Self-generated Sounds in Human Neonates

Emma Last[1], Dr Cigdem Gelegen[2]
Supervised by Dr. Kimberley Whitehead
University College London
Department of Neuroscience, Physiology, and Pharmacology

Completed as part of the Laidlaw Research and Leadership Scholarship Programme

1. Abstract

Self-generated sounds, such as cries, are some of the first sensory inputs that the developing brain receives. Eventually, infants need to learn to associate their cries with a caregiving response in order to consciously use crying as a signal. Thus, the acoustic characteristics of their cries are key to neural and behavioural development. This project examines the acoustic features of cries in neonates and discusses how the brain might encode these cries. To explore this topic, a literature review was completed on neonatal self-generated sounds in healthy infants, premature infants, and those with brain injury, and how these sounds, particularly crying, have been analysed by other research groups. Following this, existing EEG datasets containing neonatal cry recordings were analysed using MATLAB/EEGLAB, with a focus on the fundamental frequencies of each cry. Results for each sample are discussed in context with the corrected gestational age and regularity of cries, and point to how acoustic characteristics may change throughout neonatal crying episodes.

2. Background

Vocalisation in humans begins with crying: the key method which neonates use to communicate their needs to the caregiving environment. Crying in infants involves a combination of respiratory, laryngeal, and sublaryngeal movements. The sound is produced when there is a drop in pressure across the glottis which causes vibration of the adducted vocal cords in the larynx. The frequency of vocal fold vibration generates the fundamental frequency (F0), one of the predominantly analysed features of infant cries and what we hear as the pitch of the cry (1).

The assessment of cries as an insight into human health began following reports that infants with neurological disorders had cries that were different from those of normal infants (2). Since then, acoustic characteristics of infant crying have been considered a non-invasive tool for assessing neurophysiological states. Certain acoustic features of crying have been linked to diagnoses related to neurological damage, prematurity, respiratory conditions, and hearing impairments. (3, 4). The most commonly investigated characteristics include F0 and its formant frequencies, frequency band intensity, regularity of cries, and duration of cry episodes (5).

Previous studies have found the cries of healthy infants to have an F0 of around 200-600 Hz, whilst those of infants with endocrine, metabolic, and neurological pathologies are of a higher frequency (>600 Hz) (6). It is also thought that a higher F0 is linked to preterm birth, but this value restores to normal levels at term-equivalent age (7). This high F0 may be due to smaller infant body size, particularly shorter vocal folds, but there is evidence to suggest a relation to neural integrity, for example one study showed full-term infants small for their

gestational age had cries of the same F0 as those of normal birth weight, indicating more dependence on developmental maturity than the infant's body size (8).

The role of the central nervous system (CNS) in neonatal crying involves the brain stem, midbrain and limbic system, followed by later cortical involvement (9). Brain stem input impacts the contour and cross-sectional airway of the vocal tract, which determines the formant frequencies, other commonly studied characteristics of human vocalisation (10). Cranial nerves IX-XII and both the phrenic and thoracic nerves control the larynx, pharynx, chest, and upper neck muscles involved in vocalisation. F0 of vocalisations relies on CNS control of laryngeal and respiratory actions. Vagal inputs from the medulla are said to inhibit tightening of vocal folds through affecting laryngeal muscle contraction. Thus, reduced vagal activity (such as that seen in preterm infants) may result in a higher F0 due to increased tightening of the vocal folds, contributing to the effect of small body size on F0 in preterm infants (11).

A 2016 study investigated this association in spontaneous cries (measuring vagal activity through assessing respiratory sinus arrhythmia (RSA) during quiet sleep) (8). They observed significantly lower RSA and higher F0 in preterm infants whilst there was a positive association between higher RSA and mean, maximum and range of F0 in term infants, supporting the idea that differences in vagal function might be associated with the F0 of spontaneous cries through vocal fold tension in infants at an early developmental stage.

These higher pitched cries of preterm infants may impact their development of crying as a tool to communicate their needs to caregivers. For example, Frodi et al. (12) found the cries of preterm infants elicited more negative emotions than cries of term infants did, and suggested early attention to higher F0 cries in preterm infants could benefit the caregiving relationship. Furthermore, there are a few studies which suggest that high F0 cries during infancy are associated with reduced cognitive abilities later in life. One study observed that both preterm and term infants who had high F0 cries within 2 weeks of term postconceptual age were more likely to have a lower score on the Bayley Scales of Infant Development at 18 months of age than infants with lower F0 cries (13).

3. Methods

Recruitment criteria for the cohort of neonates studied included a corrected gestational age (GA at birth plus postnatal age) (CGA) between 34 + 5 and 41 + 0 weeks, and a postnatal age (PNA) of up to 10 days. Datasets consisted of sound recordings (recorded using snore sensors by SleepSense) with time-locked EEG data. The microphone was placed on the collar of the babygrow to pick up self-generated sounds, and any sounds heard during recording which weren't reflected in the microphone trace were annotated and further processed in MATLAB to extract a clearer signal. The sampling rate of the recorded signal was 2000 Hz.

Overall, 17 3-second cry samples were analysed from 8 different neonates. Datasets were analysed in MATLAB/EEGLAB. Preprocessing included applying a 6Hz high-pass filter to remove unwanted signals and extracting 3 seconds of crying. For analysis of a single isolated cry, a 3 second period was selected which contained the approximately 1-second-long cry, and 1 second before and after the cry. For analysis of intermittent crying, 3 seconds of the sequence were selected.

For each cry or cry sequence, a morlet-wavelet power spectrum was generated along with a power spectral density estimate (PSD) using Welch's method. From these figures, the fundamental frequency was estimated. Other cry features explored in this study included the duration of cries, the time between cries, the power of cry signal, and how these features changed as the crying episode progressed.

4. Results

Figures were generated for all crying examples from all neonates sampled. The figures for 6 crying samples amongst 3 babies were selected for discussion in this paper. Though those from remaining neonates can be viewed at the end of this section, they either produced a weaker signal which could not be effectively filtered out using the aforementioned methods, or contained artefacts which interfered with the clarity of the figures and were therefore deemed less suitable for analysis.

Data from Baby 724 (male, singleton, GA=29.86, PNA=40, post-menstrual age (PMA)=35.57) included four 3-second sequences of consecutive cries. Cry sequence 1 (Fig1) occurred first, followed 7 seconds later by cry sequence 2 (Fig2). Cry sequence 3 (Fig3) came 28 seconds after the second sequence. A fourth sequence (Fig4) was also analysed, with this one occurring around 1 hour after the previous ones.

One 3-second sample containing a single cry was analysed in both Baby 727 (Fig5) (corrected GA at time of EEG was 38 weeks) and Baby 816 (Fig6) (male, singleton, GA=36.86, PNA=3, PMA=37.29).

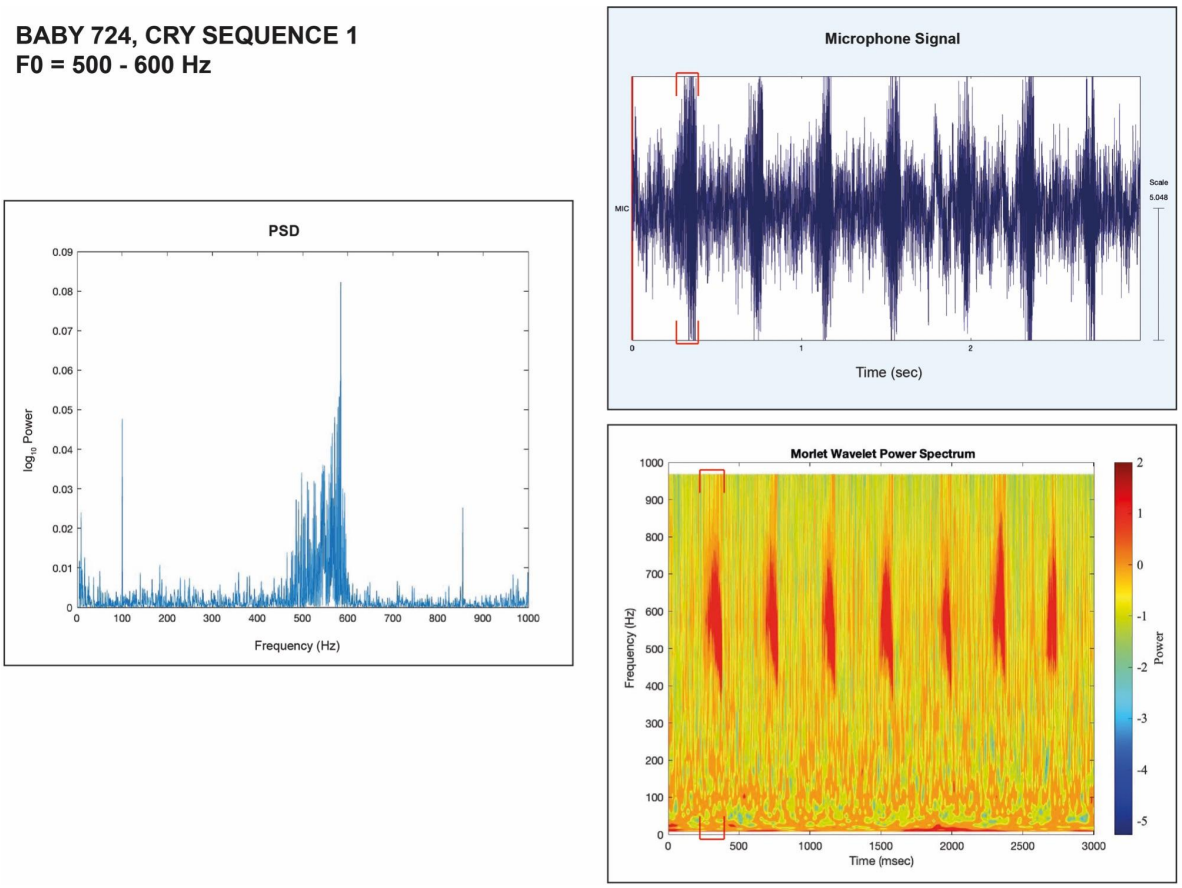

Fig1. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the first isolated cry sequence in Baby 724. The red brackets show where the first of the 7 cries in the sequence is found on both the microphone channel and power spectrum. The estimated F0 for cry sequence 1 in Baby 724 is between 500 and 600 Hz. This was estimated by examining the peaks on the PSD and supported by location of the darkest areas (red) in the power spectrum. The deeper colours appear at the time each cry occurs, confirming that it is the cry signal which produces the PSD peaks.

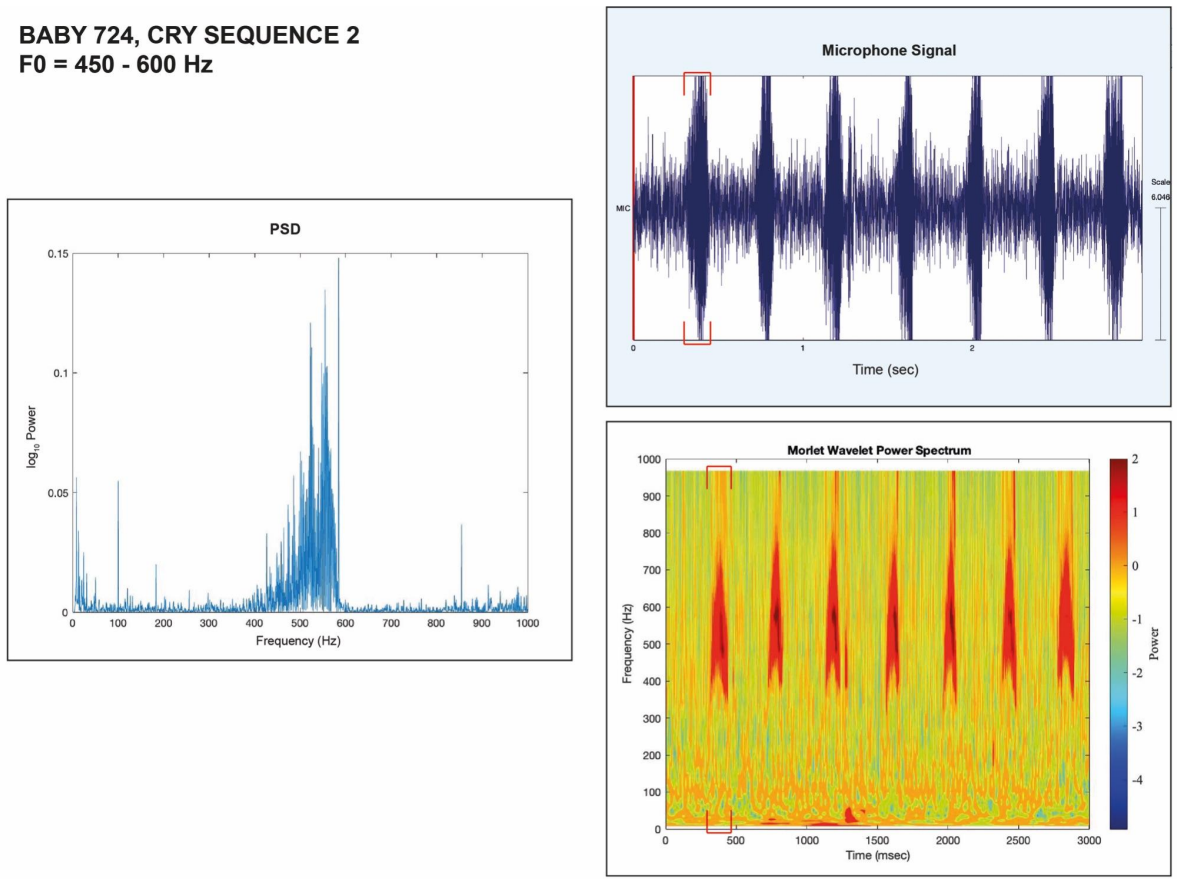

Fig2. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the second isolated cry sequence in Baby 724. The red brackets show where the first of the 7 cries in the sequence is found on both the microphone channel and power spectrum. The estimated F0 for cry sequence 2 in Baby 724 is between 450 and 600 Hz. This sequence yielded PSD peaks at slightly lower frequencies and produced a higher power signal than the previous sequence.

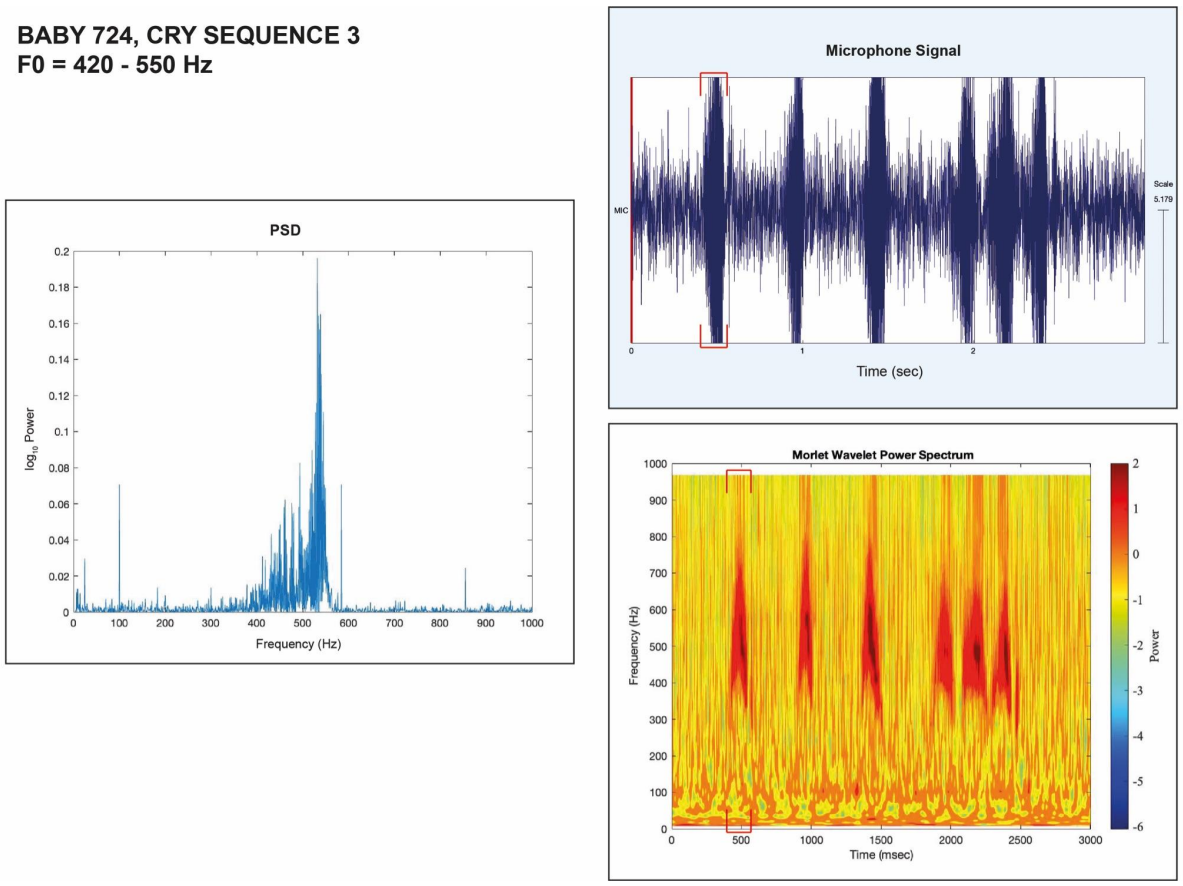

Fig3. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the third isolated cry sequence in Baby 724. The red brackets show where the first of the 6 cries in the sequence is found on both the microphone channel and power spectrum. The estimated F0 for cry sequence 3 in Baby 724 is between 420 and 550 Hz. Again, there are PSD peaks at even lower frequencies and higher powers than observed in Baby 724's previous cry sequences.

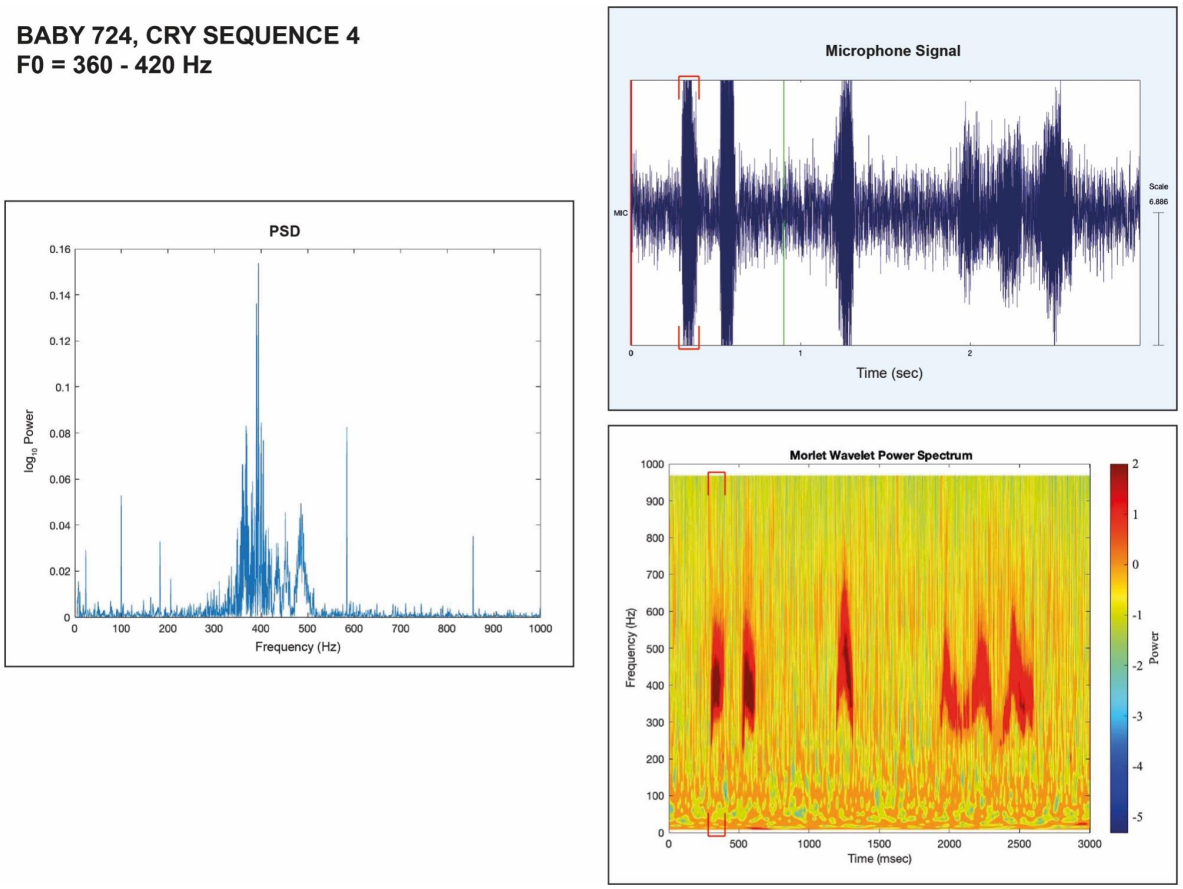

Fig4. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the fourth isolated cry sequence in Baby 724. The red brackets show where the first of the 6 cries in the sequence is found on both the microphone channel and power spectrum. The estimated F0 for cry sequence 4 in Baby 724 is between 360 and 420 Hz.

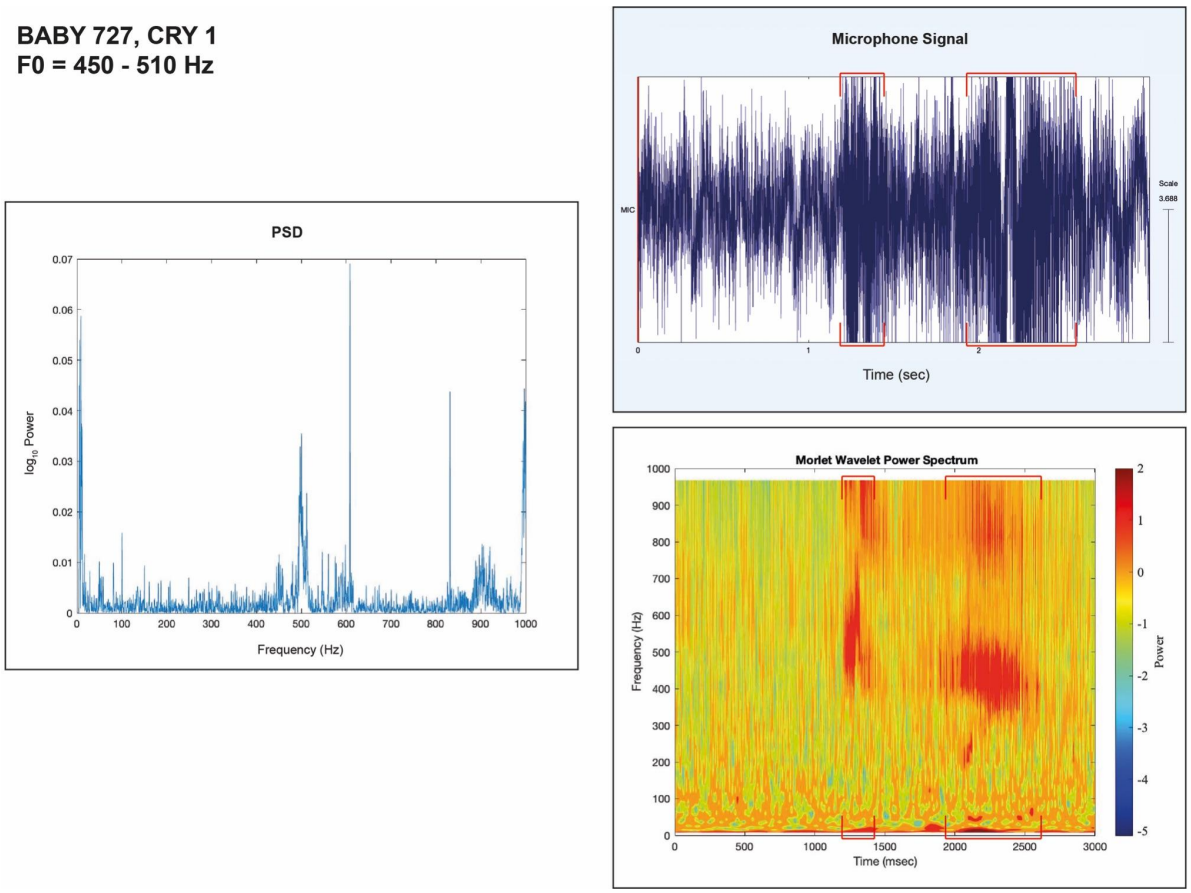

Fig5. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the isolated 3-second cry sample in Baby 727. The red brackets show where the neonate's cry is found on both the microphone channel and power spectrum. The estimated F0 for cry 1 in Baby 727 is between 450 and 510 Hz.

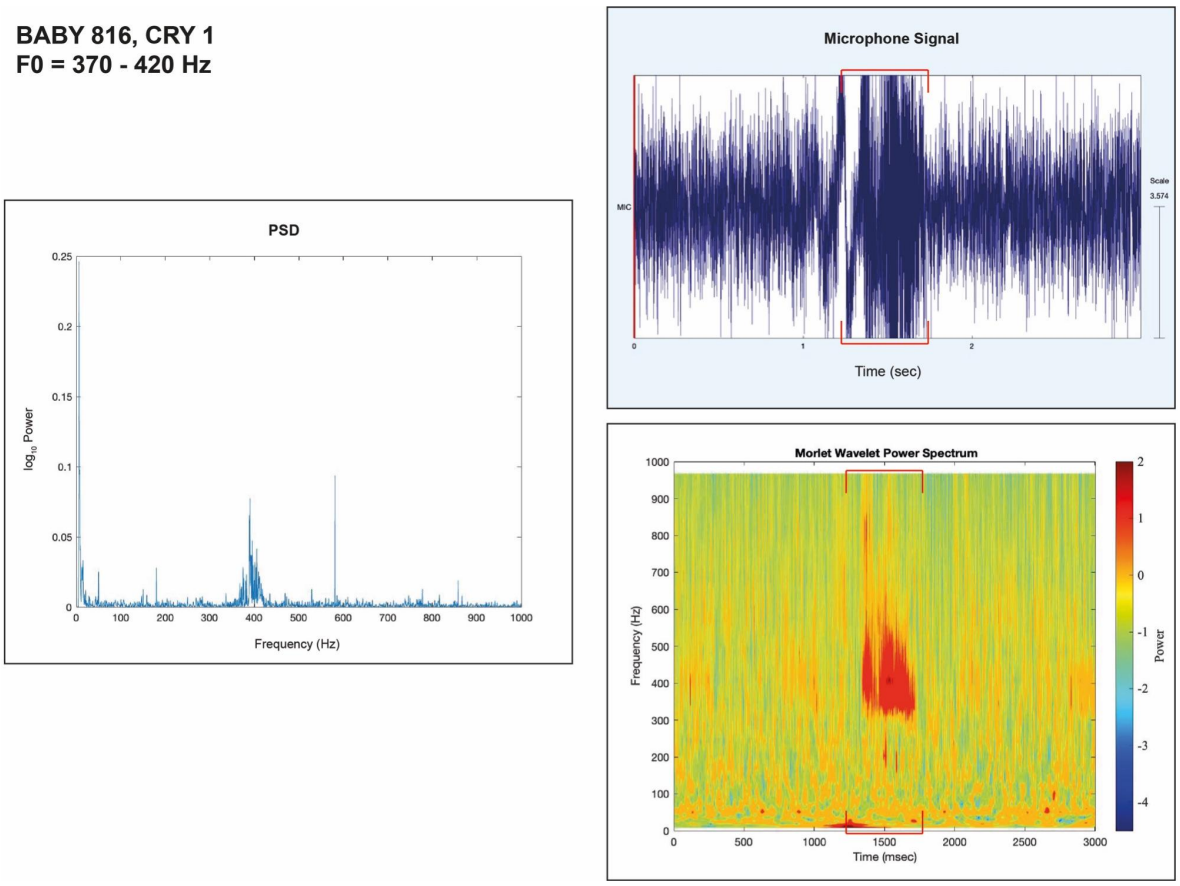

Fig6. PSD (left), signal from microphone channel in EEGLAB (top right), and Morlet Wavelet Power Spectrum (lower right) of the isolated 3-second cry sample in Baby 727. The red brackets show where the neonate's cry is found on both the microphone channel and power spectrum. The estimated F0 for cry 1 in Baby 816 is between 370 and 420 Hz.

5. Discussion

The results of cry analysis reflect the findings of previous studies, with F0s ranging from 370-600 Hz. Some cry samples in Baby 724 had a lower F0 than others (for instance, min F0 decreased from 500 Hz in cry sequence 1 to 420 Hz in cry sequence 3).

Figures 1-3 show sequences of cries that follow one another. In addition to the gradual decrease in frequency between sequences there is also an increase in power observed on the PSD in each successive sequence. Although there were no other cry sequences in the dataset to see if the same trend is observed, this may suggest a mechanism is used by the infant to adapt the acoustic characteristics of their cry to elicit a more successful caregiving response (lowering the frequency and increasing the power).

6. Future Directions

There could potentially be a corollary discharge mechanism used during infant vocalisation. A corollary discharge is a signal that encodes an intended action, which is then compared with sensory feedback (reafference) resulting from the actual action (14). It is typically discussed in context to motor commands, but has been explored as a mechanism potentially used for auditory self-monitoring. Vocalised communication involves not only sounds being heard by the caregiver, but by the infant producing them. For successful communication, the infant needs to determine whether or not the sound was self-generated, maintain perception of external sounds during vocalisation, and monitor their output for accurate production (15). Marmosets and other nonhuman primates have shown evidence of feedback-dependent vocal control in response to environmental changes or auditory feedback of self-generated vocalisations, including adapting the timing and structure of vocalisations and increasing the vocal amplitude. Marmosets have also been shown to alter the frequency contents of their vocalisations in response to auditory feedback (16, 17, 18).

In addition, higher F0 and lower intensity cries have been observed through various studies in infants with hearing impairments (19). This may be a result of inhibited auditory self-monitoring.

Future studies on whether neonates use self-monitoring in an attempt to elicit a more successful caregiving response would be interesting. There is also a need for further exploration into the reasoning behind higher frequency cries in infants with certain pathologies. For example, whether preterm infants generate higher frequency sounds due to their smaller body size (and therefore shorter vocal folds) or insufficient neural development. This also raises the question of whether preterm infants or those with neurological damage have not developed an effective system for auditory self-monitoring, and thus have higher F0 and lower power cries. The analysis of consecutive cry sequences along with EEG data in infants using a larger sample of infants might serve to answer these questions in the future.

References:

1. LaGasse, L. L., Neal, A. R., & Lester, B. M. (2005). Assessment of infant cry: acoustic cry analysis and parental perception. *Mental retardation and developmental disabilities research reviews*, *11*(1), 83–93. https://doi.org/10.1002/mrdd.20050

2. Michelsson, K., & Michelsson, O. (1999). Phonation in the newborn, infant cry. *International journal of pediatric otorhinolaryngology*, *49 Suppl 1*, S297–S301. https://doi.org/10.1016/s0165-5876(99)00180-9

3. Cacace, A. T., Robb, M. P., Saxman, J. H., Risemberg, H., & Koltai, P. (1995). Acoustic features of normal-hearing pre-term infant cry. *International journal of pediatric otorhinolaryngology*, *33*(3), 213–224. doi:10.1016/0165-5876(95)01211-7

4. Laufer, M., & Horii, Y. (1977). Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks. *Journal of Child Language, 4*(2), 171-184. doi:10.1017/S0305000900001586

5. Gabrieli, G.; Scapin, G.; Bornstein, M.H.; Esposito, G. (2019) Are Cry Studies Replicable? An Analysis of Participants, Procedures, and Methods Adopted and Reported in Studies of Infant Cries. *Acoustics*, *1*, 866-883. doi:10.3390/acoustics1040052

6. Soltis, J. (2004). The signal functions of early infant crying. The Behavioral and brain sciences. 27. 443-58; discussion 459. doi:10.1017/S0140525X0400010X.

7. Michelsson, K., Järvenpää, A. L., & Rinne, A. (1983). Sound spectrographic analysis of pain cry in preterm infants. *Early human development*, *8*(2), 141–149. https://doi.org/10.1016/0378-3782(83)90070-1

8. Shinya, Y., Kawai, M., Niwa, F. and Myowa-Yamakoshi, M. (2016), Associations between respiratory arrhythmia and fundamental frequency of spontaneous crying in preterm and term infants at term-equivalent age. Dev Psychobiol, 58: 724-733. https://doi.org/10.1002/dev.21412

9. Porges, S.W. (2007). The polyvagal perspective. *Biological Psychology, 74,* 116-143, doi: 10.1016/j.biopsycho.2006.06.009

10. Benson, J. and Haith, M., 2008. *Encyclopedia of infant and early childhood development*. 2nd ed. Elsevier Inc, p.361.

11. Porter, F. L., Porges, S. W., & Marshall, R. E. (1988). Newborn pain cries and vagal tone: parallel changes in response to circumcision. *Child development*, *59*(2), 495–505. https://doi.org/10.2307/1130327

12. Frodi, A.M., Lamb, M.E., Neff, C., Sherry, D., Leavitt, L.a., Donovan, W.L., (1978) Father's and Mother's responses to the smiles and cries of normal and premature infants. *Dev Psychol, 14,* 490-498, https://doi.org/10.1037/0012-1649.14.5.490

13. Barry M. Lester; Developmental Outcome Prediction From Acoustic Cry Analysis in Term and Preterm Infants. *Pediatrics* October 1987; 80 (4): doi: 529–534. 10.1542/peds.80.4.529

14. Poulet, J. F., & Hedwig, B. (2006). The cellular basis of a corollary discharge. *Science (New York, N.Y.)*, *311*(5760), 518–522. https://doi.org/10.1126/science.1120847

15. Levelt W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

16. Eliades, S. J., & Wang, X. (2019). Corollary Discharge Mechanisms During Vocal Production in Marmoset Monkeys. *Biological psychiatry. Cognitive neuroscience and neuroimaging*, *4*(9), 805–812. https://doi.org/10.1016/j.bpsc.2019.06.008

17. Roy, S., Miller, C. T., Gottsch, D., & Wang, X. (2011). Vocal control by the common marmoset in the presence of interfering noise. *The Journal of experimental biology*, *214*(Pt 21), 3619–3629. https://doi.org/10.1242/jeb.056101

18. Pomberger, T., Risueno-Segovia, C., Löschner, J., & Hage, S. R. (2018). Precise Motor Control Enables Rapid Flexibility in Vocal Behavior of Marmoset Monkeys. *Current biology : CB*, *28*(5), 788–794.e3. https://doi.org/10.1016/j.cub.2018.01.070

19. Mahmoudian, S., Aminrasouli, N., Ahmadi, Z. Z., Lenarz, T., Farhadi, M. (2019). Acoustic Analysis of Crying Signal in Infants with Disabling Hearing Impairment, *Journal of Voice, 33(6),* 946 https://doi.org/10.1016/j.jvoice.2018.05.016