Musical Intervals in Infants' Spontaneous Crying over the First 4 Months of Life

Lotte Armbrüster[a]   Werner Mende[b]   Götz Gelbrich[c]   Peter Wermke[d]
Regina Götz[e]   Kathleen Wermke[a]

[a]Center for Pre-Speech Development and Developmental Disorders, University Hospital Würzburg, Würzburg, Germany; [b]Berlin-Brandenburg Academy of Sciences and Humanities, Berlin, Germany; [c]Institute for Clinical Epidemiology and Biometry, Faculty of Medicine, University of Würzburg, and Clinical Trial Center, University Hospital Würzburg, Würzburg, Germany; [d]IT Center, University of Würzburg, Würzburg, Germany; [e]Institute for Special Education, Faculty of Human Sciences, University of Würzburg, Würzburg, Germany

Abstract

Introduction: Perception and memorizing of melody and rhythm start about the third trimester of gestation. Infants have astonishing musical predispositions, and melody contour is most salient for them. *Objective:* To longitudinally analyse melody contour of spontaneous crying of healthy infants and to identify melodic intervals. The aim was 3-fold: (1) to answer the question whether spontaneous crying of healthy infants regularly exhibits melodic intervals across the observation period, (2) to investigate whether interval events become more complex with age and (3) to analyse interval size distribution. *Methods:* Weekly cry recordings of 12 healthy infants (6 females) over the first 4 months of life were analysed (6,130 cry utterances) using frequency spectrograms and pitch analyses (PRAAT). A preselection of utterances containing a well-identifiable, noise-free and undisturbed melodic contour was applied to identify and measure melodic intervals in the final subset of 3,114 utterances. Age-dependent frequency of occurrence of melodic intervals was statistically analysed using generalized estimating equations. *Results:* 85.3% of all preselected melody contours (*n* = 3,114) either contained single rising or falling melodic intervals or complex events as combinations of both. In total 6,814 melodic intervals were measured. A significant increase in interval occurrence was found characterized by a non-linear age effect (3 developmental phases). Complex events were found to significantly increase linearly with age. In both calculations, no sex effect was found. Interval size distribution showed a maximum of the minor second as the prevailing musical interval in infants' crying over the first 4 months of life. *Conclusion:* Melodic intervals seem to be a regular phenomenon of spontaneous crying of healthy infants. They are suggested to be a further candidate for developing an early risk marker of vocal control in infants. Subsequent studies are needed to compare healthy infants and infants at risk for respiratory-laryngeal dysfunction to investigate the diagnostic value of the occurrence of melodic intervals and their age-depending complexification.

© 2020 The Author(s)
Published by S. Karger AG, Basel

Kathleen Wermke
Center for Pre-Speech Development and
Developmental Disorders, University Hospital Würzburg
Pleicherwall 2, DE–97070 Würzburg (Germany)
wermke_k @ ukw.de

Introduction

In the first 2 months of life, it is primarily crying and fussing that assume a predominant role in the expressive vocal repertoire of the infant [1]. For both, parents and scientists, infant crying is often recognized as a natural alarm signal, an emergency expression of a motorically helpless young infant. As a result of its helplessness, the human infant is dependent upon close social contact with its mother from birth (the term "mother" is used here to imply the infant's primary caregiver). A variety of sensory signals would help to facilitate this interaction. However, crying is the most powerful signalling available to the young infant. There are numerous publications in the popular press on the topic of infant crying, and most of them are dedicated to providing advice for the best methods to stop the crying [2–4]. Although most of these reports and guidebooks mention that crying is a normal expressive behaviour of all infants, there is a general focus on its unpleasant and disturbing character, particularly so-called excessive crying [2]. However, the observation of this rather extreme crying behaviour should not constrain the recognition of spontaneous crying within the mother-infant context as a regular behaviour of vocal development. Indeed, from a developmental perspective, crying has been recognized as a communicative behaviour that provides a foundation for later language acquisition [5–7]. For the present study, this perspective was relevant.

There is evidence that the infant phonatory system responsible for crying is reflective of the underlying neurophysiological functioning [7–13]. In the early 1960s, Scandinavian phoniatricians and paediatricians described this indicative valence by acoustically investigating infant crying from a medical diagnostic perspective [10, 12, 13]. Following the research line newly established at that time, a large number of researchers demonstrated that specific acoustic features of an infant's cry were suitable as risk markers that directly reflect dysfunctions in vocal control over the next about 30 years [see review in 2, 11]. Since then, there has been a shift in focus by examining the relationship between acoustic features of infant cry and later language development. Particularly the time function of the fundamental frequency (f_0), the cry melody, was suggested to have an indicative value for delays of early language development [14–16]. In a longitudinal study, it was found that cry melody complexity, that is slow modulation of the f_0, at about 2.5 months of age could predict the language outcome at 2 years [14]. On a perceptive level, a more recent brain physiological study

corroborated the importance of melody for language development. François et al. [17] demonstrated that melodic cues contained in songs can facilitate word form extraction. Neonatal brain responses for sung streams predicted a later expressive vocabulary: "using a longitudinal design, we provide empirical and computational evidence for a link between the enhanced brain responses to melodically enriched streams (prosodic modulations) and later expressive vocabulary measures at 18 months" [17, p. 7].

According to a melody development model by Wermke and Mende [6, 18], melody complexity increases over the first months of life. Starting with the production of a single-arc structure, melody develops towards double- and multiple (complex)-arc structures with age. These melody contour patterns are easily perceptible in an infant's spontaneous crying and set them apart from pain cries with their falling, often vibrato-like contour, disrupted by subharmonics and phonatory noise [19–21]. However, the complexity of the arc-like melody structure is not the only indicator for vocal control and maturity found in infant cry utterances. For example, very short-term variability of f_0, like jitter (period-to-period variability) and other f_0 perturbation indices, were also found to reflect constituents of vocal control [22, 23].

Efforts are under way to describe the short-time vocal flexibility of young infants by examining melodic intervals [5, 24–26]. In musical melodies, intervals reflect the difference in pitch between 2 discrete tones (e.g., semitone, full tone or third). Wermke and Mende [5] commented: "At first sight, the continuous versus discrete nature of cry melody and musical melody seems to limit structural comparisons, but both aspects can be reconciled by an appropriate choice of resolution in time and frequency […]. In music based on tonal scales, a small inventory of discrete tones is defined inside the octave space […], cry melodies do often not have a marked step-structure with stationary frequencies; they are swept over a frequency interval (FM-amplitude)" (p.155).

Nonetheless, the cry melody contour may contain interval-like episodes that can be described as a glide between 2 following f_0 plateaus (Fig. 1a) [5, 6, 25]. Such episodes, which we described as melodic intervals, can occur when the rising or falling flanks of melody arcs are not smooth but stair-like (Fig. 1a), especially investigated in an early study of 2-week-old infants by one of the authors (R.G.) [25]. An alternative variant to create a melodic interval is exemplified in Figure 1b. The displayed cry melody contains a single melodic interval whose size is defined by the frequency ratio of the 2 f_0 plateaus. In con-
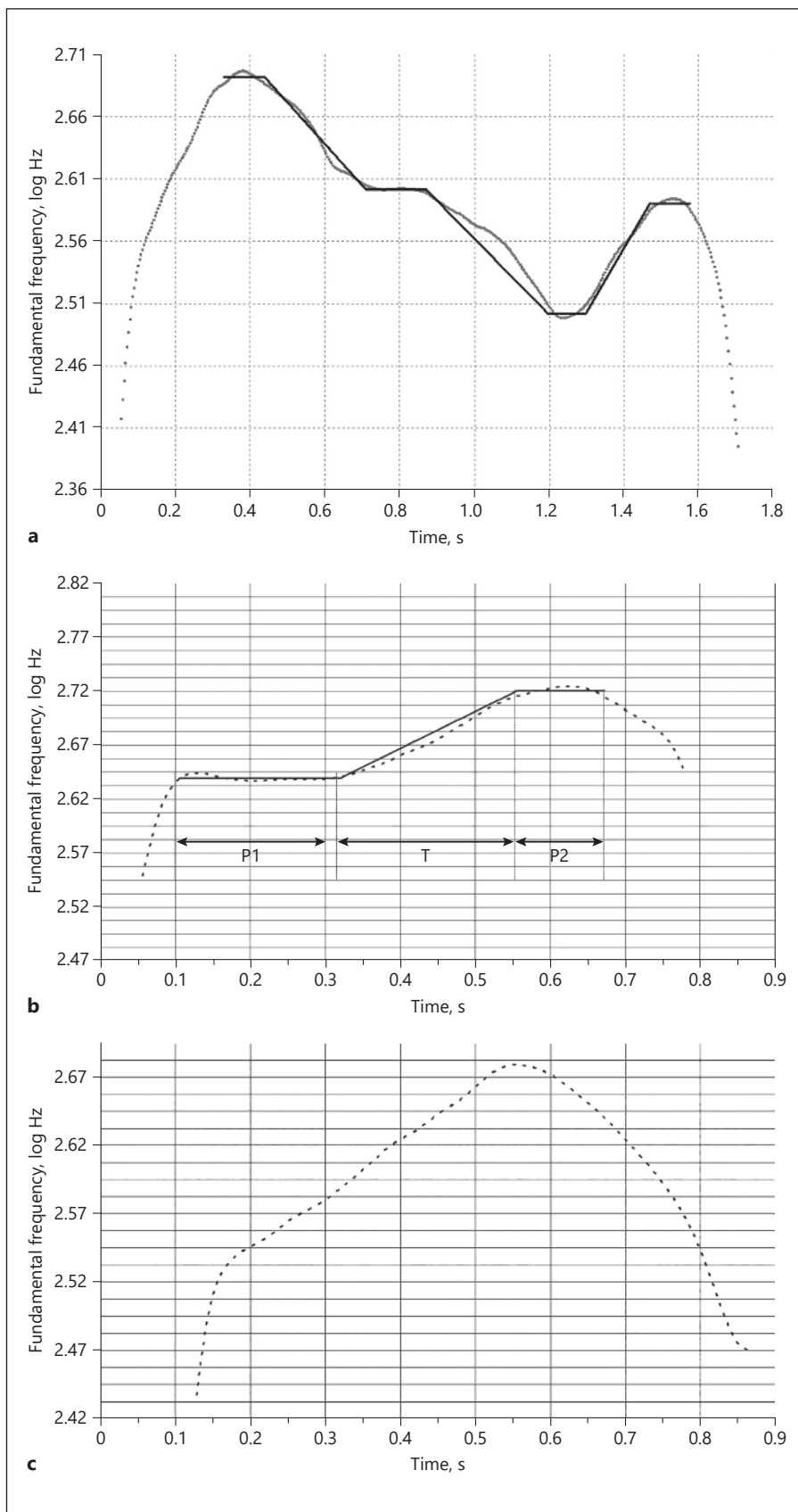
Armbrüster/Mende/Gelbrich/Wermke/
Götz/Wermke

Fig. 1. a Melody diagram exemplifying a double-arc melody structure containing 3 successive single intervals, 2 falling and 1 rising interval (complex event). **b** Melody diagram exemplifying a single-arc melody structure. A single interval was identified on the rising flank of the arc (a simple event). The first f_0 plateau (P1) is linked to the second plateau (P2) by a transition period (T). Interval size is defined by the ratio P2:P1 (here: 325.72 cents). **c** Melody diagram exemplifying a single-arc melody structure without exhibiting a melodic interval.

Musical Intervals In Infant Crying

Table 1. The original and final data set (number of cry utterances) available for each subject per week

Infant		Age, week																Total
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
A		36	34	12	32	58	44	26	17	22	18	50	29	15	9	6	9	417
	Final	7	28	8	26	46	34	20	16	13	16	24	18	15	4	2	7	284
B		31	22	26	27	23	32	33	29	75	10	47	22	19	25	16	0	437
	Final	23	15	15	16	7	22	26	15	24	7	28	17	10	17	4	0	246
C		8	22	11	16	16	15	25	18	19	20	30	18	12	29	30	65	354
	Final	6	9	3	11	9	7	18	15	12	9	16	13	6	18	19	25	196
D		25	71	80	28	81	69	48	8	59	53	26	64	16	58	48	60	794
	Final	13	21	35	11	32	30	21	4	25	18	16	18	6	9	8	30	297
E		34	0	16	38	12	17	37	40	14	48	39	16	35	24	18	25	413
	Final	10	0	9	17	9	11	15	21	4	24	17	4	18	15	16	10	200
F		15	4	35	24	25	28	32	24	34	28	46	21	29	29	26	34	434
	Final	13	1	23	9	21	22	17	20	16	16	41	14	22	19	18	23	295
G		38	3	21	32	48	23	27	15	16	18	22	17	25	39	33	69	446
	Final	22	0	11	23	36	5	14	10	12	11	13	7	18	16	12	5	215
H		36	22	7	16	14	20	37	63	20	19	18	15	31	27	9	14	368
	Final	5	17	6	15	5	12	25	40	17	11	9	7	12	15	5	2	203
I		34	68	53	44	33	36	59	29	50	51	23	38	64	41	24	28	675
	Final	17	25	21	28	8	15	16	26	21	32	5	14	21	13	23	16	301
J		40	44	39	37	45	20	40	35	19	31	90	39	34	38	57	23	631
	Final	19	32	15	11	18	10	20	9	13	20	35	21	25	19	25	18	310
K		61	29	52	40	58	54	70	46	62	24	31	52	26	19	27	48	699
	Final	20	4	9	12	25	25	25	39	31	16	15	36	20	15	11	24	327
L		21	22	43	32	41	27	0	21	33	43	25	24	37	52	19	22	462
	Final	13	4	10	14	5	24	0	20	15	13	20	21	24	37	0	20	240
Total		379	341	395	366	454	385	434	345	423	363	447	355	343	390	313	397	6,130
	Final	168	156	165	193	221	217	217	235	203	193	239	190	197	197	143	180	3,114

trast, a cry melody without a melodic interval is exemplified in Figure 1c. The defined melodic intervals, as described later in more detail, are in a certain sense similar to tone variations in tonal languages or pitch accents in non-tonal languages. Both, the slower f_0 modulations in the form of melody arcs and the sometimes emerging melodic intervals make up the overall shape of the cry melody and are related to intonation and stress patterns in speech prosody. Both are also relatively independent from keys or voice registers. Previous studies postulated that specific musical intervals, mainly semitones, occur regularly in young infants' spontaneous crying [24–26]. This raised the question of how these intervals can be interpreted. Do they reflect an infant's musical talent or musical home? Or, are melodic intervals in cry-

ing reflections of an infant's "melody play", that is his intentional "trying out" of laryngeal opportunities and their perceptual effects?

To date, it is not yet clear whether these interval phenomena are simply coincidental epiphenomena of a partially still immature vocal control mechanism [9, 27]. Hence, the present study aimed to investigate whether the occurrence and complexity of interval-like episodes is age-dependent. If melodic intervals are caused by the immaturity of the vocal system, their probability of occurrence should decrease with maturation across the observation period. Otherwise melodic intervals reflect a regular "melodic play" which is characterized by more and more complex events with increasing maturity of the auditory-vocal system across the age. In both cases, a dem-

Armbrüster/Mende/Gelbrich/Wermke/
Götz/Wermke

Table 2. Number (%) of included and excluded melodies of each infant for interval analysis

Infant	Sex	Included	Excluded	Total
A	f	284 (68.10)	133 (31.89)	417
B	f	246 (56.29)	191 (43.70)	437
C	m	196 (55.36)	158 (44.63)	354
D	f	297 (37.40)	497 (62.59)	794
E	m	200 (48.42)	213 (51.57)	413
F	f	295 (67.97)	139 (32.02)	434
G	f	215 (48.20)	231 (51.80)	446
H	m	203 (55.16)	165 (44.83)	368
I	f	301 (44.59)	374 (55.40)	675
J	m	310 (49.12)	321 (50.87)	631
K	m	327 (46.78)	372 (53.21)	699
L	m	240 (51.94)	222 (48.05)	462
Total		3,114 (50.8)	3,016 (49.2)	6,130

onstration of melodic intervals representing a robust property of healthy infants' spontaneous crying would postulate a further suitable candidate for assessing vocal control in young infants. This would be particularly helpful since occurrence of these interval phenomena can be easily perceived by parents, phoniatricians and speech-language therapists.

Subjects and Methods

Participants

The study used audio files from the archive at the Center for Pre-Speech Development and Developmental Disorders of the University Hospital of Würzburg. For this longitudinal study, vocalizations of 12 (6 female) healthy infants from middle-class monolingual families were carefully selected. Inclusion criteria were an unsuspicious pregnancy, age-adequate somatic development at birth, an unsuspicious perinatal adaptation (APGAR score of ≥8 at 5 min and ≥9 at 10 min), metabolic indices within a normal range and passing of the newborn hearing screening, that is healthy infants. The mean birthweight was 3,298 g (range 2,860–3,795 g), the mean head circumference was 34.7 cm (range 33–37.5 cm) and the mean birth length was 50.6 cm (range 48–52 cm). The mean gestational age was 39 weeks (range 37–42 weeks). Using archived files, we also confirmed that none of the children showed any developmental delay or disorder with respect to language or cognitive development during their first 3 years of life.

Vocalizations

A data set of 6,130 spontaneous cry utterances of the 12 healthy infants was used. The cries were uttered between the age of 1 and 112 days (i.e. 16 weeks; Table 1). A cry utterance (vocalization) was defined as the total vocal output during one single expiration. These utterances were available as anonymized audio (wav) files, recorded at weekly intervals. According to archived file information, original cry recordings were made in an isolated room in a quiet environment within the hospital (first week of life) and the infant's homes. The noise level in the room was perceptually judged to be low in ambient noise and acceptable for audio recordings. The cry recordings were made using a DAT recorder (Sony TCD-D100) coupled to a hand-held directional microphone (SONY ECM-MS950). Cry recording began when an infant started to fuss or at a time when the mother would normally feed the child. None of the cries were associated with the administration of a painful stimulus to the infant. Sampling frequency was 48 kHz, and the dynamic range was 16 bits. The microphone was positioned approximately 10 cm from the infant's mouth. The average duration of each infant's recording was approximately 2 min. To ensure a standardization, each infant was recorded longitudinally by the same research assistant and mostly on the same day of the week.

Method

In total, 6,130 frequency spectrograms were automatically calculated (CSL 4500, KayPentax, USA). Each cry utterance that contained a well-identifiable melody was audiovisually preselected using these frequency spectrograms [20]. Typically, young infants produce a relatively high number of cries with a disturbed melody due to subharmonics or phonatory noise that are readily visible in frequency spectrograms [19, 20, 28]. For reasons of reliability, cry utterances containing noise/subharmonic phenomena that caused a disturbed melody contour and those being very short (fussing; ≤300 ms) were excluded from the final interval analysis (Table 2). Following the preselection of undisturbed utterances, melody diagrams (f_0 contours) were calculated and analysed by using the open-source software PRAAT 6.0.37 [29]. The final data set for interval identification included 3,114 cry utterances with well-identifiable melody contours, on average 260 cries per infant (minimum of 196 cries, maximum of 327 cries per infant; Table 2).

Interval Definition and Identification

A low-pass filtering (40-Hz Gaussian filter) of the melody was performed before identifying melodic intervals [14, 24]. As shown in previous studies [24–26], a single melody contour, independent of its number of melody arcs, can exhibit either none (Fig. 1c), one (Fig. 1b) or more than one melodic interval (Fig. 1a). Corresponding to melodic intervals in music [30], here a single melodic interval was defined as consisting of 3 elements: a first f_0 plateau (P1), defining perceptually a tone (relatively stable f_0 over ≥50 ms, amplitude variation < ± a quartertone), followed by a transition (T), defined by an f_0 gliding from P1 to a subsequent higher/lower f_0 plateau (P2: identical criteria; shown in Fig. 1b). A single interval can exhibit one of 2 interval directions, namely

rising (⌐⌐) or falling (⌐⌐).

Like in musical melodies the f_0 ratio of 2 tones, the f_0 ratio of both f_0 plateaus P1 and P2 (f_0 plateau: geometric mean of all f_0 values) defines the size of the respective melodic interval. Single intervals may occur isolated or in an adjacent sequence. Here, we differentiated between "simple" and "complex" events. A simple event was defined by exhibiting

one (⌐⌐; ⌐⌐) or 2 single intervals (⌐⌐; ⌐⌐; ⌐⌐; ⌐⌐).

A complex event describes any combination of more than 2 single intervals. A complex event describes any combination of more than 2 single intervals. The minimum duration to define an f0 plateau (50 ms) was chosen according to a perceptual study

Fig. 2. Distribution of all identified interval sizes (frequency ratios) over the observation period. Musical intervals: prime ≤50 cent, minor second >50 – ≤150 cent, major second >150 – ≤250 cent, octave >1,150 – ≤1,250.
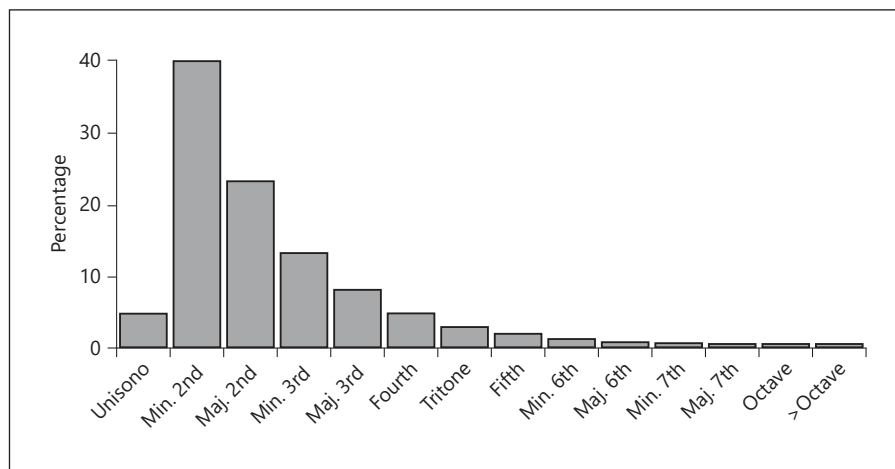
demonstrating this high temporal resolution for tone intervals in infants as young as 2–3 days [31].

Although the applied measurement criteria allowed a reliable quantitative analysis, interval identification was verified using a subset of the data including 100 cry utterances randomly selected. The estimated Cohen's kappa (SPSS 25) for the intrarater reliability (reproducibility) displayed a value of 0.895 ($p < 0.001$). The result of the interrater reliability Cohen's kappa was a value of 0.874 ($p < 0.001$). Both values express a very reliable reproducibility for interval identification. Interval identification and size calculation (ratio of frequency values P1/P2) was performed by using the lab-internal cry data analysis programme (CDAP, pw-project, author P.W.). Melody diagrams (f_0 vs. time) were produced with a logarithmic scale for f_0 and a quarter tone grid (Fig. 1b). Single intervals of the melody contour were identified by using the defined criteria for f_0 plateaus (≥50 ms duration; f_0 variation within ± a quarter tone) and manually marked by cursors (Fig. 1b). Set cursor positions were automatically documented in the Excel tables that contained the f_0/time (melody) values of the respective melody contour. Finally, the first author, who is also a musician, re-evaluated all 6,814 identified melodic intervals auditorily by listening to the respective f_0 contour, using a special PRAATscript provided by J. Mayer (http://praatpfanne.lingphon.net). This additional method was applied to test the reliability and accuracy of the identified intervals – all measured intervals were confirmed to be well perceptible.

Outcome Measures

The analysis focused on 3 outcome measures:
1. *Frequency of occurrence of melodic intervals* in cry utterances with a well-identifiable melody ($n = 3,114$). This was analysed by assigning each cry utterance and its melody structure, respectively, to one of 2 categories defined by *not containing* (Fig. 1c) or *containing at least one interval* (binary variable; Fig. 1a, b). This analysis aimed to answer the question whether cry utterances exhibit significantly more melodies with than without melodic intervals and whether there is an age or sex effect observable.
2. *Event complexity* was analysed by assigning each melody contour of the above-mentioned category "containing at least one

interval" to one of 2 subcategories to whether the contour exhibited (1) only a "simple event" (Fig. 1b) or (2) a "complex event" (Fig. 1a). This analysis aimed to answer the question whether event complexity is related to age and/or sex.
3. *Melodic interval size (f$_0$ plateau ratios)*: To evaluate the frequency of occurrence of melodic intervals, a distribution of all single intervals, independent of their occurrence in simple or complex events, identified across the observation period (1–112 days) was calculated. For assignment of interval sizes along the full scale from prime to octave (well-tempered Western music scale), ratios between 2 plateau frequencies were expressed in cent using the formula: $c = 1{,}200 \cdot 3.322038 \log_{10} (f_2/f_1)$. The musical unit cent is based on the 12-tone equal temperament. Here one octave is defined consisting of 1,200 cents, which can be divided into 12 semitones, each of them containing 100 cents. As shown in Figure 2, this dimension is helpful since further equally tempered intervals can be described by simply adding cents (e.g., the major second obtained 200 cents, the third 300 cents).

Statistical Analysis

Descriptive information on melodic intervals was presented in terms of occurrence and interval sizes. Generalized estimating equations (GEEs) were applied to analyse the dependency of occurrence of melodic intervals on age and sex. GEEs are extensions of commonly known regression models [32, 33]. Note that repeated measurements within an individual are usually statistically dependent. This means that a certain common amount of information is contained in the measurements obtained from the same individual. Therefore, for example, 100 measurements obtained from 10 individuals (10 measurements in each of them) provide less information than 100 measurements in 100 independent individuals (one measurement in each of them). Hence, if we would apply usual logistic regression (which assumes independent measurements), the calculated confidence intervals for odds ratios and probabilities would be too narrow, that is, we would suggest a higher precision of statistical estimates than the actual amount of information allows. In contrast, GEEs appropriately account for the intrasubject dependencies of measurements, thus providing estimates with adequately sized confidence intervals.
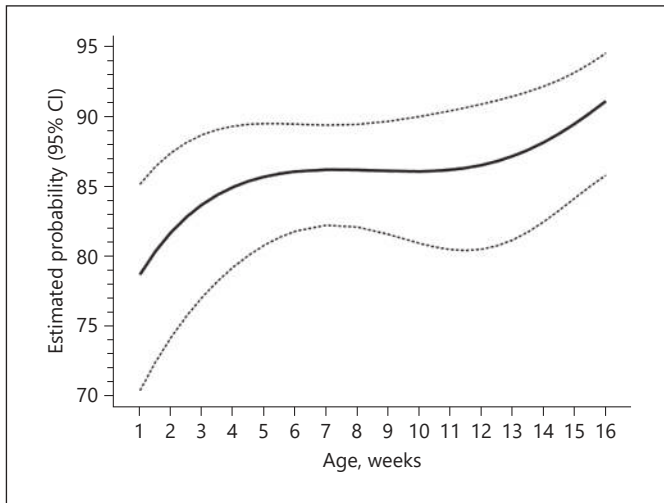
Armbrüster/Mende/Gelbrich/Wermke/
Götz/Wermke

Fig. 3. Estimated probability of occurrence of cry melodies containing at least one single melodic interval across the observation period.
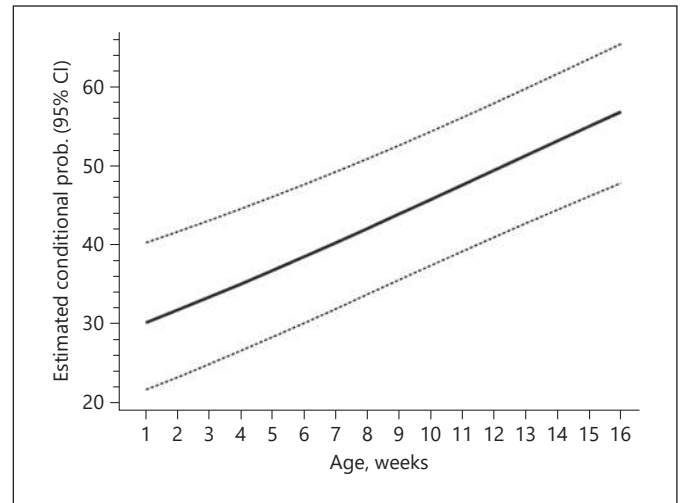

Fig. 4. Estimated conditional probability of occurrence of complex events with age.

To allow for non-linear relationships between the logit of the probability of melodic intervals and age, polynomial dependencies were considered. Since we found different patterns of dependency on age for the occurrence of melodic intervals and their complexity, we did not use an ordinal model, as it might be suitable at first glance, because proportional odds could not be assumed. Instead, we computed 2 models: one for the probability of melodic intervals and a second for the conditional probability of a complex event, given that a melodic interval did occur. Statistical analyses were carried out using SPSS version 25 (released 2017 by IBM Corp., Armonk, NY, USA).

Results

The majority (85.3%) of all analysed cry utterances ($n = 2,657$) exhibited a melody contour containing at least one single melodic interval, whereas only 457 melodies (cry utterances) did not exhibit entities of the melodic intervals defined here across the observation period. In these 2,657 cry utterances, a number of 6,814 single melodic intervals were identified over the first 16 weeks (approx. 4 months) of life. With respect to direction, 46.4% of single melodic intervals were identified as

rising (⌐), while 53.6% were falling (¬) intervals.

The result of the interval size analysis is shown in Figure 2. Each measured single interval was assigned to one class of musical intervals to demonstrate a respective distributional pattern over the observation period of 1–112 days (approx. 4 months). The minor second was found to

Table 3. Weekly probability of occurrence of cry utterances with a melody containing at least one single interval

Age, weeks	Estimate, %	95% CI
1	78.7	70.4–85.2
2	81.8	74.2–87.5
3	83.8	77.1–88.8
4	85.0	79.3–89.4
5	85.8	80.8–89.6
6	86.1	81.9–89.5
7	86.1	81.9–89.5
8	86.1	81.9–89.5
9	86.1	81.9–89.5
10	86.1	81.9–89.5
11	86.2	80.5–90.5
12	86.6	80.6–91.0
13	87.2	81.2–91.5
14	88.2	82.5–92.2
15	89.5	84.2–93.2
16	91.2	85.9–94.6

be the most prominent interval across all months (1–4 months: 43.7, 38.9, 37.7, 40.2%). The number of all other intervals decreased with increasing interval size. However, smaller intervals than semitones were also identified (shown in Fig. 2).

The first GEE model for the probability of the occurrence of melodic intervals revealed a non-linear relationship between the logit of that probability and age. The first ($p < 0.001$), second ($p = 0.01$) and third ($p = 0.03$)

Table 4. Conditional probability of occurrence of cry melodies containing complex events across the observation period

Age, weeks	Estimate, %	95% CI
1	30.2	21.8–40.3
2	31.8	23.3–41.7
3	33.5	25.0–43.1
4	35.1	26.7–44.6
5	36.8	28.4–46.1
6	38.6	30.2–47.7
7	40.3	32.0–49.3
8	42.1	33.8–51.0
9	44.0	35.6–52.7
10	45.8	37.5–54.4
11	47.6	39.2–56.2
12	49.5	41.0–58.0
13	51.4	42.8–59.8
14	53.2	44.5–61.7
15	55.1	46.2–63.6
16	56.9	47.9–65.5

Fig. 5. Age-dependent probability of melodies without, with simple and with complex events.

powers of age were significant. This relationship is shown in Figure 3. No significant relationship with sex could be detected ($p = 0.92$). As well, there was no significant interaction of sex with age ($p = 0.69$) and its square ($p = 0.44$) or cube ($p = 0.36$). The resulting probabilities are displayed in Table 3.

In the second model for complex events, only the dependency on age was significant ($p = 0.001$) but not the relationships with the square ($p = 0.29$) and its cube ($p = 0.18$; Fig. 4). Again, no significant relationships with sex ($p = 0.77$) or its interaction with age ($p = 0.92$) were observed. The resulting probabilities are displayed in Table 4. The combined results of both GEE analyses (i.e., the age-dependent fractions of melodies with no, simple or complex events) are shown in Figure 5.

Discussion

Infants learn their native language fast and effortlessly during the first years of life. Meanwhile there is mounting evidence for the preparatory function of early vocal development for subsequent developmental steps towards language [6, 18, 34–36]. Both, "emotional" and "linguistic" prosody involve the elements of f_0, rhythm and other elements as they relate to the grammatical structuring of meaning in utterances. Among other prosodic elements, melody is tied in complex ways to phonology, the systematic encoding of meaning in the sounds of language.
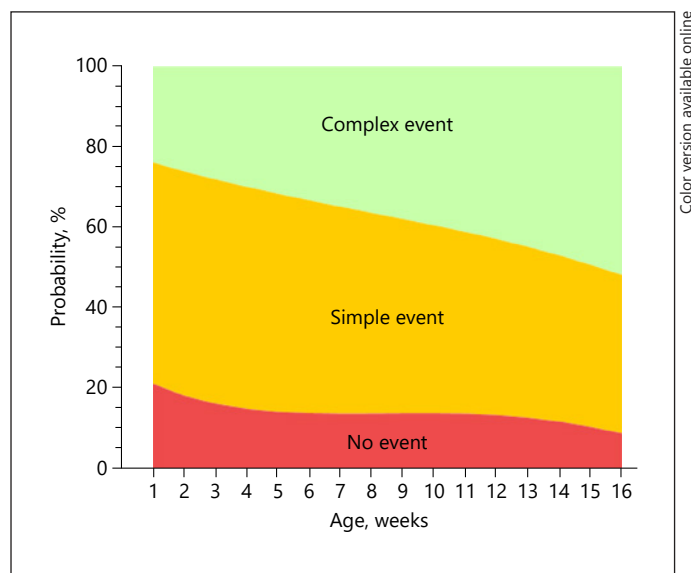
Undoubtedly, it is important to recognize features of early vocal development that contribute to the emergence of speech; this includes constituents of infant sounds which in one or the other way provide the raw material for the later speech sounds, particularly syllables [37]. However, suprasegmental "musical" characteristics (prosody) are not only as decisive for spoken languages as segmental, but also have among their precursors one that is crucial to acquire language, from both the ontogenetic as well as the phylogenetic perspective [5–7, 14, 18, 22, 31, 38–41]. This is sound melody, as time function of the f_0 ("intonation"). With regard to prosodic features of infants' sounds, melodic contours are of primary importance and, moreover, melody provides a kind of scaffolding for the developing articulatory abilities [6, 18, 41]. According to Brown's "musilanguage" hypothesis, a common source for both music and language existed in prehistoric times in a stage of "lexical tones." Discrete tones were matched to meanings and combined to "melodorhythmic" phrases [42]. Human infants' perceptual and productive preferences for musical elements of speech (prosody) support Brown's hypothesis and bridge the importance of melody from phylogenesis to ontogenesis with respect to language development.

Particularly over about the last 20 years, brain studies and behavioural experiments accumulated considerable evidence for precocious performances with respect to melody and rhythm of the fetus, newborn and young in-

Armbrüster/Mende/Gelbrich/Wermke/
Götz/Wermke

Color version available online

fant [43–47]. There is general agreement among scientists and speech-language therapists that melody contour is key to understanding infants' early steps of vocal development towards language. Behavioural studies reported the discrimination of melodic contours [46–48] and temporal speech sound patterns in newborns and 1- to 4-month-old infants [49–52].

The present study highlights the regularly occurring and age-dependently developing melody properties that define melodic intervals in spontaneous crying over the first 4 months of life. The probability of vocalization melody containing at least one single melodic interval was found to be high already in newborns' crying (78.7%; 95% CI 70.4–85.2) and significantly increased to 91% (95% CI 85.9–94.6) over the next 15 weeks. A non-linear development in 3 main phases was observed: (1) a steep increase within the first 1.5 months, (2) a relatively stable phase (consolidation) until about the end of the third month and (3) a new increase across the fourth month of life (Fig. 3). This result demonstrated that melodic intervals seem to be a regularly occurring phenomenon in spontaneous crying of healthy infants, most likely reflecting both, an increasing vocal control and vocal flexibility during the first 4 months of life [35, 53–56]. Moreover, our findings demonstrated also that interval complexity and, hence, intentional "melody play" increased in spontaneous crying with age (from 30% in week 1 to 57% in week 16).

Consequently, we found a dramatic increase in melodic intervals across the first 6 weeks that was followed by a temporarily steady state and performance consolidation over the next 6 weeks. However, interval complexity further increased during the stabilization phase and beyond (Fig. 4). This means that the development process is identical to that found for the overall contour development [18]. Here too, the development of melodic complexity runs through 3 phases in the same time pattern [18]. This suggests that both developmental processes are synchronized to generate the overall shape of the melody contour or cry intonation. In summary, the second and third months represent a first prosodic stabilization phase. Wermke and Mende [18] postulated that during this time previously trained melody patterns of different time scales now function as scaffolding for articulatory activity that becomes continuously more prominent during the next months of life [41]. This results in a synchronous development of suprasegmental elements with the emergence of segmental elements starting at about the third month of life, when typical cooing sounds with their vowel- and consonant-like elements (vocants and closants) are pro-

duced during face-to-face interactions. When babbling with its syllable-like constituents occurs over the next months, prosody again is important for further vocal development towards language [34, 57]. However, we analysed only cry vocalizations and it remains to be determined if melodic intervals also occur in non-cry vocalizations and if so, how they interact with segmental constituents of cooing and babbling. Several studies have shown a synchronous occurrence of voice phenomena that are independent of vocalization type during the first months [e.g., 6, 18, 35]. Our results contribute to the assumption that the young infant is "trying out" various control options of the vocal control system at the same time. The combinatory sound constituents will eventually result in a huge vocalization repertoire. Admittedly, the sample size of the present study was rather small and thus requires repetition with a larger corpus. This would also allow to analyse variations of interval size distributions in individual infants with age. However, this would require an automatic interval identification (or a larger "analysis crew") since the analyses were highly time consuming. Meanwhile, we encourage phoniatricians and speech-language therapists to "musically" listen to infant crying.

Infants do not only exhibit precocious abilities in melody production, but also have a "musical ear." Infants' processing of musical or music-like patterns is much like that of adults [48, 58]. Especially melodic contour seems to be a salient feature for infant listeners. Young infants have perceptive capabilities for rising/falling pitch contours [58–60] and the detection of beat [60–62]. Various studies demonstrated the ability of infants to perceive f_0 contours and to recognize melody intonation and melodic intervals [31, 58–65]. By 2 months of age, infants are able to remember short melodies and can discriminate those from similar melodies [66]. Sandra Trehub has been investigating infants' musical predispositions for many years. For example, Trehub et al. [67, 68] discovered that infants are capable of noticing a 1-semitone change within a melody, when the interstimulus intervals were brief and the standard and comparison melodies were presented in the same key. These findings could support the remarkable preference for the semitone interval in our results across the observation period. Trehub et al. [48] also found that infants are more attentive and exhibit more positive affect while listening to consonant music than to music with many dissonant intervals. The authors suggest that the "preferences could arise from the predominance of consonant intervals in the infant's environment (e.g., speech, music, environmental sounds)" [48].

Patel et al. [69] described a "predominance" for small intervals (with a maximum of 2 semitones) in cross-cultural music and speech melodies (namely 56% in English and French spoken sentences). Neuroscientists and musicologists even discuss the semitone to be a perceptual boundary in pitch-interval discrimination [70]. If true, young infants are hard-wired with this ability as our demonstration of the semitone as prevailing melodic interval in spontaneous crying seems to suggest.

The ability of a young infant to generate melodic intervals in crying with its typical high subglottal pressures and tiny vocal folds as a young infant demands a well-functioning and very fast-acting laryngeal-respiratory coordination. The developmental shift from simple to complex events (Fig. 5) seems to be the result of a very steady maturational process of vocal control without any discontinuity. It would be worthwhile to explore whether melodic intervals in infants' vocalizations could establish a risk marker for mature laryngeal, respiratory, phonatory and vocal control capacity.

Conclusion

The present quantitative analysis provided a further example for the development of prosodic elements during very early stages towards language. However, so far we cannot answer the question whether melodic intervals are a crucial property for normal vocal development or just a lovely reflection of infants' musical predispositions from birth. In any case, it seems to be important to listen to the young infants' crying from this "musical" perspective. The study further contributes to the recognition that infant spontaneous crying must be set apart from "animal call-like vocalization" and be treated as an essential vocalization type on the path towards spoken language.

Statement of Ethics

The audio files analysed here were selected from the anonymized archive of the Center for Pre-Speech Development and Developmental Disorders, University Hospital, Würzburg, Germany. Therefore, an ethical approval was not required.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Author Contributions

All authors have made substantial contributions to the conception of the work. W.M., K.W., R.G. and L.A. designed the research. K.W. supervised and coordinated recordings in several of her previous projects; P.W. is responsible for the audio file archive and data management. K.W., P.W. and L.A. analysed the cry utterances; L.A. identified and measured the melodic intervals and, together with K.W., wrote the main text. G.G. conducted and described statistical analysis. L.A., P.W. and G.G. prepared the figures. All authors jointly interpreted the results and edited the paper.

References

1 Wolff PH. The Development of Behavioral States and the Expression of Emotions in Early Infancy: New Proposals for investigation. Chicago: The University of Chicago Press; 1987.

2 Barr R, Hopkins B, Green J. Crying as a Sign, Symptom & a Signal. London: Cambridge University Press; 2000.

3 Akhnikh S, Engelberts A, van Sleuwen B, L'Hoir M, Benninga M. The Excessively Crying Infant: Etiology and Treatment. Pediatr Ann. 2014 Apr;43(4):e69–75.

4 Karp H. The happiest baby on the block: The new way to calm crying and help your newborn baby sleep longer. New York: Bantam Books; 2015.

5 Wermke K, Mende W. Musical elements in human infants' cries: in the beginning is the melody. Music Sci. 2009;13(2):151–75.

6 Wermke K, Mende W. From Emotion to Notion: The Importance of Melody. In: Decety J, Cacioppom JT, editors. The Oxford Handbook of Social Neuroscience. New York: Oxford University Press; 2011. pp. 624–48.

7 Wermke K. Neonatal crying behaviors. In: Wright JD, editor. International encyclopaedia of the social and behavioral sciences. Oxford: Elsevier; 2015. p. 475–80.

8 Wermke K, Sereschk N, May V, Salinger V, Sanchez MR, Shehata-Dieler W, et al. The Vocalist in the Crib: the Flexibility of Respiratory Behaviour During Crying in Healthy Neonates. J Voice. 2019, Online ahead of print.

9 Schweinfurth JM, Thibeault SL. Does hyaluronic acid distribution in the larynx relate to the newborn's capacity for crying? Laryngoscope. 2008 Sep;118(9):1692–9.

10 Bosma JF, Truby HM, Lind J. Cry motions of the newborn infant. Acta Paediatr. 1965; 54(s163):60–92.

11 LaGasse LL, Neal AR, Lester BM. Assessment of infant cry: acoustic cry analysis and parental perception. Ment Retard Dev Disabil Res Rev. 2005;11(1):83–93.

12 Wasz-Höckert O, Lind J, Vuorenkoski V, Partanen T, Valanne E. The Infant Cry: A Spectrographic and Auditory Analysis. London: Mac Keith Press; 1968.

13 Michelsson K. Cry analyses of symptomless low birth weight neonates and of asphyxiated newborn infants. Acta Paediatr Scand Suppl. 1971;216:1–45.

14 Wermke K, Leising D, Stellzig-Eisenhauer A. Relation of melody complexity in infants' cries to language outcome in the second year of life: a longitudinal study. Clin Linguist Phon. 2007 Nov-Dec;21(11-12):961–73.

15 Wermke K, Birr M, Voelter C, Shehata-Dieler W, Jurkutat A, Wermke P, et al. Cry melody in 2-month-old infants with and without clefts. Cleft Palate Craniofac J. 2011 May; 48(3):321–30.

16 Shinya Y, Kawai M, Niwa F, Imafuku M, Myowa M. Fundamental Frequency Variation of Neonatal Spontaneous Crying Predicts Language Acquisition in Preterm and Term Infants. Front Psychol. 2017 Dec; 8(Dec):2195.

17 François C, Teixidó M, Takerkart S, Agut T, Bosch L, Rodriguez-Fornells A. Enhanced Neonatal Brain Responses To Sung Streams Predict Vocabulary Outcomes By Age 18 Months. Sci Rep. 2017 Sep;7(1):12451.

18 Wermke K, Mende W. From Melodious Cries to Articulated Sounds: Melody at the Root of Language Acquisition. In: Fonseca-Mora MC, Gant M, editors. Melodies, Rhythm and Cognition in Foreign Language Learning. Lady Stephenson Library. Newcastle upon Tyne: Cambridge Scholars Publishing; 2016. pp. 24–47.

19 Mende W, Wermke K, Herzel H. Bifurcations and chaos in newborn infant cries. Phys Lett A. 1990 Apr;145(8-9):418–24.

20 Fuamenya NA, Robb MP, Wermke K. Noisy but effective: crying across the first 3 months of life. J Voice. 2015 May;29(3):281–6.

21 Robb MP. Bifurcations and chaos in the cries of full-term and preterm infants. Folia Phoniatr Logop. 2003 Sep-Oct;55(5):233–40.

22 Fuller BF, Horii Y. Differences in fundamental frequency, jitter, and shimmer among four types of infant vocalizations. J Commun Disord. 1986 Dec;19(6):441–7.

23 Lind K, Wermke K. Development of the vocal fundamental frequency of spontaneous cries during the first 3 months. Int J Pediatr Otorhinolaryngol. 2002 Jun;64(2):97–104.

24 Armbrüster L, Mende W, Ehlert H, Wermke G, Wermke K. Musical Intervals in Baby Sounds. 10th International Conference of Students of Systematic Musicology. London: https://sysmus17.qmul.ac.uk/wp-content/uploads/2017/08/armbruster_intervals_baby_sounds.pdf [21.05.20]; 2017.

25 Weigl R. Lärm oder Lied? Analyse musikalischer Intervalle in den Lautäußerungen von Säuglingen als komplexe vorsprachliche Fähigkeit. [Noise or tune? Analysis of musical intervals in infant cry melodies as a complex pre-speech competence.] Unpublished diploma-thesis. Germany: University Würzburg; 2012.

26 Dobnig D, Stephan S, Wermke P, Wermke K. It all starts with music – musical intervals in neonatal crying. Speech Lang Hear. 2017; 21(2):77–80.

27 Rosenberg TL, Schweinfurth JM. Cell density of the lamina propria of neonatal vocal folds. Ann Otol Rhinol Laryngol. 2009 Feb;118(2):87–90.

28 Titze IR. Vocal Fold Physiology. San Diego, California: Singular Publishing Group; 1993.

29 Boersma P, Weenink D. Praat: Doing phonetics by computer [Computer program]. Version 6.0.37. [retrieved 14 March 2018]. http://www.praat.org/. 2011.

30 Deutsch D. The Psychology of Music. Cognition and Perception. Academic Press; 1998. pp. 215–64.

31 Stefanics G, Háden GP, Sziller I, Balázs L, Beke A, Winkler I. Newborn infants process pitch intervals. Clin Neurophysiol. 2009 Feb; 120(2):304–8.

32 Dunteman GH, Ho MR. An introduction to generalized linear models. Thousand Oaks: Sage Publications; 2006.

33 Garson DG. Generalized Linear Models & Generalized Estimating Equations. Asheboro (NC): Statistical Publishing Associates; 2012.

34 Vihman M. Phonological Development. The origins of language in the child. Oxford: Basil Blackwell; 1996.

35 Robb MP, Yavarzadeh F, Schluter PJ, Voit V, Shehata-Dieler W, Wermke K. Laryngeal Constriction Phenomena in Infant Vocalizations. J Speech Lang Hear Res. 2019 Dec; 63(1):49–58.

36 Lacheret-Dujour A, Bobin-Bègue A, Devouche E, Gratier M. The perception of intention and emotion in non-cry pre-babbling infant vocalizations. 9th International Conference on Speech Prosody. Poznań, Poland; 2018 [retrieved from https://halshs.archives-ouvertes.fr/halshs-01974915/document).

37 Oller DK. The Emergence of Speech Capacity. Lawrence Erlbaum Associates; 2000.

38 Tonkova-Yampolskaya RV. Development of speech intonation in infants during the first two years of life. In: Ferguson CA, Slobin DI, editors. Studies of Child Language Development. New York: Holt; 1973. pp. 128–38.

39 D'Odorico L. Non-segmental features in pre-linguistic communications: an analysis of some types of infant cry and non-cry vocalizations. J Child Lang. 1984 Feb;11(1):17–27.

40 Delack JB, Fowlow PJ. The ontogenesis of differential vocalizations: Development of prosodic contrastivity during the first year of life. In: Waterson N, Snow C, editors. The development of communication. Chichester: Wiley; 1978. pp. 93–110.

41 Wermke K, Mende W, Manfredi C, Bruscaglioni P. Developmental aspects of infant's cry melody and formants. Med Eng Phys. 2002 Sep-Oct;24(7-8):501–14.

42 Brown S. The „musilanguage" model of music evolution. In: Wallin NL, Merker B, Brown S, editors. The origins of music. Cambridge (MA): MIT Press; 2000. pp. 271–300.

43 Hepper PG, Shahidullah BS. Development of fetal hearing. Arch Dis Child Fetal Neonatal Ed. 1994 Sep;71(2):F81–7.

44 Moon C, Cooper RP, Fifer WP. Two-day old infants prefer native language. Infant Behav Dev. 1993 Oct-Dec;16(4):495–500.

45 Mampe B, Friederici AD, Christophe A, Wermke K. Newborns' cry melody is shaped by their native language. Curr Biol. 2009 Dec; 19(23):1994–7.

46 Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertoncini J, Amiel-Tison C. A precursor of language acquisition in young infants. Cognition. 1988 Jul;29(2):143–78.

47 Granier-Deferre C, Bassereau S, Ribeiro A, Jacquet AY, Decasper AJ. A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. PLoS One. 2011 Feb; 6(2):e17304.

48 Trehub SE. Musical predispositions in infancy. Ann N Y Acad Sci. 2001 Jun;930(1):1–16.

49 Friederici AD, Friedrich M, Christophe A. Brain responses in 4-month-old infants are already language specific. Curr Biol. 2007 Jul; 17(14):1208–11.

50 Sansavini A, Bertoncini J, Giovanelli G. Newborns discriminate the rhythm of multisyllabic stressed words. Dev Psychol. 1997 Jan; 33(1):3–11.

51 Dehaene-Lambertz G, Dehaene S. Speed and cerebral correlates of syllable discrimination in infants. Nature. 1994 Jul;370(6487):292–5.

52 van Ooijen B, Bertoncini J, Sansavini A, Mehler J. Do weak syllables count for newborns? J Acoust Soc Am. 1997 Dec;102(6):3735–41.

53 Gregory A, Tabain M, Robb M. Duration and Voice Quality of Early Infant Vocalizations. J Speech Lang Hear Res. 2018 Jul;61(7):1591–602.

54 Kent RD, Mitchell PR, Sancier M. Evidence and role of rhythmic organization in early vocal development in human infants. In: Fagard J, Wolff PH, editors. The development of timing control and temporal organization in coordinated action: invariant relative timing, rhythms and coordination. Amsterdam: North Holland; 1991. p. 135–49.

55 Wermke K, Haschemi AA, Hesse V, Robb MP. Inspiratory Phonation in Baby Voice. J Voice. 2018 Mar;32(2):185–91.

56 Esling JH, Benner A, Moisik SR. Laryngeal articulatory function and speech origins. In: Little H, editor. 18th International Congress of Phonetic Sciences (ICPhS 2015) Satellite Event: The Evolution of Phonetic Capabilities: Causes constraints, consequences (pp. 2-7). Glasgow: Little, H.; 2015. pp. 2–7.

57 Davis BL, MacNeilage PF, Matyear CL, Powell JK. Prosodic correlates of stress in babbling: an acoustical study. Child Dev. 2000 Sep-Oct;71(5):1258–70.

58 Trehub SE, Bull D, Thorpe LA. Infants' perception of melodies: the role of melodic contour. Child Dev. 1984 Jun;55(3):821–30.

59 Trehub SE, Thorpe LA, Morrongiello BA. Organizational processes in infants' perception of auditory patterns. Child Dev. 1987 Jun; 58(3):741–9.

60 Carral V, Huotilainen M, Ruusuvirta T, Fellman V, Näätänen R, Escera C. A kind of auditory 'primitive intelligence' already present at birth. Eur J Neurosci. 2005 Jun;21(11): 3201–4.

61 Háden GP, Honing H, Török M, Winkler I. Detecting the temporal structure of sound sequences in newborn infants. Int J Psychophysiol. 2015 Apr;96(1):23–8.

62 Winkler I, Háden GP, Ladinig O, Sziller I, Honing H. Newborn infants detect the beat in music. Proc Natl Acad Sci USA. 2009 Feb; 106(7):2468–71.

63 Wermke K, Teiser J, Yovsi E, Kohlenberg PJ, Wermke P, Robb M, et al. Fundamental frequency variation within neonatal crying: does ambient language matter? Speech Lang Hear. 2016 Jun;19(4):211–7.

64 Chang HW, Trehub SE. Infants' perception of temporal grouping in auditory patterns. Child Dev. 1977 Dec;48(4):1666–70.

65 Trehub SE. The developmental origins of musicality. Nat Neurosci. 2003 Jul;6(7):669–73.

66 Plantinga J, Trainor LJ. Melody recognition by two-month-old infants. J Acoust Soc Am. 2009 Feb;125(2):EL58–62.

67 Trehub SE. In the Beginning: A Brief History of Infant Music Perception. Music Sci. 2010 Sep;14(2):71–87.

68 Trehub SE, Cohen AJ, Thorpe LA, Morrongiello BA. Development of the perception of musical relations: semitone and diatonic structure. J Exp Psychol Hum Percept Perform. 1986 Aug;12(3):295–301.

69 Patel AD, Iversen JR, Rosenberg JC. Comparing the rhythm and melody of speech and music: the case of British English and French. J Acoust Soc Am. 2006 May;119(5 Pt 1):3034–47.

70 Mary Zarate J, Ritson CR, Poeppel D. Pitch-interval discrimination and musical expertise: is the semitone a perceptual boundary? J Acoust Soc Am. 2012 Aug;132(2):984–93.