Infant Behavior and Development 74 (2024) 101908

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

The role of acoustic features of maternal infant-directed singing in enhancing infant sensorimotor, language and socioemotional development

Raija-Leena Punamäki [a,*], Safwat Y. Diab [a], Konstantinos Drosos [a,b], Samir R. Qouta [c], Mervi Vänskä [a]

[a] *Tampere University, Faculty of Social Sciences, Department of Psychology, Finland*
[b] *Nokia Research Center, Espoo, Finland*
[c] *Doha Institute for Graduate Studies, School of Social Sciences and Humanities, Qatar*

A R T I C L E I N F O

Keywords:
Infant-directed singing
Acoustic features
Infant sensorimotor
Languge and socioemotional development

A B S T R A C T

The quality of infant-directed speech (IDS) and infant-directed singing (IDSi) are considered vital to children, but empirical studies on protomusical qualities of the IDSi influencing infant development are rare. The current prospective study examines the role of IDSi acoustic features, such as pitch variability, shape and movement, and vocal amplitude vibration, timbre, and resonance, in associating with infant sensorimotor, language, and socioemotional development at six and 18 months. The sample consists of 236 Palestinian mothers from Gaza Strip singing to their six-month-olds a song by their own choice. Maternal IDSi was recorded and analyzed by the OpenSMILE- tool to depict main acoustic features of pitch frequencies, variations, and contours, vocal intensity, resonance formants, and power. The results are based on completed 219 maternal IDSi. Mothers reported about their infants' sensorimotor, language-vocalization, and socioemotional skills at six months, and psychologists tested these skills by Bayley Scales for Infant Development at 18 months. Results show that maternal IDSi characterized by wide pitch variability and rich and high vocal amplitude and vibration were associated with infants' optimal sensorimotor, language vocalization, and socioemotional skills at six months, and rich and high vocal amplitude and vibration predicted these optimal developmental skills also at 18 months. High resonance and rhythmicity formants were associated with optimal language and vocalization skills at six months. To conclude, the IDSi is considered important in enhancing newborn and risk infants' wellbeing, and the current findings argue that favorable acoustic singing qualities are crucial for optimal multidomain development across infancy.

The need to hear the caregiver's voice is inherent for newborn babies (Filippa et al., 2017) and already fetuses recognize maternal voice (Jardri et al., 2012; Kisilevsky et al., 2009). Subsequently, the prosodic and acoustic features of maternal infant-directed speech (IDS) and singing (IDSi) are considered vital to newborn and infant wellbeing and development (Corbeil et al., 2013; Franco et al., 2022; Saint-Georges et al., 2013; Spinelli et al., 2017), possibly predicting optimal sensorimotor, language, and socioemotional development. Favorable, prototypic or protomusical IDSi involves specific acoustic, prosodic, lexical, and rhythmic features that

* Corresponding author.
E-mail address: Raija-leena.punamaki-gitai@tuni.fi (R.-L. Punamäki).

https://doi.org/10.1016/j.infbeh.2023.101908
Received 21 February 2023; Received in revised form 30 October 2023; Accepted 10 November 2023
Available online 22 November 2023

infants prefer to IDS and adult-directed singing (Dunst et al., 2012; Saint-Georges et al., 2013; Trainor, 1996), reflecting uniquely meaningful, exaggerated, and lively mother-infant communication patterns that serve alleviating infant distress, activating and entraining affect regulation, and maintaining attention (Corbeil et al., 2016; Trehub, 2001). These IDSi features encompass (a) uniquely high, wide, and smooth fundamental pitch F0 *frequencies* and F0 *variation,* indicated by differences between lowest and highest pitch (F0) measured by Hertz (Hz), (b) greater *intensity and range in amplitude*, indicated by loudness and energy, (c) plentiful, clear, and rich *rhythm and vocal resonance,* indicated by formants, and (d) distinctive and varied (rising, falling, and sinusoid) *shapes and movements of pitch variation* over time, indicated by F0 contours. Opposite to exaggeration nature of IDSi, a prototypic feature involves (e) low *tempo* and *power,* especially in lullabies aiming at soothing the infant (Falk & Kello, 2017; Trainor, 1996).

Although these protomusical acoustic features of IDSi are theoretically well related to optimal infant development (Thiessen & Saffran, 2009; Trehub, 2001), empirical research is scarce and narrow. First, the prototypic acoustic features are mainly studied in maternal IDS, but not in IDSi, and are commonly indicated by only three pitch (F0) acoustic features, i.e., by F0 fundamental frequencies, F0 variability, and F0 contours (Spinelli et al., 2017). The current study provides a more comprehensive view by analyzing, in addition, vocal amplitude intensity, energy, and vibration, vocal rhythm and resonance, and tempo and power as indicators of acoustic features of mothers' IDSi. Second, studies are available on the role of maternal IDS acoustic features in newborn attention and prelingual skills and, to some extent, in infant language learning (Schön et al., 2008; Spinelli et al., 2017), but less is known about how IDSi acoustic features are impacting multiple domains of infant development across a longer developmental period. The current study analyses how a comprehensive range of acoustic features of mothers' singing to their six-month-old infants are associated with the sensorimotor, language, and socioemotional development in early (six months) and late (18 months) infancy.

1. Mechanisms explaining the role of infant-directed singing

Three broad mechanisms can be delineated to explain why maternal IDSi protomusical acoustic features of rich and varied pitch, high amplitude intensity and vibration, and distinctive and extensive contours contribute to optimal infant development. The first emphasizes the goals and functions of the IDSi itself in regulating infants' arousals, stabilizing attention, and helping them interpret and response to maternal emotional cues as vital part of bonding (Fernald, 1992; Saint-Georges et al, 2013; Trehub & Trainor, 1998). The second mechanism focuses on infants' developmental windows or periods when they are especially sensitive to predictable and synchronized communicative interaction that IDSi represents (Niwano & Sugai, 2002; Ullsten et al., 2018). Third, according to neurobiological mechanisms infant brain creates new connections and strengthens and reorganizes existing neural networks based on infants' auditory, sensory, motor, visual experiences, and thus the IDSi is effectively enhancing corresponding cortical processes (Hunnius & Meyer, 2020; Van Druden et al., 2023; Virtala & Partanen, 2018).

Parents employ IDSi to soothe, calm down and console stressed and fussy infants (Trehub, 2001; Fernald, 1992). Successful modulation of distress and affect attunement has immediate benefits of increasing infants' pleasure and decreasing discomfort and subsequently advances their self-regulation and socioemotional skills (Feldman, 2007). Affect attunement closely functions with innate vitality affects, connected to vital body rhythms and synchronies, such as breathing, blood pressure, and heart beating, and rhythmic processes of hunger and sleep (Ullsten et al., 2018). The IDSi is conveying a variety of emotions, facilitating infants' optimal self-regulation, stabilizing dyadic interaction, and balancing stress, which is crucial to promote socioemotional development (Corbeil et al., 2016; Lense et al., 2022; Ullsten et al., 2017; Ullsten et al., 2018). Importantly, mothers accommodate the tempo, timbre and intonation of their IDSi to follow the infants' breathing and attuning rhythm (Trainor, 1996). It might be that IDSi creates a close interplay between bodily, sensorimotor, and neuroendocrinological (e.g., cortisol stress-reactivity) regulation and affective, social, and dyadic regulation in infants' experiences, which makes IDSi an effective facilitator of an optimal cognitive and emotional development.

The IDSi function of promoting moderate arousal levels is crucial in sustaining infants' attention. Attentive state of mind facilitates learning new skills, such as imitation, orientation to novelties and new social engagements, which then underlie development of language and other cognitive skills (Franco et al., 2022; Senju & Csibra, 2008). The rhythm structure of the IDSi involves regular, repeated, and predicable beats with heightened contrasts and melodical alternations, thus providing infants primary experiences of discriminating and remembering rhythms, sense of time-aligned actions and expecting and coordinating responses (Lense et al., 2022). These experiences in turn form the building blocks of cognitive processes of memory, sense of causality, and task devotion. Importantly, in a state of sustained attention infants are capable to take advantage of multiple and complex converging cues, such as lyrics paired with consistent melodies, thus explaining IDSi contributing to infants' language learning (Thiessen & Saffran, 2009). Synchrony is considered vital in safe and optimal dyadic interaction (Feldman, 2007), and the IDSi is a perfect training platform for reciprocal, shared, and partaking interaction. In these actions infants further improve their still immature attentional, communicative, and motoric capacities. Apparently protomusical acoustic features of IDSi help create optimal receptive state of mind, attention, and reciprocal responsiveness that all are necessary precondition to multidomain developmental skills.

The regular pulse of music intensifies emotional sharing and synchrony between mother and infant, thus proving vital element for bonding and socioemotional development (Lense et al., 2022; Nakata & Trehub, 2004). Infants learn to recognize and express a variety of emotions in dyadic interaction, and in 2–6 months they can recognize and from seven months express basic emotions of fear, sadness, anger, and joy (Ruba & Repacholi, 2020; Leppänen & Nelson, 2009). Vocal expressions of specific and discrete emotions form a core of human communication, and IDSi provides, again, an entrancing interaction to train further more complex emotions, such as envy, shame, guilt, or satisfaction (Russell, Bachorowski, & Fernández-Dols, 2003). It is noteworthy that IDSi acoustic features of high pitch, rhythmic variations, vocal richness and vivacity, and arousing waves correspond exactly the speech characteristics that reflect positive emotions and happiness (Juslin & Laukka, 2003). Exaggerated and high pitch levels of IDSi may thus not only be attuning and bonding due to their greater salience to infants, but also because they communicate positive affect (Fernald, 1992). Characteristic to dyadic interaction boosting optimal cognitive and emotional development are moments of meeting involving shared happiness and joy (Mäntymaa et al., 2015).

The IDSi presents a salient developmental window or sensitivity period during the first year, evidenced by infants' readiness and preference to it's wide pitch variations, amplitude intensity and vibration, and distinctive and extensive contours. Effectiveness of protomusical acoustic features of IDSi in enhancing later development is based on the fact that sensory, motoric, cognitive and communicative capacities are still immature and infants demand caregivers' engagement in order to survive and learn (Schore, 2003). Thus, accurate timing and goodness of fit between infants' unique developmental needs and parental care and communication are decisive for optimal cognitive, social and emotional development (Fonagy et al., 2002). In the early 4–7 weeks infants are highly responsive to IDS or motherese involving heightened pitch, exaggerated pitch contours, and slow tempo (Papoušek et al., 1990), and in 5–10 months they are especially receptive and attracted to the protomusical acoustic features of IDSi (Falk & Audibert, 2021; Nakata & Trehub, 2004). The preference change reflects developmental needs for more complex, rhythmically varied, and multi-intensity amplitudes, tempos, and time-aligned repetitive vocal stimuli (Thiessen & Saffran, 2009). Mothers seem intuitively to modify and fine-tune their singing to match infants' preferences, needs, and age-specific developmental sensitivity (Ullsten et al., 2018). In optimal dyadic interaction mothers and infants structure their actions by enhancing communicative values and cueing redundancy at each protomusical pitch and beat (Lense et al., 2022). Infants' emerged developmental gains, such as ability to predict rhythms and tempo and to synchronize vocal emotional exchanges, open further developmental windows underlying cognitive skills of memory, differentiation, categorization, synchronization, and prediction, as well as socioemotional skills of affect recognition, expression and regulation, and playful mutual communication.

Unique to the first year developmental period is infants' ability to synthetize sensory and motor developmental gains, i.e., they are capable of receiving information derived from one sensory modality, for instance auditory, when hearing IDSi, and to translate it into other sensory modalities, for instance to motion,. This kind of a modality is based on audiovisual mirror neurons (Giudice, Manera, & Keysers, 2009), which also can be an essential mechanism of IDSi, characterized by wide pitch variability, high amplitude intensity, vibration, and distinctive F0 contours, for enhancing imitation, prelingual emotion expression, and shared social bonding. Lense et al. (2022) suggest that IDSi entrainment process observed in one physical systems, such as eye-looking matching with time-aligning to the singing rhythm, can scaffold other developmental systems, such as in bonding and increasing socioemotional capacity.

Concerning the neurobiological mechanism, the maternal IDSi has specific benefits in organizing newborn and infant neural networks and enhancing their brain development (Virtala et al., 2018). Brain neurological and physiological developments in turn are underlying infants' linguistic and other communicative advances (Hunnius & Meyer, 2020). Fetal brain prepares for vocal sensitivity as all neurons are formed by birth, and newborns are thus ready to social communication. Infants' head size, total brain volume, and thickness of the cerebral cortex increase rapidly, through dendric differentiation, axonal growth, and synaptogenesis (Parsons et al., 2010). In tandem, the number of synapses and connectivity between brain networks increases in all neocortical areas, apparently related to the growth-spurt in the formation of coticocortial circuitry (Kostovic ́et al., 2008). Protomusical IDSi is expected to increase connectivity in multiple brain areas, thus not only to activate auditory cortex through processing of incoming sound waves, rhythms and distinctive F0 contours. The favorable protomusical IDSi can also activate infants' visual cortex due to them watching maternal movements, sensory cortex due to their amodal and synthetizing processes and tactile feedbacks, and motor cortex through dyadically communicative moves of arms, legs, and tongues (Kostovic ́et al., 2008; Van Drunen et al., 2023). Intensive emotional sharing of IDSi is activating and sophisticating the amygdala-related brain circuits, and expanding and distinctive rhythms and F0 contours increase hippocampal volume to enhance memory skills. Formation of these neural circuities is a part of the process of producing and reorganazing of cortical synapses, which is underlying cognitive and behavioral development (Kostović et al., 2019). Importance of IDSi to language development has also been dedicated to findings that neural networks subserving language and music perception are partly overlapping (Koelsch & Siebel, 2005).

2. Evidence of impact of IDS on infant development

Compared to the IDSi, research is abundant on the role of the IDS in early child development. Studies have confirmed beneficial impact of prototypic IDS on pre-linguistic attentional and language development. The IDS characterized by high fundamental F0 frequency, wide F0 variability, and distinctive F0 contours are associated with infants' high attention, advanced pre-linguistic learning, and language acquisition (Cristia, 2013; Gratier & Devouche, 2011; Roberts et al., 2013). A nationally representative British study specified that high F0 variability of IDS predicted infants' high attention skills, indicated by abilities to focusing the gaze (watching) according to salient and alternating stimuli (Roberts et al., 2013). Furthermore, the IDS with synchronized pitch duration and rich F0 contours (shapes of falling, rising, U-shapes and bell-shapes) promoted infants' engagement in dyadic imitation and vocal repetition (Gratier & Devouche, 2011), and IDS with high F0 variability was associated with infants' larger productive vocabularies (Porritt et al., 2014).

Studies show favorable impact of IDS involving prototypical acoustic features (high fundamental F0 frequency and wide F0 variability) on preterm newborns' improved attention, muscle tone, and sucking intensity (Butler et al., 2014). Importantly, Niwano and Sugai (2002) found that mothers attune their F0 contour patterns according to their infant's prelinguistic development, and that the successful synchrony was decisive for enhancing infant vocal responsivity: falling contour shape was associated with high speed of infants' responses at early (3–5 months) and rising contour shape with high speed responses at later (7–9 months) development.

Concerning later language development, maternal IDS with wide F0 variability and rich rhythmic qualities predicted optimal word learning among 18 month old infants (Han, De Jong, & Kager, 2022). Further, maternal IDS characterized by flat and narrow F0 contours associated with toddlers' language delay (D'Odorico & Jacob, 2006), and low fundamental F0 frequency, narrow F0 variability, and flat F0 contours were associated with impaired linguistic learning among infants (Kaplan et al., 1999).

Intervention studies among mothers of preterm infants provide evidence about benefits of maternal singing beyond the non-singing. Yet, these studies do not tell about the role of acoustic features or vocal quality of the maternal IDSi. A systematic review confirmed that maternal IDSi is associated with favorable neonatal development among preterm newborns, indicated by increased physiological and behavioral stability (Filippa et al., 2017). A meta-analysis of the efficacy of music therapy, typically using IDSi to enhance infant development, found favorable effects on infant heart rate, respiratory rate, oral feeding volume, and stress level among preterm newborns (Yue et al., 2021). In a prospective study among full-term infants, Persico et al. (2017) showed that maternal lullaby singing in pregnancy predicts a lower levels of infant crying and colic behavior. Virtala and Partanen (2018) reported beneficial impact of maternal singing on auditory cortical processing (syllable discrimination) among preterm infants. For example, parental singing with its high resonance, repetition, and rich vocal variation, repertories, and rhythmicity can promote infants' early auditory neural processing, which explains the important role of IDSi in children's language development (Virtala & Partanen, 2018). Thus, maternal singing seems to provide a valuable way to improve emotional and physiological basis for infant development (Corbeil et al., 2016; Ullsten et al., 2017; 2018).

3. Aims of the study

The prospective study examined whether and how protomusical acoustic features of maternal singing to their six month old infants may associate with their sensorimotor, language, and socioemotional development. Based on research on IDS, we hypothesized that IDSi characterized by *high fundamental F0 frequency*, *high and wide F0 variability*, *rich and variable vocal amplitude and vibration, distinctive and extensive F0 contours of shapes and movements*, *high repertoires of vocal resonance of formants*, and *rich rhythmicity,* but *low vocal energy, tempo, and power* would be associated with infants' optimal sensorimotor, language, and socioemotional skills at six months and 18 months. Statistical functions of maternal voice analysis reflect the variation (e.g., means, minimum, maximum, ranges, and standard deviation, SD), shape, movement, and course (e.g., Interquartile Range (IQR), differences between wave pitches, differences between the 1st, 2nd and 3rd quartiles), and extremes and biases (kurtosis, skewness, and percentiles) of the acoustic features.

4. Method

4.1. Participants

The vocal material for this study was gathered as part of a Palestinian research project (the Gaza Infant Study). The participating mother-infant dyads lived in Gaza Strip that is a military occupied Palestinian territory under an international boycott and Israeli military siege since 2007 (UN-OCHA, 2022). The study involves three-time waves: the original sample (T1) consisted of 502 mothers, recruited at delivery in maternity units representing the four regions of the Gaza strip (North, Middle, South, and Gaza City), 392 mothers participated when the infants were six months (T2; $M = 6.2$, $SD = 0.4$), and 386 when the infants were 18 months (T3; $M = 18.0$, $SD = 1.4$). Between T1 and T2 data collection, 110 (21.9 %) mothers dropped out, because their home address had changed due to bombarded homes and subsequent family displacements ($n = 90$), death of the baby ($n = 13$), and withdrawal for personal reasons ($n = 7$). The drop-out rate was independent of infant gender, birthweight, or age and education of the mother and father. However, participation at T2 was related to a longer gestation weeks, $t\,(502) = 2.50$, $p < .01$ and better newborn health, $t\,(502) = 7.65$, $p < .01$. The drop-out rate between T2 and T3 periods was five mothers.

The vocal material was collected during the T2 home visits by recording the mothers singing to their infants a song of their own choice. Of the 392 mothers participating at T2, 236 (60.2 %) agreed to perform a song in addition to answering to a questionnaire interviews. The singing and non-singing (n = 156, 39.8 %) groups did not differ in maternal or paternal age, number of children, newborn birthweight, gestational weeks, or infant weight at six or 18 months (non-significant t-test values ranged between 0.086 and 1.068, $p =$ ns.), or in child gender, father employment, and place of residence (non-significant $\chi2$ test values ranged between 0.122 and 0.132, $p =$ ns.). However, in the singing group fewer mothers worked or studied outside home (7.3 %) than in the non-singing group (14.1 %), $\chi2\,(1, 501) = 5.81$, p = .016, Cesarean delivery was more common (17.9 %) than the non-singing (9.9 %) group, $\chi2\,(1501) = 6.12$, p = .013, and more (67 %) of newborns had excellent health, as compared to the non-singing group (38.6 %), $\chi2\,(4, 495) = 44.206$, p < .004.

Of the 236 recorded maternal infant-directed songs, 219 were included in the acoustic analysis. The loss of 17 songs was due to lack of singing (speaking), poor technical quality, and high noise during the recording.

4.2. Study procedure

In maternity clinics the nurses on duty presented the idea of the study and its prospective nature to mothers and asked voluntary mothers to join. Mothers willing to participate signed an informed consent. The inclusion criteria were voluntariness and being pregnant of first trimester during the 2014 War on Gaza, a 54-day long, intensive military operation by the Israeli army. It involved shelling and bombing from air, land, and sea, and curfew of the Gaza Strip, resulting in extensive human and material losses and displaced families (UN-Human Rights Council, 2014).

Ten Palestinian fieldworkers, with bachelor's degrees and former research experience, made home visits to interview the mothers from June to October of 2015 (T2) and from August to November of 2016 (T3). They attended comprehensive training on the research

Table 1
Acoustic Features. Definitions and Variables Extracted by OpenSMILE Algorithms of Voice Quality and Measured by Statistical Functions (Variables).

Acoustic features	Definitions	Variables[a])
Fundamental F0 frequency	Acoustic: Lowest periodic cycle component of the acoustic waveform, measured in Hz and reflecting reflects pitch F0[b]) Prosodic: The rate of vibrations of the vocal cords within the larynx, reflects pitch F0[b]) variations of the voice Width of pitch changes: The difference between Max and Min F0 peaks over the whole utterance	F0 Mean F0 Kurtosis F0 Maximum value (Max)[d]) F0 Minimum value (Min)[d]) F0 Range (change or D F0 = F0 Max – F0 Min F0 Skewness F0 Standard deviation. SD
F0 Variability	Jitter is the fast variations in F0 or Difference of Differences of Periods (DDP). The DDP is the average absolute difference between consecutive differences between consecutive periods (divided by the average period)	Jitter DDP Mean Jitter DDP IQR differences between 1st and 2nd waves[c) d)] Jitter DDP IQR differences between 1st and 3rd[d)] Jitter DDP IQR differences between 2nd and 3rd[d)] Jitter DDP Kurtosis Jitter DDP Linre2 Jitter DDP LinreA Jitter DDP LinreQ Jitter DDP Percentage99[d]) Jitter DDP Quartile 1 Jitter DDP Quartile 2[d]) Jitter DDP Quartile 3[d]) Jitter DDP Skewness Jitter DDP Standard deviation. SD
Vocal intensity and energy	Acoustic: The acoustic intensity is *perceived* as the energy and loudness of the sound. Prosodic: Vocal intensity refers to the glottal flow waveform (e.g. peak flow or speed) in terms of lung pressure and phonation threshold pressure. Energy is assessed through the root-mean-square RMS of statistical functions.	RMS Energy Mean RMS Energy Kurtosis[d]) RMS Energy Maximum value (Max) RMS Energy Minimum value (Min)[d]) RMS Energy Range (change or D F0 =F0 Max – F0 Min RMS Energy Skewness RMS Energy Standard deviation. SD
Amplitude and vibration	Acoustic: The amplitude of the vibrations (i.e. the size of the oscillations of the vocal folds) affecting loudness and energy of the voice. The greater the amplitude of the vibrations, the greater the amount of energy carried by the wave and the more intense is the sound. Loudness is indicated through Pulse Code Modulation (PCM) of discrete amplitudes Prosodic: Energy and loudness refer to the frequency of vibration of the vocal folds is perceived as the voice.	PCM Loudness Mean PCM Loudness IQR differences between 1st and 2nd PCM Loudness IQR differences between 1st and 3rd PCM Loudness IQR differences between 2nd and 3rd PCM Loudness Kurtosis PCM Loudness Linre2[d]) PCM Loudness LinreQ[d]) PCM Loudness Percentage1 % PCM Loudness Percentage99 % PCM Loudness Quartile 1 PCM Loudness Quartile 2 PCM Loudness Quartile 3 PCM Loudness Skewness PCM Loudness Standard deviation
F0 Contours: Shape and movement	Acoustic: The F0 contour refers to changes over time in the curse of an utterance or to overall shape of voice production in terms of its pitch (F0) variation over time. Movement in time is indicated by upper and lower envelopes (waves) Prosodic: The F0 contour is a realization of the vocal fold oscillation with slowly varying frequencies. Their dynamics are governed by a combination of the length and elasticity of vocal folds, laryngeal muscle tension, and subglottal air pressure. F0 Envelope: Shape and structures of the movements classified on the basis of their direction and slope. The envelope of the smoothed fundamental frequency contour refers to a rapidly varying signal is a smooth curve outlining its extremes in amplitude ([e.g. flat/unitonal indicating no changes in F0, rising (F0 increases) or falling (F0 decreases).	F0 Contour Lower Envelope F0 Contour Envelope Mean F0 Contour Envelope IQR differences between 1st and 2nd F0 Contour Envelope IQR differences between 1st and 3rd F0 Contour Envelope IQR differences between 2nd and 3rd F0 Contour Envelope Kurtosis F0 Contour Envelope Percentage1 % F0 Contour Envelope Percentage99 F0 Contour Envelope Quartile 1 F0 Contour Envelope Quartile 2 F0 Contour Envelope Quartile 3 F0 Contour Envelope Skewness F0 Contour Envelope Standard deviation. SD F0 Contour Upper Envelope

(*continued on next page*)

Table 1 (*continued*)

Acoustic features	Definitions	Variables[a])
		F0 Contour Final Mean F0 Contour Final IQR differences between 1st and 2nd F0 Contour Final IQR differences between 1st and 3rd F0 Contour Final IQR differences between 2nd and 3rd F0 Contour Final Kurtosis F0 Contour Final Percentage1 %[d]) F0 Contour Final Percentage99 F0 Contour Final Quartile 1 F0 Contour Final Quartile 2 F0 Contour Final Quartile 3 F0 Contour Final Skewness F0 Contour Final Standard Deviation. SD
Resonance and timbre via formant	Acoustic: A Formant is a concentration of acoustic energy around a particular resonating frequency in the singing wave. Each formant occurs at a different frequency. such as first (F1) formant 500 Hz, second (F2) formant 1500 Hz and third (F3) formant 2500 Hz. Each formant corresponds to a resonance in the human vocal tract (each bump in the frequency response curve). This study uses the F1 Formant. Prosodic: Formants can be detected as ridges or peaks in the spectrum of a song signal. Resonances are frequencies where sound waves are amplified by some point of the vocal tract.	The first Formant F1 at 500 Hz Formant F1 Mean Formant F1 Kurtosis[d]) Formant F1 Maximum value (Max)[d]) Formant F1 Minimum value (Min)[d]) Formant F1 percentage1 % Formant F1 % 99[d]) Formant F1 Quartile 1[d]) Formant F1 Quartile 2 Formant F1 Quartile 2 – Q - 1 Formant F1 Quartile 3[d]) Formant F1 Quartile 3[d]) Q – 1 Formant F1 Quartile 3 – Q - 2 Formant F1 Skewness Formant F1 Standard deviation. SD
Rhythmicity via bandwidth	Rhythmicity refers to vocal properties of duration, speed, intensity. word or vocal stressing, pausing, spectral balances, and regularity in tempo Bandwidth of a voice signal is the difference between the higher/ upper and lower frequency, measured in Hertz.	Bandwidth Mean Bandwidth Kurtosis Bandwidth Maximum value (Max)[d]) Bandwidth Minimum value (Min)[d]) Bandwidth percentage 1 % Bandwidth percentage 99[d]) Bandwidth Quartile 1 Bandwidth Quartile 2[d]) Bandwidth Quartile 2 – Q - 1 Bandwidth Quartile 3 Bandwidth Quartile 3 – Q - 2 Bandwidth Skewness Bandwidth Standard deviation. SD
Vocal tempo and power via attack-time	Acoustic: Attack is the initial impulse required to create oscillation for tempo in utterances. The beginning is defined as change in slope from baseline. Tempo refers to the rise in amplitude (loudness) over time for a given phoneme. To ensure reproducibility and power, the end is defined by maximum amplitude. Prosodic: Attack-time is analogous to how much an utterance is expressing 'punch' and stress Attack-time slope can also change direction.	Attack-time Mean[d]) Attack-time IQR differences between 1st and 2nd [d]) Attack-time IQR differences between 1st and 3rd [d]) Attack-time IQR differences between 2nd and 3rd [d]) Attack-time Kurtosis Attack-time Percentage1 %[d]) Attack-time Percentage99[d]) Attack-time Quartile 1[d]) Attack-time Quartile 2[d]) Attack-time Quartile 3[d]) Attack-time Range Attack-time Skewness Attack-time Standard Deviation. SD

Notes [a]) Variables extracted by OpenSMILE algorithms of voice quality and indicated by statistical functions. Here named variables
[b]) F0 = pitch, referring to the frequency at which vocal chords vibrate in voiced sounds
[c]) IQR Interquartile Range (IQR) differences between 1st and 2nd. 1st and 3rd. and 2nd and 3rd waves to indicate shape of the movement and structure
classified on the basis of the direction and slope of movements
[d]) Variables that are not showing multivariate normality. based on graphical information of the probability-probability (P-P) plots and histograms. They are omitted in further analyses.

tasks, ethics, interview procedures, and recording mothers singing to their six-month-olds. A local research team supervised the fieldwork through weekly meetings and guidance. All study procedures were conducted in the Arabic language. The average duration of the home visits was 90 min.

Mothers undersigned an informed consent that emphasized the participant rights, such as its voluntary nature and possibility to withdraw from the study any time without specifying a reason. The Palestinian Health Research Council and the Helsinki Committee for Ethical Approval reviewed the study and approved its research tools and procedures (PHRC/HC/126/14. The data that support the findings of this study are available from the corresponding author upon request.

4.3. Measures

Demographic and obstetric information (T1). Mid-wives collected the Standard European and US birth register information at birth, including birthweight, length of gestation, mode of delivery, newborn health, and the neonatal intensive care unit (NICU). Mothers also reported their education (no formal, basic, college, polytechnic or university education), family employment (mother: working at home or as professional; father: unemployed, worker, or professionally), family size, and parental age.

Acoustic features of infant-directed singing (T2). Fieldworkers presented the same instruction to mother to sing to their six-month-olds ("Many mothers sing to their babies. We would like to record infants' directed songs. Would you like to volunteer just for free singing to your baby. Please take your time. We are thankful for you letting us record and collect your singing"). They recorded the singing by Zoom H1 Handy Recorder device.

The acoustic analyses were performed by OpenSMILE-software for automatic extraction of features from audio signals and for classification of speech and music signals (SMILE: Speech and Music Interpretation by Large-space Extraction; Eyben et al., 2010; Eyben et al., 2013). The tape-recorded singing stimuli were digitalized at 44 kHZ with 16-bit resolution. The software provides statistical functions (e.g., means, skewness, standard deviation, and interquartile ranges, IQR) based on simple moving average (SMA) The SMA is the unweighted mean of the previous sequences of vocal data, and can efficiently ensure the variation and nature aligned with the fluidity of the acoustic features of maternal singing.

Table 1 presents the acoustic features, definitions, and variables (statistical functions, such as means, ranges, standard deviations, SD, and Interquartile Ranges, IQR). They are (1) *Fundamental F0 frequency*, indicated by mean, kurtosis, maxim, range, skewness, and SD; (2) *F0 Variability*, indicated by jitter (fast variation) values of Difference of Differences of Periods (DDP), assessed by statistical functions of mean, IQR differences between 1st and 2nd, 1st and 3rd, and 2nd and 3rd vocal waves, kurtosis, linear regressions (linre2, linreA, and linreQ), percentiles (1.0/99.0), 1st, 2nd and 3rd quartiles, skewness, and SD; (3) *Vocal intensity and energy,* assessed by the root-mean-square (RMS) values of mean, kurtosis, maxim, minimum, range, skewness, and SD; (4) *Vocal amplitude and vibration*, indicated by loudness of a signal and assessed through pulse code modulation (PCM) of discrete amplitudes by mean, IQR differences between 1st and 2nd, 1st and 3rd, and 2nd and 3rd vocal waves, kurtosis, percentiles (1.0/99.0), 1st, 2nd and 3rd quartiles, skewness, and SD; (5) *F0 Contour of the shape and movement*, indicated by the direction and slope of movements or pitch (F0) variations over time, and assessed by F0 waves (envelopes), separating lover envelopes and upper envelopes. The variables had corresponding names of F0 Contour LowEnvelope and F0 Contour UpperEnvelope, both assessed by mean, IQR differences between 1st and 2nd, 1st and 3rd, and 2nd and 3rd waves, kurtosis, percentiles (1.0/99.0), 1st, 2nd and 3rd quartiles, skewness, and SD; (6) *Formants of vocal resonance and timbre* are frequencies of amplified sound waves in the vocal tract, emerging as ridges or peaks in the spectrum of a voice signal. Formants are assessed by mean, kurtosis, maximum, minimum, percentiles (1.0/99.0), IQR differences between 1st, 2nd and 3rd quartiles, quartiles 1, 2, and 3, skewness, and SD; (7) *Rhythmicity through bandwidths,* referring to duration, stress, quality, and repetition of sounds, is indicted by bandwidth (a voice signal expressing the difference between the higher/upper and lower frequencies) and assessed by mean, kurtosis, maximum, minimum, percentiles (1.0/99.0), IQR differences between 1st and 2nd, 1st and 3rd, and 2nd and 3rd waves, 1st, 2nd and 3rd quartiles, skewness, and SD, and (8) *Vocal tempo and power,* indicated by attack-times referring the initial impulse required to create oscillation, tempo in vocal utterances, or rise in amplitude over time, assessed by mean, IQR differences between 1st and 2nd, 1st and 3rd, and 2nd and 3rd waves, kurtosis, maximum, minimum, percentiles (1.0/99.0), quartiles 1, 2, and 3, range, skewness, and SD of time.

Infant developmental skills at six months (T2). A 50-item questionnaire, based on the Minnesota Child Development Inventory (MCDI; Kopparthi et al., 1991), was applied to assess infants' sensorimotor, language-vocalization, and socioemotional development at six months. The sensorimotor and language-vocalization scales included items describing infant behavior and responses reflecting increasingly demanding skills. The *sensorimotor development* -scale combines 12 gross-motor and 15 fine-motor items. The gross-motor scale begins with descriptions, such as "Can roll from her/his stomach on to her/his back without help" and ends with "Can sit independently seeking support by hands." Fine-motor part begins with descriptions, such as "Baby keeps both hands tightly clenched" and ends with "Baby can beat/hit/bang two blocks or other toys several times against each other" Seven fine-motor performances were visualized by drawings in addition to the verbal descriptions. The *language-vocalization* -scale begins with descriptions, such as "Baby's vocalization begins to resemble utterances, e.g., ala, ila" and ends with "Baby produces utterances that sound as real words, e.g., expressions such as umm, abu (mother, father)." *Socioemotional* scale involves 13 items, for instance, "The baby makes eye contact" or "The Baby identifies and recognizes the parents, for example, smiles and is excited at the sight of parents." Mothers evaluated each item on a 3-point scale whether the infant had performed the task (1 = not observed; 2 = observed once or occasionally; 3 = observed many times, routinely).

The MCDI has been validated and corresponds with Bayley Sensory Motor Skills scale for 8- to 16-month old infants in North American (Kopparthi et al., 1991) and Egyptian (Baheiri, 2013) samples. The current scale has been found reliable and validated for the evaluation of early sensorimotor, communication and language and socioemotional development among Finnish (Lyytinen et al.,

2000) and Palestinian (Qouta et al., 2021) infants.

Infant developmental skills at 18 months (T3) were measured using three scales of the 50-item Arabic version of the Bayley Scales of Infant Development, BSID-III (Baheiri, 2013): *sensorimotor development* (11 items, e.g., "Standing without support", "Reaching items with a goal-directed attempt"); *cognitive-language development* (22 items, e.g., "Looking at the pictures in the book", "Building a

Table 2
Sociodemographic. Obstetric and Newborn Characteristics of Participants at Birth (T1).

	Participants[a]	
	%	n
Mother's age (years)		
16–24	41.6	91
25–29	30.6	67
30–34	17.4	38
35–39	8.2	18
40–44	2.3	5
Father's age (years)		
18–24	15.5	34
25–29	32.4	71
30–34	26.9	59
35–39	15.5	34
40–44	5.0	11
45–70	4.6	10
Number of children		
First child	30.6	67
1–3	50.2	110
4–6	15.5	34
7–10	3.6	8
Mother education		
No formal education	2.1	4
Basic education	25.0	55
Gymnasium/College	34.4	76
Polytechnic	32.1	70
University	6.4	14
Mother employment		
Works at home	92.7	203
Works in profession	5.9	13
Student	1.4	6
Father employment		
Unemployed	23.6	51
Worker	41.2	89
Blue collar: teacher. officer	32.4	70
High professional (doctor. engineer)	2.8	6
Type of residence		
Urban area	45.9	100
Village	21.6	47
Refugee camp	32.6	71
Child's sex		
Girl	53.4	117
Boy	46.6	102
Gestational age (weeks)		
<37	3.4	7
37–40	96.6	201
Birth weight (gr)		
<2500	3.4	7
2500–3499	55.2	112
3500–4499	39.9	81
>4500	1.5	3
Newborn health		
Excellent	67.4	145
Good	31.6	68
Health problems[b]	1.0	2
Birth defect		
ICD 10 diagnosis	4.6	10
Not defect	95.4	209
Type of delivery		
Normal vaginal	82.6	180
Caesarean	17.4	38

Notes:
[a] N = 236; Numbers differed due to missing data;
[b] Combines reasonable and severe problems (both n = 1); 7 dead newborns not included.

tower of two cubics or more", "Imitating words"), reflecting increasingly advanced skills; and *socioemotional development that* incorporates capacity to use a range of emotions and expressions in social engagements (11 items, e.g., "Being responsive to people", "Succeeds to be soothed when upset", "Trusting in others without fear"). A psychologist assessed the 18-month old infants individually for their sensorimotor and cognitive-language skills. Each task that a toddler performed was scored on a scale (yes—succeeded, performed; or no—did not perform) or counting the number of specified performances. The fieldworkers observed the baby and interviewed the mother about the occurrence of toddlers' socioemotional skills using a nine-point scale (1 =not at all; 9 =completely).

Translation. All measures were in Arabic, BSID-III was originally in Arabic and the fine-motor, gross-motor, language and socioemotional scales at T2 were translated and validated in earlier Arab studies (Baheiri, 2013; Qouta et al., 2021).

4.4. Data analysis

We used IBM SPSS statistics 27to examine demographic distributions and to factor analyze variables of acoustic features and infant developmental skills. The OpenSmile acoustic feature extractor provides a large number of variables or statistical functions. (See Table 1) In order to use Structural Equation Modelling (SEM) with efficient power, we had to reduce the number of observed and latent variables to be acceptable for modelling the available data (N = 219). We used a free-available calculator of SEM, Soper (2022), for the sufficiency of the sample size of 219 of mothers with high quality IDSi material. The calculation is based on statistical power of Cohen's $d = 0.80$, anticipated effect size 0.30, and $\alpha = 0.05$ level of significance (Wolf et al., 2013). The calculation with ten latent variables (4 acoustic features and 2 × 3 infant developmental skills), consisting of 42 observed variables (27 for acoustic features and 15 for infant developmental skills) resulted in a recommended minimum sample size of 190. The calculated sample size was slightly smaller than the rule of thumb suggestion of SEM sample size of five cases per variable (Nunally, 1978), amounting to N = 210 in our SEM Model with 42 observed variables.

We proceed with three stages to reduce the number of acoustic features variables (statistical functions) and to detect their conceptually unified dimensions. First, we omitted variables that did not show multivariate normality, based on graphical information of the probability-probability (P-P) plots and histograms with normality curves. Second, we conducted an exploratory factor analysis with varimax rotation on the remaining acoustic feature variables. The criteria for omitting variables were small communality value (<.50) of a variable explained by the factors, low and unstable factor loadings (<.40) of a variable, and variable loadings on two different factors (Hair et al., 2006). Third, the measurement models in SEM would further detect non-significant standardized regressions to be omitted.

In order to construct latent variables for SEM of infant development at six and 18 months, we conducted confirmatory factor analyses with varimax rotation separately on corresponding sensorimotor, language, and sosioemotional items. Expecting 2–3 factors was based on recommendations of the number of observed variables contributing to latent variables in SEM (Hayduk & Littvay, 2012). Measurement models of latent variables of infant developmental skills further test the effectiveness of the emerged factors.

To test the research hypotheses about the impact of maternal IDSi on infant development, we applied Structural Equation Models (SEM; AMOS 26.0, Arbuckle, 2011). It included four latent exogenous acoustic features variables, of which the observed variables are based on factor analyses with direct effects with outcome variables of infant developmental skills. These skills were indicated by three latent variables at six months (sensorimotor, language-vocalization, and socioemotional) and three latent variables at 18 months (sensorimotor, language-cognitive, and socioemotional), of which the observed variables are based on factor analyses.

The choice of the SEM with latent acoustic features and developmental variables was due to it (a) providing tools to analyze measurement models of highly correlating observed variables (e.g., statistical functions of means, minimum, maximum, range, and SD typically correlate), thus avoiding the risk of multicollinearity with large measurement errors, (b) simultaneous assessing of multiple relationships between acoustic features of maternal IDSi and specific infant developmental skills at two salient developmental periods, and (c) conceptualizing the acoustic features as latent variables, which is novel in voice analyses and helps aggregating observed acoustic feature variables in a model representing core characteristics of maternal IDSi. Latent variables are superior, for instance, to sum variables, because the contributing observed variables have own unique weights and the measurement error is controlled to guarantee full reliability (Fan et al., 2016).

The criteria for model fitness were comparative fit index (CFI >.90; a measure of relative model fit), Tucker-Lewis index (TLI >.90; a measure of non-normed fit index), the root mean square error of approximation (RMSEA <.06; indicating parsimony of the model), and Cmin/DF (values 1–3 measured of absolute model fit) (Brosseau-Liard & Savalei, 2014).

5. Results

5.1. Descriptive statistics

Table 2 presents the demographic and obstetric factors of mothers and their infants. The average age of the mothers was 26 years ($SD = 6.0$). A quarter had basic education, a third high school or college and also a third polytechnic education. Only 2 % had no formal education. Majority of mothers currently worked at home (93 %) and the rest worked as blue collar staff, such as teachers or nurses or were students. Of fathers 41 % were workers and a third blue collar employees, and yet about a quarter were unemployed. Almost a half (46 %) of families lived in urban areas and a third (33 %) in refugee camps.

About a third of mothers had their first child, and a half had 1–3 children previously. There was slightly more girls (53 %) than boys in the sample. Regarding obstetric data, the average duration of gestation was 39.23 weeks ($SD = 1.77$) and birthweight 3348.38 g ($SD = 527.96$). The prevalence of preterm delivery (< 37 weeks) was 3 %, low birth weight (< 2.500 g) 3 %, and birth defects 5 %. A majority (83 %) had vaginal normative delivery.

5.2. Preliminary analyses

Preparation for the SEM analysis involved data reduction and constructing observed composite variables. First, Table 1 presents the original names (statistical functions) of acoustic variables, and we have marked by [d]) those variables that did not show multivariate normality variance and were omitted from further analyses. It is noteworthy that only in F0 contours: Vocal shapes and movements (lower envelope or waves) all variables satisfied the multivariate normality criteria, while in most acoustic features 1–7 single variables were non-normal.

Second, Supplement Table I presents the results of factor analysis, conducted in order to detect conceptually salient dimensions of acoustic features variables and further reduce the number of single variables (statistical functions). The explorative factor analysis resulted in five factors that explained 66.51 % of their variation. Before naming the dimensions, a number of variables were omitted due to low communality, low factor loading, or double loading (in two factors). Results show that all variables featuring Fundamental F0 frequency and Vocal intensity and energy had to be omitted according to the above mentioned criteria, and consequently the factor V was not valid. The contents of detected factors or acoustic feature dimensions are: *I F0 Variability*, covering four variables of jitterDDP and one F0 contour variable (upper envelope skewness); *II Vocal amplitude and vibration*, covering six PCM loudness variables; *III Resonance and rhythm*: Formants and bandwidth, covering three formant variables and five bandwidth variables, and *IV Vocal shapes and movement*: F0 Contours, covering four F0 contour lower envelope and four upper envelope variables.

We conducted confirmatory factoring to detect dimensional structure of sensorimotor, language, and sosioemotional developmental skills at six and 18 months. The analyses prepared for constructing latent variables the observed variables of which are based on these extracted factors. Results showed three-factor and two-factor solutions, presented in Supplement, Table II a for six months and Table 2 b for 18 months.

Table 3

Loading Estimates of Observed Variables of Acoustic Features of Maternal Infant-directed Singing Statistical Functions for Corresponding Latent Constructs of the Structural Equation Model (SEM) [a].

Observed variables for latent constructs	Unstd β	SE	Std β	t-tests
F0 Variablity				
jitterDDP LinregerrQ	.016	.001	.911	14.389****
jitterDDP Linregc2	.052	.004	.854	13.588****
jitterDDP Skewness	-1.643	.129	-.812	-12.693****
jitterDDP_Standard deviation	.053	.004	.917	14.658****
F0finEnv_Skewness[a])	1.000		-.671	
Amplitude and vibration				
PCM loudness IQR 1–2	.259	.022	.880	11.618****
PCM loudness IQR 2–3	.266	.025	.970	10.613****
PCM loudness Percentile 1.0	.097	.015	.500	6.643****
PCM loudness Percentile 99.0	.977	.118	.689	8.300****
PCM loudness Quartile 1	.349	.040	.713	8.732****
PCM loudness Skewness[a])	1.000		-.596	
Resonance and rhythm: Formants and bandwidths		.		
Formant_0 Quartile_2	6.907	.578	.891	11.923****
Formant_0 Skewness	-.003	.000	-.543	-6.491****
Bandwidth_Kurtosis	.101	.012	.791	8.242****
Bandwidth_0 Quartile 2	6.706	.735	.997	9.128****
Bandwidth_0 Quartile 2–1	2.935	.306	.979	9.579****
Bandwidth_0 Quartile 3	6.322	.693	.997	9.126****
Bandwidth_0 Skewness	-.012	.001	-.873	-8.272****
Formant_0 Mmean[a])	1.000		.521	
Shapes and movements: F0 Contours				
F0finEnv IQR 1–2[c])	.520	.133	.440	3.002***
F0finEnv IQR 2–3	1.036	.315	.782	3.993****
F0finEnv Kurtosis	-.104	.029	-.632	-3.618****
F0final IQR 1–2[c])	.810	.226	.715	3.583****
F0final IQR 2–3	.927	.310	.507	3.000***
F0final Kurtosis	-.146	.041	-.595	-3.564****
F0final Percentile 99.0	.919	.084	.342	10.939****
F0finEnv Percentile 99.0[a])	1.000		.304	

Notes: ****$p < .0001$, ***$p < .001$. [a] N = 219. [a]) The value of 1 of the unstandardized regression estimate (Unstd b) refers to parameters fixed to the measurement models.

[b]) The model fitness indices are interpreted concerning the total model including latent variable measurement models and predicting coefficients, reported in Table 4

[c]) F0finEnv -variable name refers to lower envelopes (sound waves), and F0final-variable name to upper envelopes.

5.3. Measurement models for acoustic features and infant development

Table 3 presents the measurement models of observed variables of the exogenous latent variables of acoustic features based and named according to the results of the factor analysis. All observed variables loaded statistically significantly on the constructed latent variables of F0 Variability, Vocal amplitude and vibration, Resonance and rhythmicity: Formants and bandwidth, and F0 Contours: Vocal shapes and movements. All had estimation loadings $\beta > .50$, except two variables having $\beta > .30$. The measurement models are part of SEM, and fit indices concerning the whole final model with latent variables are presented in Table 4.

Supplement Table III presents the measurement models of observed variables of the outcome variables of infant developmental skills. Results show that all observed variables loaded statistically significantly on the constructed latent variables based on confirmatory factors analyses of sensory, language and socioemotional variables at six and 18 months. All variables had estimation loadings $\beta > .50$, except four having $\beta > .30$. The measurement models are part of SEM, and fit indices concerning the whole final model with latent variables are presented in Table 4.

5.4. Acoustic features of maternal IDSi associating with infant development

Fig. 1 illustrates the SEM results of acoustic features of maternal IDSi associating with infant developmental skills at six and 18 months, and Table 4 summarizes the non-standardized (Unstd b) and standardized (Std β) estimates, standard errors (SE), t-tests, explained variance (percentages based on R^2) and fit indices. The fit indices showed good fit of the SEM model with data, but the CFI did not reach 0.95, which makes the model open to criticism.

Results reveal, as hypothesized, that maternal IDSi acoustic features of high and rich vocal amplitude and vibration were significantly associated with six month olds infants' optimal sensorimotor (β = ⊡38, p = .006), language and vocalization (β = ⊡25, p = .004) and socioemotional (β = ⊡22, p = .001) development. Further, high and wide F0 variability was significantly associated with optimal language and vocalization (β = ⊡26, p = .006) and socioemotional (β = ⊡20, p = .005) infant development, and marginally with sensorimotor (β = ⊡19, p = .060) development six months. Also, high repertoires of vocal resonance and rich rhythmicity (Formants) were associated significantly with six months olds' optimal language and vocalization development (β = ⊡17; p = .050). Against our hypothesis distinctive and extensive shapes and movements (F0 contours) did not associate with optimal development of six-month-old infants. Also, high repertoires of vocal resonance and rich rhythmicity (Formants) did not associate with optimal sensorimotor or socioemotional development.

When infants were 18 month olds, only maternal IDSi acoustic features of high and rich amplitude and vibration significantly predicted optimal sensorimotor (β = ⊡18, p = .025), language and cognitive (β = ⊡29, p = .002), and socioemotional (β = ⊡14, p = .050) development, as hypothesized. Further, extensive and distinct shapes and movements (F0 contours) predicted marginally sensorimotor (β = ⊡21, p = .063) and language and cognitive (β = ⊡20, p = .099) infant development at 18 months. Against our hypotheses, high and wide F0 variability and high repertoires of vocal resonance and rich rhythmicity (Formants) did not predict any

Table 4
Structural Equation Model (SEM) on Predictors (Latent Constructs [a]) of Acoustic Features of Maternal Infant-directed Singing) of Infant Development (Latent Constructs [a]) of Sensorimotor, Language, and Socio-Emotional) at six and 18 months: Parameter Estimates (Coefficients) Associations and Model Fit Indices [a]).

	Infant development at 6 months											
	Sensorimotor				Language and vocalization				Sosioemotional			
	Unstd β	SE	Std β	t-tests	Unstd β	SE	Std β	t-tests	Unstd β	SE	Std β	t-tests
F0 Variablity	.209	.116	.189	1.804+	.484	.175	.257	2.774**	.437	.155	.200	2.822**
Amplitude and vibration	.846	.307	.382	2.757**	.946	.332	.251	2.853**	.963	.297	.220	3.245***
Resonance and rhythmicity: Formants and bandwidth	.015	.021	.059	0.684	.072	.037	.170	1.958*	.014	.032	.027	0.421
Shapes and movements: F0 Contours	.007	.005	.171	1.726	.012	.008	.174	1.530	-.008	.007	.109	1.285
Explained variance (%; R^2)	18 %				16 %				08 %			
	Infant development at 18 months											
	Sensorimotor				Language and cognition				Sosioemotional			
	Unstd β	SE	Std β	t-tests	Unstd β	SE	Std β	t-tests	Unstd β	SE	Std β	t-tests
F0 Variablity	.034	.076	.062	0.449	.061	.093	.038	0.652	.176	.345	.038	0.509
Amplitude and vibration	.326	.146	.180	2.235*	.577	.186	.291	3.104**	1.295	.663	.141	1.958*
Resonance and rhythmicity: Formants and bandwidth	.017	.016	.085	1.074	.021	.020	.092	1.037	.037	.073	.035	0.502
Shapes and movements: F0 Contours	.007	.004	.211	1.858+	.007	.004	.197	1.624+	.008	.014	.050	0.584
Explained variance (%; R^2)	10 %				17 %				03 %			

Model fit indices
χ^2 = 1698,020. df = 749. p = .0001; Cmin/df = 2267, CFI = .91; TLI = .90; RMSEA = .,076 [90 % CI:.071–.081]
Notes: N = 219; +p < .10. *p < .05. ** p < .01. *** p < .001 [a]) The fix indices are from total SEM models with the measurement models of latent constructs, reported for acoustic variables in Table 3 and for infant development variables in Supplement Table II a and Table II b.

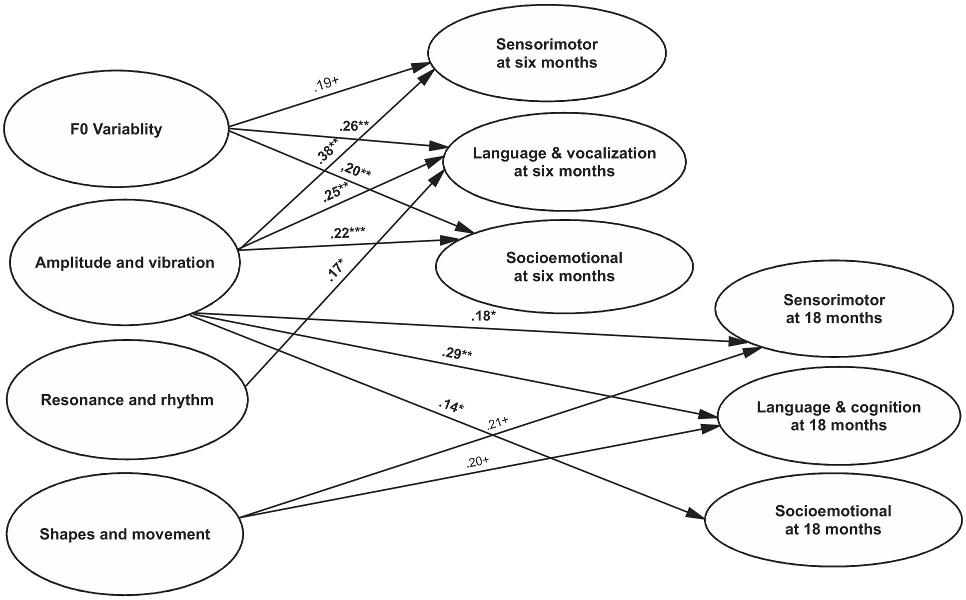

Fig. 1. Model I of Acoustic Features of Maternal Infant-Directed Singing and Infant Development at six and 18 Months.

developmental skills at 18 months.

6. Discussion

We wanted to learn about acoustic features and qualities of mothers singing to their infants and were curious about their role in early and later infant development. Research has documented beneficial protomusical features or 'acoustic signatures' of mother-infant vocal interaction involving certain exaggerated and vivacious qualities, reflected in uniquely high pitch fundamental F0 *frequency*, high, wide, and smooth F0 v*ariability*, and varied shapes and movements of F0 c*ontours* (Falk & Audibert, 2021; Falk & Kello, 2017; Spinelli et al., 2017; Ullsten et al., 2018). Our analysis covered a more comprehensive set of acoustic features of maternal IDSi, including also *vocal amplitude and vibration*, *vocal resonance* and *rhythmicity* and *vocal tempo and timbre*. Our results partly accord with earlier studies on IDS, as in line with the hypotheses, maternal IDSi characterized by wide *F0 variablity* and high and rich vocal *amplitude and vibration* was associated with infants' optimal sensorimotor, language, and sosioemotional development, and also high and wide vocal *resonance and rhythmicity* with language skills at six months. High and rich vocal amplitudes and vibration predicted significantly sensorimotor, language and cognitive, and socioemotional developmental skills when infants were 18 months. Against to the hypotheses, distinguished and extensive shapes and movements of maternal ISDi did not play a role in infants' developmental skills at six months, and high F0 variability and wide and vivid resonance and rhythmicity at six months. The results suggest that proto-musical features of maternal IDSi can have a comprehensively positive influence on infants' development. Yet, as the indices of parsimony of the SEM models was moderate, the findings must be interpreted by caution and should be replicated in more multi-cultural mother-infant interactions.

6.1. Specific developmental roles of variations, vibrations, and shape diversity of IDSi

According to our results, only high and rich amplitude and vibration of IDSi, indicating intensive, pulsative, vivacious, and lively singing associated with both concurrent and predicted later optimal sensorimotor, language, and socioemotional infant development. These acoustic features thus had a comprehensive developmental impact, whereas wide F0 variability, indicating rich, varied, and multitudinous tone and pitch changes was important only for six-month-olds. The IDSi, reflecting high repertoires of resonance and rhythmicity was associated only with better language and vocalizing skills at early infancy, but predicted more comprehensive, although marginal, impact infants' later developmental skills. Thus the more complex features of extensive and distinguished shapes and movements in maternal singing were favorable to 18-month-old infants.

Researchers understand certain IDSi patterns of acoustic features as evolutionally functional and individually adaptive (Russell, Bachorowski, & Fernández-Dols, 2003). An example of individual adaptivity is infants' preference of prototypic maternal IDS with high, diverse, wide and rich variations, vibrations, pulses, and extensive shapes and movements, as compared to adult-directed speech (Dunst et al., 2012; Trainor, 1996). Infants invite mothers to synchronize their singing with their developmentally salient communication responses, which in turn strengthens both favorable qualities of vocal acoustic expressions and mutually rewarding interactional processes. An example of the beneficial role of more complex, extensive, and distinguished shapes and movements is the F0 contours in the IDSi predicting, although marginally, infants' later optimal sensorimotor and languge-cognitive development at 18 months. These preferences and benefical impacts reflect also intensified neural network and synaptic connectivity in brain development (Kostović et al., 2008). The phenomenally intensive learning during the first year is based on experience- and action-dependent development (Schore, 2003), where neurobiological, social, emotional, and cognitive achievements are based on infants' experiences, including dyadic communication in the IDSi (Alviar et al., 2023).

Our results of the salience of protomusical IDSi acoustic features in language development concur with earlier evidence of rich, wide and exaggerated IDSi contributing to optimal pre-linguistic and language achievements (Franco et al., 2022; Schön et al., 2008; Tierney et al., 2017). From birth infants are sensitive and active in attending to vocal acoustic messages, and at three-months they learn specifical social functions of acoustic features, show selective responses, and prefer maternal voice in both IDS and IDSi (Cooper et al., 1997). Around five months infants communicate with canonical babbling, proceeding at six months to recognize and understand few frequent words, such as mother (Tincoff & Jusczyk, 2012), use word information in cognitive processes, such as categorization (Vouloumanos & Waxman, 2014), and finally at 18 months are communicating words with gestures. Infants' sensitivity to prosodic, vocal, and accoustic stimuli serve especially cognitive developemnt of learning and memory during the first months, thus enhancing language development (Franco et al., 2022; Newman, & Berstein, 2016; Thiessen & Saffran, 2009). In tandem, language learning corresponds with intensifying with neural connections and thickening brain areas (Broca's and Wernicke's areas), responsible for processing both musical elements of harmony and melody and language elements of grammar and vocabulary (Koelsch & Siebel, 2005; Trollinger, 2010).

The benefits of the IDSi acoustic features on infant development were, to some extent, skills and age specific. Distinguished and extensive shapes and movements of F0 contours in the IDSi predicted, although marginally, optimal sensorimotor and language-cognitive development when infants were 18-months olds. We may carefully speculate that complex and wealthy structure of IDSi shapes and movements do not serve pre-lingual infants, who communicate with babbling and train to shift towards recognizing single intimate words. The extensive, distinctive, and intensively over-time changes in singing provides infants with experiences of discriminating and remembering rhythms, sense of time-aligned actions and expecting, and coordinating responses, which seem beneficial later in development, indicating goodness of fit with emergence of more complex and sophisticated cognitive capacities (Lense et al., 2022; Thiessen & Saffran, 2009). Older infants benefited from complex, multi-shaped and extensive IDSi, as they had already acquired increased verbal communication skills, through word recognition into expressing multiple words and sentences at 18 months (Kuhl, 2004).

6.2. Comprehensive view of infant development

Our study is maybe the first to analyze the role of IDSi in enhancing infant development conceptualized as a multidomain phenomenon, whereas earlier studies have focused on separate outcomes of prelingual responsiveness, infant language skills (Schön et al., 2008), motor (Van Drunen et al., 2023), cognitive such as memory (Thiessen & Saffran, 2009) or socioemotional (Lense et al., 2022) development. Some research is also available on acoustic IDSi features of infant and mother interaction, including attachment (Falk & Audibert, 2021; Milligan et al., 2003). Our results suggest rather comprehensive influences of IDSi on salient developmental domains in infancy. It would be desirable to analyze how quality of early IDSi would predict toddler or school age development.

Our results show beneficial impacts of maternal singing characterized with high and rich amplitudes and vibrates in promoting socioemotional skills at both six and 18 months, and high and rich variability singing facilitating socioemotional skills at six months. They concur with theoretical views of the bonding function of IDSi (Lense et al., 2022). The IDSi has multiple functions in underlying mother-infant emotion sharing, emotion recognition, expressing, and regulation, as well as distress attunement (Ullsten et al., 2018). Interventions with newborns at risk show that music, especially singing, is successful in modulating infants affect, pain and distress (Trehub, Ghazban, & Corbeil, 2015; Ullsten et al., 2017; Virtala & Partanen, 2018). Affects in turn serve the basic neurological foundation of emotions and feelings, together with multimodal senses of touch, sight, smell, and taste, in addition to hearing (Hart, 2011). Infancy is the core period of creating a dyadic attachment bond that influences the nature and quality of later essential social relationships (Schore, 2003), and therefore the high-quality IDSi can play a role in socioemotional development.

Research of maternal vocal communication is largely focusing on infant language development, because musical elements promote learning of cognitive skills of categorizing, causality, initiation, expectation, attention, and memory (Papadimitriou et al., 2021; Schön et al., 2008). Our study setting allowed to analyze relative importance of acoustic features of IDSi in associating and predicting language skills as compared to sensorimotor and socioemotional skills. At six months, the results do not confirm the argument about the predominant role of ISDi in language development, based on the share of the IDSi acoustic features explaining the variation of three developmental domains. The IDSi acoustic features equally explained sensorimotor (18 %) and language and vocalization skills (16 %) among six month old infants. However, we may cautiously conclude that the IDSi acoustic features were somewhat more important for cognitive-language skills (explained 17 %) as compared to sensorimotor skills (10 %) at 18 months. In both developmental periods the acoustic features of IDSi explained significantly, but relatively less importantly infants' socioemotional skills, 8 % at six and only 3 % at 18 months.

6.3. Methodological questions concerning acoustic features of maternal singing

All IDSi acoustic features, such as fundamental F0 frequency, F0 variability, amplitude vibration, intensity and energy, resonance and rhythmicity formants and bandwidths, and vocal shapes F0 contour, and attack-time tempo correlated significantly with each other in our study (correlation Tables available from authors). In order to detect the most salient acoustic features and their underlying structure or dimensionality, we performed exploratory factor analyses and proceeded with confirmatory SEM measurement models. In addition to conceptualization, the modelling was necessary to limit the great number of single acoustic variables (statistical functions) due to case-variable -ratio. The emergence of acoustic feature structure of the IDSi showed some noteworthy differences from earlier IDS studies. Strange enough, the fundamental F0 frequency dimension did not work in our modelling, as its observed variables (e.g., F0 mean, F0 kurtosis, F0 range or F0 standard deviation) did not significantly load on the corresponding latent variable. The statistical misfit was unexpected as the fundamental F0 frequency is considered one of the core acoustic features of IDS, and it is commonly empirically studied in addition to F0 variability and F0 contours (Spinelli et al., 2017).

Also, the acoustic feature of vocal tempo and power did not work in our modelling as attack-time observed variables, indicating vocal tempo oscillation or phonemic pulse changes, did not load on corresponding latent variable. It is a pity, because our aim was to extend the scope of indicators of IDSi acoustic features from the three commonly assesses pitch frequency, variability, and contours. Thus loss of total salient acoustic phenomenon reflecting vocal tempo and power for statistic reason is narrowing the information that we provided.

Our study is the first to apply SEM and latent acoustic constructs to study associations between features of maternal IDSi and infant development. The measurement models of four IDSi latent constructs, reflecting various aspects and characteristics were relatively pure, as indicated by statistical indices. This line of statistics may encourage future research in this important area of vocal, prosodic, and acoustic communication in early months of life.

Our study evokes questions about universal vs. culturally-bound nature of ISDi acoustic features. The participants, Palestinian mothers and infants, represent Arab Middle-Eastern Islamic culture, and we could not detect other similar studies in Middle-Eastern countries. Cross cultural comparisons of IDS show similarities across various language groups, representing e.g., Asian (Japanese and Chinese), European (German, French, Italian and British), and North American (English) cultures. They confirm cross-cultural or language-related consistencies in the patterns of prosodic and acoustic modifications in parents speaking to infants, characterized by high fundamental F0 frequency, wide F0-variability, distinguished F0 contours, shorter utterances, and long pauses (Fernald, 1992; Grieser & Kuhl, 1988; Sulpizio et al., 2018). Hilton et al., (2022) detected similar acoustic differences between infant- and adult directed vocalizations (both speech and music) across cultures in a large comparison of about 1600 recordings from 21 societies. These comparison studies did not include Middle Eastern language groups, such as Arabic, Farsi, Kurdish or Hebrew, but universalist nature of motherese is commonly evidenced. There are no cultural comparisons available concerning acoustic features IDSi (although open-access corpus of the speech and song by Hilton et al., 2022 would allow it). It would be important to replicate our results concerning the role of IDSi in multidomain infant development in other cultural contexts in order to understand it's culturally-bond and universal nature.

6.4. *Limits of the study*

The problems of the study relate to assessment of infant development, large number of single acoustic variables, and concentrating only on acoustic features in infant development. We used a standard objective assessment (Baley II) of infants' sensorimotor, language and cognitive, and socioemotional development at 18 months, but mothers reported these developmental skills of their six-moth-olds, which of open to criticism. Parent-reported child characteristics can be biased due to maternal mood and wellbeing, social expectations and desirability, although mothers are real experts in observing early development and intimate dyadic interaction. Yet, the measurement models of infants' sensorimotor, language, and socio-emotional skills were statistically sufficient and pure at both assessment times, indicating thus validity of mother-reported method.

The OpenSmile method provides tools comprehensively to document acoustic features of IDSi. Yet the plentitude caused problems for statistical analysis that forced us to reduce the large number of variables of statistical functions (ranging from 7 for fundamental F0 frequencies to 13 for F0 contours). We could not find earlier research or guidelines how to optimally combine extensive vocal pitch information described by statistical functions (e.g., variability and extensions with ranges, skewness or standard deviation and changes in rhythm and tempo by differences in quartiles or IQRs). We started by omitting statistically deficient acoustic variables (See Table 1 for violation of multivariate normativity), then proceed with a factor analysis to explore possible underlying structure or dimensions of all acoustic variables, omitting non-fit variables, and finally tested measurement models of fitting observed variables on latent constructs of main dynamic acoustic features. The economizing the data can be criticized and alternative ways of statistics of acoustic analyses are welcome.

Finally, our analyses focused solely on the role of IDSi acoustic features in infant development. Critics can legitimately note that many other mother, infant, family relations and society -related factors also influence both the acoustic features of ISDi and the child development. For instance, research shows maternal mental health to influence IDS, as depressive mothers speak with flat and emotionless voice (Kaplan et al., 1999) and infants' attachment behavior evokes different communication patters (Falk & Audibert, 2021; Milligan et al., 2003). Ample evidence shows that maternal mental health problems are associated and predict infant developmental problems, and thus our research setting would have benefited from including mental health measures. Critics may also note that we have not studied the role factors such as family economics or parenting goals and philosophies that may influence both maternal vocal communication and infant development. Our study could thus not contribute to research analyzing both factors explaining IDSi acoustic features and factors related to family processes that may mediate or moderate the impacts of IDSi on infant development.

Author note

We are grateful to the Palestinian families for their participation and to the Academy of Finland (275197) and Jacobs Foundation

(5585) for the financing of the study. We like to thank our excellent fieldworkers who visited families in very difficult post-war conditions. Without their commitment this study could not be realized.

CRediT authorship contribution statement

Raija-Leena Punamäki: Conceptualization, Data curation, Formal análisis, Funding acquisition, Visualization, Writing - original draft, Writing - review & editing. **Safwat Y. Diab:** Conceptualization, Data curation, Project administration, Writing - original draft, Writing - review & editing. **Konstantinos Drosos:** Data curation, Formal análisis, Methodology, Resources; Software. **Samir R. Qouta:** Conceptualization, Project administration, Writing - original draft, Writing - review & editing. **Mervi Vänskä:** Investigation, Methodology, Supervision; Validation, Writing - original draft, Writing - review & editing.

Data availability

The authors do not have permission to share data.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.infbeh.2023.101908.

References

Alviar, C., Sahoo, M., Edwards, L. A., Jones, W., Klin, A., & Lense, M. (2023). Infant-directed song potentiates infants' selective attention to adults' mouths over the first year of life. *Developmental Science*. , Article e13359. https://doi.org/10.1111/desc.13359

Arbuckle, J.L. (2011). IBM SPSS Amos 20 user's guide. Amos development corporation, SPSS Inc, 226–229.

Baheiri, A. R. (2013). *Standardization of Bayley Scales of Infant Development* (Fourth edition). Cairo: El Nahda & El Massreia.

Brosseau-Liard, P. E., & Savalei, V. (2014). Adjusting incremental fit indices for nonnormality. *Multivariate Behavioral Research, 49*, 460–470. https://doi.org/10.1080/00273171.2014.933697

Butler, S. C., O'Sullivan, L. P., Shah, B. L., & Berthier, N. E. (2014). Preference for infant-directed speech in preterm infants. *Infant Behavior and Development, 37*(4), 505–511. https://doi.org/10.1016/j.infbeh.2014.06.007

Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior & Development, 20*(4), 477–488. https://doi.org/10.1016/S0163-6383(97)90037-0

Corbeil, M., Trehub, S., & Peretz, I. (2013). Speech vs. singing: Infants choose happier sounds, 372–372 *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00372.

Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing delays the onset of infant distress. *Infancy, 21*(3), 373–391. https://doi.org/10.1111/infa.12114

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass, 7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

D'Odorico, L., & Jacob, V. (2006). Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language & Communication Disorders, 41*(3), 293–311. https://doi.org/10.1080/13682820500342976

Dunst, C., Gorman, E., & Hamby, D. (2012). Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning, 5*(1), 1–13.

Eyben, F., Weninger, F., Gross, F., & Schuller, B. (2013). Recent developments in opensmile, the munich open-source multimedia feature extractor. *Proceedings of the 21st ACM International Conference on Multimedia*, 835–838.

Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: the munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM International Conference on Multimedia*, 1459–1462.

Falk, S., & Audibert, N. (2021). Acoustic signatures of communicative dimensions in codified mother-infant interactions. *The Journal of the Acoustical Society of America, 150*(6), 4429–4437. https://doi.org/10.1121/10.0008977

Falk, S., & Kello, C. T. (2017). Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition, 163*, 80–86. https://doi.org/10.1016/j.cognition.2017.02.017

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: An updated review. *Ecological Processes, 5*(1), Article 19. https://doi.org/10.1186/s13717-016-0063-3

Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry, 48*(3–4), 329–354. https://doi.org/10.1111/j.1469-7610.2006.01701.x

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262–279). Cambridge University Press.

Filippa, M., Panza, C., Ferrari, F., Frassoldati, R., Kuhn, P., Balduzzi, S., & D'Amico, R. (2017). Systematic review of maternal voice interventions demonstrates increased stability in preterm infants. *Acta Paediatrica, 106*(8), 1220–1229. https://doi.org/10.1111/apa.13832

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalisation, and the development of the self*. New York, NY: Other Press.

Franco, S. C., Spinelli, M., Kozar, I., & Fasolo, M. (2022). Singing to infants matters: Early singing interactions affect musical preferences and facilitate vocabulary building. *Journal of Child Language, 49*(3), 552–577. https://doi.org/10.1017/S0305000921000167

Giudice, M. D., Manera, V., & Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Developmental Science, 12*(2), 350–363. https://doi.org/10.1111/j.1467-7687.2008.00783.x

Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Developmental Psychology, 47*(1), 67–76. https://doi.org/10.1037/a0020722

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*(1), 14–20. https://doi.org/10.1037/0012-1649.24.1.14

Hair, J., Black, B., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis* (6th edition). Upper Saddle River, NJ: Prentice-Hall.

Han, De Jong, N. H., & Kager, R. (2022). Prosodic input and children's word learning in infant- and adult-directed speech. *Infant Behavior & Development, 68*, Article 101728. https://doi.org/10.1016/j.infbeh.2022.101728

Hart, S. (2011). *The impact of attachment: Developmental neuroaffective psychology*. WW Norton & Company.

Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology, 12*, 159–164. https://doi.org/10.1186/1471-2288-12-159

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., & Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour, 6*(11), 1545–1556. https://doi.org/10.1038/s41562-022-01410-x

Hunnius, S., & Meyer, M. (2020). *New Perspectives on Early Social-Cognitive Development*. Elsevier.

Jardri, R., Houfflin-Debarge, V., Delion, P., Pruvo, J., Thomas, P., & Pins, D. (2012). Assessing fetal response to maternal speech using a noninvasive functional brain imaging technique. *International Journal of Developmental Neuroscience, 30*(2), 159–161. https://doi.org/10.1016/j.ijdevneu.2011.11.002

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129* (5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kaplan, P. S., Bachorowski, J. A., & Zarlengo-Stauss, P. (1999). Child-directed speech produced by mothers with symptoms of depression fails to promote associative learning in 4-month-old infants. *Child Development, 70*, 560–570.

Kisilevsky, B. S., Hains, S. M., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development, 32*(1), 59–71. https://doi.org/10.1016/j.infbeh.2008.10.002

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences, 9*(12), 578–584. https://doi.org/10.1016/j.tics.2005.10.001

Kopparthi, R., McDermott, C., Sheftel, D. N., Lenke, M. C., Getz, M., & Frey, M. (1991). The Minnesota Child Development Inventory: validity and reliability for assessing development in infancy. *Journal of Developmental and Behavioral Pediatrics, 2*(4), 217–222. https://doi.org/10.1097/00004703-199108000-00001

Kostović, I., Judaš, M., & Petanjek, Z. (2008). Structural development of the human prefrontal cortex. In C. A. Nelson, & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience. Cambridge* (pp. 213–235). MIT Press, A Bradford Book.

Kostović, I., Sedmak, G., & Judaš, M. (2019). Neural histology and neurogenesis of the human fetal and infant brain. *NeuroImage, 188*, 743–773. https://doi.org/10.1016/j.neuroimage.2018.12.043

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831–843. https://doi.org/10.1038/nrn1533

Lense, M. D., Shultz, S., Astésano, C., & Jones, W. (2022). Music of infant-directed singing entrains infants' social visual behavior. *Proceedings of the National Academy of Sciences, 119*(45), Article e2116967119. https://doi.org/10.1073/pnas.2116967119

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience, 10*(1), 37–47. https://doi.org/10.1038/nrn2554

Lyytinen, P., Ahonen, T., Eklund, K., & Lyytinen, H. (2000). Assessment of vocal and motoric development in infancy. *Jyväskylä: Niilo Mäki Instituutti (In Finnish)*.

Mäntymaa, M., Puura, K., Luoma, I., Latva, R., Salmelin, R., & Tamminen, T. (2015). Shared pleasure in early mother-infant interaction: Predicting lower levels of emotional and behavioral problems in the child and protecting against the influence of parental psychopathology. *Infant Mental Health Journal, 36*(2), 223–237. https://doi.org/10.1002/imhj.21505

Milligan, K., Atkinson, L., Trehub, S. E., Benoit, D., & Poulton, L. (2003). Maternal attachment and the communication of emotion through song. *Infant Behavior and Development, 26*(1), 1–13. https://doi.org/10.1016/S0163-6383(02)00165-0

Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development, 27*(4), 455–464. https://doi.org/10.1016/j.infbeh.2004.03.002

Newman, R. M. L., & Berstein, R. N. (2016). Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *Journal of Child Language, 43*(5), 1158–1173. https://doi.org/10.1017/S0305000915000446

Niwano, K., & Sugai, K. (2002). Intonation contour of Japanese maternal infant-directed speech and infant vocal response. *The Japanese Journal of Special Education, 39* (6), 59–68. https://doi.org/10.6033/tokkyou.39.59_2

Nunally, J. C. (1978). *Psychometric theory*. New York: McGraw Hill.

Papadimitriou, A., Smyth, C., Politimou, N., Franco, F., & Stewart, L. (2021). The impact of the home musical environment on infants' language development. *Infant Behavior and Development, 65*, Article 101651. https://doi.org/10.1016/j.infbeh.2021.101651

Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development, 13*(4), 539–545. https://doi.org/10.1016/0163-6383(90)90022-Z

Parsons, C. E., Young, K. S., Murray, L., Stein, A., & Kringelbach, M. L. (2010). The functional neuroanatomy of the evolving parent–infant relationship. *Progress in Neurobiology, 91*(3), 220–241. https://doi.org/10.1016/j.pneurobio.2010.03.001

Persico, G., Antolini, L., Vergani, P., Costantini, W., Nardi, M. T., & Bellotti, L. (2017). Maternal singing of lullabies during pregnancy and after birth: Effects on mother–infant bonding and on newborns' behaviour. Concurrent Cohort Study. *Women and Birth, 30*(4), 214–220. https://doi.org/10.1016/j.wombi.2017.01.007

Porritt, L. L., Zinser, M. C., Bachorowski, J. A., & Kaplan, P. S. (2014). Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. *Language Learning and Development, 10*(1), 51–67. https://doi.org/10.1080/15475441.2013.802962

Qouta, S. R., Vänskä, M., Diab, S. Y., & Punamäki, R.-L. (2021). War trauma and infant motor, cognitive, and socioemotional development: Maternal mental health and dyadic interaction as explanatory processes. *Infant Behavior and Development, 63*, Article 101532. https://doi.org/10.1016/j.infbeh.2021.101532

Roberts, S., Fyfield, R., Baibazarova, E., van Goozen, S., Culling, J. F., & Hay, D. F. (2013). Parental speech at 6 months predicts joint attention at 12 months. *Infancy, 18*, E1–E15. https://doi.org/10.1111/infa.12018

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*(1), 329–349. https://doi.org/10.1146/annurev.psych.54.101601.145102

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., & Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and cognition?(A systematic review). *PloS One, 8*(10), Article e78103.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106*(2), 975–983. https://doi.org/10.1016/j.cognition.2007.03.005

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*(9), 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Soper, D.S. (2022). A-priori Sample Size Calculator for Structural Equation Models [Software]. Available from 〈https://www.danielsoper.com/statcalc〉.

Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review, 44*, 1–18. https://doi.org/10.1016/j.dr.2016.12.001

Sulpizio, K., Dalsasso, K., Asakawa, M., Bornstein, T., Doi, M. H., Esposito, G, H., & Shinohara, K. (2018). Discriminating between mothers' infant- and adult-directed speech: Cross-linguistic generalizability from Japanese to Italian and German. *Neuroscience Research, 133*, 21–27. https://doi.org/10.1016/j.neures.2017.10.008

Thiessen, E. D., & Saffran, J. R. (2009). How the melody facilitates the message and vice versa in infant learning and memory. *Annals of the New York Academy of Sciences, 1169*(1), 225–233. https://doi.org/10.1111/j.1749-6632.2009.04547

Tierney, A., White-Schwoch, T., MacLean, J., & Kraus, N. (2017). Individual differences in rhythm skills: Links with neural consistency and linguistic ability. *Journal of Cognitive Neuroscience, 29*(5), 855–868. https://doi.org/10.1162/jocn_a_01092

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy, 17*(4), 432–444. https://doi.org/10.1111/j.1532-7078.2011.00084.x

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development, 19*(1), 83–92. https://doi.org/10.1016/S0163-6383(96)90046-6

Trehub, S. E. (2001). Musical Predispositions in Infancy. *Annals of the New York Academy of Sciences, 930*(1), 1–16. https://doi.org/10.1111/j.1749-6632.2001.tb05721.x

Trehub, S. E., Ghazban, N., & Corbeil, M. (2015). Musical affect regulation in infancy. *Annals of the New York Academy of Sciences, 1337*(1), 186–192. https://doi.org/10.1111/nyas.12622

Trehub, S. E., & Trainor, L. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research, 12*, 43–78.

Trollinger, V. L. (2010). The brain in singing and language. *General Music Today, 23*(2), 20–23. https://doi.org/10.1177/1048371309353878

Ullsten, A., Eriksson, M., Klässbo, M., & Volgsten, U. (2018). Singing, sharing, soothing – biopsychosocial rationales for parental infant-directed singing in neonatal pain management: A theoretical approach. *Music & Science, 1*, 1–13. https://doi.org/10.1177/2059204318780841

Ullsten, A., Hugoson, P., Forsberg, M., Forzelius, L., Klässbo, M., Olsson, E., Volgsten, U., Westrup, B., Ådén, U., Bergqvist, L., & Eriksson, M. (2017). Efficacy of live lullaby singing during procedural pain in preterm and term neonates. *Music and Medicine, 9*(2), 73–78. https://doi.org/10.47513/mmd.v9i2.546

UN-Human Rights Council. (2014). The United Nations Independent Commission of Inquiry on the 2014 Gaza conflict. Retrieved from: http://www.ohchr.org/EN/HRBodies/HRC/CoIGaza/Conflict/Pages/ReportCoIGaza.aspx.2014.

Van Drunen, L., Schultz, B. G., Becht, A. I., Schaefer, R. S., & Wierenga, L. M. (2023). *How music alters brain plasticity: A longitudinal twin study on sensorimotor synchronization and brain developmental patterns*. Available SSRN. https://doi.org/10.2139/ssrn.4415030.

Virtala, P., & Partanen, E. (2018). Can very early music interventions promote at-risk infants' development? Very early music interventions in at-risk infants. *Annals of the New York Academy of Sciences, 1423*(1), 92–101. https://doi.org/10.1111/nyas.13646

Virtala, P., Partanen, E., Tervaniemi, M., & Kujala, T. (2018). Neural discrimination of speech sound changes in a variable context occurs irrespective of attention and explicit awareness. *Biological Psychology, 132*, 217–227. https://doi.org/10.1016/j.biopsycho.2018.01.002

Vouloumanos, A., & Waxman, S. R. (2014). Listen up! Speech is for thinking during infancy. *Trends in Cognitive Sciences, 18*(12), 642–646. https://doi.org/10.1016/j.tics.2014.10.001

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913–934. https://doi.org/10.1177/0013164413495237

Yue, W., Han, X., Luo, J., Zeng, Z., & Yang, M. (2021). Effect of music therapy on preterm infants in neonatal intensive care unit: Systematic review and meta-analysis of randomized controlled trials. *Journal of Advanced Nursing, 77*(2), 635–652. https://doi.org/10.1111/jan.14630