# GENERAL INTRODUCTION

One of the first actions a human infant performs after birth is the utterance of a birth cry. From then on, vocalization is a pervasive behavioral pattern in infants which occurs in a broad variety of different situations. Parents pay intense attention to the vocal utterances of infants. Underlying this attention is the assumption that infants' preverbal vocalizations provide listeners with information about the emotional and physical state of the child. In other words, preverbal vocalizations are seen either as communicative signals or as symptoms which point to the physical state and health of the infant.

The aim of this work is to explore aspects of both the communicative function of preverbal vocalizations and the use of preverbal vocalizations for diagnostic purposes. Specifically, I investigated whether preverbal infants express different emotions in their utterances, and whether the early diagnosis of hearing impairment can be improved by analyzing preverbal vocalizations. To lay the foundations for these explorations, I first of all characterized the whole preverbal vocal repertoire based on acoustic parameters and described the development of infant vocalizations during the first year of life.

Before going into detail, I will present some theoretical and practical background which lays the foundations of my investigations and leads to the specific questions I worked on.

## Vocalizations as communicative signals

Communication can be defined as interactive behavior that solves, or that is supposed to solve, a given problem through the use of signals (Todt & Kipper 2003). Communication between infants and their caretakers should exist, because infants are dependent on their caretakerrs (usually their mothers) for nutrition, thermoregulation, transport, and protection for a long time. Thus, their survival may be crucially dependent upon their ability to signal their needs and states to their mothers. The ability to raise offspring successfully, in turn may depend significantly on the mothers' ability to recognize and respond appropriately to her infant's signals (Maesteripieri & Call 1996).

A signal can be defined as any feature of an individual or its behavior shaped by natural selection for use in communication (Krebs & Davies 1993, p. 349). Besides facial expressions (Izard et al. 1980), the most prominent signals young human infants are able to produce are vocalizations (e.g. Oller 2000; Papoušek 1994). Vocal signals, in general, are

useful to transport urgent and fast changing messages. Furthermore, vocalizations have the potential to attract attention even when the visual contact between sender and receiver is interrupted (Krebs & Davies 1993; Todt 1986). This is an advantage that makes vocal communication very important, especially for human infants which are unable to re-establish (visual) contact to their preferred caretaker by themselves, when put aside or left to other persons.

**Vocalizations as an expression of emotions**

In spite of their immense intuitive relevance, there is only limited knowledge about what kind of information in detail is transmitted via vocal utterances. Ever since Darwin's (1872) work on the expression of emotions in humans and animals, vocalizations are believed to be expressions of emotions. Emotions have been explained to be evolved, adaptive mechanisms that facilitate an organism to cope with important events affecting its well-being (Darwin 1872; Scherer 1993). Although there is an ongoing debate about the nature of emotions, there is a growing consensus among theorists that emotion needs to be viewed as a multi-component entity (Frijda 1986; Lazarus 1991; Scherer 1984). The three major components of emotion are: neurophysiological response patterns (in the central and autonomic nervous systems), motor expression (in face, voice and gesture) and subjective feelings. Many theorists also include the evaluation or appraisal of the antecedent event and the action tendencies generated by the emotion as additional components of the emotional process (see Frijda 1986; Lazarus 1991; Scherer 1984; Smith & Ellsworth 1985). The structure of vocal expressions of emotions is supposed to be determined by different influences (Johnstone & Scherer 2000). First, physiological processes, such as respiration and muscle tone may influence vocalizations. For example, increased muscle tension produced by ergotropic arousal can, as a by-product, affect breathing patterns, the functioning of the vocal folds or the shape of the vocal tract. Second, vocalizations can be interpreted by the social environment and thus have the potential to become a signal. Once expressive motor behavior starts to serve as a signal, functionally based movements will be shaped selectively for the purposes of communication, through processes such as ritualization, formalization, or symbolization. Additionally, the transmission conditions of a given environment (e.g. noise) can shape the structure of vocalizations (Johnstone & Scherer 2000). The subjective component of emotions, that is the perception of feelings is only provable for humans of an age at which they can report about their inner life. However, the conscious experience of feelings is not absolutely necessary for the basic functions of emotions. Therefore it is

possible to study emotion similarly in organisms that have defensible consciousness (humans) and in those for which consciousness cannot be proven (human infants and non-human animals (e.g. LeDoux 1994; 1996).

Support for this view comes also from studies concerning the neurobiology of vocal production. Panksepp (1994) notes that cognition depends on the neocortex and hippocampus. However, neurobiological studies show that vocal expression of emotions does not require cortical structures on all levels of complexity of vocal communication. The control of genetically determined vocal reactions (for example, shrieking to pain in humans) seem to be limited to the brain stem. Cortical structures only become important on the higher levels of vocal communication, that is the voluntary initiation and inhibition of innate vocal reactions, and the learning of new vocal patterns (Jürgens 1992; 2002).

Since the proof of consciousness is neither indispensable for basic emotional processes nor for basic vocal expressions of emotions, I will use the term 'emotion' regardless of whether referring to humans or non-humans and will not principally distinguish 'emotion' from 'affective state' or 'internal state', but use these terms synonymously.

**Vocal expression of emotions in adults**

Since the estimation of the emotional state is easier in human adults, the investigation of the expression of emotions in adult vocalizations constitute an important background for studying infant vocalizations. Therefore, I will give a brief description of the vocal expression of emotion in adults.

The studies which deal with the vocal expression of emotions in adults focus on two major questions: Do listeners identify emotions from vocal cues, and which are the specific acoustic structures encoding specific emotions? (For reviews, see Bezooyen 1984; Frick 1985; Murray & Arnott 1993; Scherer 1979). The identification of emotions has been studied, for example, with sentences spoken by professional actors with different emotions (Banse & Scherer 1996; Bezooyen 1984), with single words (Hammerschmidt et al. submitted; Leinonen et al. 1997) or with natural speech uttered under different levels of stress (e.g., recordings of pilots in crashing airplanes and utterances during stressful psychological tests; Ruiz-Miranda et al. 1998; Williams & Stevens 1972). Ratings of the vocalization by listeners showed that they were able to infer vocally expressed emotions much better than by chance (Banse & Scherer 1996; Murray & Arnott 1993). However, to assign specific emotions to particular vocalizations seems to be a difficult task. First, there is different recognition accuracy for different emotions. Accuracy, for example, is much higher for hot anger than for

disgust (Banse & Scherer 1996). Second, there is confusion between emotions that share certain properties. For example, hot anger and panic fear are often confused by listeners, because they are of same intensity, and anxiety and panic fear are often confused because they are of similar quality (Banse & Scherer 1996). It is also difficult to assign a specific acoustic structure to a specific emotion. For example, anger and fear seem to be characterized by an increase in level and variability of the fundamental frequency, increase in high-frequency energy, and range of the fundamental frequency, but joy is characterized by nearly the same acoustic features. Hammerschmidt et al. (submitted) showed that an increase in aversion and an increase in intensity of an emotional state is mainly characterized by similar structural changes of the vocalizations. This may lead to misinterpretations, especially in cases of single calls or words.

**Vocal expression of emotions in infants**

While it is relatively easy to estimate which emotion is experienced by human adults, because they can report their emotions introspectively or can be asked to express different emotions by will, such estimations are difficult for preverbal infants. It is a point of discussion whether in young infants the emotional system is as differentiated as in older children or adults (see Chapter 1, Discussion). However, investigating the expression of emotions in infant vocalizations, one is dependent on assumptions about their internal state. These assumptions can be based on contextual information (e.g., certain situations in which certain emotions are probable), or on comparing the facial expressions of infants with the facial expressions of emotions of adults (e.g., Ekman & Oster 1979; Izard et al. 1980)). A further possibility to judge the emotional state of infants is to use their parents' ratings (Wasz-Höckert et al. 1968), since in a communicative process the receivers of a signal should be able to decode the transported information, if there is any information transmitted.

The few studies that analyzed whether differences in the acoustic structure of infant vocalizations code differences in the underlying emotional state focused mostly on infant cries, since the cry is the most prominent call type of infants ( e.g. Keller & Schölmerich 1987; Muller et al. 1974; Murry et al. 1975; Porter et al. 1986; Wasz-Höckert et al. 1968). These investigations, however, revealed inconsistent results. Some studies found that cries acoustically and functionally fall into distinct categories (birth, hunger, pain and pleasure) that can be decoded by listeners (Keller & Schölmerich 1987; Wasz-Höckert et al. 1968). Other studies suggest that infant crying is a graded signal, mirroring a continuum ranging from arousal to urgency (Brennan & Kirkland 1982; Porter et al. 1986; Protopapas & Eimas 1997;

Zeskind et al. 1985), and that listeners are not able to identify the eliciting stimulus (Muller et al. 1974; Murry et al. 1975, 1977) or the accompanying emotional state (Papoušek 1992). Infant vocalizations other than cries were hardly studied at all with respect to expression of emotions.

**Infant vocalizations as diagnostic tool**

Besides of their communicative function, infant vocalizations are also in focus because of their potential usefulness for diagnostic purposes and risk assessment. Wasz-Höckert et al. (1968) describes that cries from infants with ‚cri du chat' syndrome and Down's syndrome show other acoustic features than cries of healthy infants. Variations in cry characteristics have been documented in infants with hyperbilirubinemia, encephalitis, meningitis, aphyxia, as well as various forms of brain damage. Additionally, cry characteristics have been associated with factors occurring during gestation that place the infant at risk for later handicap, such as premature birth, low birth weight, undernutrition or maternal substance abuse during pregnancy (for an overview, see Lester & Boukydis 1992).

Several studies investigate the influence of hearing impairment on the preverbal vocalizations of human infants (e.g. Locke 1993; Murai 1963; Nakazima 1962; Oller 1980; Papoušek 1994; Stark 1980; Tonkova-Yampol'skaya 1969). There are different ways how a reduced or lacking auditory input can influence the vocal output. If vocal patterns have to be learned by hearing conspecifics, these vocal patterns will be produced imperfectly or not at all by hearing-impaired individuals. Another possibility is that the emergence of certain call types is not dependent on learning, but follows an innate developmental program. Inborn call types should be affected by hearing deficiency if the fine adjustment of the acoustic structure (e.g., loudness, melody contour) or the usage of such call types requires auditory control or must be learned from social partners. There is evidence for both, absence and change of utterances related to hearing impairment in infancy.

The emergence of spoken language clearly requires auditory learning (Pinker 1998). Without therapy, hearing deficiency leads to partial or complete loss of spoken language (Diller et al. 2001; Yoshinaga-Itano et al. 1999). But hearing impairment has an effect not only on the ability to speak, but also the precursors of spoken language, the preverbal vocalizations. Especially one type of vocalization, the canonical babbling, whose onset is between 7-10 months of age in normally hearing infants, is affected by auditory deficiency. In hearing-impaired infants, canonical babbling appears later than in normally hearing infants or does not appear at all. The phonemic inventory in babbling as well as its temporal structure

also differs between hearing-impaired and normally hearing infants (Eilers & Oller 1994; Kent et al. 1987; Oller 2000; Oller et al. 1985; Papoušek 1994; Stoel-Gammon 1988; Stoel-Gammon & Otomo 1986; for more details, see Chapter 2, Introduction).

In the past, research on infant vocalizations focused to a great extent on language development, asking when and how the properties necessary for matured referential speech arise (e.g., Locke 1993; Murai 1963; Nakazima 1962; Oller 1980; Papoušek 1994; Stark 1980; Tonkova-Yampol'skaya 1969). Therefore, the methods for investigating differences in the utterances of normally hearing and hearing-impaired infants were mainly taken from the linguists, describing preverbal utterances phonetically. However, phonetic descriptions imply well-formedness of syllables in infancy, which is not the case at least in the first six months (in normally hearing infants) (Oller 1978; Oller 2000). Thus, investigating infant vocalizations using purely phonetic descriptions may disguise differences between vocalizations ability that can not be grasped using the linguistic terminology. This problem can be evaded by inspecting the acoustic structure of the vocalizations.

One of the scarce studies investigating whether the acoustic structure of infant vocalizations is affected by hearing impairment was conducted by Möller & Schönweiler (1999), who showed that differences in the acoustic structure of crying could be useful in detecting hearing impairment. The acoustic structure of other call types than crying never has been studied with respect to hearing impairment.

Concerning the usage of certain call types by normally hearing and by hearing-impaired infants, it has been reported that hearing-impaired infants produce some call types more often than their hearing peers (Clement & Beinum 1995). But a study by Oller et al. (1985) delivered contradictory results; so until now, this remains unclear and demands further investigation.

Summarizing, there is limited knowledge about the extent hearing impairment influences the acoustic structure and the usage of preverbal utterances. The importance of gaining such knowledge is given by the fact, that infant hearing impairment is relatively widespread (1-2 in 1000 infants in Germany; Garvel & Tocci 1998). Early therapy, including hearing aids and aural rehabilitation training, enhances the therapeutic success (Yoshinaga-Itano et al. 1999).

More knowledge about the consequences of hearing impairment on the preverbal utterances of infants could possibly lead to an earlier diagnosis and more effective therapy. Furthermore, if the acoustic structure of vocalizations in normally hearing infants transports information about the infant's emotional state to its parents, and if this process is dependent

on auditory control, communication between hearing-impaired infants and their caregivers might be negatively influenced long before the age, spoken language emerges normally. For the sake of infants and parents, one should be aware of these possible problems.

**Concept of this study**

This thesis provides a longitudinal study of the preverbal vocalizations of 7 normally hearing and 7 hearing-impaired infants, based mainly on multiparametric acoustic analyses. The infants were regularly recorded in their familiar surrounding in 11 different situations of normal infant life. The normally hearing infants were recorded by their parents, to obtain a comprehensive vocal repertoire of the infants. Each recording session lasted one week. For several reasons (see Chapter 2), it was necessary to alter the recording method for the hearing-impaired infants so that I visited the families at home and made the recordings in the course of one day. The same situations were recorded as in the normally hearing infants. During each recording of an individual situation, the emotional state of the infant was noted. The ratings of the emotional states of all infants were based on the judgment of their parents. The parents could choose between the terms joy, contentment, interest, surprise, unease, anger and pain to describe the emotional state of their infants. These seven emotions were assumed to occur regularly enough in normal infant life without intervention from outside to allow systematic recordings. Since the estimation of the infant's emotional state might have been influenced by rating biases of the parents, we made a cross-check with the situations in with the recordings were made, assuming that certain contexts evoke certain emotions.

Chapter 1 describes the preverbal utterances of the normally hearing group. Since in the literature, there is no description of the whole vocal repertoire of preverbal infants, in acoustic terms, I first carried out a comprehensive acoustic classification of the preverbal vocal repertoire of infants. As a second step, the emergence of the individual call types and the development-related changes in acoustic structure of four of the call types during the course of the first year was investigated. Third, I analyzed whether there are emotion-related changes in acoustic structure of these four call types and, fourth, it was tested, whether the usage of the vocal repertoire differed in different emotional contexts.

In Chapter 2, the vocalizations of the normally hearing infants are compared with those of the impaired ones, focusing on three questions: (1) Are there different call types in the vocal repertoire of normally hearing and hearing-impaired infants? (2) Do the individual call types emerge at the same age in both groups? (3) Does hearing impairment influence the

acoustic structure of the preverbal call types common to both, hearing-impaired and normally hearing infants?

In Chapter 3, the comparison of hearing-impaired and normally hearing infants is continued, concentrating on the question of whether different emotions influence the preverbal vocalizations of the both study groups in the same way. It was asked ( 1) if different emotions influence the acoustic structure of individual preverbal vocalizations similarly in both study groups, and (2) if there are differences in the composition of call sequences uttered in different emotional contexts related to hearing ability.