Abstract: The tracking of resonance frequencies and the analysis of their interaction with the fundamental frequency (F0) allows a description of (pre-)articulatory activity in very young infants. Subjects are six healthy infants. Spontaneous cries were recorded weekly from the 4th until the 20th week. For resonance frequency estimation a spectral parametric technique was applied, which was based on autoregressive models whose order is adaptively estimated on subsequent signal frames [1]. Cry melodies exhibiting different degrees of complexity (e.g. single-arc-melodies, multiple-arc-melodies) were selected for analysis. We found that resonance (formant) tuning occurs much earlier than expected. Here we demonstrate the early occurrence of a tuning between resonance frequencies and the cry melody in infants from 8 weeks onward. A more intense tuning between the melody and the lower resonance frequencies was found beginning about the 2nd / 3rd month. This tuning is interpreted as an early articulatory activity in infant’s crying. In a broader perspective it is attributed to a language-related behaviour preparing formant tuning in speech. Medical applications are seen for infants with disturbances of the vocal tract transfer function, e.g. infants with cleft-lip-palate.

Keywords: cry melody, vocal tract resonance, formant analysis, pre-speech development

I. INTRODUCTION

In a preceding paper [2] we have outlined both, the high control capacity of mechanisms underlying laryngeal sound production in infants, and the interaction between laryngeal (melody) and pharyngeal (resonance frequencies) activity. The results of this earlier study provide good reasons to consider in more detail the resonance properties of the infant’s vocal tract during the earliest phases of pre-speech development.

The hypothesis that cry melody patterns might be direct precursors of melodic features of speech is not new [3-8]. Meanwhile there is good evidence that the development of certain cries (mitigated cries) serves as a preparatory activity for language acquisition [9-11]. Tuning processes between the cry melody and resonance frequencies need a certain training - period before they are at disposal for intentional use, e.g. imitating surrounding speech sounds at the babbling age. Starting about the fourth month of life a rapid expansion of non-cry vocalizations (marginal babbling) occurs, including many vowel-like sounds and near-syllables [12, 13]. So, we should expect that intentional articulatory activity is developed well before this age.

II. METHODOLOGY

SUBJECTS: We investigated six healthy, term-born German infants. All infants were without clinical history of pre- and postnatal illness and free of clinical signs of developmental or hearing disorders.

DATA ACQUISITION: Spontaneous cries of all six infants were recorded in weekly intervals from the 4th to 20th week. Cries were recorded in home environment by trained persons using a SONY-DAT-recorder (TCD-D100). The sampling frequency was 48 kHz and the amplitude resolution was 16 Bit.

DATA ANALYSIS: A set of 100 harmonic cries with a high signal-to-noise ratio was selected for analysis out of a total amount of 2000 recorded cries. Cry analysis was performed in a first step by an evaluation of broad-band and narrow-band spectrograms made with a CSL-4300-Model (Kay Elemetrics Corp., NJ/ USA). In-depth data processing was performed by means of a software tool developed on a PC under Matlab 5 environment at the Dept. of Electronics and Telecommunications, Faculty of Engineering, University of Firenze, Italy. Fundamental frequency F0 is estimated by means of a robust two-step procedure [14]. As for formant estimation, the parametric AutoRegressive (AR) approach is applied. This method is particularly suited for newborn infant cries, which are characterised by higher resonance frequencies than those of adults. Many criteria have been defined for finding the best model order p, including both the estimated variance $\sigma^2$ and the model.
complexity $p$ in one statistics. The DME (Dynamic Mean Evaluation) model order selection criterion is applied here to the decreasing sequence of variance values, on subsequent data frames of varying length [1]. For comparisons, also fixed model orders were tested. We got better formant tracking than traditional approaches [1, 15].

In the figures presented here resonance frequencies estimated with the AR-method are shown, as we found this method producing the most coherent resonance tracks. Note that the resonance frequencies in infant cries (roughly seen as spectrographic amplitude enhancements) are in most cases not yet identical to formants of speech sounds. We call these resonances “R1”, “R2” and “R3”, because it is not yet known how they are related to the vowel formants in later speech.

In order to visualize the interaction between melody and time varying resonance tracks we made a special diagram. This diagram contains a background pattern with the melody and the corresponding harmonics ($F_0 = \text{first harmonic}$) together with the resonance tracks. This representation is well-suited to assess relations between resonance frequencies and harmonics of the melody up to the 7th harmonic of the melody.

III. RESULTS

Here we present typical examples of melodies and the corresponding spectral resonance functions during crying for the age period 8 – 14 weeks. The selected examples demonstrate also developmental changes of tuning processes. In the oral presentation we will present developmental sequences of the mentioned tuning processes for all infants. We will show both, the lack of such tuning in crying of the youngest infants and the step by step development of melody–resonance–tuning in older infants. We found between three and four main resonance frequencies up to 10 kHz within the age range under investigation.

During the first weeks of life, the resonance frequencies (particularly R1) were relatively constant without movements over the central part of the cry. At the age of about 8 weeks already a partial tuning between the first resonance frequency (R1) and the melody is observable.

We could observe relatively longer periods of a strong resonance, where the resonance tracks take a course closely following a certain harmonic of the melody. In contrast to this, there were relatively fast transitions of resonance frequencies from one harmonic to the other. This fact allows us to conclude that there exists a longer time period of coupling of the resonance movement and the melody, which can be interpreted as the action of a neuro-physiological tuning mechanism between cry melody and resonance frequencies.

Fig. 1 displays the first two resonance tracks (R1, R2) together with the cry melody and its harmonics. At the maximum of the first melody arc (at about 0.85 sec), R1 and R2 show a strong resonance peak around the 5th harmonic.

![Fig. 1: First two resonance frequencies (black points) of a mitigated cry from a healthy infant at the age of 8 weeks. They are displayed together with the cry melody (lowest line) and its harmonics. R1 and R2 show a conspicuous convergence toward a punctuated resonance (rectangle) at the 5th harmonic of $F_0$ (bold line).](image)

R1 is moving step-wise from a resonance near the second harmonic at the beginning of the cry to the third harmonic. For about 30 ms R1 is fairly well-tuned with the third harmonic and is then moving toward the 5th harmonic at the maximum of the first melody arc. R2 is relatively constant, exhibiting a resonance tuning at the 7th harmonic for about 200 ms (0.4 – 0.6 sec); then R2 moves to the 6th harmonic. Note that R2 seems to support the R1-melody-tuning by a short down shift to the 5th harmonic exactly at the time point of the maximum of the first melody arc. The coordinated action of R1 and R2 seems to stabilize the melody at its apex and produces a punctuated resonance. This is interpreted as a pre-articulatory training process, which indicates an active neuro-physiologically controlled tuning. However, the infant did not repeat the tuning in the second melody arc. Although at the age of 10 weeks already more complex melodies (multiple-arc-melodies) are produced, we selected for reasons of comparability again cries consisting of a double-arc melody (Fig. 2a, b). In contrast to the punctuated resonance in Fig. 1 (8th week), the resonance track “R1” is coupled to the 6th harmonic over the whole cry (lasting resonance). Only in the transition region the melody-arc-tuning is lost, but immediately with the beginning of the second melody-arc the resonance tuning at the 6th harmonic occurs again. R2 shows an independent course in relation to the melody in both cries (Fig. 2a, b). Both cries of the infant from the same day exhibit the same tuning phenomenon between the melody and R1. The regular recurrence of tuning
events supports the assumption that the observed tuning is not by chance, but a controlled behaviour.

Fig. 2a, b: Resonance frequencies (black points) of two double-arc cries (a, b) from a healthy infant at the age of 10 weeks. In both cries the resonance frequency “R1” is coupled to the 6th harmonic of F0 (bold line). In the transition regions (within vertical lines) between both melody-arcs the tuning is lost (tuning is indicated by a rectangle).

Fig. 3: Enlargement of the region 3.2–3.5 log Hz of Fig. 2a (source spectrum corrected).

In Figure 3, the frequency region between 3.2–3.5 log Hz is zoomed in, in order to demonstrate the good tuning. Note that in Figure 3 a rather precise coincidence of melody and R1 in the strongest resonance regions results when the necessary frequency correction is made taking into account the slope of the laryngeal source spectrum. We did that after we observed that in the strong resonance regions (i.e., a close coupling of melody and resonance) the resonance frequency is mostly situated a certain ratio (approximately 13%) down a harmonic of the melody (Fig. 1, 2a, b, 4). We believe that this small but significant discrepancy of frequency is due to the spectral amplitude slope of the harmonics of the laryngeal source signal.

In Fig. 4b, a selected example of a cry from an older infant (14th week) demonstrates a well-developed tuning between the first two resonance frequencies with harmonics of the melody (Fig. 4b).

IV. DISCUSSION

The results of previous studies [e.g. 16, 17] of resonance frequencies of infant cries are difficult to compare to the present results, because former studies provide only averaged values for resonance frequencies (formants) in pre-speech utterances. In contrast, the present study provides time functions of the resonance frequencies and investigates the interaction between these and harmonics of the melody. Our approach (tested in a preliminary study with twins [2]) allows investigating an (pre-) articulatory activity at a very early age and enables us to characterize developmental processes directed toward language acquisition.

We could confirm our former results concerning developmental changes, but the more fine-grained analysis applied here (coupling analysis of resonance frequency and melody) allows discovering the tuning process in infant’s crying even earlier in life. In the preceding study [2] we found a coupling of the lower resonance frequencies to the melody or its harmonics in infant cries beginning at the age of 15 – 17 weeks of life. In the present study we observed tuning processes between the resonance frequencies and the melody, at least during short parts of the cry, much earlier (8th week). At the age of 8 weeks we observed already coordinated actions of R1 and R2 during short parts of the melody. This behaviour seems to stabilize the melody at its apex. Later stages of development are mainly characterized by longer well-tuned times between the first two resonance frequencies and the melody. Already two weeks later the resonance frequency “R1” is coupled to a harmonic over the whole cry. Only in the transition
region between two melody-arcs the tuning is lost. At the age of 14 weeks, a well-developed tuning between the first two resonance frequencies with the melody regularly occurs and reflects a neuro-physiological maturation.

V. CONCLUSION

The regularly occurring tuning as observed is probably a result of four factors: Firstly, the preceding “training” at earlier weeks. Secondly, the anatomical restructuring of the supra-laryngeal vocal tract at about 3 months [4], thirdly, a better control of sub-glottal air pressure at about 3 months of life [21], and fourthly, a co-ordination between laryngeal and pharyngeal activity. At this age also more voluntary phonation occurs and the infant has to co-ordinate and exercise sub-glottal air pressure and laryngeal-pharyngeal control. The “training” idea, suggested by Philip Lieberman already in 1986, is confirmed by our results of the present study. These tuning processes have undoubtedly a preparatory function for intentional articulatory activities at later ages. The analysis presented here supports strongly the assumption of a continuous development from the first infant’s sound productions to speech. So, this findings further support that cry development is an integral part within pre-speech development.

REFERENCES


