Sex hormone influence on human infants’ sound characteristics: melody in spontaneous crying

Kathleen Wermke1, Johannes Hain2, Klaus Oehler4,†, Peter Wermke1,3 and Volker Hesse5,6

1Center for Prespeech Development and Developmental Disorders, 2Department of Mathematics (Statistics), and 3IT Center, University of Wuerzburg, Wuerzburg, Germany
4Sana Klinikum Lichtenberg, Laboratory Medicine, Berlin, Germany
5German Center for Growth, Development and Health Encouragement during Childhood and Youth, Children’s Hospital Berlin Lindenhof, Germany
6Charité—University Medicine, Institute for Experimental Paediatric Endocrinology, Berlin, Germany

The specific impact of sex hormones on brain development and acoustic communication is known from animal models. Sex steroid hormones secreted during early development play an essential role in hemispheric organization and the functional lateralization of the brain, e.g. language. In animals, these hormones are well-known regulators of vocal motor behaviour. Here, the association between melody properties of infants’ sounds and serum concentrations of sex steroids was investigated. Spontaneous crying was sampled in 18 healthy infants, averaging two samples taken at four and eight weeks, respectively. Blood samples were taken within a day of the crying samples. The fundamental frequency contour (melody) was analysed quantitatively and the infants’ frequency modulation skills expressed by a melody complexity index (MCI). These skills provide prosodic primitives for later language. A hierarchical, multiple regression approach revealed a significant, robust relationship between the individual MCIs and the unbound, bioactive fraction of oestradiol at four weeks as well as with the four-to-eight-week difference in androstenedione. No robust relationship was found between the MCI and testosterone. Our findings suggest that oestradiol may have effects on the development and function of the auditory—vocal system in human infants that are as powerful as those in vocal-learning animals.

1. Introduction

Sex steroids are powerful and enduring signalling molecules in the body and represent a class of hormones that plays an important role in the development of the brain and related gender-specific behaviours [1–5]. Neuroendocrine theories of brain development hold that testosterone is the predominant factor mediating sex-specific cortical growth and the ensuing lateralization of hemispheric function [1–3]. Structural and functional cerebral asymmetries in language-relevant areas of the brain have been found at an early age [4].

Although initial reports are somewhat inconsistent, young girls score higher than young boys on vocabulary comprehension and production [5]. New findings reveal evidence for gender differences in speech and language-relevant gestures of young children [6,7].

Recently, Bowers et al. [8] postulated that Foxp2 protein levels were associated with the more communicative sex, namely girls in humans and males in rodents. Moreover, Bowers and co-workers [9] found that sex steroid hormones, both androgens and oestrogens, played a crucial role in Foxp2 expression, at least in rodents. Foxp2 is a gene for which a link to speech and even language
has been found [10,11], and it has been proposed that it contributes to the evolution of human speech and language by adapting cortico-basal ganglia circuits [12].

The rationale of this study rests on the well-known influence of sex steroid hormones on brain development, including language-relevant brain regions, and increasing evidence for sex differences in early language development. It also relies on the following assumptions: (i) Foxp2 expression seems to be associated with vocal learning in vertebrates and, particularly, language development in humans [13]; (ii) Foxp2 shows a sex difference in expression [8,14] and (iii) Foxp2 may be affected by sex steroid hormones. In our opinion, the investigation of young infants’ vocalizations upon the early postnatal surge of sex steroids (‘mini-puberty’) provides a suitable approach to acquiring new insights and could advance our understanding of the mechanisms mediating sex differences in vocal development and language acquisition.

Mini-puberty, a phenomenon that is also common in other species [15], describes the postnatal activation of the hypothalamic–pituitary–gonadal axis, beginning in humans about one week after birth, when the inhibitory effect of maternal oestrogen wanes. In boys, a rapid fall in testosterone concentrations after delivery is typical, followed by a second surge peaking at about two months (50 days) before it again falls to pre-pubertal levels [16]. The serum level of testosterone in boys is as high at about the second to fourth months of life as it will be in late puberty and adulthood, respectively [17]. In girls, a continuous testosterone decline after delivery is typical, whereas oestradiol secretions increase during the first three months [18]. The precursor hormone androstenedione peaks in both sexes at about two months and thereafter also decreases continuously to pre-pubertal levels [19].

This study questioned whether the high exposure to sex hormones at two months influences the melodic patterns of spontaneous crying. Human infants exhibit a characteristic continuous development from simple melodies (fundamental frequency contour) in early emotive utterances to complex melodies that provides prosodic primitives for later language [20]. Two-month-old infants with a high percentage of complex melodies in their crying were found to have better...
language outcomes at 2.5 years [21]. An individual infant’s ‘melody performances’, i.e. the share of cries with complex melodies among all cries uttered during a certain time interval, can be assessed using the melody complexity index (MCI). Based on prior longitudinal studies, the MCI is taken as an early measure of language development [21].

2. Material and methods

(a) Participants and procedure

Eighteen healthy, full-term German infants (nine female) with normal hearing from a monolingual family background (middle-class families) were investigated twice, at about four weeks (mean age 29 ± 2 days, range 25–33 days) and eight weeks (mean age 58 ± 3 days, range 53–66 days; see also the electronic supplementary material).

The research took place at the Children’s Hospital Lindenhof, Berlin, within the framework of a broader study after receiving approval from the research ethics committee of the Humboldt University of Berlin. Pregnant mothers were contacted and asked to participate in the research according to the informed consent procedures of the institution. All of the parents gave written informed consent. Spontaneous crying was recorded in the home environment, averaging two 30 min samples taken at four and eight weeks, respectively. Blood samples for hormone assessment were taken at the hospital within 1 day of collection of the crying samples. Non-fasting peripheral venous blood samples were taken from an antecubital vein or the dorsal venous network of the hand. Total serum testosterone (T), oestradiol (E2), their precursor androstenedione (A), their inter-rater reliability of this interactive coding procedure was calculated using a randomly selected sub-sample of 400 cry utterances (see the electronic supplementary material).

(b) Cry recordings and analysis

In total, 4500 cries (utterance of an infant during a single expiration) were digitally (48 kHz sampling frequency, 16 bit) recorded at about feeding time using a TASCAM (DA-P1) recorder and an Earthworks microphone (TC20). Frequency spectrograms were calculated for all cries (CSL 4400; KAYPENTAX).

Based on visual inspections of the spectrograms, cries containing broad regions of phonatory noise were excluded, since the fundamental frequency cannot be reliably determined in those signals (see the electronic supplementary material).

Finally, 3104 single cries were selected, and their melody analysed using PRAAT v. 5.2 [22]. Inter-individual variability in utterance numbers per infant (girls: 178, range 73–232; boys: 167, range 30–424), caused by the infants’ individual spontaneous ‘talkativeness’, was taken into account by a statistical bootstrap analysis (see the electronic supplementary material).

(c) Calculation of the melody complexity index and statistical analysis

As outlined elsewhere in more detail [20,21], newborns’ voiced cries with an identifiable melody mainly exhibit a single ascending–descending arc. This changes dramatically over the first three months, when these simple, single-arc melodies are increasingly replaced by complex, that is, double- or multiple-arc melodies. The cry melody was quantitatively analysed, and then utterances were subdivided into those with only a simple (single-arc) melody and those with a complex (multiple-arc) melody (figure 1). The inter-rater reliability of this interactive coding procedure was tested using a randomly selected sub-sample of 400 cry utterances (see the electronic supplementary material).

For each infant, the MCI was calculated, expressing the share of cries with complex melodies among all the cries uttered either within a cry bout (figures 1 and 2) or isolated during the two recording sessions; fussing was excluded.

A hierarchical, multiple regression approach was used to investigate associations between an infant’s melodic skill (MCI) and certain serum hormone levels (oestradiol (E2), testosterone (T), androstenedione (A)). The bioactive sex steroids that are not bound to SHBG, i.e. the ratios E2/SHBG and T/SHBG, were used as a proxy for the total amount of steroids affecting brain development, bearing the following caveats in mind: (i) the values are not actual hormone levels measured in the brain, in which local concentrations of steroids may differ; and (ii) steroids bound to binding globulins, like α-feto-protein, are detected in the brain and may confer distinct signalling effects by acting at the membrane or via select transportation into cells [23] (see the electronic supplementary material for more details).

Finally, a bootstrap analysis was conducted to show the stability of the regression coefficients with respect to the inter-individual variation of cry numbers (see the electronic supplementary material).

3. Results

While testosterone at four and eight weeks and oestradiol at eight weeks differed significantly between the sexes, oestradiol at four weeks (p = 0.05), androstenedione and the MCI did not (table 1).

Significant predictors of the individual MCIIs (omitted variables: see the electronic supplementary material, table S4) were the ratio E2/SHBG at four weeks (effect size 0.378), i.e. the SHBG-unbound serum fraction of oestradiol; and A_diff, i.e. the differences in androstenedione between weeks four and eight (Cohen’s d effect size 0.393) (table 2). There was a positive association between E2/SHBG and the MCI (standardized linear regression coefficient: 0.767, t_{17} = 4.63, p < 0.0001) (figure 2). Girls deviate from the regression line more than boys; the sex-specific r^2 values are 0.452 (girls) and 0.756 (boys).

4. Discussion

The current findings are the first to show in human infants that vocal patterns might be associated with serum concentrations.
of bioactive oestradiol. Oestradiol is seen as the most biologically prevalent and active compound of all oestrogens, and as exerting wide-ranging effects on the developing brain [24–27].

There are still huge gaps in our understanding of the molecular interactions that mediate language-relevant neural circuitries in human infants. Although the sample size does constrain the generalizability of our results, the robust prediction of the individual infants’ melodic skills by the circulating unbound oestradiol fraction seems to at least suggest mechanisms that resemble oestradiol function in vocal learning in other species, in particular song birds [23,28]. Human infants are able to vocally communicate long before vocabulary and grammar are established, by making extensive use of melody patterns [20].

The strong association identified between oestradiol and the earliest predictor (MCI) of a child’s later language competence points to the exploitation of the same regulatory important sex hormones by human infants when learning to speak, by song birds when learning to sing and, probably, by other vocal learners when it comes to producing their conspecific sounds [29,30].

A further important conclusion is that gender seems to be a potential predictor of early language development and that sex should be regarded as an essential confounding factor in infant studies of vocal communication.

Acknowledgement. We are grateful to all the parents and infants who participated in the study, as well as to C. J. Partsch for his support in the androgen assessment. The paper is dedicated to our former colleague Klaus Oehler, who passed away during its completion. The authors would also thank two anonymous reviewers for their helpful suggestions.

Funding statement. The study was partially supported by the German Research Foundation (DFG; WE-1724/4-1) and the Max Planck Institute for Human Cognitive and Brain Sciences as part of Research Group 381—Early Language Development and Specific Language Disorders (FR-519/18-1).

### Table 1. Sex differences. Mean/s.d. or median/ranges (non-normally distributed data) are reported, as are the p-values of the unpaired t-test or the Mann–Whitney test.

<table>
<thead>
<tr>
<th></th>
<th>males (n = 9)</th>
<th>females (n = 9)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean/median</td>
<td>s.d./range</td>
<td>mean/median</td>
</tr>
<tr>
<td>E2_4w (pmol l⁻¹)</td>
<td>627.20</td>
<td>462.7 – 1335.5</td>
<td>913.70</td>
</tr>
<tr>
<td>E2_8w (pmol l⁻¹)</td>
<td>359.85</td>
<td>296.8 – 632.5</td>
<td>629.60</td>
</tr>
<tr>
<td>E2/SHBG_4w</td>
<td>7.57 x 10⁻³</td>
<td>3.08 x 10⁻³</td>
<td>1.05 x 10⁻²</td>
</tr>
<tr>
<td>E2/SHBG_8w</td>
<td>3.10 x 10⁻³</td>
<td>1.52 x 10⁻³</td>
<td>7.19 x 10⁻³</td>
</tr>
<tr>
<td>T_4w (pmol l⁻¹)</td>
<td>6865.00</td>
<td>3779 – 27 597</td>
<td>520.00</td>
</tr>
<tr>
<td>T_8w (pmol l⁻¹)</td>
<td>8980.00</td>
<td>5443 – 25 586</td>
<td>312.00</td>
</tr>
<tr>
<td>T/SHBG_4w</td>
<td>0.088</td>
<td>0.029</td>
<td>0.007</td>
</tr>
<tr>
<td>T/SHBG_8w</td>
<td>0.064</td>
<td>0.024</td>
<td>0.003</td>
</tr>
<tr>
<td>A_4w (pmol l⁻¹)</td>
<td>1768.3</td>
<td>540.7</td>
<td>1473.6</td>
</tr>
<tr>
<td>A_8w (pmol l⁻¹)</td>
<td>1341.7</td>
<td>415.2</td>
<td>1321.8</td>
</tr>
<tr>
<td>A_diff (pmol l⁻¹)</td>
<td>426.6</td>
<td>507.2</td>
<td>266.1</td>
</tr>
<tr>
<td>MCI</td>
<td>0.578</td>
<td>0.10</td>
<td>0.653</td>
</tr>
</tbody>
</table>

### Table 2. The results of the two-step hierarchical multiple regression model (forward selection), including bootstrap analysis. B, unstandardized regression coefficients; s.e., standard error; β, standardized regression coefficients; CI, confidence intervals (lower to upper bound). Model assumptions: Durbin–Watson statistic: D = 2.017; Shapiro–Wilk test for normality of standardized residuals: p = 0.544.

<table>
<thead>
<tr>
<th>step</th>
<th>predictor variables included</th>
<th>B</th>
<th>s.e.</th>
<th>β</th>
<th>95% CI (Bootstrap 95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>step 1</td>
<td>constant</td>
<td>0.309</td>
<td>0.064</td>
<td>0.171 – 0.477</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>(R² = 0.635)</td>
<td>E2/SHBG_4w</td>
<td>32.452</td>
<td>6.581</td>
<td>0.797</td>
<td>18.338 – 46.567</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>step 2</td>
<td>constant</td>
<td>0.322</td>
<td>0.057</td>
<td>0.200 – 0.445</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>(R² = 0.738)</td>
<td>E2/SHBG_4w</td>
<td>28.552</td>
<td>6.030</td>
<td>0.701</td>
<td>15.525 – 41.578</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>A_diff</td>
<td>7.716 x 10⁻⁵</td>
<td>3.400 x 10⁻⁵</td>
<td>0.336</td>
<td>0.371 – 15.061 x 10⁻⁵</td>
<td>0.041</td>
</tr>
</tbody>
</table>
References


