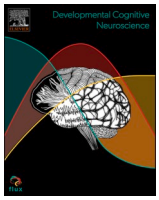
Developmental Cognitive Neuroscience 81 (2026) 101782

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

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



Do babies perceive cries as speech?

Gaia Lucarini [a,b], Irene de la Cruz-Pavía [c,d], Jessica Gemignani [a,b], Caroline Nallet [a,b,e], Alexandre Lapillonne [f], Judit Gervain [a,b,g,*]

[a] Department of Developmental and Social Psychology, University of Padua, Padua 35131, Italy
[b] Padova Neuroscience Center, University of Padua, Padua 35131, Italy
[c] Faculty of Social and Human Sciences, University of Deusto, Bilbao 48007, Spain
[d] Basque Foundation for Science Ikerbasque, Bilbao 48009, Spain
[e] Independent speech and language therapist, Besançon 25000, France
[f] Department of Neonatology, Hôpital Necker - Enfants Malades, Université Paris Cité, Paris 75015, France
[g] Integrative Neuroscience and Cognition Center, Université Paris Cité & CNRS, Paris 75006, France

ARTICLE INFO

Keywords:
Newborns
Cry perception
Speech perception
Adults
fNIRS

ABSTRACT

Speech perception and production are linked in adults. Since infants show sophisticated speech perception abilities before they produce language, the link has been assumed to emerge with development. However, crying is a communication signal that infants produce from birth and their cry melodies are modulated by prenatally heard speech. Cries and speech may thus be developmentally continuous, and the production-perception link may be most evident early in life for cries. We therefore examined whether newborns perceive cries similarly to speech. French neonates exposed to French cries and French speech showed greater neural responses to cries than to speech in the right temporal region. By contrast, Italian adults unfamiliar with French responded more strongly to speech than to cries. Newborns' heightened response to a communicative signal that is less familiar and acoustically less complex, but one they can produce suggests that vocal production and auditory perception may be linked from birth.

1. Introduction

Infants learn language with amazing ease and efficiency. Around their first birthday, they utter their first words and by the time they turn 3 years, they are proficient conversational partners. But for the first few months of life, their main communication tool is crying. Infants cry immediately upon being born, and often get their needs met through crying even before they can smile or babble. Cries, as speech, are thus interpreted by caregivers as communicating information.

With development, distinct crying patterns emerge (Wermke et al., 2002) alongside other vocalizations such as cooing (Kent, 2021), transitioning from largely reflexive productions to intentionally communicative ones (Fogel and Thelen, 1987). These changes are scaffolded by increasing exchange opportunities with caregivers as infants grow (Yoo et al., 2018). Not only do adults promptly respond to cries to meet the needs of their babies, but newborns themselves, upon hearing other newborns cry, respond by crying and showing facial expressions of distress (Dondi et al., 1999; Martin and Clark, 1982; Sagi and Hoffman, 1976). This phenomenon is known as "emotional contagion", and while it is debated whether this stems from true empathy or a simpler reaction to stimuli with a negative valence, e.g. increased arousal (Ruffman et al., 2017), it is undeniable that newborns treat cries as highly salient and relevant auditory stimuli.

Interestingly, the melodies of cries may already be modulated by the speech newborns hear prenatally in the womb (Mampe et al., 2009; Manfredi et al., 2019), even if currently available evidence is not conclusive (Van Niekerk et al., 2025). If the contours of newborns' cries already reflect the intonational contours of the prenatally heard language, then speech and cries may share some underlying mechanisms and may be developmentally continuous such that sound patterns and production mechanisms may be shared between the two.

Speech has been argued to be special because it is our species-specific communicative signal the human vocal tract has evolved to produce (Liberman, 1984; Liberman and Mattingly, 1985; Vouloumanos et al., 2010; Werker, 2018). Behavioral and neuroimaging evidence show that infants are born with a bias to attend to and process speech more readily than other auditory signals including low-pass filtered, whistled, backward, or sinewave speech, and with a distinct neural response to speech

* Correspondence to: DPSS, Università degli Studi di Padova, via Venezia 8, Padova (PD) 35131, Italy
E-mail address: judit.gervain@unipd.it (J. Gervain).

https://doi.org/10.1016/j.dcn.2026.101782
Received 29 October 2025; Received in revised form 3 July 2026; Accepted 8 July 2026
Available online 9 July 2026

(May et al., 2018; Spence and DeCasper, 1987; Vouloumanos and Werker, 2007). Importantly, this preference is initially broad, extending to other primate vocalizations, specifically monkey calls, before narrowing down to human speech by 3 months of age (Vouloumanos et al., 2010; Vouloumanos and Werker, 2007). Given that cries are vocal communicative signals, and critically, they are the only ones that newborns themselves can produce, cries might initially fall within this broadly tuned "speech" category. We put to test the exciting hypothesis that cries, as a precursor to speech, might be perceived and processed by newborns in a manner similar to speech.

Even though the causal links are debated (Liberman et al., 1967; Liberman and Mattingly, 1985; Studdert-Kennedy et al., 1970), considerable empirical evidence suggests that speech perception and speech production may be linked in adults, since during production, the speaker experiences the position of their articulators as well as the sounds they generate, which can, therefore, all be combined in a single representation (Guenther, 2016; Ito et al., 2009; Yeung and Scott, 2021). According to the Motor Theory of Speech Perception (Liberman and Mattingly, 1985), speech sounds are represented in the brain as motor commands. Although a strictly motor-based account of speech perception is not supported empirically (Galantucci et al., 2006; Lotto et al., 2009), neuroimaging evidence suggests that the neural systems supporting speech perception and speech production are indeed tightly coupled in adults (Cheung et al., 2016; Liebenthal and Möttönen, 2018; Poeppel and Assaneo, 2020; Skipper et al., 2017).

In early development young infants do not yet produce speech, but their speech perception abilities are very sophisticated. This developmental lag has led to the often implicit assumption that the perception-production link emerges later in development, and thus cannot be a driving force behind early language acquisition. Consequently, the perception-production link has traditionally received little attention in language acquisition research. Recently, however, studies have started addressing this question experimentally in older infants (3 months being the youngest age tested to date; Bruderer et al., 2015; Choi et al., 2021, 2023; Frota et al., 2026; Masapollo et al., 2016; Polka et al., 2014, 2025; Szmytke, 2026; Vilain et al., 2019; Werker, 2018). These studies find that infants' speech perception is deeply intertwined with their emerging sensorimotor and vocal abilities, and that this link strengthens with age and with infants' growing speech production abilities (DePaolis et al., 2011, 2013; Frota et al., 2026; Lorenzini and Nazzi, 2022; Majorano et al., 2014). Even before they can produce linguistic sounds, infants' tongue and mouth movements appear to shape how they perceive and process phonemes (Bruderer et al., 2015; Vilain et al., 2019). As they gain vocal production skills, they start to show preferences for speech with infant-like acoustic properties (Polka et al., 2014, 2025). Finding evidence that newborn infants process speech and cries similarly, namely that cries recruit the cortical regions typically engaged in speech processing, would constitute the strongest and earliest evidence that a production-perception link may exist from the beginning of human development.

To test this, we have presented prenatally French-exposed newborn infants with sentences in French as well as with cries produced by other, prenatally French-exposed newborns, and measured their neural responses using near-infrared spectroscopy (NIRS) in a standard block design (Fig. 1A). NIRS targeted the language areas, covering the frontal, parietal and temporal cortices (Fig. 1B-C) (Benavides-Varela and

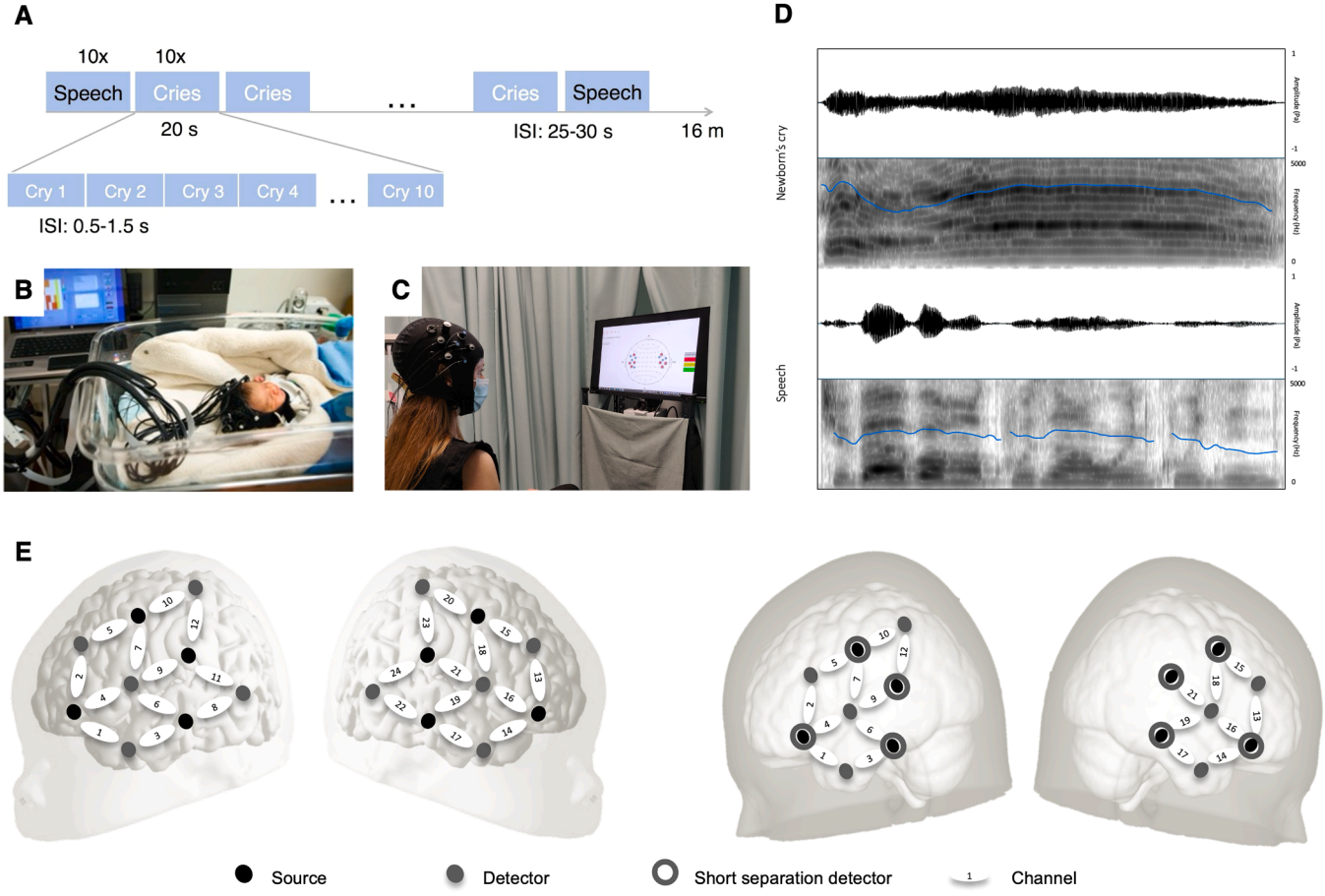


Fig. 1. The experimental paradigm. (A) The experimental design. (B) A newborn infant participating in the study. (C) An adult participating in the study. (D) Waveform and spectrogram of a French newborn's cry and a French sentence. The blue line represents the pitch contour. (E) The newborn (left) and adult (right) channel layout used in the study.

Gervain, 2017; Gervain et al., 2008; Martinez-Alvarez et al., 2023; May et al., 2018). We seek to establish whether processing prenatally heard speech and newborn cries recruits overlapping or distinct neural mechanisms, and whether motor areas are involved in their processing.

Speech and cries differ in several theoretically relevant dimensions. Both are familiar to newborns, but differ in the degree and nature of this familiarity. Exposure to speech starts between the 20th−24th week of gestation, while experience with cries begins only at birth. Perceptually, newborns have thus substantially more familiarity with speech than cries. Newborns' experience with these signals is also qualitatively different: exclusively perceptual for speech, while their experience with crying combines sensorimotor and auditory perceptual feedback. In turn, speech is acoustically more complex than cries, but newborns themselves can only produce cries. Thus we can tease apart the differential roles of these factors, as familiarity (Abboub et al., 2016; May et al., 2011, 2018; Sato et al., 2012) and complexity (Bouchon et al., 2015; Vouloumanos and Werker, 2007) predict stronger responses to speech in both newborns and adults, while production experience predicts stronger responses to cries in newborns, but stronger response to speech in adults (Liberman and Mattingly, 1985). If cries and speech recruit overlapping cortical regions, this would suggest that the neural mechanisms supporting speech perception extend to other human vocalizations. Stronger neural activation in response to speech would be consistent with a newborn bias for speech over other acoustic signals (Vouloumanos and Werker, 2007). However, if cries, despite their lesser familiarity and complexity, trigger stronger neural responses than speech, it would constitute strong evidence that production experience outweighs potential biases for speech despite its greater familiarity and complexity, strongly supporting the existence of a production-perception link from birth. Finally, similar levels of activation in both vocal signals in overlapping language areas would also suggest that infants' production abilities modulate auditory perception from birth. The localization of the responses is also relevant. We expect activation in temporal regions in response to speech (Gervain et al., 2008; Peña et al., 2003; Skeide and Friederici, 2016), and potentially also to cries, as temporal regions engage in auditory processing. Importantly, the most compelling evidence for a production-perception link would be the selective recruitment of the frontal regions covering the motor cortex. If production experience shapes humans' auditory processing from birth, we predict that newborns will show frontal activation only in response to cries (i.e. the signal they can produce), while adults will exhibit activation in the same region in response to both speech and cries.

As a comparison, we also tested adults' responses to the same stimuli and experimental setup. Adults can produce speech, and while they can also produce cries, adult cries are acoustically different from newborn cries, and adults only very rarely use cries to communicate, especially as the sole means of communication. We tested Italian adults who do not speak or understand French so that they process the stimuli at the acoustic/phonological level just like newborns do, lacking the knowledge to process syntax and semantics. The phonetic overlap between Italian and French ensured, however, that the speech stimuli were comparable at the phonetic level, keeping adults' and newborns' processing demands equivalent. In adults, who are potential caregivers and thus the primary addressees of newborn cries, these may elicit a stronger response than speech due to their heightened emotional valence. However, we expect greater responses to speech than to cries, as speech is more familiar to adults than cries, adults more frequently produce and use speech for communication than cries, and speech has greater acoustic complexity. Specifically, we expect activation to speech in an unfamiliar language to be greater than to cries in the bilateral temporal areas. Since the language used is unfamiliar to our adult participants, we expect no or reduced left lateralization and little involvement of the left inferior frontal regions such as Broca's area, since the speech stimuli are only processed acoustically and phonologically, but not morpho-syntactically or lexically (Friederici, 2011, 2012; Poeppel, 2014).

2. Methods

2.1. Participants

Twenty-five healthy full-term newborn infants, who had been exposed to French prenatally, were included in the final analysis (15 females; mean age: 2.05 days; age range: 1–4 days, mean gestational age: 39 weeks 6 days, gestational age range: 38 weeks – 41 weeks 5 days, mean weight: 3320 g, weight range: 2860–4100 g, Apgar score five minutes after birth: minimum 8). No preterm infants were included. Data collection was carried out between January 2020 and April 2023. Forty-five additional infants were tested but excluded from analysis due to poor data quality/an insufficient number of valid trials (< 40%) (11), fussiness, crying (22), less than 80% exposure to French during the pre- and postnatal period (5) and technical error (7). We note that the higher than usual rejection rate is due to emotional contagion, i.e. newborns' tendency to start crying and produce facial expressions of distress when they hear other newborns cry (Dondi et al., 1999; Martin and Clark, 1982; Sagi and Hoffman, 1976), causing strong movement artifacts in the NIRS data. Nevertheless, the final achieved sample size has proven to yield robust and replicable results in a recent meta-analysis of a large body of infant NIRS studies on speech perception (Gemignani et al., 2023).

All parents gave written informed consent before the study, which was approved by the CPP Sud Est V ethics committee (N° ID RCB: 2018-A01136–49, 20/12/2018). The protocol followed all institutional, French national and international regulations, such as the Declaration of Helsinki, and respected European data protection and privacy laws (GDPR). In addition, 27 Italian-speaking non-parent adults with no knowledge of French took part in the study and entered the final analysis. Another 11 adult participants were tested but excluded from analysis due to poor data quality (< 50%) (11). Rejection due to poor data quality was performed in batch prior to statistical analysis, following the same criteria for all adults and infants (see Data Analysis). Participants gave written informed consent before the study, which was approved by the Ethical Committee of the University of Padova (N° 4403, 26/10/2021). The protocol followed all institutional, Italian national and international regulations, such as the Declaration of Helsinki, and respected European data protection and privacy laws (GDPR).

2.2. Stimuli and design

Stimuli consisted of cries recorded from 10 French newborns (none of which participated in the study) and spoken French sentences produced by 10 adult female French natives. To test newborns' cry perception in its broadest definition, we deliberately included a wide variety of cries rather than restricting them to a single category (e.g. tiredness) (34). Likewise, the spoken sentences varied in both structure and length. Each newborn was paired with one adult talker. Each cry was then matched in length with a sentence (Fig. 1D; mean cry length: 1095 ms, $SD = 0.49$, mean sentence length: 1094 ms, $SD = 0.49$, comparisons between the newborn and adult members of a pair used independent samples *t*-tests or Man-Whitney tests depending on the normality of distribution of the samples: all newborn-adult pairs, $p \geq .635$). Intensity was normalized across all stimuli. Cries were carefully selected so that they did not contain pauses, ensuring a continuous acoustic signal comparable to that of the sentences. Cries had significantly higher mean pitch than speech, as is to be expected due to the fact that infants have higher f0 than adult females (mean pitch of cries: 430 Hz, $SD = 65.5$, and of sentences: 233 Hz, $SD = 31.1$; pair comparisons used independent-sample *t*-tests or Man-Whitney tests depending on the normality of distribution of the samples: all pairs $p \leq .001$).

The experiment, identical for newborn and adult participants, consisted of a total of 20 blocks (Fig. 1A), 10 per condition. Every block of cries comprised 10 cries, each produced by a different newborn. Similarly, every block of speech comprised 10 sentences, each produced by a

different French speaker. No sentence or cry was repeated during the study. The order of presentation of the blocks was intermixed and counterbalanced, such that no more than two blocks of the same condition occurred consecutively, and block order varied across participants. Stimulus presentation within blocks was also randomized across participants. Blocks were separated by silent intervals jittered between 25 and 30 s. Within blocks, cries/sentences were also separated by silences jittered between 500 and 1500 ms, yielding blocks of approximately 20 s (range 19.43–20.41 s). The whole experiment lasted about 16 min.

2.3. Procedure

Newborns were tested in a quiet room of the maternity ward of the Hôpital Necker-Enfants Malades (Paris, France), in the presence of at least one of their parents. During the study, infants lied in their cribs asleep or in quiet rest, while their brain responses were measured with a NIRx NIRScout 16–16 machine (source-detector separation: 3 cm; pulsated LED lights at two wavelengths of 760 nm and 850 nm; sampling rate: 15.625 Hz). The optical probes were inserted into a stretchy cap (EasyCap, Brain-Products GmbH, Germany) placed on the infants' head using surface landmarks (nasion, and the preauricular points), covering the language areas in the bilateral temporal, frontal, and parietal cortices (12 channels/hemisphere, Fig. 1B,E).

As determined by the localization analysis conducted in Abboub et al. (2016) using age-appropriate structural magnetic resonance imaging and stereotaxic atlases, channels 1, 2, 4, 5 on the LH and 13, 14, 15, 16 on the RH were on average positioned over the frontal area, channels 3, 6, 8, 11 on the LH and 17, 19, 22, 24 on the RH over the temporal area, channel 9 on the LH and 21 on the RH were positioned in the frontier between the temporal and parietal areas, and channels 7, 10, 12 on the LH and 18, 20, 23 on the RH were on average positioned over the parietal area. Stimuli were presented using E-Prime, and delivered through two speakers placed at an angle of 30° and approximately 1 m from the newborn's head, and elevated to the height of the crib. The computer running E-Prime sent time stamps to the NIRS machine.

Adult participants were tested in a laboratory at the University of Padua, Italy. Their brain responses were measured with a NIRx NIRSport 2 8–16 machine (same characteristics as the NIRScout 16–16 used to record the newborns' brain activity, sampling rate: 10.173 Hz). The optical probes (8 sources, 7 detectors, and 8 short-distance detectors) were inserted into a stretchy cap (EasyCap, Brain-Products GmbH, Germany) placed on the adults' head using surface landmarks (nasion, and the preauricular points), covering the language areas in the bilateral temporal, frontal, and parietal cortices (10 channels on the LH, 8 channels on the RH, Fig. 1C,E). To be able to correct for the contamination of the hemodynamic response by surface blood flow, short-separation detectors were placed around each source, at a fixed distance of 8 mm as is standard in adult NIRS setups (Brigadoi and Cooper, 2015). The NIRx setup requires one detector to be used exclusively for the short-separation channels, reducing the number of regular, long-separation channels (channels 20 and 23 in the RH are not present in adults). Due to newborns' thin tissues, short-separation channels are not recommended in infants (Emberson et al., 2016). The infant and adult montages are depicted in Fig. 1.

2.4. Data processing and statistical analyses

Changes in the concentration of oxygenated hemoglobin (oxyHb) and deoxygenated hemoglobin (deoxyHb) were calculated from the absorption of near-infrared light as metabolic indicators of neural activity following a validated pipeline (Gemignani and Gervain, 2021). Data were band-pass filtered between 0.01 and 0.7 Hz. Movement artifacts, defined as concentration changes larger than 0.1 mmol x mm over 0.2 s, were removed by rejection block-channel pairs where artifacts occurred. For the non-rejected blocks, a baseline was linearly fitted between the means of the 5 s preceding the onset of the block and the 5 s starting 40 s after the onset of the block (25 s of stimulation plus 15 s of resting period), and then subtracted from the time series of the block. This baseline detrending step is performed prior to block averaging and serves to remove slow drifts and inter-trial offsets unrelated to task-evoked hemodynamic responses. Similar linear detrending approaches are commonly used in fNIRS preprocessing pipelines (Yücel et al., 2025).

In addition, for adult data, short-separation channel correction was carried out over optical densities, following Saager and Berger (2005), before the above preprocessing routine was run. The data quality of short-separation channels was first assessed by computing their coefficient of variation (CV) subject by subject for each wavelength. The CV is defined as the standard deviation of the timeseries divided by its average. Then, short-separation channels displaying a CV smaller than 10% were selected and averaged within each hemisphere. This CV threshold lies between commonly used cutoffs: 7.5% as a strict criterion (e.g. Hocke et al., 2018; Zimeo Morais et al., 2017) and 15% as a more lenient one (e.g. Piper et al., 2014), and corresponds to a signal-to-noise ratio of 10, which is consistent with conservative quality-control values reported in widely employed preprocessing routines (Nguyen et al., 2021). Afterwards, the resulting hemisphere-wise short-separation time series was employed to remove the physiological confound from the respective long-separation channels using the method proposed by Saager and Berger (2005): data from the short-separation channel is first fit to the long-separation channel by least-squares regression, then the result of the fit is removed from the long-separation channel by subtraction. This method has been shown to be effective in removing physiological confounds due to extracerebral tissues from adult NIRS data, thus reducing the detection of false positives and false negatives (Tachtsidis and Scholkmann, 2016).

Newborns were included in the analysis if at least 40% of their data was found valid after pre-processing (on average 59.38% of data was retained, 57.90% for speech, 60.87% for cries). Adults were included in the analysis if at least 50% of their data was found valid after pre-processing (on average 90.55% of data was retained, 91.38% for speech, 89.71% for cries).

After pre-processing, statistical analyses were carried out over the concentration changes of oxyHb and deoxyHb. We conducted a cluster-based permutation analysis (Maris and Oostenveld, 2007) to identify clusters of channels showing a significant difference in concentration changes between the two conditions or between each condition and the zero baseline. Specifically, in a first step paired-sample *t*-tests were conducted at each time point and channel. Critical t-values were 2.064 ($df = 24$, newborns) and 2.056 ($df = 26$, adults), two-tailed. Temporally and spatially adjacent samples, lasting at least 2 s and statistically significant (Marino et al., 2026; Martinez-Alvarez et al., 2023), were grouped into candidate clusters, and cluster-level statistics were computed by summing their t-values. Statistical significance was assessed via 1000 random permutations of condition labels to generate a null distribution. For each candidate cluster, the p-value was determined as the proportion of permutations yielding a more extreme cluster statistic than observed (minimum resolvable p value = 0.001). This non parametric statistical analysis, described in detail for neural time series dataset in Maris and Oostenveld (2007), overcomes the multiple comparisons problem, which typically challenges channel-by-channel statistical tests. It is now standardly used in NIRS studies (Abboub et al., 2016; Cabrera and Gervain, 2020; Ferry et al., 2016; Gemignani and Gervain, 2024; Giordano et al., 2021; Mahmoudzadeh et al., 2013).

The comparison against the baseline can identify clusters of channels in which cries and speech respectively elicit a significant brain response relative to the baseline. These effects may be hidden when directly comparing the two conditions, if their processing triggers responses of similar amplitudes in similar channels. The cluster-based permutation analysis defines regions of interest (ROIs) and time windows of interest in a non-arbitrary, data-driven manner, while preventing multiple

comparisons (Benavides-Varela and Gervain, 2017; Mahmoudzadeh et al., 2013).

Given the overall differences in brain maturation, neurovascular coupling, the much greater attenuation of the hemodynamic response by adults' thicker tissues, and greater noise/variability in infants than in adults, a direct comparison of the two age groups was not performed.

3. Results

Adults and newborns' grand average hemodynamic responses to Speech and Cries are shown in Fig. 2.

3.1. Newborns

A cluster-based permutation analysis (Abboub et al., 2016; Mahmoudzadeh et al., 2013; Maris and Oostenveld, 2007) comparing newborns' oxyHb responses in the two conditions showed greater activation to Cries than to Speech in the right temporal region (channel 19, $p = .001$). Importantly, this was not because they failed to perceive or respond to the Speech stimuli. The cluster-based permutation test comparing activations for Speech to a zero baseline revealed significantly greater responses to Speech than to baseline in the left temporal (channels 3 and 6, $p < .001$) and right temporo-parietal (channels 20 and 23, $p < .001$) areas. A similar analysis showed greater responses to Cries than to baseline in the left fronto-temporo-parietal (channels 4 and 9, $p < .001$) and right temporal (channels 19 and 22, $p < .001$) areas.

Converging results were found in a similar analysis over deoxyHb. The cluster-based permutation analysis revealed a right temporo-parietal cluster (channels 18, 19 and 21; $p < .001$) where deoxygenated hemoglobin concentration changes differed significantly between Cries and Speech. In addition, the permutations showed that the deoxyHb activation was significantly more negative than the zero baseline both in response to Speech (in the form of a right temporal cluster comprising channels 17 and 19; $p < .001$) and to Cries (in the right temporal channel 17, $p < .001$).

3.2. Adults

As expected, but in contrast with the newborn results, adults showed a greater response to Speech than to Cries in the left temporo-parietal (channels 3, 6, 9 and 12; $p < .001$) and right temporal (17 and 19, $p < .001$) areas in a cluster-based permutation test over oxyHb. A similar analysis comparing Speech to a zero baseline revealed two left clusters, one temporal (channels 3, 6, and 9, $p < .001$) and one temporo-parietal (channels 2, 5, 7, and 9, $p < .001$), and two right clusters, one fronto-temporal (channels 14, 16, 17, 19, and 21, $p < .001$) and one fronto-parietal cluster (channels 13 and 15, $p < .001$) where the oxyHb response to Speech was significantly greater than the baseline. The oxyHb response to Cries was also significantly higher than the baseline, as indicated by a left fronto-parietal cluster (channels 2, 5, 10 and 12, $p < .001$) and two right clusters, a temporal (channels 17 and 19, $p < .001$) and a fronto-parietal one (channels 13 and 15, $p < .001$).

The cluster-based permutation analysis over the concentration changes of deoxyHb yielded converging results. Thus, brain activity significantly differed between Cries and Speech in a left temporo-parietal cluster (channels 4, 6, 7 and 9, $p < .001$). DeoxyHb responses were also significantly more negative than baseline in both conditions. The comparison of Speech to the baseline revealed a left fronto-parietal cluster containing channels 2 and 5 ($p < .001$) and a temporo-parietal cluster involving channel 9 ($p < .001$) as well as a right fronto-temporal cluster containing channels 14, 17, and 19 ($p < .001$), while the comparison of Cries to baseline revealed a right temporal cluster (channels 17, 19, and 21, $p < .001$).

4. Discussion

Our study investigating adults and newborn infants' neural processing of speech and newborn cries showed opposite patterns of results in the two populations. As predicted, and despite the strong emotional valence of the newborn cries, the adult brain responded more strongly to speech than to cries in the bilateral temporal areas. This predominantly auditory and phonological activation corresponds to processing speech sounds in an unfamiliar language, the grammar and vocabulary of which is unknown to the listener, making syntactic, lexical and semantic processing unavailable (Friederici, 2011, 2012; Poeppel, 2014).

If speech had the same privileged role for newborns as it does for adults, we would expect a similar pattern of results in the neonate brain. But this is not what we observed: the neonate brain responded more to newborn cries than to speech. Importantly, this was not because they failed to perceive or process the speech stimuli. In line with a large body of literature showing sensitivity to speech and language in the newborn brain (Gervain et al., 2008; Homae et al., 2007; Mahmoudzadeh et al., 2013; May et al., 2018; Minagawa-Kawai et al., 2011; Peña et al., 2003), we also observed significant activation in response to speech in the left temporal and right parietal areas. These findings dovetail with the left lateralized or bilateral temporo-parietal activation typically observed in newborns for the prenatally heard language (May et al., 2011; Peña et al., 2003; Sato et al., 2012; Vannasing et al., 2016).

Importantly for our research question, newborns' brain responses to cries were greater than their responses to speech, in particular in the right temporal cortex. What may be the reason for this heightened response to cries in newborns, but not in adults? As hypothesized initially, while speech in the prenatally heard language is a more complex signal and more familiar to newborns than cries, cries are the only vocal, communicative signals newborns produce themselves. The ability to produce cries, even if reflexive at birth, may thus underlie newborns' increased neural responses to them, suggesting that a link between production and perception might already be present at birth. This interpretation is supported by our findings, as speech triggered a strictly temporal activation in the left hemisphere, while cries triggered parietal and frontal activation, where the pre-motor and motor cortices are found. Adults, by contrast, showed activation in the same region bilaterally in response to both cries and speech. Our results thus suggest production-related effects, with activation of motor regions only in response to those signal(s) with which each age group has production experience. If newborns' responses to cries were primarily driven by novelty, we would expect enhanced responses to cries only in temporal areas, typically involved in auditory processing, rather than selective activation of motor regions in response only to those signal(s) with which each age group has production experience. While we cannot rule out that arousal might have enhanced newborns' response to cries, as reflected by a greater than usual attrition rate, the involvement of motor regions and opposite activation patterns found in the two populations support the role of production experience in newborns' processing of cries and speech, in line with theories positing the production-perception link to be innate and biologically driven (Liberman and Mattingly, 1985).

While our results provide the first evidence for a potential production-perception link at birth, further studies are needed to provide converging evidence and confirmation. Future studies examining other types of vocal productions that newborns and young infants can produce will further disentangle the role of familiarity, acoustic complexity, emotional valence and other related factors. For instance, cries produced by newborns prenatally exposed to a different language may be particularly relevant, since cry melody may already be modulated by prenatal experience (Mampe et al., 2009; Manfredi et al., 2019). Thus, cries produced by newborns with a different prenatal exposure may sound "non-native" and less familiar. Coos and other reflexive vocalizations are also interesting cases in point. Additionally, future studies could also investigate the production-perception link more

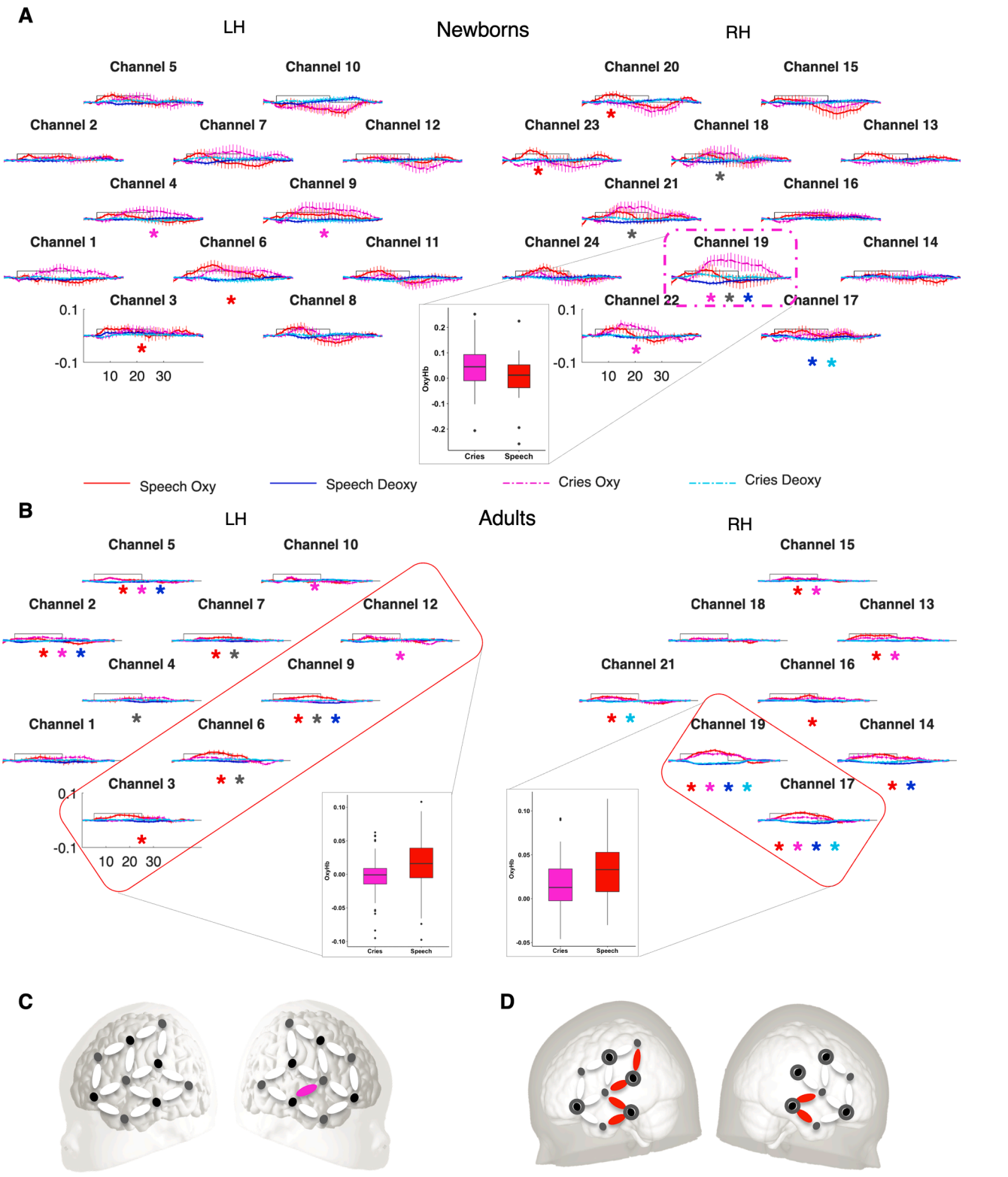


Fig. 2. The grand average hemodynamic responses obtained in the study. (A-B) Channels are plotted following the probe placement displayed in Fig. 1D. The x-axis represents time in seconds; the y-axis shows concentration changes in mmol x mm. The rectangle along the x-axis indicates the time of stimulation. OxyHb and deoxyHb concentration changes in response to the Speech condition are shown in red and blue, respectively (continuous line). OxyHb and deoxyHb concentration changes in response to the Cry condition are shown in magenta and cyan, respectively (dashed line). Boxes represent clusters of channels in which the OxyHb concentration changes were significantly different between the two conditions. The boxes encircle the cluster of channels in which oxyHb concentration significantly differed between cries and speech conditions, while the gray asterisks, placed below individual channels, depict channels with significantly different deoxyHb concentrations in cries vs speech conditions. Other asterisks indicate significant differences between the condition indicated by the asterisk's color vs. baseline. (CD) The location of significant channels indicated in magenta when Cry > Speech, and in red when Speech > Cry on the newborn and adult channel layouts.

directly, testing newborn infants' responses to their own cries.

Newborn cries also elicited significant activation in a right temporal area in both newborns and adults. Right lateralized activation has been linked to the processing of music and language prosody, or more generally to the processing of slowly and/or spectrally modulated sounds in adults (Hickok and Poeppel, 2007; Poeppel, 2014; Zatorre and Belin, 2001), young infants (Homae et al., 2007) and newborns (Martinez-Alvarez et al., 2023). As communicative signals, newborn cries convey both emotional prosody, e.g. distress, physiological needs etc., and, as some studies suggest (Mampe et al., 2009; Manfredi et al., 2019), the linguistic prosody of the native language the newborns heard in utero. The right-lateralized activation found in response to cries may, therefore, be linked to the processing of the prosodic contours of cries. It is important to note that the overall higher pitch of cries does not in itself provide an alternative explanation for our findings, since we would expect similar patterns of activation in adults and newborns. Since prenatal language experience may already modulate the intonational contours of cries at birth (Mampe et al., 2009; Manfredi et al., 2019), future studies should investigate newborns' processing of non-native cries and speech.

By sharing neural and functional foundations with speech, cries may provide a bridge from simple vocal output to the acquisition of structured language. Its relevance for language development is mirrored by studies on laughter, another early-produced vocal signal. Like cries, laughter is constrained in its production and conveys meaning, signaling affective states and shaping social interactions (Ginzburg et al., 2020). Such vocalizations may thus scaffold infants' sensitivity to communicative signals.

In sum, this study provides the first demonstration that unlike adults, newborn infants show heightened brain responses to cries, a communicative signal that is less familiar and less complex than speech, but one that the infants themselves can produce, suggesting that a link between vocal production and auditory perception may be present early in life.

Funding

This work was supported by the ERC Consolidator Grant "BabyRhythm" no. 773202, the FARE grant no. R204MPRHKE from the Italian Ministry for Universities and Research, the "SYNPHONIA" project (Grant Nos. CLP: PNRR-MAD 2022-12376739 and CUP: C63C22001320007) funded by the European Union Next Generation EU (Grant. No. PNRR M6C2, Investment 2.1 "Enhancement and strengthening of biomedical research in the NHS"), the PRIN grant no. 2022WX3FM5 (CUP: C53D23004290006) from the Italian Ministry for Universities and Research to J Gervain; the Grant RYC2021–03395-I funded by MICIU/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR, the Basque Foundation for Science Ikerbasque to IdlCP; and the Marie Curie Individual Fellowship EF-ST "BabyMindReader" 101031716 awarded by the European Commission to JG.

CRediT authorship contribution statement

Judit Gervain: Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Caroline Nallet:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Alexandre Lapillonne:** Resources. **Irene de la Cruz-Pavía:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Jessica Gemignani:** Writing – review & editing, Writing – original draft, Funding acquisition, Data curation. **Gaia Lucarini:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis.

Declaration of Competing Interest

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

The authors wish to thank the staff of the Maternity Ward of the Hôpital Necker - Enfants Malades for their help with testing as well as the participating newborns and their parents, and the participating adults.

The authors created part of the graphical abstract in BioRender. Lucarini, G. (2026) https://BioRender.com/clut710.

Data availability

All data and code needed to evaluate the conclusions in the paper are available at https://osf.io/tve9w/.

References

Abboub, N., Boll-Avetisyan, N., Bhatara, A., Höhle, B., Nazzi, T., 2016. An exploration of rhythmic grouping of speech sequences by French- and German-learning infants. Front. Human. Neurosci. 10. https://doi.org/10.3389/fnhum.2016.00292.

Benavides-Varela, S., Gervain, J., 2017. Learning word order at birth: a NIRS study. Dev. Cogn. Neurosci. 25, 198–208. https://doi.org/10.1016/j.dcn.2017.03.003.

Bouchon, C., Nazzi, T., Gervain, J., 2015. Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain. PLOS One 10 (10), e0140160. https://doi.org/10.1371/journal.pone.0140160.

Brigadoi, S., Cooper, R.J., 2015. How short is short? Optimum source–detector distance for short-separation channels in functional near-infrared spectroscopy. Neurophotonics 2 (2), 025005. https://doi.org/10.1117/1.NPh.2.2.025005.

Bruderer, A.G., Danielson, D.K., Kandhadai, P., Werker, J.F., 2015. Sensorimotor influences on speech perception in infancy. Proc. Natl. Acad. Sci. USA 112 (44), 13531–13536. https://doi.org/10.1073/pnas.1508631112.

Cabrera, L., Gervain, J., 2020. Speech perception at birth: the brain encodes fast and slow temporal information. Sci. Adv. 6 (30), eaba7830. https://doi.org/10.1126/sciadv.aba7830.

Cheung, C., Hamilton, L.S., Johnson, K., Chang, E.F., 2016. The auditory representation of speech sounds in human motor cortex. eLife 5, e12577. https://doi.org/10.7554/eLife.12577.

Choi, D., Dehaene-Lambertz, G., Peña, M., Werker, J.F., 2021. Neural indicators of articulator-specific sensorimotor influences on infant speech perception. Proc. Natl. Acad. Sci. USA 118 (20), e2025043118. https://doi.org/10.1073/pnas.2025043118.

Choi, D., Yeung, H.H., Werker, J.F., 2023. Sensorimotor foundations of speech perception in infancy. Trends Cogn. Sci. 27 (8), 773–784. https://doi.org/10.1016/j.tics.2023.05.007.

DePaolis, R.A., Vihman, M.M., Keren-Portnoy, T., 2011. Do production patterns influence the processing of speech in prelinguistic infants? Infant. Behav. Dev. 34 (4), 590–601. https://doi.org/10.1016/j.infbeh.2011.06.005.

DePaolis, R.A., Vihman, M.M., Nakai, S., 2013. The influence of babbling patterns on the processing of speech. Infant. Behav. Dev. 36 (4), 642–649. https://doi.org/10.1016/j.infbeh.2013.06.007.

Dondi, M., Simion, F., Caltran, G., 1999. Can newborns discriminate between their own cry and the cry of another newborn infant? Dev. Psychol. 35 (2), 418.

Emberson, L.L., Crosswhite, S.L., Goodwin, J.R., Berger, A.J., Aslin, R.N., 2016. Isolating the effects of surface vasculature in infant neuroimaging using short-distance optical channels: a combination of local and global effects. Neurophotonics 3 (3), 031406. https://doi.org/10.1117/1.NPh.3.3.031406.

Ferry, A.L., Fló, A., Brusini, P., Cattarossi, L., Macagno, F., Nespor, M., Mehler, J., 2016. On the edge of language acquisition: inherent constraints on encoding multisyllabic sequences in the neonate brain. Dev. Sci. 19 (3), 488–503. https://doi.org/10.1111/desc.12323.

Fogel, A., Thelen, E., 1987. Development of early expressive and communicative action: reinterpreting the evidence from a dynamic systems perspective. Dev. Psychol. 23 (6), 747–761. https://doi.org/10.1037/0012-1649.23.6.747.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392.

Friederici, A.D., 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn. Sci. http://www.sciencedirect.com/science/article/pii/S1364661312000794.

Frota, S., Severino, C., Pejovic, J., Paulino, N., Vigário, M., Noiray, A., 2026. Articulating while listening supports the emerging perception-production link in early infancy. Sci. Rep. https://doi.org/10.1038/s41598-026-42240-5.

Galantucci, B., Fowler, C.A., Turvey, M.T., 2006. The motor theory of speech perception reviewed. Psychon. Bull. & Rev. 13 (3), 361–377. https://doi.org/10.3758/BF03193857.

Gemignani, J., de la Cruz-Pavía, I., Martinez, A., Nallet, C., Pasquini, A., Lucarini, G., Cavicchiolo, F., Gervain, J., 2023. Reproducibility of infant fNIRS studies: a meta-analytic approach. Neurophotonics 10 (02). https://doi.org/10.1117/1.NPh.10.2.023518.

Gemignani, J., Gervain, J., 2021. Comparing different pre-processing routines for infant fNIRS data. Dev. Cogn. Neurosci. 48, 100943. https://doi.org/10.1016/j.dcn.2021.100943.

Gemignani, J., Gervain, J., 2024. Brain responses to repetition-based rule-learning do not exhibit sex differences: an aggregated analysis of infant fNIRS studies. Sci. Rep. 14 (1), 2611. https://doi.org/10.1038/s41598-024-53092-2.
Gervain, J., Macagno, F., Cogoi, S., Peña, M., Mehler, J., 2008. The neonate brain detects speech structure. Proc. Natl. Acad. Sci. USA 105 (37), 14222–14227. https://doi.org/10.1073/pnas.0806530105.
Ginzburg, J., Mazzocconi, C., Tian, Y., 2020. Laughter as language. Glossa. A J. General. Linguist. 5 (1). https://doi.org/10.5334/gjgl.1152.
Giordano, V., Alexopoulos, J., Spagna, A., Benavides-Varela, S., Peganc, K., Kothgassner, O.D., Klebermass-Schrehof, K., Olischar, M., Berger, A., Bartha-Doering, L., 2021. Accent discrimination abilities during the first days of life: an fNIRS study. Brain Lang. 223, 105039. https://doi.org/10.1016/j.bandl.2021.105039.
Guenther, F.H., 2016. Neural Control of Speech. The MIT Press.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402.
Hocke, L.M., Oni, I.K., Duszynski, C.C., Corrigan, A.V., Frederick, B. deB., Dunn, J.F., 2018. Automated processing of fNIRS Data—a visual guide to the pitfalls and consequences. Algorithms 11 (5), 67. https://doi.org/10.3390/a11050067.
Homae, F., Watanabe, H., Nakano, T., Taga, G., 2007. Prosodic processing in the developing brain. Neurosci. Res. 59 (1), 29–39.
Ito, T., Tiede, M., Ostry, D.J., 2009. Somatosensory function in speech perception. Proc. Natl. Acad. Sci. 106 (4), 1245–1248. https://doi.org/10.1073/pnas.0810063106.
Kent, R.D., 2021. Developmental functional modules in infant vocalizations. J. Speech Lang. Hear. Res. 64 (5), 1581–1604. https://doi.org/10.1044/2021_JSLHR-20-00703.
Liberman, A.M., 1984. On finding that speech is special. In: Gazzaniga, M.S. (Ed.), Handbook of Cognitive Neuroscience. Springer US, pp. 169–197. https://doi.org/10.1007/978-1-4899-2177-2_9.
Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. Psychol. Rev. 74 (6), 431–461.
Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech perception revised. Cognition 21 (1), 1–36. https://doi.org/10.1016/0010-0277(85)90021-6.
Liebenthal, E., Möttönen, R., 2018. An interactive model of auditory-motor speech perception. Brain Lang. 187, 33–40. https://doi.org/10.1016/j.bandl.2017.12.004.
Lorenzini, I., Nazzi, T., 2022. Early recognition of familiar word-forms as a function of production skills. Front. Psychol. 13, 947245. https://doi.org/10.3389/fpsyg.2022.947245.
Lotto, A.J., Hickok, G.S., Holt, L.L., 2009. Reflections on mirror neurons and speech perception. Trends Cogn. Sci. 13 (3), 110–114. https://doi.org/10.1016/j.tics.2008.11.008.
Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., Grebe, R., Wallois, F., 2013. Syllabic discrimination in premature human infants prior to complete formation of cortical layers. Proc. Natl. Acad. Sci. USA 110 (12), 4846–4851.
Majorano, M., Vihman, M.M., DePaolis, R.A., 2014. The relationship between infants' production experience and their processing of speech. Lang. Learn. Dev. 10 (2), 179–204. https://doi.org/10.1080/15475441.2013.829740.
Mampe, B., Friederici, A.D., Christophe, A., Wermke, K., 2009. Newborns' cry melody is shaped by their native language. Curr. Biol. 19 (23), 1994–1997. https://doi.org/10.1016/j.cub.2009.09.064.
Manfredi, C., Viellevoye, R., Orlandi, S., Torres-García, A., Pieraccini, G., Reyes-García, C.A., 2019. Automated analysis of newborn cry: relationships between melodic shapes and native language. Biomed. Signal. Process. Control. 53, 101561. https://doi.org/10.1016/j.bspc.2019.101561.
Marino, C., Gemignani, J., Varnet, L., Zanetto, L., Baraldi, E., Gervain, J., 2026. Singing to the newborn brain uncovers early traces of specialized neural networks. Commun. Psychol. https://doi.org/10.1038/s44271-026-00451-6.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods 164 (1), 177–190.
Martin, G.B., Clark, R.D., 1982. Distress crying in neonates: species and peer specificity. Dev. Psychol. 18 (1), 3.
Martinez-Alvarez, A., Benavides-Varela, S., Lapillonne, A., Gervain, J., 2023. Newborns discriminate utterance-level prosodic contours. Dev. Sci. 26 (2), e13304. https://doi.org/10.1111/desc.13304.
Masapollo, M., Polka, L., Ménard, L., 2016. When infants talk, infants listen: pre-babbling infants prefer listening to speech with infant vocal properties. Dev. Sci. 19 (2), 318–328. https://doi.org/10.1111/desc.12298.
May, L., Byers-Heinlein, K., Gervain, J., Werker, J.F., 2011. Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? Front. Psychol. 2. https://doi.org/10.3389/fpsyg.2011.00222.
May, L., Gervain, J., Carreiras, M., Werker, J.F., 2018. The specificity of the neural response to speech at birth. Dev. Sci. 21 (3), e12564. https://doi.org/10.1111/desc.12564.
Minagawa-Kawai, Y., Cristià, A., Vendelin, I., Cabrol, D., Dupoux, E., 2011. Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. Front. Psychol. 2, 135. https://doi.org/10.3389/fpsyg.2011.00135.
Nguyen, T., Hoehl, S., Vrtička, P., 2021. A guide to parent-child fNIRS hyperscanning data processing and analysis. Sensors 21 (12), 4075. https://doi.org/10.3390/s21124075.
Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., Mehler, J., 2003. Sounds and silence: an optical topography study of language recognition at birth. Proc. Natl. Acad. Sci. 100 (20), 11702–11705. https://doi.org/10.1073/pnas.1934290100.
Piper, S.K., Krueger, A., Koch, S.P., Mehnert, J., Habermehl, C., Steinbrink, J., Obrig, H., Schmitz, C.H., 2014. A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. NeuroImage 85, 64–71. https://doi.org/10.1016/j.neuroimage.2013.06.062.
Poeppel, D., 2014. The neuroanatomic and neurophysiological infrastructure for speech and language. Curr. Opin. Neurobiol. 28, 142–149. https://doi.org/10.1016/j.conb.2014.07.005.
Poeppel, D., Assaneo, M.F., 2020. Speech rhythms and their neural foundations. Nat. Rev. Neurosci. 21 (6), 322–334. https://doi.org/10.1038/s41583-020-0304-4.
Polka, L., Alonso-Arteche, M.F., Phillips, N.K., Moradi, S., Ménard, L., Masapollo, M., 2025. Infants' attraction to infant vocalizations – a catalyst for infant development. Infant. Behav. Dev. 81, 102150. https://doi.org/10.1016/j.infbeh.2025.102150.
Polka, L., Masapollo, M., Ménard, L., 2014. Who's talking now? Infants' perception of vowels with infant vocal properties. Psychol. Sci. 25 (7), 1448–1456. https://doi.org/10.1177/0956797614533571.
Ruffman, T., Lorimer, B., Scarf, D., 2017. Do infants really experience emotional contagion? Child. Dev. Perspect. 11 (4), 270–274. https://doi.org/10.1111/cdep.12244.
Saager, R.B., Berger, A.J., 2005. Direct characterization and removal of interfering absorption trends in two-layer turbid media. JOSA A 22 (9), 1874–1882. https://doi.org/10.1364/JOSAA.22.001874.
Sagi, A., Hoffman, M.L., 1976. Empathic distress in the newborn. Dev. Psychol. 12 (2), 175–176. https://doi.org/10.1037/0012-1649.12.2.175.
Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., Uchida-Ota, M., Konishi, Y., Maki, A., 2012. Cerebral hemodynamics in newborn infants exposed to speech sounds: a whole-head optical topography study. Human. Brain Mapp. 33 (9), 2092–2103.
Skeide, M.A., Friederici, A.D., 2016. The ontogeny of the cortical language network. Nat. Rev. Neurosci. 17 (5), 323–332. https://doi.org/10.1038/nrn.2016.23.
Skipper, J.I., Devlin, J.T., Lametti, D.R., 2017. The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. Brain Lang. 164, 77–105. https://doi.org/10.1016/j.bandl.2016.10.004.
Spence, M.J., DeCasper, A.J., 1987. Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. Infant. Behav. Dev. 10 (2), 133–142. https://doi.org/10.1016/0163-6383(87)90028-2.
Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., Cooper, F.S., 1970. Motor theory of speech perception: a reply to Lane's critical review. Psychol. Rev. 77 (3), 234–249. https://doi.org/10.1037/h0029078.
Szmytke, M.2026. The role of sensorimotor engagement in speech perception and production in infancy [Doctoral Dissertation, University of Warsaw]. 42970c3b-a28e-4f29-87ee-b6c294f6e261.
Tachtsidis, I., Scholkmann, F., 2016. False positives and false negatives in functional near-infrared spectroscopy: Issues, challenges, and the way forward. Neurophotonics 3 (3), 031405. https://doi.org/10.1117/1.NPh.3.3.031405.
Van Niekerk, E., Junge, C., Chen, A., 2025. The influence of prenatal language exposure on the use of pitch in newborns' vocalisations: a systematic review. First Lang., 01427237251372203 https://doi.org/10.1177/01427237251372203.
Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., Wallois, F., Lepore, F., Béland, R., Lassonde, M., Gallagher, A., 2016. Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. Neuropsychologia 84, 63–69. https://doi.org/10.1016/j.neuropsychologia.2016.01.038.
Vilain, A., Dole, M., Lœvenbruck, H., Pascalis, O., Schwartz, J.-L., 2019. The role of production abilities in the perception of consonant category in infants. Dev. Sci. 22 (6), e12830. https://doi.org/10.1111/desc.12830.
Vouloumanos, A., Hauser, M.D., Werker, J.F., Martin, A., 2010. The tuning of human neonates' preference for speech. Child. Dev. 81 (2), 517–527.
Vouloumanos, A., Werker, J.F., 2007. Listening to language at birth: evidence for a bias for speech in neonates. Dev. Sci. 10 (2), 159–164. https://doi.org/10.1111/j.1467-7687.2007.00549.x.
Werker, J.F., 2018. Perceptual beginnings to language acquisition. Appl. Psycholinguist. 39 (4), 703–728. https://doi.org/10.1017/S0142716418000152.
Wermke, K., Mende, W., Manfredi, C., Bruscaglioni, P., 2002. Developmental aspects of infant's cry melody and formants. Med. Eng. & Phys. 24 (7), 501–514. https://doi.org/10.1016/S1350-4533(02)00061-9.
Yeung, H.H., Scott, M., 2021. Postural control of the vocal tract affects auditory speech perception. J. Exp. Psychol. General. 150 (5), 983–995. https://doi.org/10.1037/xge0000990.
Yoo, H., Bowman, D.A., Oller, D.K., 2018. The origin of protoconversation: an examination of caregiver responses to cry and speech-like vocalizations. Front. Psychol. 9, 1510. https://doi.org/10.3389/fpsyg.2018.01510.
Yücel, M.A., Luke, R., Mesquita, R.C., Von Lühmann, A., Mehler, D.M.A., Lührs, M., Gemignani, J., Abdalmalak, A., Albrecht, F., De Almeida Ivo, I., Artemenko, C., Ashton, K., Augustynowicz, P., Bajracharya, A., Bannier, E., Barth, B., Bayet, L., Behrendt, J., Khani, H.B., Zemanek, V., 2025. fNIRS reproducibility varies with data quality, analysis pipelines, and researcher experience. Commun. Biol. 8 (1), 1149. https://doi.org/10.1038/s42003-025-08412-1.
Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11 (10), 946.
Zimeo Morais, G.A., Scholkmann, F., Balardin, J.B., Furucho, R.A., De Paula, R.C.V., Biazoli, C.E., Sato, J.R., 2017. Non-neuronal evoked and spontaneous hemodynamic changes in the anterior temporal region of the human head may lead to misinterpretations of functional near-infrared spectroscopy signals. Neurophotonics 5 (01), 1. https://doi.org/10.1117/1.NPh.5.1.011002.