and Social Awareness: Declaring and Requesting in the Second Year", *Journal of Child Language* , Vol. 23, pp. 307 – 336

Franco F. (1997). "The Development of Meaning in Infancy: early communication and social understanding". In S. Halla (ed),*The Development of Social Cognition*. London: UCL Press

Galligan R. (1987): "Intonation with single words: Purposive and grammatical use", *Journal of Child Language*, Vol. 14, pp. 1 – 21

Goldstein M.H. and West M.J. (1999). “Consistent Responses of Human Mothers to Prelinguistic Infants: The Effect of Prelinguistic Repertoire Size”. *Journal of Comparative Psychology*, Vol. 113 (1), pp. 52 – 58

Golinkoff,R. M. (1986). “I beg your pardon? The preverbal Negotiation of Failed Messages”. *Journal of Child Language*, Vol. 13 (3), 455 - 476

Golinkoff, R.M. (1983). “*The Transition from Prelinguistic to Linguistic Communication”,* London, UK: Earlbaum

Gómez, J. C., Sarriá, E., & Tamarit, J. (1993). “The comparative study of early communication and theories of mind: Ontogeny, phylogeny and pathology”. In S.Baron-Cohen, H.Tager-Flusberg, & D.Cohen (Eds.), “*Understanding other minds: Perspectives from autism”* Oxford, UK: Oxford University Press, pp. 397–426

Gómez J.-C. (2007). “Pointing Behaviours in Apes and Human Infants: A Balanced Interpretation”. *Child Development*, Vol. 78 (3), pp. 729 – 734

Goodall, J. (1986). “*The Chimpanzees of Gombe: Patterns of Behaviour”*. Cambrdige, MA: Harvard University Press

Gouzoules, H. and Gouzoules, S. (1999). “Agonistic screams differ among four species of macaques: the significance of motivation-structural rules”, *Animal Behaviour,* Vol. 59 (3), pp. 501-512

Griebel, U. and Oller, K.D. (2009). “Evolutionary Forces Favoring Contextual Flexibility”. In D.K. Oller and U. Griebel U. *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication (Vienna Series in Theoretical Biology),* Cambridge, MA: MIT Press

Halliday M.A.K. (1975). “*Learning How to Mean – Explorations in the Development of Language”*. London, UK: Edward Arnold

Hammerschmidt, K. and Fischer, J. (1998). “The vocal repertoire of Barbary macaques: a quantitative analysis of a graded system”. *Ethology*, Vol. 194, pp. 203-216

Harding C.G. (1982): *“Development of the Intention to Communicate”, Human Development*, Vol. 25, pp. 140 – 151

Hobaiter, C. and Byrne, R.W. (2011). “The gestural repertoire of the chimpanzee”. *Animal Cognition*, Vol. 14(5), pp. 745 – 767

Iverson, J.M. and Goldin-Meadow, S. (2005). “Gestures Paves the Way for Language Development”. *Psychological Science*, Vol. 16(5), pp. 367-371

Janik V.M., Slater P.J.B. (2000). “The different Roles of Social Learning in Vocal Communication”. *Animal Behaviour* , Vol. 60, pp. 1 – 11

Janik V.M, Slater P.J.B. (1997). “Vocal Learning in Mammals”, *Advances in the Study of Behaviour*, Vol. 26, pp. 59 – 99

Kaminski, J. (2011). Communicative cues among and between human and nonhuman primates: Attending to specificity in triadic gestural interactions. In M. Boos, M. Kolbe, P. Kappeler and T. Ellwart (Ed.) *Coordination in Human and Primate Groups*, Heidelber, DE: Springer-Verlag

Keller, H. (2007). *“Cultures of Infancy”,* London, UK: Lawrence Earlbaum Associates
Kita, S. (2003). *“Pointing : Where Language, Culture and Cognition Meet”*. London, UK: Lawrence Earlbaum Associates

Kitchen, D.M., Cheney, D.L., & Seyfarth, R.M. (2005). “Male chacma baboons (*Papio hamadryas ursinus*) discriminate loud call contests between rivals of different relative ranks”. *Animal Cognition,* Vol. **8**, pp. 1-6

Kuhl, P. (2004): “Early language acquisition: cracking the speech code”. *Nature Reviews Neuroscience*, Vol. 5, pp. 831-843

Kuhl, P. K. and Meltzoff, A. N. (1996). “Infant vocalisations in response to speech: Vocal imitation and developmental change”. *Journal of the Acoustical Society of America*, Vol. 100, pp. 2425-2438.

Laporte, M.N.C. and Zuberbühler, K. (2011). “The development of a greeting signal in wild chimpanzees”. *Developmental Science*, Vol. 4 (5), pp. 1220 – 1234

Leavens, D.A. and Hopkins, W. D. (1999). “The Whole-Hand Point: The Structure and Function of Pointing from a Comparative Perspective”, *Journal of Comparative Psychology,* Vol. 113 (4), pp. 417 – 425

Leger, D.W. (1993). “Contextual sources of information and responses to animal communication signals”, *Psychological Bulletin*, Vol. 113, pp. 295 – 304

Leopold, W. (1939): “*Speech development of a bilingual child I: Vocabulary growth in the first two years”*, Evanston, IL: Northwestern University Press

Leroy, M., Mathiot, E. and Morgenstern, A. (2009). “Pointing Gestures, Vocalisations and Gaze: Two Case Studies”, in J. Zlatev, M.J. Falck, C. Lundmark and M. Andren, *Studies in Language and Cognitio*, Cambridge, UK: Cambridge Scholar Publishing, pp. 386 – 404

Lester, B.M. and Boukydis, C.F.Z. (1989). “No Language but a Cry”. In H. Papoušek, and U. Jurgens (Eds): *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Cambridge, UK: Cambridge University Press

Lester, B.M., Garcia-Coll, C.T. and Valcarel, M. (1989). "Perception of infant cries in adolescents and older mothers", *Journal of Youth and Adolescents*, Vol. 18, pp. 231 – 243

Leung, E.H. and Rheingold, H.L. (1981). "Development of pointing as a social gesture", *Developmental Psychology*, Vol. 17 (2), pp. 215 -220

Lewis, M.M. (1936). "*Infant speech: A study of the beginning of language*", New York: Harcourt Brace

Liebermann, P. (1991). "Uniquely Human – Speech, Thought and Selfless behaviour", Harvard: Harvard University Press

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & De Vos, C. (in press). "A prelinguistic gestural universal of human communication". *Cognitive Science*

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). "Prelinguistic infants, but not chimpanzees, communicate about absent entities". *Psychological Science*, Vol. 20, pp. 654-660

Liszkowski, U. (2008). "Before L1: A differentiated perspective on infant gestures". *Gesture*, Vol. 8(2), pp. 180-196

Liszkowski, U., Albrecht, K., Carpenter, M., & Tomasello, M. (2008)." Infants' visual and auditory communication when a partner is or is not visually attending". *Infant Behavior and Development*, Vol. 31(2), pp. 157-167

Liszkowski, U. (2006)." Infant pointing at twelve months: Communicative goals, motives, and social-cognitive abilities". In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: culture, cognition and interaction*. New York, NY: Berg ,pp. 153-178

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). "Twelve-month-olds point to share attention and interest". *Developmental Science*, Vol. 7(3), pp. 297-307.

Lock, A. (1980). *"The Guided Reinvention of Language".* London, UK: Academic Press

Locke, J.L. (1993). *"The Child's Path to Spoken Language"*. Cambridge, MA: Harvard University Press

Lynip, A.W. (1951). "The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant", Genetic Psychology Monographs, Vol. 44, pp. 221-262

Macedonia, J.M., (1993). "The Vocal Repertoire of the Ringtailed Lemur". *Folia Primatologica*, Vol. 61 (4), pp. 186 – 217

Mampe, B., Friedrici, A.D., Christophe, A., and Wermke, K. (2009). "Newborns' Cry Melody Is Shaped By Their Native Language". *Curren Biology*, Vol. 19 (23)

Marler, P. , Evans, C.S. and Hauser, M.D. (1989). "Animal signals: Motivational, referential or, both?. In In H. Papoušek, and U. Jurgens (Eds): *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Cambridge, UK: Cambridge University Press

Masataka, N. (2003): *"The Onset of Language"*, Cambridge, UK: Cambridge University Press

Masur, E.F. (1980). "The Development of communicative gestures in mother-infant interactions", *Papers and Reports on Child Language Development*, Vol. 19, pp. 121 – 128

McCune, L., Vihman, M.M., Roug-Hellichius, L. , Delery, P.B. and Gogate, L. (1996). "Grunt communication in human infants (*homo sapiens*)". *Journal of Comparative Psychology,* Vol. 110, pp. 27-37

Mitani, J.C. and Nishida, T. (1993). "Contexts and social correlates of long-distance calling by male chimpanzees", *Animal Behaviour*, Vol. 45, pp. 735-746

Nathani S., Oller K.D. (2001). "Beyond ba-ba and gu-gu: Challenges and strategies in coding infant vocalisations". *Behaviour Research Methods, Instruments and Computers*, Vol.33 (3), 321 – 330

Oller, D. K., Niyogi, P., S. Gray, J. A. Richards, J. Gilkerson, D. Xu, U. Yapanel, S. F. Warren (2010). "Automated Vocal Analysis of Naturalistic Recordings from Children with Autism, Language Delay and Typical Development." *Proceedings of the National Academy of Sciences,* Vol. 107(30), pp. 13354-13359

Oller, D.K and Griebel, U. (2008). "Contextual Flexibility in Infant Vocal Development and the earliest Steps in the Evolution of Language" in D.K. Oller and U. Griebel (eds) *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication* (Vienna Series in Theoretical Biology). Cambridge MA: MIT Press

Oller, D.K. (2000). *"The Emergence of the Speech Capacity"*. Mahwah, NJ: Lawrence Erlbaum Associates

Oppenheimer, JR (1973). "Social and communicatory behavior in the Cebus monkey". In C. R. Carpenter *Behavioral Regulators of Behavior in Primates*. Lewisburg: Bucknell University Press, pp. 251-271.

Owren M.J. and Golstein M.H. (2009). "Scaffolds for Babbling: Innateness and Learning in the Emergence of Contextually Flexible Vocal Production in Human Infants", in D.K. Oller and U. Griebel *Evolution of Communicative Flexibility:*

Complexity, Creativity, and Adaptability in Human and Animal Communication (Vienna Series in Theoretical Biology), Cambridge MA: MIT Press

Owren, M.J., Rendall, D. and Bachorowski, J.-A. (2005). “Conscious and Unconscious Emotion in Nonlinguistic Vocal Communication”, in L.F. Barrett, P. Niedenthal and P. Winkelman (eds), *Emotion and Consciousness*, New York, NY: Guilford Publications

Owren, M. J., & Rendall, D. (2001). “Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signalling”. *Evolutionary Anthropology*, Vol. 10, pp. 58-71

Papaeliou C.F. and Trevarthen C. (2006): “Prelinguistic pitch patterns expressing ‘communication’ and ‘apprehension’”, *Journal of Child Language* , Vol. 33, pp. 163 – 178

Papaeliou, C.F., Minadakis, G. and Cavouras, D. (2002). “Acoustic Patterns of Infant Vocalisations Expressing Emotions and Communicative Functions”, *Journal of Speech, Language and Hearing Research,* Vol. 45, pp. 311-317

Papoušek M. (2007). “Communication in Early Infancy: An Arena of Intersubjective Learning”. *Infant Behaviour and Development,* Vol. 30, 258 – 266

Papoušek M. (1992). “*Vom ersten Schrei zum ersten Wort – Anfänge der Sprachentwicklung in der vorsprachlichen Kommunikation*”. Kempten: Verlag Hans Huber

Papoušek, H. (1989). “Early Ontogeny of Vocal Communication in Parent-Infant Interactions”. In H. Papoušek, and U. Jurgens (Eds): *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Cambridge, UK: Cambridge University Press

Papoušek,H. and Bornstein, M.H. (1989). “Didactic Interactions: Intuitive parental support of vocal and verbal development in human infants” in H. Papoušek, and U. Jurgens (Eds): *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Cambridge, UK: Cambridge University Press

Papoušek H. and Papoušek M. (1989): “Determinants of Responsiveness to Infant Vocal Expression of Emotional States”, *Infant Behaviour and Development*, Vol. 12, pp. 505 – 522

Pika, S. Liebal, K., Call, J. and Tomasello, M. (2007). “The gestural communication of apes”. In K. Liebal, C. Mueller and S. Pika (eds), *Gestural Communication in Nonhuman and Human Primates*, Amsterdam, NL: John Benjamins

Pollick A.S., Gouzoules H., de Waal F.B.M. (2005): “Audience effects on food calls in captive brown capuchin monkeys, Cebus apella” *Animal Behaviour*, Vol.70 (6), pp. 1273 – 1281

Pozzi, L., Gamba, M. and Giacoma, C. (2010). “Use of artificial neural networks to classify primate vocalisations: a pilot study on black lemurs”, *American Journal of Primatology*, Vol. 72, pp. 337-348

Protopapas, A. and Eimas, P.D. (1997). “Pereptual Differences in Infant Cries Revealed By Modifications of Acoustic Features”. *Journal of the Acoustic Society of America*, Vol. 102(6), pp. 3723- 3734

Ratner N., Bruner J. (1977): “Games, Social Exchange and the Acquisition of Language”, *Journal of Child Language,* Vol. 5, pp. 391 – 401
Rendall, D., Owren, M.J. & Ryan, M.J. (2009). “What do animal signals mean?” *Animal Behaviour,* Vol. 78, pp. 233-240

Scheiner E., Hammerschmidt K., Jürgens U. and Zwirner P. (2002). “Acoustic Analysis of Developmental Changes and Emotional Expression in the Preverbal Vocalisations of Infants”. *Journal of Voice*, Vol. 16 (4), pp. 509 – 529

Schieffelin, E. and Ochs, B. (1979). *“Developmental Pragmatics”*. New York, NY: Academic Press

Schrader, L. and Hammerschmidt, K. (1997): “Computer-aided analysis of acoustic parameters in animal vocalisations: a multi-parametric approach”. *Bioacoustics*, Vol. 7, pp. 2427-265

Seyfarth, R.M. & Cheney, D.L. (1986). ”Vocal development in vervet monkeys”. *Animal Behaviour,* Vol 34, pp. 1640-1658.

Seyfarth, R.M., Cheney, D.L. & Marler, P. (1980). “Vervet monkey alarm calls: Semantic communication in a free-ranging primate”. *Animal Behaviour,* Vol. 28, pp. 1070-1094

Slocombe, K. and Zuberbühler, K. (2010): “Vocal Communication in chimpanzees” In E.V. Lonsdorf, S. Ross and T. Matzusawa (Eds), “*The Mind of the Chimpanzee Ecological and Empirical Perspectives”*, Chicago, IL: University of Chicago Press, pp. 192-207

Slocombe K. , Zuberbühler K. (2006). “Food Associated Calls in Wild Chimpanzees: Responses to Food types or food preference”. *Animal Behaviour*, Vol.72 (5),pp. 989 – 999

Slocombe, K. and Zuberbühler, K. (2005). “Agonistic screams in wild chimpans (*pan troglodytes schweinfurthii)* vary as a function of social role”, *Journal of Comparative Psychology*, Vol. 119 (1), pp. 67-77

Snowdon, C.T. (2009). “Plasticiy of Communication in Nonhuman Primates”. *Advances in the Study of Behaviour*, Vol. 40, 239 – 276

Snowdon, C.T. (2008): “Contestually Flexible Commuication in Nonhuman Primates”. In D.K. Oller and U. Griebel *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication (Vienna Series in Theoretical Biology),* Cambridge MA: MIT Press

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief in two-year-olds. *Psychological Science,* 18 (7)*, 587- 592*

Sperber, D. and Wilson, D. (1995): *“Relevance – Communication and Cognition”*, Oxford, UK: Blackwell Publishing

Stark, R.E. (1980). “Stages of Speech Development in the First Year of Life”, in G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds), *Child Phonology*, New York, NY: Academic Press

Stoel-Gammon, C. (1998). “Role of babbling and phonology in early linguistic development”, In A.M. Wetherby, S.F. Warren and J. Reichel (Eds), *Transitions in Prelinguistic Communication,* Baltimore, MD: Paul H. Brooks Publishing

Struhsaker, T. T. (1967)."Auditory communication among vervet monkeys (Cercopithecus aethiops)". In S. A. Altmann (Ed.), *Social communication among primates*, Chicago, IL: University of Chicago Press, pp. 281 – 324

Taine, M. (1877). "On the Acquisition of Language by Children". *Mind*, Vol. 2(6), p. 252 – 259

Titze, I.R. (1994). *"The Principles of Voice Production".* London, UK: Prentice Hall
Tomasello, M. (2008). "*The Origins of Human Communication"*, Cambridge, MA: MIT Press

Tomasello, M., Carpenter, M. and Liszkowski, U. (2007). "A new look at infant pointing". *Child Development*, Vol. 78, pp. 705-722

Tomasello, M., Call, J. and Hare, B, (2003). "Chimpanzees understand psychological states – the question is which ones and to what extent". *Trends in Cognitive Sciences,* Vol. *7*, pp. 153-156.

Tomasello, M. and Zuberbühler, K. (2002). "Primate Vocal and Gestural Communication", in C. Allen,M. Bekoff and G. Burghardt, (eds.). *The Cognitive Animal*, MIT Press, pp. 293-299.

Tomasello, M. (1999). *"The Cultural Origins of Human Cognition"*. Cambridge, MA: Harvard University Press.

Tomasello, M. and Call, J. (1997): *"Primate Cognition",* Oxford, UK: Oxford University Press

Vihman, M.M. (1996). "*Phonological Development – The Origins of Language in the Child"*. Oxford, UK: Blackwell Publishers

Vihman, M.M. and De Boysson-Bardies, B. (1994). "The nature and origins of ambient language influence on infant vocal production and early words". *Phonetica*, Vol. 51, pp. 159-169

Vihman, M.M, Ferguson, C.A. and Elbert, M. (1986). "From Babbling to Speech: Common Tendencies and Individual Differences". *Applied Psycholinguistics*, Vol. 7, pp. 3-40

Warlaumont, A.S., Oller, D.K. and Buder, E.H. (2010). "Data-driven automatic analysis of human infant vocalisations using neural network tools". *Journal of the Acoustic Society of America*, Vol. 127 (4), pp. 2563-2577

Winter, P., D.W. Ploog and J. Latta (1966). "Vocal repertoire of the squirrel monkey (Saimiri sciureus), its analysis and significance". J*ournal of Experimental Brain Research*, Vol. 1,pp. 359-384

Wolff, P. (1969). "The natural history of crying and other vocalisations in early infancy". In B. Foss (ed.), *Determinants of Infant Behaviour*, London, UK, Methuen, pp. 81-109

Zuberbühler, K. and Wittig, R.M. (2011): "Field experiments with nonhuman primates: a tutorial". In J.M. Setchell and D.J. Curtis (Eds), *Field and Laboratory Methods: A Practical Guide*, Cambridge, UK: Cambridge University Press, pp. 207-224

Zuberbühler,K. (2006). "Language Evolution: The origin of meaning in primates", *Current Biology*, Vol. 16 (4), pp. 123 -125

Zuberbühler, K. (2003). "Referential signalling in non-human primates: Cognitive precursors and limitations for the evolution of language". *Advances in the Study of Behaviour*, Vol. 33, pp. 265 – 307

Zuberbühler, K.(2000). "Referential labelling in wild Diana monekys", *Animal BehaviourI,* Vol. 59, pp. 209-220

Zuberbühler, K., Noë, R., and Seyfarth, R. M. (1997). “Diana monkey long-distance calls: messages for conspecifics and predators”. *Animal Behaviour*, Vol. 53, pp. 589-604

Appendix: Ethical approval for empirical studies

Page 1 of 1

Main Identity

From: "psyethics" <psyethics@st-andrews.ac.uk>
To: <vk29@st-andrews.ac.uk>
Cc: "Juan Carlos Gomez" <jg5@st-andrews.ac.uk>; "Klaus Zuberbuhler" <kz3@st-and.ac.uk>
Sent: 02 December 2008 10:35
Subject: PS5112_Kersken Ethical Approval

01 December 2008

Ethics Reference No: *Please quote this ref on all correspondence*	**PS5112**
Project Title:	Prelinguistic vocalisations in infants: Investigating context specificity and co-occurrence with gestural behaviour
Researchers Name(s):	Verena Kersken
Supervisor(s):	Juan-Carlos Gomez & Klaus Zuberbuhler

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 1 December 2008. The following documents were reviewed:

1. Ethical Application Form 01.12.2008
2. Consent Form 01.12.2008
3. External Permissions 01.12.2008
4. Letters to Parents/Children/Headteacher etc… 01.12.2008

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for completion within the stated time period. Projects, which have not commenced within the time given must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

On behalf of the Convenor of the School Ethics Committee OR Convener of UTREC

02/12/2008

05 August 2010

Ethics Reference No: *Please quote this ref on all correspondence*	PS6728
Project Title:	Non-linguistic Vocalisations in Human Infants – A Cross-Cultural Comparison
Researchers Name(s):	Verena Kersken
Supervisor(s):	Prof. Klaus Zuberbuhler and Dr Juan-Carlos Gomez

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 12th July 2010. The following documents were reviewed:

1.	Ethical Application Form	05/08/2010
2.	Participant Information Sheet	05/08/2010
3.	Consent Form	05/08/2010
4.	Debriefing Form	05/08/2010
5.	Letters to Parents/Children/Headteacher etc.	05/08/2010
6.	Application for research permit from Ugandan Government	05/08/2010
7.	Enhanced Disclosure Scotland and Equivalent	05/08/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Approval will take effect only after you have registered this research project with the UNCST, and is subject to any further approval required as part of this registration process.

Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Uganda National Council For Science and Technology

(Established by Act of Parliament of the Republic of Uganda)

UNCST

Your ref: .... *Date :* ....

Our ref: ....SS 2404... **20/12/2010**

Ms. Verena Angela Kersken
C/O Budongo Conservation Field Station
P.O Box 362
Masindi

Dear Ms. Kersken,

RE: RESEARCH PROJECT, "NONLINGUISTIC VOCALISATIONS IN HUMAN INFANTS-A CROSS CULTURAL COMPARISON"

This is to inform you that the Uganda National Council for Science and Technology (UNCST) approved the above research proposal on **August 16, 2010.** The approval will expire on **March 16, 2011.** If it is necessary to continue with the research beyond the expiry date, a request for continuation should be made in writing to the Executive Secretary, UNCST.

Any problems of a serious nature related to the execution of your research project should be brought to the attention of the UNCST, and any changes to the research protocol should not be implemented without UNCST's approval except when necessary to eliminate apparent immediate hazards to the research participant(s).

This letter also serves as proof of UNCST approval and as a reminder for you to submit to UNCST timely progress reports and a final report on completion of the research project.

Yours sincerely,

Leah Nawegulo
for: Executive Secretary
UGANDA NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY

LOCATION / CORRESPONDENCE
Plot 3/5/7, Nasser Road
P.O. Box 6884
Kampala, Uganda

COMMUNICATION
TEL: (256) 414-250 499, (256) 414 705 500
FAX: (256)414-234 579
EMAIL: uncst@starcom.co.ug
WEBSITE: http://www.uncst.go.ug

28 January 2010

Ethics Reference No: *Please quote this ref on all correspondence*	PS6104
Project Title:	How do listeners perceive the vocalizations of prelinguistic infants?
Researchers Name(s):	Verena Kersken
Supervisor(s):	Dr Juan Carlos Gomez, Prof. Klaus Zuberbuhler

Thank you for submitting your application which was considered at the School Ethics Committee meeting on the 16th December 2009. The following documents were reviewed:

8. Ethical Application Form	16/12/2009
9. Participant Information Sheet	16/12/2009
10. Consent Form	16/12/2009
11. Debriefing Form	16/12/2009
12. External Permissions	16/12/2009

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

7 April 2011

Ethics Reference No: *Please quote this ref on all correspondence*	PS6104 (Amendment)
Project Title:	How do listeners perceive the vocalizations of prelinguistic infants?
Researchers Name(s):	Verena Kersken
Supervisor(s):	Dr Juan Carlos Gomez and Professor Klaus Zuberbuhler

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 6th April 2011. The following documents were reviewed:

13. Ethical Amendment Form 07/04/2011
14. Advertisement 07/04/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.