STUDY OF CRY PATTERNS IN INFANTS AT HIGH RISK FOR AUTISM

S. Orlandi[1], L. Bocchi[1], C. Manfredi[1], M. Puopolo[2], A. Guzzetta[3] S. Vicari[4] and M.L. Scattoni[2]
[1] Dept. of Electronics & Telecommunications, Università degli Studi di Firenze, Firenze, Italy
[2]Department of Cell Biology and Neurosciences, Istituto Superiore di Sanità, Rome, Italy
[3] Department of Developmental Neuroscience, Stella Maris Scientific Institute, Pisa, Italy
[4]Department of Neuroscience, Children's Hospital Bambino Gesù, Rome, Italy

Abstract: **Autism Spectrum Disorders (ASD) show an increasing prevalence in children, and are often undiagnosed up to the third year of life. Analysis of cry patterns appears a promising approach for allowing an early ASD detection and diagnosis. In this work we compare the main acoustic parameters collected by recording infant crying in control subjects with the parameters obtained in high risk subjects, namely infant siblings of children with ASDs. Results confirm previous finding obtained using home video recordings, and indicate a weaker coupling between fundamental frequency and first resonance frequencies in high risk subjects.**
Keywords: **Autism Spectrum Disorders, cry analysis, fundamental frequency, resonance frequencies.**

I. INTRODUCTION

Cry constitutes the first communication channel available to newborns for fulfilling their needs and attracting attention of the caregivers. Hearing a crying baby produces, in the adults, a reaction aimed at activating parental caretaking, ensuring newborn survival and comfort.

Cry involves activation of both the newborn and listener neural system, increasing the reciprocal attention level. It is produced when the newborn perceives a negative stimulus, from an internal or external source, and it involves a coordinated effort of several brain regions, mainly brainstem and limbic system.

For this reason, cry can be candidate as an early sign of potential problems and pathologies involving the neural system, and it should be included and analyzed during the evaluation of newborn state.

Recently, there has been a large interest in the analysis of cry features in children with Autism Spectrum Disorders (ASD), mainly because of the major role played by brainstem and limbic system, both areas compromised in children with ASD [1,2], in the production of infant crying.

Autism is a neurodevelopmental disorder characterized by impairments in social and communication development, and by restricted and repetitive behavior. Typically children are not diagnosed before two years of life despite 50% of parents of children with ASDs report that they suspected a problem before their child was 1 year of age [3], thus a more precocious diagnosis seems possible.

Recent epidemiological studies reported prevalence rates in the general population of 58-67/10.000, suggesting that ASDs affect many families and represent a serious public health problem. Over the years, interventions have focused on enhancing developmental skills and on ways of ameliorating behavioral difficulties by teaching more effective communication skills. There are studies demonstrating that early intensive behavioral intervention initiated at preschool age and sustained for 2-3 years results in substantial improvements for a large subset of ASD children. Gains are found in IQ, language, and educational placement. Although a few pharmacological treatments can reduce some associated symptoms, early behavioral interventions remain the most effective treatment for the symptoms of autism. Thus, early identification of ASDs allows the possibility of early intervention, for the ultimate purpose of optimizing quality of life and functional independence.

Previous studies on the properties of cry in autistic children involved the spectrographic analysis of the sound signal and of the modulation of the acoustic wave, reporting the presence of significant differences between controls and subjects later diagnosed with ASD. More in detail, crying episodes in ASD subjects have shorter duration, less modulation, and lack of regular peaks than crying recorded in control subjects [4]. Moreover, fundamental frequency is lower than in control subjects, and structural properties appear atypical [5,6].

II. METHODS

Acquisition protocol

The project will recruit a set of about 200 control subjects and a set of high risk subjects to be followed prospectively (tentatively, 20 subjects). By allowing accurate and detailed assessments of behavioral measures at fixed time points, prospective studies offer theoretical advantages to detect early modifications of ASD, while avoiding biases.

Presently, no diagnostic tool is available for the early detection of autistic children. Recent advances in early detection research have resulted from prospective studies carried out on high risk infants. We planned to recruit later-born infant siblings of children diagnosed with

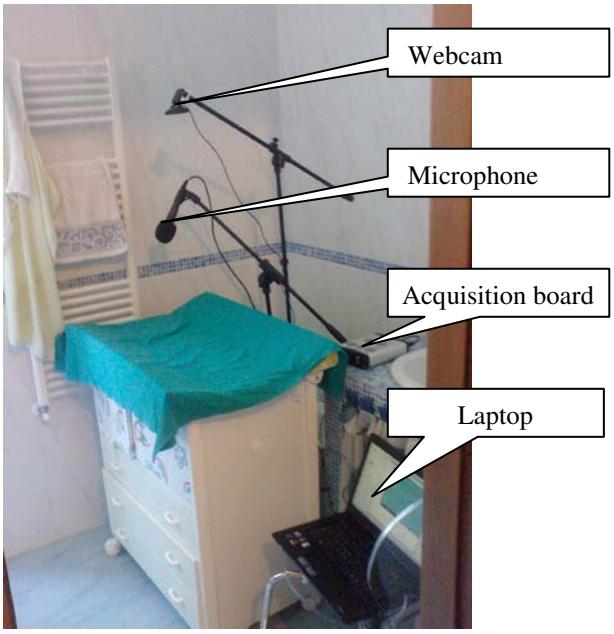

Fig. 1 Acquisition system in a typical setup

ASD. These infant siblings are themselves at especially high risk for an autism or ASD diagnosis [7] and this population is arguably the most clearly defined high risk group available [8,9].

The planned acquisition procedure is totally not invasive, minimizing the ethical issues involved in the recruitment of control subjects and high risk infants.

Each subject is involved in a set of measures, scheduled every six weeks, starting a few days after birth up to the 24th week of life (hence, each one will undergo to five testing sessions). In this work, we present results obtained during the 2nd testing session, recorded approximately during the 8th week of life.

Informed consent has been obtained from all parents. The study protocol has been approved by the local ethical committee (Istituto Superiore di Sanità, IRCCS Fondazione Stella Maris, and IRCCS Pediatric Hospital Bambino Gesù).

The acquisition system has been designed for being used in the patient home, minimizing the discomfort for the involved subjects and the impact of the external environment on children habits. Hence, the basic requirement is the ease in transporting and assembling the system. According to this requirement, the proposed system, shown in Fig. 1, includes a laptop which is connected to an high speed USB video camera (Logitech HD pro webcam C910), able to provide a 1280x1024 pixel video stream and an external audio acquisition device (Tascam US-144-MK2, as the quality of the audio card embedded in the laptop is not adequate to the recording specifications) and a professional microphone (Shure SM58).

Signal processing

In the present work, we focus on the analysis of the audio track, performed using BioVoice, custom software developed in Matlab language. The detail of the algorithms used in BioVoice for processing cry recordings have been already described in [10], however we summarize only part of the elaboration relevant for this study in the following paragraphs. The first processing step aims at detecting each crying episode in order to label it for further elaboration. This allows speeding up the elaboration by removing unnecessary data. Detection of crying episodes is performed using a Voiced/Unvoiced detection procedure. In the next step, ach detected crying episode undergoes a detailed analysis, where the following features are extracted from the signal:

- length and average amplitude of the episode
- fundamental frequency F_0
- resonance frequencies, mainly first and second resonance frequency (F_1 and F_2, respectively)

F_0 estimation and voiced/unvoiced detection

The fundamental frequency, F_0, was estimated with a two-step procedure. Simple inverse filter tracking (SIFT) was applied first [11,12], to signal time windows of short and fixed length. The window length was chosen as $M = 3F_s/F_{min}$, where F_s is the signal sampling frequency and F_{min} is the minimum allowed F_0 value for the signal under consideration (for newborn cry: F_{min} = 150 Hz).

In the second step, F_0 was adaptively estimated inside $[F_l, F_h]$. This allowed for a more precise F_0 estimation. A variable window length for analysis was applied, inversely proportional to the changing F_0. Very short time windows, ranging from 5 to 15 ms, were thus obtained, locally dependent on the signal variability. Over each time window, the signal was band-pass filtered (for newborn cry the range was settled to 150–900 Hz) with the Mexican hat continuous wavelet transform, and the signal periodicity was extracted by means of the average magnitude difference function (AMDF) approach. In case of fast and abrupt F_0 changes, this procedure was shown to increase the robustness of the F_0 estimation, giving enhanced results with respect to standard methods [12].

In order to disregard voiceless parts of the signal, a *voiced/unvoiced decision* (V/UV) was applied. It was based on the approach proposed previously in [13] and was suitably modified for our purposes here. Basically, a signal frame is selected as voiced if voicing evidence, γ_{max}, defined as the amplitude of the autocorrelation function on that frame, is larger than a threshold value. A number of controls made on adjacent frames have been added to ensure continuity of the detected pitch in order to exclude possible wrong V/UV choices [13]. For a newborn cry, it was commonly found that $\gamma_{max} \geq 0.06$.

Resonance frequencies

Even if vowel frequencies cannot be found in newborn cries, resonance frequencies (RFs) reflect important acoustical characteristics of the infant vocal tract. For RFs estimation and tracking, a robust parametric technique is used, obtained by peak picking in the power spectral density (PSD) plot. This was evaluated on the same adaptive time windows as previously described. For PSD estimation, autoregressive (AR) models were used. The model order q varied according to the signal characteristics. The "modified covariance method" was applied, as it was shown to give the best results for the reduction of spectral line splitting and bias of the frequency estimations [12].

The relation $q \cong 0.5F_s$ (in kHz) was found to be optimal for obtaining an enough detailed spectrum, while preventing spectral smoothing and consequent loss of spectral peaks. Co-ordinates of PSD maxima on each time window, as well as their mean and std value on the whole signal, were also evaluated. Thus, details were given about RFs evolution in time as related to energy. The first three RFs, are extracted by the BioVoice software, however in the present work we focused only on the first two of them, F_1 and F_2.

III. RESULTS

Cry episodes detection

The performance of the detector has been assessed by qualitative inspection of the audio signal superimposed on the output of the voice detector. Visual analysis of the resulting waveform indicates a substantially correct extraction of cry episodes.

An example of the detected voice segments is shown in Fig. 2, where a signal frame containing four crying episodes is shown.

Fundamental frequency analysis

A comparison of the characteristics of cry episodes of a control subject and a high risk subject of the same age show some interesting differences. As already described in the introduction, results confirmed the lower frequency range of the fundamental frequency in the high risk newborn with respect to control subjects. Moreover, we studied possible the relationship between F_0 and F_1 during cry episodes. In control cases, we often observed a strong coupling between the two variables. As shown in Fig. 3, a large number of frames shows a linear relation coupling F_0 and F_1: for instance, in the case shown in figure, when F_0 is in the range from 400 to 550Hz, we have found $F_1 = 3\ F_0$ in almost all frames. As shown in the figure, most frames indicate a strong coupling between the two frequencies, accordingly to a linear relation with an angular coefficient which can be expressed as ratio of small numbers (the other alignments which may be seen in the figure correspond to angular coefficients equal to 2+1/4, 2+1/2 and 2+3/4).

By contrast, the scatter plot relating F_0 and F_1 in a high risk subject is usually similar to the one reported in Fig. 4, where the coupling between the two frequencies F_0 and F_1 is weaker.

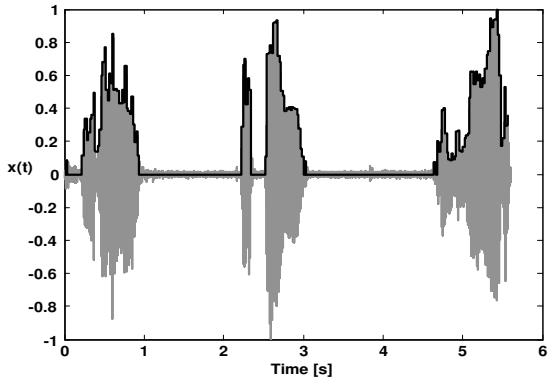

Fig. 2. Detection of cry episodes (black line) and audio signal (gray)

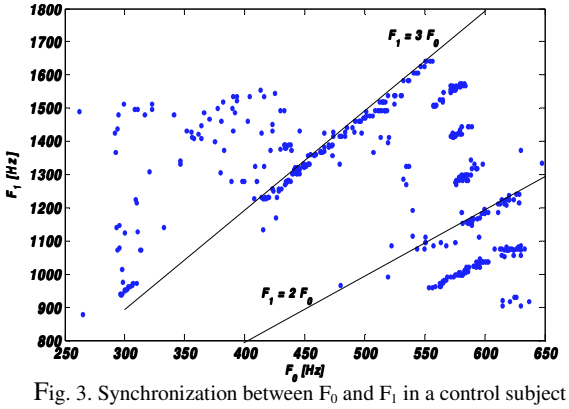

Fig. 3. Synchronization between F_0 and F_1 in a control subject

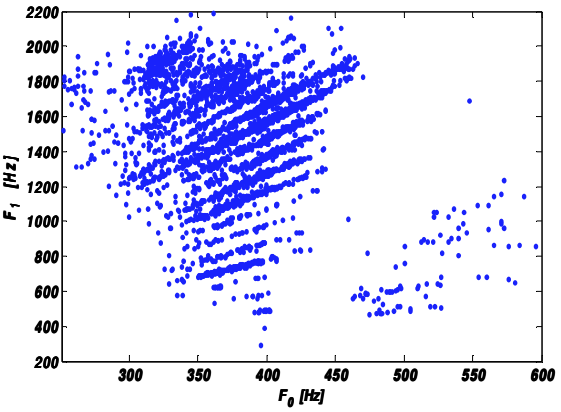

Fig. 4. Synchronization between F_0 and F_1 in a high risk subject

Melody

In each cry episode, the fundamental frequency presents a well-defined trend. Four typical patterns have

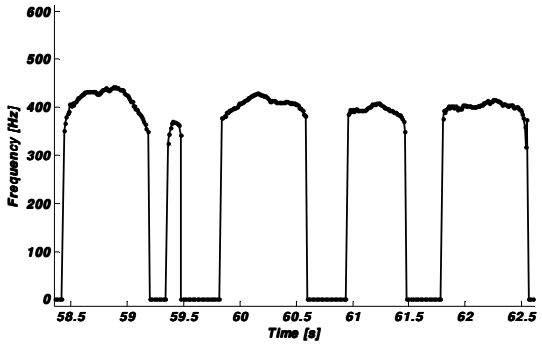

Fig. 5. Typical fundamental frequency trends. First peak is clearly symmetrical, the last one is a plateau. The two central ones are not clearly defined.

been observed in newborns [14,15], namely the symmetrical pattern (frequency rising and falling around a central peak), the rising pattern (frequency peak appears near the end of the episode), the falling pattern (frequency peak appears at the beginning), and the plateau (with an almost constant frequency). A cry recording of a high risk newborn, 2 months old, with samples of symmetric and plateau patterns, is reported in Fig. 5.

IV. Conclusion

Autism Spectrum Disorders (ASD) are often undiagnosed up to the third year of life. Analysis of cry patterns appears a promising approach for allowing its early diagnosis and treatment. In this work main acoustic parameters obtained in control subjects with the parameters found in a small group of high risk subjects are compared. The results of these first experiments indicate that appreciable differences can be found. At present, the sample size, especially as concerns high risk subjects, is too small to assess the statistical significance of these differences. Work is in progress in order to increase the sample size and to define best acoustic parameters suited for an early non-invasive diagnosis of autism spectrum disorders.

Acknoledgement

Supported by the Italian Ministry of Health Grant (GR3), Young Researcher 2008, "Non-invasive tools for early detection of Autism Spectrum Disorders".

References

[1] Amaral DG, Schumann CM, Nordahl CW. „Neuroanatomy of autism.". Trends Neurosci. 31(3) pp. 137-45. (2008)

[2] Schulkin, J. "Autism and the amygdala: An endocrine hypothesis". Brain and Cognition, 65, 87–99 (2008).

[3] E. Werner, G. Dawson, J. Osterling, N. Dinno, "Brief report: recognition of autism spectrum disorder before one year of age: retrospective study based on home videotapes". Journal of autism developmental disorders, 30, pp. 157-162, 2002.

[4] P. Venuti, G. Esposito, Z. Giusti, "A qualitative analysis of crying and vocal distress in children with autism", Journal of Intellectual Disability Research, 48, 2004.

[5] P. Venuti , Z. Giusti , F. La Femina , G. Esposito, F. Domini "Qualitative analysis of children's autistic cry observed by homevideo", Infant Cry Workshop 8th International Workshop, Padova, 2002.

[6] G. Esposito, P. Venuti, "Typical and atypical expression of distress: A study on cry", Int. J. of Public Health, 1(2), pp. 141-150, 2009.

[7] L. Zwaigenbaum, S. Bryson, T. Rogers, W. Roberts, J. Brian, P. Szatmari, "Behavioral manifestations of autism in the first year of life", Int J Dev Neurosci, 2005 Apr-May, 23(2-3) pp. 143-52.

[8] L. Zwaigenbaum, S. Bryson, C. Lord, S. Rogers, A. Carter, L. Carver, K. Chawarska, J. Constantino, G. Dawson, K. Dobkins, D. Fein, J. Iverson, A. Klin, R. Landa, D. Messinger, S. Ozonoff, M. Sigman, W. Stone, H. Tager-Flusberg, N. Yirmiya , "Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants", Pediatrics, 2009 May, 123(5) pp. 1383-91.

[9] Zwaigenbaum L. "Advances in the early detection of autism", Curr Opin Neurol. 2010 Apr, 23(2) pp. 97-102.

[10] C. Manfredi, L. Bocchi, S. Orlandi, L. Spaccaterra, G.P. Donzelli, "High-resolution cry analysis in preterm newborn infants", Medical Engineering & Physics, 31(5) pp. 528-532, 2009.

[11] J.D. Markel, "The SIFT algorithm for fundamental frequency estimation", IEEE Trans Audio Electroac 20, pp. 367–377, 1972.

[12] S.L. Marple, "Digital spectral analysis with applications", Prentice Hall, Englewood Cliffs, NJ, USA, 1987.

[13] L. M. Van Immerseel, J. P. Martens, "Pitch and voiced/unvoiced determination with an auditory model" J. Acoust. Soc. Am. 91 (6), pp. 3511-3526, 1992.

[14] K. Wermke, W. Mende, "Musical elements in human infants' cries: in the beginning is the melody", Musicae Scientiae, 13, pp. 151-173, 2009.

[15] K. Wermke, M. Birr, C. Voelter, W. Shehata-Dieler, A. Jurkutat, P. Wermke, A. Stellzig-Eisenhauer, "Cry Melody in 2-Month-Old Infants With and Without Clefts", Cleft Palate-Craniofacial Journal, May 2011, 48(3), pp. 321-330.