INTERACTION PATTERNS BETWEEN MELODIES AND RESONANCE FREQUENCIES IN INFANTS' PRE-SPEECH UTTERANCES

Kathleen Wermke[1], Werner Mende[2], Anne Kempf[1], Claudia Manfredi[3], Piero Bruscaglioni[4], Angelika Stellzig-Eisenhauer[1]

[1]Center for Pre-Speech Development & Developmental Disorders, Department of Orthodontics, Julius-Maximilians-University Wuerzburg, Germany; [2]Berlin-Brandenburg Academy of Science, Berlin, Germany; [3]Dept. of Electronics and Telecommunications, Faculty of Engineering, University of Firenze, Italy; [4]Dept. of Physics, Faculty of Mathematics, Physics and Nat. Sci., University of Firenze, Italy

INTRODUCTION

Melody, the time function of the fundamental frequency is a key quantity for characterizing infants' utterances during the first months of life. However, additional quantities, describing the voluntary control of the vocal tract become increasingly important during pre-speech development. From a physiological point of view laryngeal phonation and vocal tract based articulation are anatomically different and independently controlled systems. For speech acquisition these two systems have to interact systematically. For instance, melody movements increasingly influence resonance frequencies of the vocal tract and vice versa.

A prominent phenomenon of a melody - resonance frequency interaction is the increasing occurrence of periods of close parallel synchronous movements of melodies and lower resonance frequencies. We call such movements 'tuning periods' when a resonance frequency moves inside a close neighborhood, a 'vicinity tube', of one of the lower harmonics and when this situation lasts for at least 20 ms. Vicinity is given by our analysis bandwidth. Tuning develops either by movements of resonance frequencies to track the melody, or conversely, by movements of the melody to approach certain resonances, or by both processes taking place simultaneously. Here, the term 'tuning' is used in a purely descriptive sense, without considering the mechanisms that are producing this phenomenon as well as without considering what is cause or effect. It is not yet clear, how much mechanical couplings inside the vocal system contribute, and how much intentional neuro-physiological control contribute to the production of longer tuning periods.

Another prominent feature of melody-resonance interaction is resonance frequency transition. Here, transition is defined as coherent and smooth movements of a resonance frequency from the vicinity of one harmonic to the vicinity of another harmonic. The duration of a transition depends on several factors, e.g., neuro-physiological maturity and integrity of the underlying control systems, achieved training level, type of utterance or syllable-structure.

Such well-formed and smooth transitions and the above-mentioned coherent tuning periods represent probably the predispositions for fast couplings (and de-couplings) of phonation and articulation and for acquiring the necessary flexibility for producing fast sequences of phonological structures in later fluent speech. At about babbling age, phonation and articulation are acting completely in concert. The production of babble - syllables is characterized by fast formant transitions which lay already within the time range of fluent speech (e.g. Oller 2000).

METHODS

Data Acquisition. Spontaneous cry utterances were recorded from the 8[th] week until the 25[th] week of life in home environment by trained persons using SONY-DAT-recorders (TCD-D100, 48 KHz/16 bit, mono) and SONY-microphones (ECM-MS957).

Data Analysis. We selected for analysis a set of 800 voiced utterances with a high signal-to-noise ratio out of our cry database. Beside broad- and narrow-band spectrograms high resolution melody computations were made using a CSL-Speech Lab 4400/ MDVP (KAY Elemetrics) in combination with a post-processing and interactive removal of outliers and macro-pitch errors. Our basic resolution in time was down to one pitch period and the frequency resolution was about 3 Hz at 500 Hz. An additional low-pass (Gaussian~40 Hz) filtering of the melody was applied to reduce the short-time variability of the melody. Resonance frequency estimation was performed by means of the LPC analysis tool of the CSL (adaptive coefficient estimation, frame length 10-20 ms, frame step 5 ms). If necessary, an interactive frame by frame check was done with different frame lengths and polynomial degrees in order to exclude critical cases with overlapping resonances. Our standard analysis bandwidth (10 ms frame ÷ 100 Hz) in resonance frequency estimation corresponds to a theoretical relative error band of ±20% at 500 Hz or ±0.2 / # for the resonance in the vicinity of the #-[th] harmonic. However, in analyzing pre-speech data with mean fundamental frequencies not higher than 350 Hz and in case of a smooth sequence of consecutive resonance points (from LPC-frames) we got a much lower uncertainty than formally given by the frame-window related analysis bandwidth. For statistical evaluations of the transition times and tuning durations the analysis was confined to signals with fundamental frequencies under 450 Hz in order to avoid uncertainties of the LPC coding algorithms for higher F0. Recordings with overlapping broad resonances were checked using the bandwidth graphics provided by the CSL-speech Lab and resonances which could not be separated were excluded from further analysis Here, the focus was set to the two lowest resonance tracks in the frequency range up to 4000 Hz (R1, R2).

For a quantitative characterization of melody-resonance interactions we measured and evaluated two quantities, tuning times and transition times. Tuning time was defined as the duration of a parallel synchronous movement of a resonance within the vicinity tube of a harmonic that lasts longer than 20 ms. Transition time between two tuning phases was defined as the residence time of a resonance in the intermediate space between the two vicinity tubes.

DIAGRAMS

The displayed time-frequency diagrams contain the melody and its harmonics on a linear frequency scale. The diagrams are designed in order to visualize the interaction of resonance tracks with the harmonics of the melody. The maximum intensity of a resonance frequency is drawn as a point in the diagram. Consecutive points yield frame by frame the resonance track synchronous to the melody.

In the diagrams a resonance frequency point with a distance ≤ 100 Hz from the nearest harmonic at a given time point is marked red. The (red-coloring) vicinity tube around each harmonic coincides precisely with our analysis bandwidth. Only periods of more than four successive red points (>20 ms) are considered as tuning periods. Resonance points in the intermediate space deviating more than 100 Hz from both

Models and analysis of vocal emissions for biomedical applications. 4th international workshop. October 29-31, 2005 – Firenze, Italy.
Edited by C. Manfredi. ISBN 88-8453-320-1 (online)

adjacent harmonics were marked by blue points. So, transitions consist of a relatively smooth sequence of blue points meeting at both ends the tuning periods at two different harmonics.

Note that in the black and white printed version red and blue points appear grey and black in the diagrams.

SUBJECTS

Subjects were eight infants, six healthy infants and two infants suffering from a cleft-lip-palate (CLP-infants). All infants were term-born German infants.

The six healthy infants were participants of the German Language Development Study (www.glad-study.de), established at the Children's Hospital Lindenhof / Charité, Berlin. These infants were without neurological or developmental disorders. Within the GLaD-study, regular medical and developmental check-ups were carried out. The status of language acquisition at 24 months was assessed by standardized tests for German children. Children below the critical values in both tests at two years, were regarded as SLI-at-risk infants (N=2) in the present investigation. Children above the critical values in both tests form the normal language development group (N=4). Specific language impairment (SLI) is defined as an impairment of oral language despite of having normal intelligence, and adequate learning environment, and despite of having physical or emotional or hearing problems (Bishop & Edmundson 1987). Hence, four of the six healthy infants had a normal language outcome, whereas two of them were retarded at two years (SLI-at-risk infants).

The two CLP-infants, suffering from a unilateral cleft-lip-palate were treated at the Department of Orthodontics of the University Würzburg. CLP-infants exhibit different resonance conditions mainly caused by an open oro-nasal space and a velum dysfunction. The CLP-infants had no other malformations and no neurological disorders.

RESULTS

The interactive tracking method of resonance frequencies and the analysis of their interaction with the melody using a special visualization concept allowed an assessment of the stepwise development of articulatory activity in very young infants. Two characteristic patterns of interaction of melody and resonance could be detected and investigated, namely tuning and transition phenomena.

Developmental changes exhibiting an increasing perfection of tuning between melody and the two lowest resonance frequencies were found. Moreover, increasingly faster resonance transitions were observed.

Averaging over the observation period, the mean duration of transitions between tuning phases for the four healthy infants with normal language development was 61.5 ms (range 20 – 148 ms). Compared to these infants, the two infants exhibiting an at-risk state for SLI at 24 months showed significantly longer mean transition times (mean 75 ms; range 35 – 470 ms). Distributions of observed transition times and developmental changes will be shown at the conference. In contrast to SLI-at-risk infants, the two CLP-infants produced resonance frequency transitions as fast as the normal infants, but only when carrying a special palatal plate, which separates the oral space from nasal space.

From our observation period, representative examples are presented for typical interaction patterns between melodies having low resonance frequencies.

As a representative example of early melody-resonance-interaction, in Figure 1 a mitigated cry of a healthy infant about 9 weeks old is displayed. The cry has a relatively simple melody in form of a single rising-falling arc and shows already longer periods of tuning with stable and coherent tracks of R1 and R2 following the melody. This seems to be a first trace of intentional 'playing' with the resonances. At this early age, a coupling between melody and resonances occurred regularly in healthy infants. But, the resonances R1 and R2 often still show a swing in behavior at the beginning of an utterance. This is interpreted as a sign for a yet immature control capacity for phonatory - articulatory couplings.

The example in Figure 1 demonstrates also a behavior observed in many other utterances, a tendency to correct larger deviations from a preceding tuning not in the direction to the instantaneously nearest harmonic, but to force a restoration toward the former tuning situation. This points to an underlying neuro-pysiological re-directional control process instead to mechanical forces. R1 and R2 seem to stabilize each other, which is particularly effective in cases where the resonances have octave distance (Fig. 1).

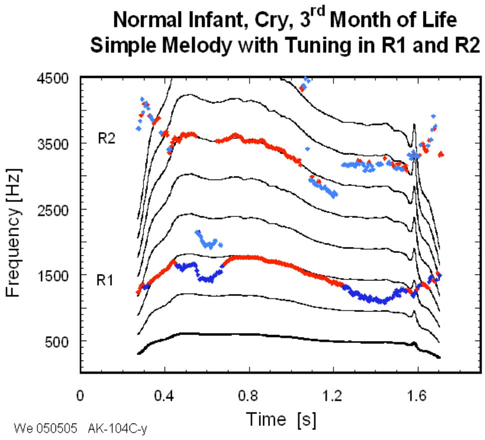

Fig.1: This cry contains relatively long periods of tuning with stable and coherent tracks of R1 and R2 following the melody. A kind of 'playing' with the resonances possibly released by nearby laying strong harmonics is observable. At this age, many utterances still show a tendency to correct stronger deviations from the tuning condition not by moving resonances in direction to the nearest harmonic but to force a back-movement to the former tuning situation.

In Figure 2, a more developed interaction feature in form of coherent and well-formed resonance transitions of R1 and/or R2 is displayed.

Figure 3 demonstrates that resonance transitions occur not only between consecutive harmonics, but also over two harmonics. The observed strong and parallel movements of R1 and R2 as well the aptitude to climb two harmonics upwards or downwards straightforwardly in only one step are interpreted as an advanced transition feature.

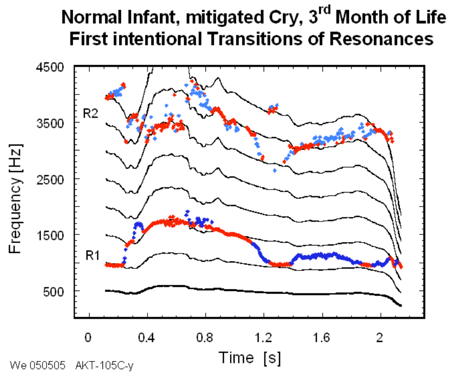

Fig.2: An important feature of articulatory development: In R1 and R2 a transition from one harmonic to another (3rd – 2nd and 7th – 6th) occurs at the same time. R2-transitions exhibit higher fluctuations and are less stable. After a long tuning period of about 600 ms a relatively soft, but coherent R1-transition downward follows. After a short tuning period at the 2nd harmonic the cry ends with a free running R1.

A subsequent stage of pre-speech development is shown in Figure 4. In this babbling utterance complex interactions in form of fast transitions and accurate tunings occur and produce fast changes of the resonances with short residence times on certain harmonics. The interaction processes in R2 proceed essentially parallel to R1, but often with somewhat lower stability and coherency.

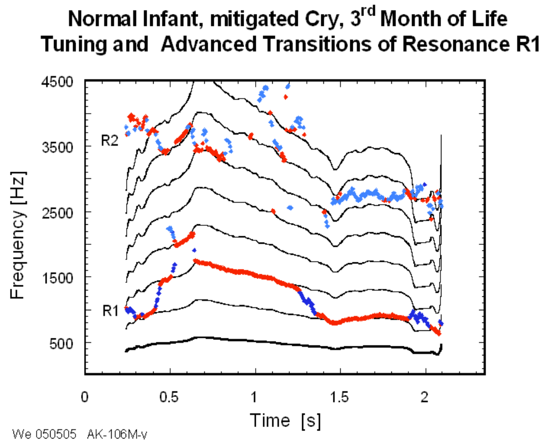

Fig. 3: A two-step R1-transition upwards is followed by a one-step transition downwards. After a longer perfect tuning a further downward R1-transition occurs. The melody consists of two arcs with the frequency maximum of the first arc occurring about 80 ms later than the R1-resonance maximum.

The two investigated SLI-at-risk infants exhibited a kind of poverty of melody complexity and a sparseness of resonances in their pre-speech utterances. Particularly, R1-transitions were prolonged and unstable. Compared to the healthy infants with normal language development, the SLI-at-risk infants had a significantly longer mean transition time of 75 ms (range 35 – 470 ms). Here a typical example for such a more instable and longer lasting transition is displayed (Fig. 5).

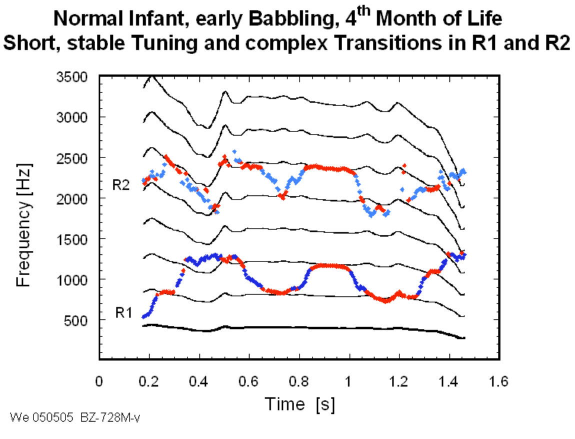

Fig. 4: In syllable-like utterances fast transitions in R1 and R2 alternate with tuning intervals.

In Figure 6 an example is displayed for demonstrating the observation that CLP-infants carrying a palatal plate are capable of exhibiting transition processes comparable to those found in healthy infants with normal language development. However, it was found that CLP-infants more often show instable R2 or higher resonances.

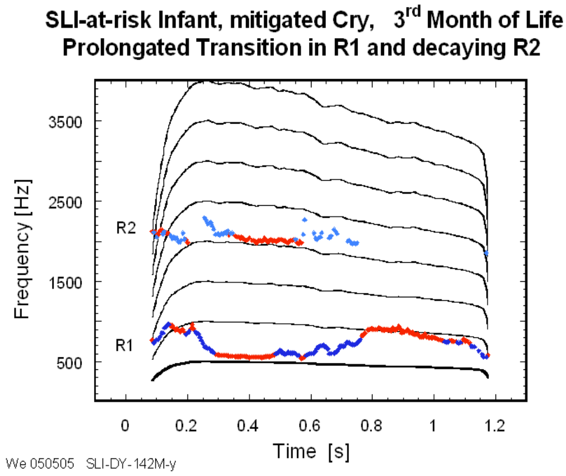

Fig. 5: In SLI-at-risk infants prolonged and unstable transitions of R1 and R2 were regularly observed. Here, the second R1-transition is considerably prolonged and unstable, even in this relatively simple transition from the first to the second harmonic. The development of R2 is also disturbed.

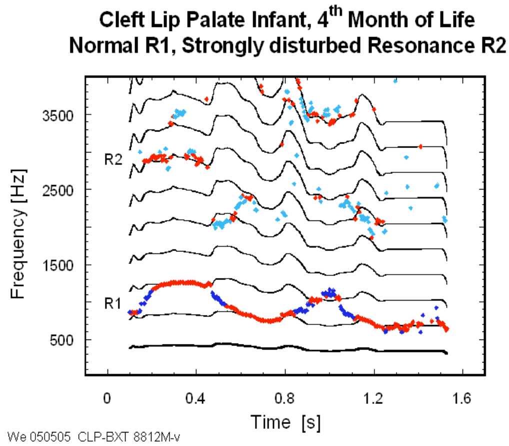

Fig. 6: This mitigated cry consists of a vibrato-like melody modulation while the lowest resonance frequency (R1) moves two times up and down, like in non-cleft infants at this age. The movement of R1 is adjusted to the melody by transitions from the second to the third harmonic of the melody wile exhibiting strong tuning phases at both harmonics.

DISCUSSION

The present investigation confirmed and complemented earlier results concerning a systematic training and a first establishment of articulation activity during infants' crying at a very early age (Wermke et al. 2002, 2005). The establishment of tuning between melody and resonance frequencies and a mastering of fast transitions seems to require a training period before being at disposal for intentional use in vocal production. At about the fourth month of life, a rapid expansion of the infant's pre-speech repertoire occurs. Utterances contain more vowel-like elements and near-syllables. So, it was anticipated that voluntary articulatory activity has to be trained step by step well before this age. The presented results strongly supported this hypothesis.

Here, characteristic interaction phenomena, namely transitions of resonance frequencies between harmonics of the melody could be identified for the first time in infants younger than four months. The mastering of such advanced transitions is an essential prerequisite for performing fast and accurate shifts between vowel formants in speech. Based on our experience, we interpret the systematicity in the production of advanced transitions as a training phase enabling the brain to establish a prospective time organization in vocal sound production necessary for fluent speech: Vocal tract articulators often have to be pre-adjusted to anticipated, but nevertheless in parallel threads planned melody movements. The observed time organization, were resonance movements often lead the course of the melody, points to higher cerebral control mechanisms underlying sound production already in young infants (see Fig. 3 for the 80 ms time lead of resonances).

It is hypothesized that the observed interaction patterns are not produced by chance. Particularly, the investigated transitions of lower resonance frequencies share many features with later formant transitions necessary for vowel articulation in speech-like babbling and word production. The observed developmental changes suggest that the fast resonance transitions are necessary presuppositions for the extremely short transitions between the phonological units in later speech. The identification of such transitions as well as the demonstration of increasingly stronger tunings between phonation and articulation supports the idea of learning and training of language-related features during infants' crying and points to a continuous developmental path from crying to word production. This hypothesis is also supported by the finding, that fast switches between tuning and resonance transition are generally learnt in healthy infants with normal language outcome within a short time span. In contrast to them, infants being at risk for developing SLI exhibit the displayed training phases over a much longer time span; sometimes they need even several months. Moreover, the longer mean resonance transition times observed in SLI-at-risk infants point to a relation of these early articulatory activities to later language performance. Concerning processing auditive information in the brain, a disturbed time organization is reported for SLI-at-risk infants and is interpreted as an important component in developing a Specific Language Impairment (e.g. Tallal 1989, Jusczyk 1997, Weber et al. 2004). The present findings suggest the idea that also in controlling vocal production a disturbed time organization may characterize SLI-at-risk infants.

In CLP-infants, the finding that a palatal plate seems to enable the infants to produce transition phenomena comparable to those in healthy non-cleft infants is very important for treatment strategies to minimize later speech and language impairments caused by the malformation. Both infants received a special palatal plate during the first days after birth and the plates were only removed by the parents to clean them after feeding. For a next investigation to be conducted, it is hypothesized that CLP-infants who were not supplied with a palatal plate will exhibit more deviations in the observed melody-resonance-interaction patterns.

However, in face of the high variability between infants, the cases studied here were not yet sufficient to and further work is necessary to confirm the formulated hypotheses and to generalize the presented results..

CONCLUSION

The present paper provided time functions (tracks) of the resonance frequencies and investigated the interaction between these tracks and harmonics of the melody. This approach allowed to investigate pre-articulatory activity at a very early age and to observe early developmental processes directed toward speech and language acquisition. An increasing unfolding of tuning between melody and lower resonance frequencies as well as resonance frequency transitions were found in utterances of 2- to 3-months-old infants. This vocal behavior was interpreted as an early articulatory activity in infant's crying. In a broader perspective, it is seen as a language-related behavior, preparing formant tuning and focalization in later speech. Far reaching medical applications are seen for infants with disturbances of the time organization of vocal production and for CLP-infants. .

References

1. Oller, D. K. The Emergence of the Speech Capacity. Lawrence Erlbaum Associates, Mahwah, NJ, 2000.
2. Wermke, K., Mende W., Manfredi C., Bruscaglioni, P. Developmental aspects of infants' cry melody and formants. *Med. Eng. Phys.* 2002;24: 501-514
3. Wermke K., Mende W., Manfredi C., Bruscaglioni, P. & Stellzig-Eisenhauer A. Tuning phenomena of melodies and resonance frequencies (formants) in infants' pre-speech utterances. *Poster presented at the 149th Meeting of the ASA*, Vancouver 16-20 May 2005
4. Bishop D.V.M., Edmundson A. Specific language impairment as a maturational lag: Evidence from longitudinal data on language and motor development. Dev Med Child Neurol. 1987;29:442-459
5. Tallal P., Ross R., Curtiss S. Familial aggregation in specific language impairment. J Speech Hear Disord. 1989;54:167-173
6. Jusczyk P. (ed.). The discovery of spoken language. Bradford, Cambridge, MA, 1997
7. Weber Ch., Hahne A., Friedrich M., Friederici A.D. Discrimination of word stress in early infant perception: Electrophysiological evidence. Cog Brain Res. 2004;18:149-161

Acknowledgement

We thank Peter Wermke for his engaged work in all aspects of data analysis. The data characterizing the developmental state of the six healthy subjects were kindly provided by Volker Hesse, head of the pediatric clinic of the Krankenhaus Lichtenberg, teaching hospital of the Charité, Berlin. The status of language outcome of these six infants were kindly made available by Zvi Penner and Petra Schulz, principal investigators of the research group „language production and comprehension“ at the Charité – University Medical Center Berlin, Department of Audiology and Phoniatrics, headed by Prof. Dr. M. Gross.

The study was supported in part by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) (WE-1724/4-1) as part of the Research Group 381 - Frühkindliche Sprachentwicklung und spezifische Sprachentwicklungsstörungen. Further support was provided by the Gruter Institute of Law and Behavioral Research Portola Valley,USA.