CHAPTER 3

THE IMPORTANCE OF AUDITORY FEEDBACK FOR THE EXPRESSION OF EMOTIONS IN SINGLE CALLS AND IN CALL-SEQUENCES OF NORMALLY HEARING AND HEARING-IMPAIRED INFANTS

Abstract

The aim of the study was to investigate the influence of hearing impairment on the preverbal vocalizations of infants by comparing utterances of normally hearing (NH) and profoundly hearing impaired (HI) infants. First, we focused on the acoustic structure of the three most common call types. Second, we examined the composition of call sequences. In both cases, we analyzed whether there are general differences in vocalizations between NH and HI infants, and whether different emotional states affect the vocal production of NH and HI infants in the same way. Concerning the acoustic structure, we found that only one call type, the cry, showed subtle, but significant differences related to hearing ability. Emotion-related changes in call structure were the same for both study groups. In contrast, sequence composition was more affected by hearing impairment than the structure of single calls: Independent from the emotional state, HI infants produced some call types (babbling and short cry) less often, and others (coo/wail and croak) more often within a sequence than NH infants. The composition of call sequences uttered by NH infants changed according to emotional context, while there were hardly emotion-related changes in HI infants. These results indicate that the acoustic structure of preverbal call types is to a great extent predetermined, while the composition of call sequences is influenced by auditory input.

Introduction

Vocal communication in humans can be seriously affected by the lack of auditory input. We have only few knowledge about whether the communication between parents and their infants is negatively influenced when infants are hearing-impaired, but it is well known that the speech abilities of children who received reduced or no auditory input in infancy are poor or completely absent. Hearing impairment in human infants is relatively widespread (1-2 of 1000 infants in Germany; Garvel & Tocci 1998). To develop spoken language despite hearing deficiency, therapeutic interventions are required (Diller et al. 2001; Yoshinaga-Itano et al. 1999) , and it is known that the early start of therapy is one of the decisive factors for its success (Diller et al. 2001; Yoshinaga-Itano et al. 1999). For early diagnosis and effective therapy, it might be important to know, whether and to what extent the early, preverbal vocalizations are influenced by hearing impairment. A number of studies, therefore, compared the vocal ontogeny of normally hearing and hearing-impaired infants (e.g. Maskarinec et al. 1981; Stoel-Gammon 1988; Stoel-Gammon & Otomo 1986). These studies mostly focused on the development of phonation and articulation with respect to features of matured speech. Thus, based on linguistic methodology, preverbal vocalizations were mainly characterized using phonetic descriptions. Since phonetic descriptions imply well-formedness of syllables in infancy, which is not the case in at least the first six months (in normally hearing infants), Oller (1978; 2000) recommended another approach to investigate infant vocalizations. His approach apart from the description of some phonological features, includes additionally a description of the acoustic structure of infant vocalizations.

Studies using this new approach led to the insight that especially one type of preverbal vocalization, the canonical babbling, is useful to judge auditory function. Canonical babbling is characterized by true consonant-vowel repetitions with regular timing between the consonant and vowel portions of the syllable; it is emerging between 7 and10 months of age in normally hearing children. The canonical babbling of infants with severe to profound hearing impairments differs from that of normally hearing infants in a number of ways (Eilers & Oller 1994; Oller 1980; Oller et al. 1985). In hearing-impaired infants, the onset of canonical babbling is much later (about 11-49 month of age), the variety of phonemes used is reduced, hearing-impaired infants babble less often and the transitions between the consonant and the vowel are significantly longer than those of normally hearing infants.

In our longitudinal study (Scheiner et al. in press), using a methodological approach based on the acoustic structure of the vocalizations, we found similar results to those of Oller and colleagues. The emergence of babbling was highly dependent on the auditory function. However, we found no differences in the emergence of other preverbal utterances. Both, normally hearing and hearing-impaired infants, showed a similar repertoire and the same time of emergence of preverbal utterances, with the exception of babbling. In addition, the preverbal utterances of normally hearing and hearing-impaired infants showed only minor changes in the acoustic structure during the first year of life. Significant differences in the acoustic structure of NH and HI vocalizations were only found in one of three examined call types, in the call type cry. Relations between hearing ability and the acoustic structure of cries were also found by Möller and Schönweiler (1999). These authors found that HI cries had a longer call duration, lower energy in the bands 2-4 kHz and 6.4-9.5 kHz and a more complex melody contour than NH cries. No differences between HI and NH cries were found in fundamental frequency and in tonality (percentage of harmonic to nonharmonic time segments). Like Möller and Schönweiler, we found no differences in fundamental frequency or tonality but slight differences in the melody contour (Scheiner et al. in press). In addition, we found a trend towards higher values in acoustic parameters describing energy distribution in the cries of HI infants. However, while Möller and Schönweiler exclusively inspected distress cries, we included also cries uttered in positive emotions in our analysis, and these ‘positive’ cries showed the largest increases in energy parameters. In other words, compared to the prominent emotion-related differences in the energy distribution found in the vocalizations of NH infants (Scheiner et al. 2002), the differences between utterances of NH and HI infants appeared to be subtle in the case of cries and non-existent in the case of the other call types (Scheiner et al. in press).

First inspections of emotion-related variations in acoustic structure in the vocalizations of HI infants suggested that there are no general differences in the way, how different emotions influence the acoustic structure of individual call types in NH and in HI infants (Scheiner et al. in press). However, until present we did not compare the emotion-related changes in the structure of NH and of HI vocalizations in detail. Therefore, one question in the present paper is whether emotion-related differences in the acoustic structure of individual call types are similar for normally hearing and hearing-impaired infants.

Infants usually do not utter single vocalizations, but streams of vocalizations. Until now, there is only minor knowledge about whether the detailed composition, that is, the sequential and temporal organization of call sequences, is influenced by auditory deficiency.

An influence of hearing impairment on the temporal organization of call sequences is indicated by the finding that rhythmic patterns in cry bouts and babbling seem to differ between NH and HI infants. Möller and Schönweiler (1999) found that cry bouts of HI infants had lower rhythmic frequencies than cry bouts of NH infants. As already mentioned above, Eilers and Oller (1994) found that in babbling sequences the transients between the consonant and the vowel are significantly longer in HI infants than in NH infants.

The few studies which investigated the influence of hearing impairment on the frequency of producing certain call types revealed inconsistent results. Oller and colleagues (1985) showed that the relative frequency of various vocalizations was the same in one deaf and eleven normally hearing infants, while Clement and Beinum (1995) found that HI infants produced some call types more often than their hearing peers. However, Clement and Beinum as well as Oller and colleagues did not compare call sequences uttered in different emotional states. In both studies the infants were either engaged in face-to-face interactions with their mothers or they played with toys while being recorded, suggesting that they experienced some kind of positive emotion. Other investigations of call sequences have shown that the composition of call sequences uttered by NH infants differs according to the emotional context their parents ascribe to them (Scheiner et al. 2002). Call sequences that NH infants uttered in negative emotional context (anger and unease) were characterized by higher rates of cry, hic and ingressive vocalizations than sequences uttered in positive emotional context (joy, contentment and interest). Positive emotions, on the other hand, showed a significantly higher rate of babble, laugh and raspberry. Until now, we did not directly compare sequences of NH and HI infants. Therefore, the second aim of this study is to investigate, whether there are differences in the composition of sequences uttered by NH and HI infants and to analyze whether sequence composition is influences by different emotional states in the same way in NH and HI infants. In brief, this paper focuses on the questions of whether the encoding of emotions in the acoustic structure of individual preverbal vocalizations is similar in NH and HI infants and whether sequence composition is influenced by hearing impairment.

Material and Methods

Subjects

Normally hearing (NH) infants

The 7 infants selected to participate in the investigation consisted of 5 boys and two girls, all members of middle-class families. All parents of the infants were native German speakers. All infants were born at term and healthy. Contact was made through two cooperating pediatricians, who asked the parents in the third medical check-up of the infants at week four to six whether they are interested to participate in the study. Then they were visited by one of us in order to obtain their informed consent. Thereafter the infants were examined in the Department of Phoniatrics and Pedaudiology of the Georg-August University of Göttingen to make sure that they were normally hearing. The examinations included a complete otorhinolaryngological status, sound field audiometry, tympanometry, acoustic reflex threshold and measurement of transient evoked otoacoustic emissions.

Hearing-impaired (HI) infants

The group of hearing-impaired infants also consisted of five boys and two girls. The parents of one boy spoke Turkish; the parents of two boys were hearing impaired but capable to speak German; the parents of the other four infants were native German speakers. All infants were born at term and had no further anomalies besides of hearing impairment. Contact was made through cooperating physicians, who procured contact to the parents of the HI infants after making the diagnosis of profound hearing impairment (about100 dB ore more hearing loss on both ears, one infant had a hearing loss of 80 dB on the right ear and >100 dB on the left ear). All infants were provided with hearing aids on both ears soon after diagnosis and received aural rehabilitation training. The first recording was made before the provision with the first hearing aids, afterwards all infants had hearing aids which they wore more or less regularly. Two infants, HI 1 and HI 6 were provided with a cochlear implant on one ear after the study, because their hearing capacities were not satisfactory with hearing aids. HI 1 got its cochlear implant after the study, HI 6 after the 5th recording.

Vocal recordings

The vocal recordings used in this study originate from longitudinal recordings made within the scope of a comprehensive study carried out to broadly investigate the vocalizations of normally hearing and profoundly hearing-impaired infants (Scheiner et al. 2003; Scheiner et al. in press).

The vocalizations of the infants were recorded 6-8 times during the course of one year. For detailed description of the original recording schedule see, Scheiner et al. (in press).The recordings were made with Sony WM TCD-100 DAT recorders and Sennheiser directional microphones (K6 power module and ME64 recording head). To obtain a comprehensive vocal repertoire of the infants, the parents of the NH infants themselves recorded their children in familiar surrounding, after an introduction into the recording method. Each session lasted one week and contained recordings of vocalizations from 11 defined situations of normal infant life (Scheiner et al. 2003). The parents were instructed orally and in written form, how to record the situations. Each of the situations had to be recorded twice during the course of one week. For each recorded situation, the parents had to name the emotion they assumed their infant expressed, choosing between joy, contentment, interest, surprise, unease, anger and pain.

Though we had planned to ask the parents of the HI infants to record their children in exactly the same way, we were forced to alter the recording method. Some of the parents were not able to record the vocalizations of their children, partially due to their own hearing impairment, partially due to the stress induced by the diagnosis of hearing impairment of their infant and the following frequent appointments with physicians, therapists and hearing aid acousticians . We, therefore, changed the recording method for all HI infants in the way that one of us visited the families at home and made the recordings in the course of one day. The same situations were recorded as in the NH infants, and again the parents named the emotion they assumed their infant expressed.

Acoustic analysis

In order to extract acoustic parameters correlating with the emotional state, we carried out a multi-parametric analysis. First, the vocalizations were inspected for quality and then digitized, using RTS 2.0 (Engineering Design, Belmont, Mass., USA). Only calls of good quality and low background noise were used. Depending on the quality of the recordings, we selected 20-30 calls from each recording (more specifically, we chose this number of calls from each recording of the 11 defined situations; since each situation had to be recorded twice, we optimally acquired 20-30 calls out of each of 22 recordings per infant and month). This resulted in a total sample size of about 31,400 (NH: n = 16,300; HI: n = 15,100) vocalizations. If the recording of a situation contained more than 20 calls of good quality and low background noise, half of the calls digitized were chosen from the beginning of the recording and the other half from the end. Sampling frequency was 30 kHz. For each call, we calculated two fast Fourier transformations (1,024 pts; Signal 3.0, Engineering Design) at a frequency range of 4 and 12 kHz (frequency resolution 10 and 29 Hz, respectively). Time resolution was 10 ms in both cases. The resulting frequency-time spectra were analyzed with LMA 9.2 (developed by K. Hammerschmidt). LMA is a software tool to extract different sets of call parameters from acoustic signals (Hammerschmidt 1990; Hammerschmidt et al. 2000; Schrader & Hammerschmidt 1997). We used the spectra with the better frequency resolution (frequency range: 4 kHz, frequency resolution: 10 Hz) to calculate the fundamental frequency and parameters related to fundamental frequency and its variations. For the calculation of parameters describing the energy distribution, we used the spectra with the higher frequency range (frequency range: 12 kHz, frequency resolution: 29 Hz). Parameter calculations were carried out in the same way for the vocalizations of NH and HI infants and are described in detail in Scheiner et al. (2002).

Call Types

The categorization and analysis of the vocal repertoire of the NH infants is described in detail in our previous publication (Scheiner et al. 2002). We classified 11 expiratory and one inspiratory call types (namely: cry, short cry, coo/wail, moan, whoop/squeal, babble, hic, laugh, groan, croak, raspberry and ingressive vocalization (IV); see Scheiner et al. 2002; Scheiner et al. in press). Utterances which did not fit in this classification (mostly utterances mixed out of two or more call types) were put into a rest group.

TABLE 1. *Arrangement of age groups. The figures in columns HI 1-7 refer to the number of recordings made in the corresponding age group.*

Age groups (No.)	Age, weeks	HI 1	HI 2	HI 3	HI 4	HI 5	HI 6	HI 7	Infants recorded, n
1	9-16	1	0	0	0	0	0	0	1
2	17-24	1	1	1	1	2	1	1	7
3	25-32	1	1	1	1	0	1	1	6
4	33-40	2	1	2	1	1	1	1	7
5	41-48	1	1	2	1	2	1	1	7
6	49-56	1	2	1	0	1	1	2	6
7	57-64	0	1	1	1	1	0	1	5
8	65-72	0	1	0	1	0	0	0	2
9	73-80	0	0	0	1	0	1	0	2

.

Age Groups

Since the ages of the HI infants were not exactly the same in corresponding recording sessions, due to various organizational reasons, we formed age groups for better comparability (Table 1). Each age group spanned 8 weeks.

Differences in call structure related to emotion

We examined the acoustic structure of three of the most frequent call types (coo/wail, moan and cry), for which we had enough calls out of the emotional contexts joy, contentment, unease and anger. The other emotions (interest, surprise and pain) were too rare to allow a systematic analysis. All three call types selected have an essentially tonal structure.

From our previous studies (Scheiner et al. 2002; Scheiner et al. in press), we knew that infant vocalizations, even of the same call type, have a high variance. This is especially true when they are uttered under different emotional conditions. Furthermore, not all of the infants showed the whole range of emotions during one recording session. We were forced, therefore, to take utterances out of more than one session for analysis of differences in call structure related to hearing ability and emotion.

We did not hesitate to pool vocalizations recorded at different ages of the NH infants, because we knew from the previous study that age has only a minimal influence on the acoustic structure of these call types in NH infants (Scheiner et al. 2002). In HI infants not only age possibly influenced the call structure, but also the experience with hearing aids. For obvious reasons, the HI infants were supplied with hearing aids as soon as possible after their impairment was recognized (see above). For that reason, we had only one recording of each infant without hearing aids.

In order to find out whether pooling of vocalizations of the HI infants before and after provisioning with hearing aids was an acceptable procedure, we conducted an initial test on whether the supply with hearing aids had an influence on the call structure. Age group 2 (17-24 weeks, no hearing aids) was tested against age group 5 (41-48 weeks, wearing hearing aids for 13-36 weeks), with respect to the call types coo/wail, cry and moan. Separate tests were conducted for the call types out of positive and negative emotions to reduce possible influences of emotions on call structure. For the call type moan, we did not have enough recordings in negative emotional context for all infants; moan, therefore, was tested only for positive emotions. Multivariate repeated measurement tests (GLM, SPSS 10) showed that there were no significant differences between the first recording (HI infants without hearing aid; age group 2) and age group 5 (infants had a longer experience with their hearing aids, except for 2 infants, see above) in any of the call types (cry-positive: $F = 4.83$, $p = 0.8$; cry-negative: $F = 12.05$, $p = 0.217$; coo/wail-positive: $F = 2.06$, $p = 0.488$; coo/wail-negative: $F = 0.43$, $p = 0.822$; moan-positive: $F = 2.03$, $p = 0.488$). The F-values, in contrast, point to an identical structure of the vocalizations uttered before and after wearing hearing aids. Therefore, we decided to pool several recordings of the HI infants as well to have a more even sample for the following tests.

To establish a balanced data set, we only used HI calls of age groups 2, 4 and 5 (Table 1). So, HI calls were used from 17 to 48 weeks of life (mean 30.1 weeks). The calls of the NH infants included in this analysis were uttered between 7 and 58 weeks of life (mean 33.5 weeks). Calls were balanced with respect to the infants and to the four different emotions, but not with respect to age group (see above).

With this balanced data set, we performed a principal component analysis to reduce the number and the correlation between the different acoustic measurements. The principal component analysis performed on the 88 original acoustic variables generated 16 factors with an eigenvalue greater than 1. These 16 factors explained 76.2% of the total variance. The varimax rotation found interpretable results for the first eight factors. Therefore, we used these eight factors, which all had an explained variance above 3, for further statistical tests. A description of the eight factors is given in Appendix 1

Based on the factor loadings, we calculated the means per call type, emotion and infant. With these means, we tested the general hypothesis of differences in acoustic structure related to hearing ability and related to the four emotions. For these tests, we used a multivariate general linear model test (GLM, repeated measures, SPSS 10), and did subsequent univariate tests in case the multivariate tests were significant.

For investigating in more detail which emotions can be differentiated by acoustic structure, we conducted further univariate tests (GLM repeated measures, SPSS 10) comparing (1) positive (joy and contentment) and negative emotions (unease and anger) and (2) each possible pair of single emotions. All tests were conducted (a) for all infants ($n=14$) and (b) for NH and HI infants separately.

Differences in call sequence composition

For the investigation of call sequences, we used only recordings made between the age of 29 weeks to 39 weeks of six NH and six HI infants (we had not enough sequences in each emotion for NH 5 and HI 6). In this age normally hearing infants are in the developmental stage of canonical babbling (Oller 1978), that means, all of the preverbal call types are established. We limited the analyses on this 10 week period to reduce confusion with age-related changes in vocal repertoire.

Since infants show high intersession variability in the production of certain call types (Oller et al. 1985), that is, they often do not produce every call type, they are able to produce, in every recording session, we tried to take three sequences per infant out of the emotions joy, contentment, unease and anger, each. Of HI 5, two contentment sequences, one unease sequence and two anger sequences are missing. Sequences uttered in situations where the parents ascertained one of the other emotions to their infants were not used, because we had not enough recordings for each of the infants within this time period.

The start of a sequence was defined as the start of the first infant utterance in a recorded situation. The end of a sequence was defined at 60 seconds after the start or, if the recording ended earlier, at the end of the recording. Most sequences (126 of 139) had a duration of 60 seconds, 13 sequences had a shorter duration, ranging from 30 to 58 seconds (mean 44.4 s; Table 2).

***TABLE 2**. Overview of sequences*

	NH				HI			
	positive		negative		positive		negative	
N sequences	36		36		34		33	
	calls (n)	**dur (s)**	**calls (n)**	**dur (s)**	**calls (n)**	**dur (s)**	**calls (n)**	**dur (s)**
total (all sequences)	836	2156	905	2026	796	2040	1062	1915
min	5	56	5	30	8	60	4	35
max	56	60	77	60	67	60	92	60
mean	23,2	59,8	25,1	56,3	23,4	60	32,2	58
med	22,5	60	22,5	60	19	60	30	60

Shows the number of analyzed sequences in positive emotions (joy & contentment) and negative emotions (unease & anger), the total number of calls uttered in these sequences, as well as the total duration of these sequences. Additionally, the minimum, maximum, mean and median number of calls within a sequence, and the minimal, maximal, mean and median duration of the sequences are given.

For each sequence, we counted (1) the total number of calls, (2) the number of each call type, (3) the number of different call types, and (3) the number of transitions from one call type to another (e.g., the sequence ‘cry-cry-cry’ contains no transition, while the sequence ‘cry-hic-cry’ contains two transitions).

Following this, we related the number of each of the 12 call types, the number of different call types (=call type heterogeneity), and the number of transitions to the total number of calls within the sequence (to correct for vocal activity). All further analyses were made with these coefficients (in the following, we also use the term 'variables' for these coefficients).

For the statistical analysis, we first used stepwise discriminant functions to identify the most important variables describing (a) the emotion-related and (b) the hearing-related differences between call sequences. Discriminant function analysis identifies a linear combination of quantitative predictor variables that best characterize the differences among groups (Bortz 1993). Variables are combined into one or more discriminant functions. Variables that fail a tolerance test, i.e., variables that represent an almost linear combination of other variables, do not enter the analysis.

To find differences in call sequences related to emotion a stepwise discriminant function analysis was carried out to identify the most important variables describing differences between the four emotions joy, contentment, unease , and anger. For this test, we did not differentiate between the sequences of hearing and hearing-impaired infants. The mean values per infant and emotion of the resulting variables (proportion of cry and proportion of laugh) were used to carry out a multivariate repeated General Linear Model (GLM, SPSS 10) test.

Subsequently, we conducted univariate tests (GLM repeated measure, SPSS10) to test for emotion-related differences (1) between sequences uttered in positive and negative emotional contexts, (2) between sequences uttered in individual positive and negative emotions (joy/unease; joy/anger; contentment/unease; contentment/anger), and (3) between sequences uttered in the two positive emotions (joy/contentment) and in the two negative emotions (unease/anger). All these tests were carried out for NH and HI infants separately.

To find differences in call sequences related to hearing ability, we carried out a second stepwise discriminant function analysis to identify the most important variables describing differences between sequences of the hearing and of the hearing-impaired infants. (In this test, we did not differentiate between sequences out of the four emotional contexts). The mean values per infant and emotion of the relevant variables (proportion of coo/wail and babble, call type heterogeneity, and transitions) were used to carry out a multivariate repeated General Linear Model (GLM) test, in which emotion was the within-subject factor and hearing ability the between-subject factor.

We then calculated for each variable and infant (1) the mean value across all for emotions, (2) the mean values for positive emotions (joy and contentment) and negative emotions (unease and anger), and (3) the mean values for each single emotion. These means were used to carry out univariate tests (GLM, SPSS 10), testing for differences in the composition of call sequences related to hearing ability.

Results

Differences in call structure related to emotion

The multivariate test including four emotions in a repeated measurement design revealed significant differences in the acoustic structure between NH and HI infants only for cry ($F=6.57$, $p=0.026$, Hoteling's Trace, further HT), whereas coo/wail and moan did not show significant differences (coo/wail: $F=1.02$, $p=0.518$; moan: $F=0.94$, $p=0.557$). The same test showed significant within-subject differences regarding the four emotions for all three call types (coo/wail: $F=3.38$, $p<0.000$, HT; moan: ($F=5.74$, $p<0.000$, HT; cry: $F=3.63$, $p<0.000$, HT). Subsequent univariate tests (GLM repeated measures, SPSS 10) revealed significant emotional differences in four factors for coo/wail (F3, F5, F6, F7), in five for moan (F3, F4, F5, F6, F7) and in six factors for cry (F2, F3, F4, F5, F6, F7). For a more detailed investigation, we conducted univariate tests (GLM repeated measures, SPSS 10) comparing either both positive with both negative emotions or each possible pair of single emotions. All tests were conducted either for all infants (n=14) together or for NH and HI infants separately. The results of the tests are shown in Table 3.

For coo/wail the most important factor (F3) showed significant differences between the positive and negative emotions in all cases. The fundamental frequency had higher values for negative emotions. Subsequent posthoc tests showed four or three pairwise significant differences out of the six possible emotional pairs. Only in one case, a positive/positive pair (joy, contentment) revealed significance. Moan showed an inconsistent picture. Only for F3, we got a significant general positive/negative difference - but only for HI infants. In all other cases, the results of the pairwise comparison were inconsistent related to the structural changes, that is, the structural changes in various positive/negative pairs were not uniform. Cry showed differences in six factors. It also showed the most coherent picture. We found differences only between positive and negative emotions, but not between positive and positive or negative and negative emotions. Cries uttered under negative emotions had a significant energy shift toward higher frequencies (F2: e.g. NH joy = -0.063 ± 0.406, NH anger = 0.403 ± 0.393). This was also true for HI infants (see Table 4).

TABLE 3. *Differences in call structure related to emotion.*

Factors	**Explained variance (%)**	***Emotions***								
		Coo/wail			***Moan***			***Cry***		
		both	**NH**	**HI**	**Both**	**NH**	**HI**	**both**	**NH**	**HI**
F1:, Peak frequency (PF), frequency range	20.3									
F2: Distribution of frequency amplitudes (DFA)	9.8							**pos/neg↑** (3)	**pos/neg↑** (2)	**pos/neg↑** (1)
F3: Fundamental frequency	9.3	**pos/neg↑** (4)	**pos/neg↑** (3)	**pos/neg↑** (3,1+)	**Pos/neg↑** (4)		**pos/neg↑** (2)	**pos/neg↑** (2)		(1)
F4: Energy in the high frequencies	5.4				(2)	(2)	(1)	(2)	**pos/neg ↑** (2)	
F5: Trend & modulation of PF	5.3	(2)		(2)	(2)	(1)	(1)	(2)		**pos/neg↑** (2)
F6: Trend & modulation of the first dominant frequency band	3.7	**pos/neg↑** (2)	**pos/neg↑** (2)		(3)	(1)	(1)	(1)		(1)
F7: Duration, tonality	3.4	**pos/neg↑** (1)		(1)	(2)			**pos/neg↑** (2)	**pos/neg↑** (2)	**pos/neg↑** (2)
F8: Location of maximum of DFA or PF	3.4									

'Pos/neg' stands for significant difference between the two positive and the two negative emotions. The direction of the arrows marks the direction of changes from positive to negative emotions. Arrow upwards (↑) means that the negative emotion had the higher factor values. The values in parenthesis give the number of significant pairwise tests between the single emotions. The maximum possible number is six. A plus behind the number stands for differences in positive/positive emotion. We found no significant differences between negative emotions and only one between positive emotions. In all other cases (no plus sign), the pairwise differences are between positive and negative emotions. Univariate repeated measurement test, GLM, SPSS 10).

Additionally, cries with negative emotional context had higher values in the factors F3 and F7. F3 is mainly characterized by the fundamental frequency and their harmonics. Accordingly, cries uttered under negative emotions had a higher fundamental frequency. F7 contains parameters describing the duration and the tonality of a call. To examine, whether the duration-related or the tonality-related parameters were responsible for the increase in F7, we checked the mean values of the parameters enclosed in this factor. We found a significant increase of duration in negative emotions in both infant groups (mean (positive/negative) NH: 838.79 ms/1310.43 ms; HI: 1015.44 ms/1264.09 ms), but no significant increase for the parameters describing the tonality of a call. This means that cries uttered in negative emotions were longer in duration. Further details are given in Table 4.

As mentioned above, cry was the only call type, in which we found differences between NH and HI infants. Independent of the kind of emotion, HI infants had higher peak frequency (F1) and an earlier location of the maximum of the dominant frequency amplitude and peak frequency.

An important result of the present analysis, holding for all three vocalization types, was the overall high variability (see standard deviation in Table 4), independent of whether we focused on different emotions or hearing ability.

TABLE 4. *Mean and standard deviation of the factor loadings of 'cry' in the different emotions for NH and HI infants*

Cry	Normally hearing			
	joy	contentment	anger	unease
F1	0,368 (0,365)	0,303 (0,326)	0,479 (0,53)	0,675 (0,445)
F2	-0,063 (0,406)	0,025 (0,495)	0,403 (0,393)	0,472 (0,475)
F3	0,021 (0,829)	0,103 (0,665)	0,43 (0,468)	0,334 (0,491)
F4	0,265 (0,219)	-0,158 (0,193)	0,309 (0,268)	0,013 (0,234)
F5	0,057 (0,284)	0,297 (0,174)	0,092 (0,436)	0,139 (0,174)
F6	0,247 (0,605)	0,335 (0,335)	0,172 (0,444)	0,773 (0,399)
F7	0,141 (0,363)	-0,009 (0,259)	1,068 (0,537)	0,149 (0,401)
F8	-0,065 (0,32)	-0,059 (0,239)	0,033 (0,245)	0,061 (0,291)
	Hearing-impaired			
	joy	contentment	anger	unease
F1	0,882 (0,572)	0,843 (0,717)	0,758 (0,533)	0,775 (0,679)
F2	-0,118 (0,382)	0,061 (0,294)	0,293 (0,183)	0,27 (0,45)
F3	0,058 (0,348)	0,331 (0,604)	0,292 (0,582)	0,446 (0,475)
F4	0,261 (0,312)	-0,015 (0,37)	0,185 (0,231)	0,076 (0,381)
F5	-0,335 (0,25)	0,231 (0,292)	0,552 (0,417)	0,244 (0,343)
F6	0,07 (0,682)	0,606 (0,488)	0,522 (0,737)	0,684 (0,41)
F7	0,215 (0,253)	0,402 (0,579)	0,786 (0,433)	0,572 (0,36)
F8	-0,24 (0,361)	-0,318 (0,316)	-0,241 (0,342)	-0,363 (0,182)

General differences in the composition of call sequences

As mentioned in the Introduction, besides babbling, all call types can be produced by NH as well as by HI infants. In addition, all call types were produced in positive as well as in negative emotions. We analyzed 139 sequences uttered by six NH and six HI infants given in the same four emotional contexts we used for the description of the structural differences. Table 5 gives the number of infants which uttered at least once the respective call type. As already mentioned, the clearest differences between NH and HI infants was in the use of babbling.

The discriminant function analysis was carried out with all variables to find the best set of variables for the differentiation between sequences produced in the four emotional contexts. The discriminant function analysis came up with two variables, the proportion of cry and laugh. A subsequent multivariate repeated measurement test (GLM, SPSS 10) with the four emotions as within-subject factor revealed significant differences (HT=1.70; F=7.92; p=0.000). A second discriminant function analysis was carried out to find the best set of variables to differentiate between sequences uttered by NH or HI infants. This discriminant function analysis found four variables (proportions of coo/wail, babble and whoop/squeal, and call type heterogeneity). A subsequent multivariate repeated GLM test with emotion as within-subject factor and hearing ability as between-subject factor found significant differences related to hearing ability (HT=3.78; F=6.61; p=0.016). These general tests showed that the composition of call sequences is influenced by both, emotion and hearing ability. Therefore, we first describe how sequence composition was affected by different emotions. Afterwards, we will focus on differences in sequence composition related to hearing ability.

TABLE 5. *Number of infants uttering a certain call type in positive and negative emotions, respectively, in the analyzed sequences*

call type	NH (n=6)			HI (n=6)		
	all emotions	positive emotions	negative emotions	all emotions	positive emotions	negative emotions
coo/wail	6	6	6	6	6	6
moan	6	6	6	6	6	5
cry	6	6	6	6	5	6
short cry	6	6	6	4	2	4
babbling	6	6	5	1	0	1
whoop/squeal	5	5	2	5	5	4
hic	6	6	6	6	3	6
laugh	6	6	6	6	5	5
groan	6	6	6	6	6	6
croak	3	3	0	5	4	5
raspberry	5	5	1	2	2	1
IV	6	5	6	6	5	6
rest	6	6	6	4	4	3

Differences in call sequences related to emotion

The univariate tests revealed that NH infants showed significant differences between call sequences uttered in positive and negative emotions, whereas in HI infants no differences between positive and negative emotions could be found (Table 6). Sequences from NH infants uttered under negative emotions were characterized by higher amounts of cry and ingressive vocalizations (IV), while sequences uttered in positive emotions contain more laughs and croaks. The tests comparing single emotions supported the finding from the test between positive and negative emotion. In contrast to NH sequences, we found in HI sequences only few emotion-related differences, and these did not achieve global significance (Table 3). In NH infants, we did not find the same amount of laugh and croak when comparing contentment with unease or anger, as we found when comparing joy with unease or anger (Table 6). Furthermore, comparing individual emotions, NH infantsshowed more transitions for unease and anger when compared with contentment than when compared with joy (Table 6).

TABLE 6. *Differences in the composition of sequences uttered in positive and negative emotions.*

	NH					**HI**				
	pos/neg	**j/u**	**j/a**	**c/u**	**c/a**	**pos/neg**	**j/u**	**j/a**	**c/u**	**c/a**
Variable										
coo/wail										*↓
moan				*↓						
cry	*↑	*↑	*↑	*↑	*↑			*↑		
short cry										
babbling										
whoop/squeal										
hic									*↑	*↑
laugh	*↓	*↓	*↓						*↑	
groan					*↑					
croak	*↓	*↓	*↓							
raspberry										
IV	*↑		*↑		*↑					
heterogeneity										
transitions				*↑	*↑				*↑	

The first column refers to the variables. Each variable was related to the total number of calls per sequence. Pos/neg: test on differences between sequences uttered in positive (joy, contentment) and negative emotions (unease, anger). J/u, j/a, c/u, c/a: tests of differences between individual positive and negative emotions (j=joy, c=contentment, u=unease, a=anger). ↑Arrow upwards refers to higher frequency of the respective variable in negative emotions. Significant results (p<0.05) of the univariate tests (GLM repeated measures, SPSS 10) are indicated by an asterisk. The test were conducted separately for NH and HI infants.*

When comparing sequences uttered in the two positive emotional contexts (joy and contentment), more differences were found for NH infants than for HI infants. In NH infants, sequences uttered in the context of joy were characterized by higher frequencies of hic, laugh and croak and a lower frequency of moan when compared with contentment. In HI infants only the amount of whoop/squeal differed between joy and contentment. No variables were found to differentiate between sequences uttered in the two negative emotions unease and anger.

Differences in call sequences related to hearing ability

The multivariate test revealed significant differences between sequences of NH and HI infants. Subsequent univariate tests (GLM, SPSS 10) were conducted, using different means per infant and relative frequency of each call type (for all emotions, for positive/negative emotions, and for individual emotions). Regarding the frequency of individual call types within sequences, the tests revealed that in general the sequences of HI infants were characterized by higher rates of coo/wail and croak, and lower rates of babbling and short cry than those of NH infants. Testing per infant for relative frequency of each call type over all emotions we found that HI sequences were characterized by significantly ($p < 0.05$) higher rates of coo/wail and significantly lower rates of babbling. Comparing the mean values for positive emotions, we found in HI sequences significantly higher rates of croak and significantly lower rates of babbling and short cry. Comparing the mean values for negative emotions, HI sequences were characterized by significantly higher rates of coo/wail and croak and by significantly lower rates of short cry. The described differences between NH and HI sequences were found also in individual emotional contexts (Figure1). In those cases, in which they did not reach significance level, they nevertheless showed the same trend.

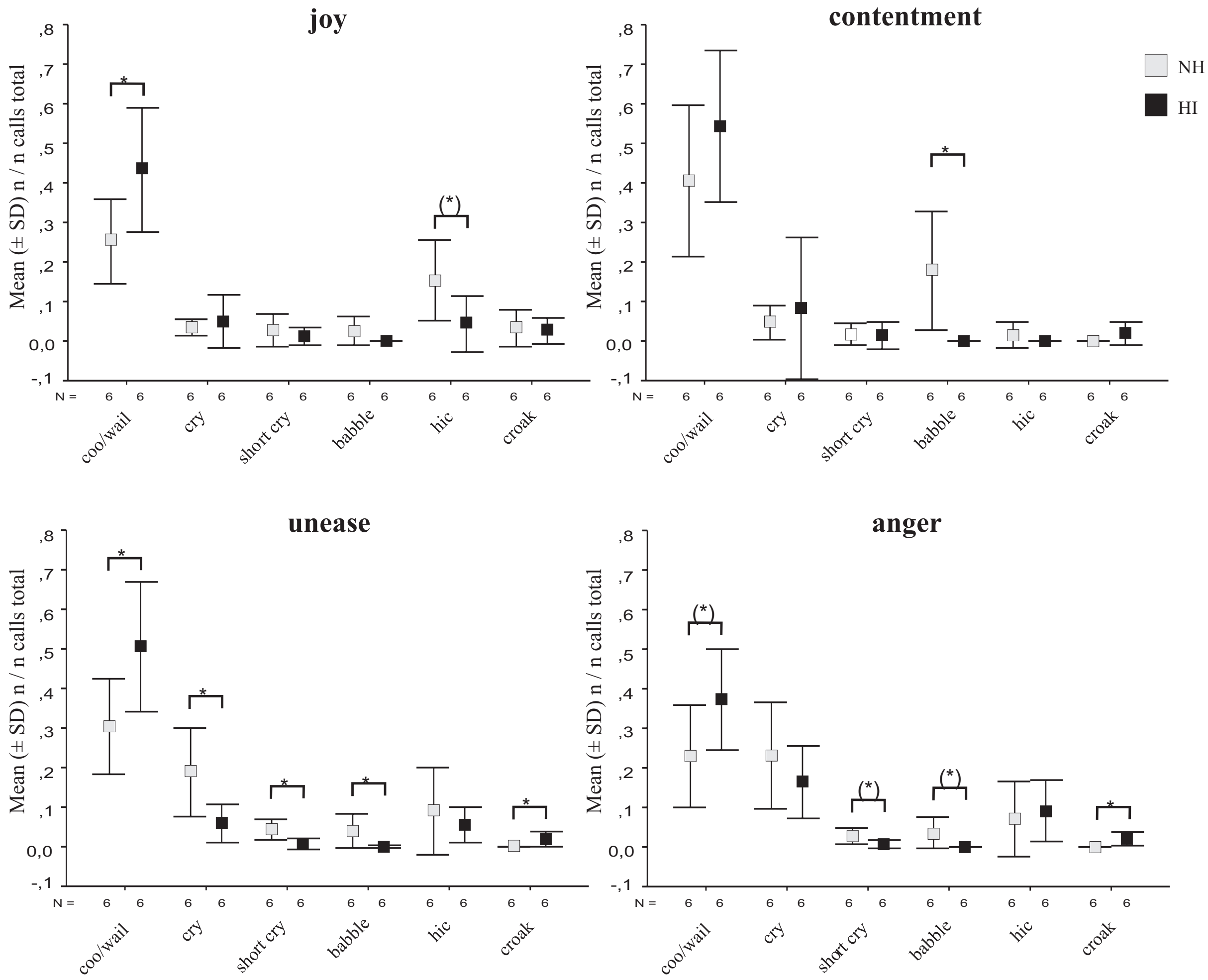

FIGURE 1. Comparison of NH and HI sequences uttered in individual emotional contexts (joy, contentment, unease and anger). Mean and standard deviation of the relative number of the most important call types. *: p<0.05; (*): p<0.1 (GLM, univariate, SPSS 10).

Additionally, we measured the call type heterogeneity (relative number of different call types within a sequence) and the relative number of transitions (changes from one call type to another). The univariate tests (GLM, SPSS 10) again were conducted using different means for all emotions, for positive/ negative emotions, and for individual emotions. They indicated that HI sequences were characterized by lower call type heterogeneity, while transition rate did not differ from NH sequences. Call type heterogeneity became significant when HI and NH sequences were compared for all emotions together. When comparing single emotions, only anger reached significance (Figure 2). A trend (p< 0.1) for lower heterogeneity, however, was also found in sequences expressing joy (Figure 2).

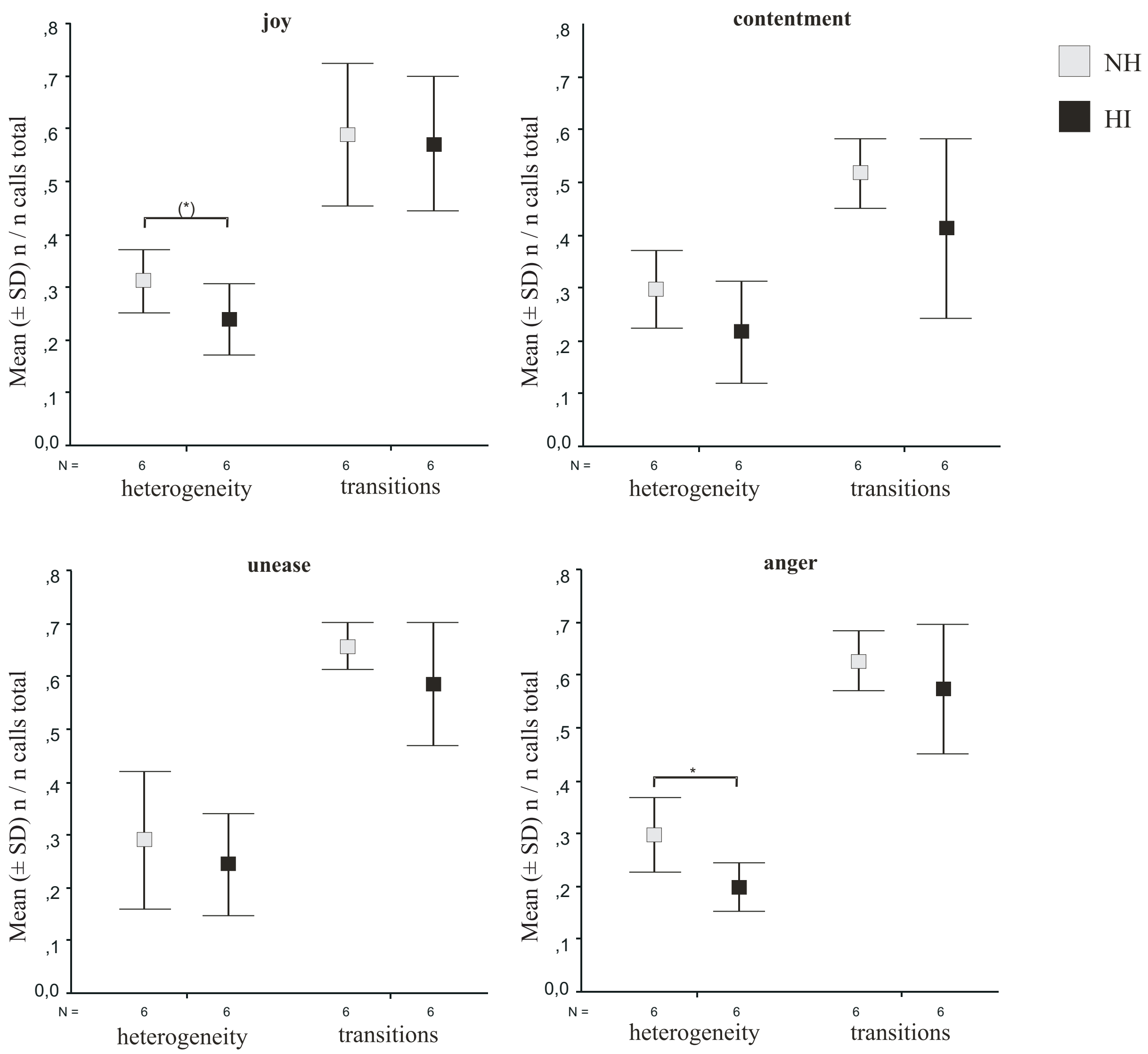

FIGURE 2. Comparison of NH and HI sequences uttered in individual emotional contexts (joy, contentment, unease and anger). Mean and standard deviation of call type heterogeneity and the rate of transitions. *: p<0.05; (*): p<0.1 (GLM, univariate, SPSS 10).

Discussion

The present study compared preverbal vocalizations of normally hearing (NH) and hearing-impaired (HI) infants. In the acoustic structure of individual call types, no general differences related to the emotional context could be found. In the composition of call sequences, in contrast, we found significant differences between NH and HI infants, relating to emotional context. These results indicate that the acoustic structure of single vocalizations is relatively independent of auditory feedback, whereas the composition of call sequences is not.

Call structure

Regarding the acoustic structure of individual call types, only one of the three call types analyzed, the call type cry, depended on hearing ability. HI infants uttered cries that tended to have a higher peak frequency and different melody course. These results are in line with earlier findings from Möller and Schönweiler (1999)and Scheiner and colleagues (2003). In contrast to these more subtle differences between NH and HI infants, we found marked differences in the acoustic structure of vocalizations uttered in positive and negative emotional contexts. Here, NH and HI infants showed the same changes. Vocalizations uttered in negative emotional states, in general, showed an increase in energy in the higher frequencies of the spectrum and had a longer duration Thus, the important information whether an infant feels good or bad is encoded in the acoustic structure of individual call types. Furthermore, encoding this information seems to be independent of hearing ability. The latter finding indicates that one of the most important functions of infants' vocal signalling, namely to signal their needs and states to their mothers (Maesteripieri & Call 1996) is not seriously disturbed by hearing impairment.

In an earlier analysis, done only with NH infants (Scheiner et al. 2002), we found that it is not possible to differentiate between specific positive emotions (joy and contentment) and specific negative emotions (unease and anger) on the basis of the acoustic structure. As it was the case for NH infants, there were hardly any differences in the vocalizations of HI infants expressing joy and contentment and there were no differences in the expression of unease and anger. Combining the vocalizations of NH and HI infants improved the significant differences between positive and negative emotions, but did not improve the possibility to differentiate between single positive or single negative emotions. This indicates that the impossibility to separate vocalizations uttered in positive or negative emotions is not caused by the sample size

There are several possible reasons for the low success to distinguish between specific positive or specific negative emotions. One reason could be that in young infants the emotional system is not as differentiated as in older children or adults. This point is still under discussion (for an overview, see Strongman 1996). Most authors agree that from the very beginning, there is a differentiation into at least two emotional states, aversive and non-aversive (e.g., Giblin 1981; Lewis 1993; Sroufe 1979), while other emotions develop successively. Other authors argue that more than two emotional states can be distinguished in early infancy (Izard & Malatesta 1987; Malatesta-Magai et al. 1991). Our data on the acoustic structure of individual vocalizations support the assumption that there are at least two different emotional states in young infants. A second reason for the low success to separate emotions with the same valence might be that not each vocalization uttered in a given emotional context is typical for this context. A third reason could be that infants are able to express various emotions, but the parents' ratings are not absolutely reliable. Uncertainty might arise, for instance, from interferences between the parents' own mood or expectations and the infant's behavior, or from incoherences in emotional labeling. To examine the latter possibility, we did a cross-check analysis with the vocalizations of NH infants (Scheiner et al. 2002). We analyzed the same calls , but instead of testing differences in acoustic structure related to emotional categories, we tested for differences related to different (emotion-eliciting) situations. This analysis produced the same results as the analysis based on the emotional ratings. Therefore, it seems unlikely that a mismatch between the infants' emotions and the parents' estimations is the main reason for the low dicriminability of related emotions in individual call types (Scheiner et al. 2002).

Composition of call sequences

Regarding the composition of call sequences, we found substantial differences between the sequences NH and HI infants uttered in the age between 29 and 39 weeks.

Our study showed that the composition of sequences uttered by NH infants in positive and negative emotions differed significantly in different emotional contexts, while there were hardly any emotion-related changes in sequence composition of HI infants. This was also true, when comparing sequences uttered in different positive emotions.

Prelinguistic sounds are often considered to be biologically predetermined (Bloom et al. 1993), and the development of vocal learning is assumed to follow an internal program of physiological and cognitive maturation. This is supported by our investigations. The emergence and the acoustic structure of most call types (except babbling) seems to be only minimally influenced by hearing deficiency (Scheiner et al. in press). Thus, all call types found in sequences of NH infants can be produced by HI infants as well. In other words, the given differences between NH and HI infants concerning sequence composition are not due to the fact that the HI infants are not able to produce the respective call types. The lower emotion-related differentiation in sequence composition of HI infants possibly is due to the reduced voluntary control over the succession of call types, caused by their reduced auditory feedback.

Another explanation might be that NH infants somehow accommodate the rate of specific call types to the respective emotional contexts, while HI infants fail to do so.

Feedback from parents, for example, may provide reliable cues about the consequences of vocalizing and serve as a source for learning in the infant (Goldstein & West 1999; Papoušek 1994). Several studies showed that infant vocal production is affected by the reactions of adults. Ramey and Ourth (1971), for example, found that immediate social reinforcement enhances the vocal production of infants, while delayed reinforcement does not. Weisberg (1963) compared contingent and noncontingent social reinforcement and found that contingency had an increasing effect on the vocalization rate, while noncontingent reinforcement had no effect. Other studies (Bloom 1988; Bloom et al. 1987; Masataka 1993) did not find a relation between contingency and total rate of infant vocalizations, but they found that contingent maternal stimulation increases the frequency of specific call types. These studies used combined social stimuli (simultaneous smiling, touching, and speaking to the infants) to influence infant vocal production. Therefore, the specific role of the acoustic stimulation in these studies can not be determined. An exception is the study of Haugan and McIntire (1972), who compared the effects of vocal imitation, tactile stimulation, and food on the vocal behavior of infants. Their results showed that vocal imitation by adults was the most effective reinforcer to enhance infant vocal production.

To sum up, there is evidence that social stimulation influences the vocal behavior of infants. Out of various stimuli, auditory stimuli seem to be the most effective. Therefore, it is possible that infants use parental vocal feedback as the main source to accommodate their sequence composition to the specific situation, in order to reach their goals effectively. This might explain, why there are less emotion-related differences in sequences of infants with hearing deficiency.

A second result of this study was that, if compared directly, NH and HI sequences differ in their composition, independent of the emotional context. Sequences of HI infants are characterized by higher rates of coo/wail and croak, less babbling and less short cries, compared to NH infants. Additionally, there are indications that the sequences of HI infants show less call type heterogeneity, that is, less variability.

It is known that speech utterances of HI children (Most 1994) and HI adults (Letowsky et al. 1993) often show longer durations compared to utterances of NH persons . It is possible that the reduced rate of short cries in the present study reflects a comparable mechanism. In our study, we defined short cries by their shorter duration, based on the bimodal distribution of all cries (Scheiner et al. 2002). Consequently, prolonged short cries would have been counted as cries.

That babbling was rare in HI sequences is not astonishing, since it is well known that the emergence of babbling is highly dependent on hearing ability (Eilers & Oller 1994; Oller 1980; Oller et al. 1985). The differences in the rate of short cry, coo/wail and croak have not yet been described in detail in other studies. However, only few other studies investigated call sequences of preverbal infants at all. One of these studies was conducted by Oller and collaborators (1985). They compared the vocal repertoire of one deaf baby (recordings at the age 8, 11, 12 & 13 months) with the vocalizations of 11 normally hearing infants (age 4-6 months). The recordings were made while the infants played, and while they were engaged in face-to face interactions with the experimenter. Oller and his colleagues found that the relative number of most call types (except babbling) was not markedly different in the deaf baby from that in the hearing sample. Unfortunately, this study is based on the vocalizations of only one deaf infant, and therefore is not representative, since infant vocal behavior is highly variable (Oller & Eilers 1992). Another study, carried out by Clement and Beinum (1995), revealed contradictory results. In that study the vocalizations of six NH and six HI infants, recorded from 2.5 to 7.5 months of age, were compared. Clement & Beinum found that HI infants produced a certain category of vocalizations more often than NH infants. However, Clement and Beinum used only two categories of call types, which they named 'utterance' and 'non-utterance'. Therefore, it remains unclear whether they also found higher rates of coo/wail and croak for the HI infants, as we did.

Locke and Pearson (1992) suggested that HI infants vocalize more than NH infants, because of the extra effort HI infants spend to get auditory feedback. However, the results of this study did not show an overall higher rate of utterances in HI infants, but a higher rate of specific vocal patterns, namely coo/wail and croak, compared with NH infants. Coo/wail, as well as croak, are call types which fit into a group of infant vocalizations named 'protophones' by Oller (2000). Oller distinguishes between protophones and fixed signals. Fixed signals, for example, cries and laughs, have a relatively fixed acoustic structure. Beside their voluntary production, they can be elicited by specific stimuli as well as produced involuntarily. Protophones, in contrast, have no biologically specified values as signals. They are generally not elicited by sign stimuli, and no emotional states are specifically associated with them. According to Oller (2000), protophones emerge directly after birth and develop throughout the first year of life, thereby reflecting the infants' growing ability to produce speech-like sounds. The protophones of each developmental stage can be seen as precursors of the protophones of the following stage. The last stage of protophone development is represented by the emergence of canonical babbling. The higher rate of coo/wail and croak found in the sequences of HI infants in the present study might be due to the fact that the NH infants replaced part of the vocal types of the expansion stage by the higher developed canonical babbling, while the HI infants continued to produce protophones of the stage before.

To sum up, this study revealed that emotions are encoded in the acoustic structure of single vocalizations as well as in the composition of call sequences. The encoding of emotions in call structure is to a great extent independent of the auditory input, while the composition of call sequences seems to be influenced by auditory learning. It remains to be clarified, whether the emotion-related changes in the infants' vocal productions are indeed salient for adult listeners. Furthermore, more detailed investigations on the mechanisms of auditory learning in preverbal infants are clearly needed.

Appendix 1: Description of factors revealed by factor analysis

Factor	Description
F1: Peak frequency (PF), frequency range	F1 combines measures of PF (frequencies with the highest amplitude) and frequency range (maximum difference between highest and lowest frequency above the noise level).
F2: Distribution of frequency amplitudes (DFA)	F2 combines different DFA measures (statistical distribution of frequency amplitudes in the spectrum).
F3: Fundamental frequency (F0)	F3 combines mean F0 together with the level of the 2nd and 3rd dominant frequency bands (DFB), which correspond to the 2nd and 3rd harmonic in the tonal parts.
F4: Energy in the high frequencies	F4 combines different measures describing the increase of energy in higher frequency parts (e.g. ratios between higher DFB's, percentage of higher DFB's).
F5: Trend & modulation of PF	F5 combines trend and modulation measures. Trend measures are calculated on the basis of the linear trend of the PF. Modulation measures are calculated on the basis of the difference between the original and average curves of PF.
F6: Trend & modulation of the first dominant frequency band (DFB)	F6 combines trend and modulation measures. Trend measures are calculated based on the linear trend of the 1st DFB. Modulation measures are calculated based on the difference between the original and average curves of the 1st DFB.
F7: Duration, tonality	F7 combines the measures duration, harmonic-to noise-ratio (HNR) and percentage of tonal parts. The main loadings are duration and max. HNR.
F8: Location of maximum of DFA or PF	F8 combines measures describing the relative position of PF and DFB maximums; calculated as a coefficient between call duration and location of PF or DFB ((1/duration [ms])* max location [ms]).