Resilience and vulnerability of speech neural tracking to early auditory deprivation

Alessandra Federici[1], Marta Fantoni[1], Francesco Pavani[2,3], Giacomo Handjaras[1], Evgenia Bednaya[1], Alice Martinelli[1,4], Martina Berto[1], Emiliano Ricciardi[1], Elena Nava[5], Eva Orzan[6], Benedetta Bianchi[7] & Davide Bottari*[1]

1. MoMiLab, IMT School for Advanced Studies Lucca, Italy;

2. Centro Interdipartimentale Mente/Cervello – CIMEC, University of Trento, Italy;

3. Centro Interuniversitario di Ricerca "Cognizione Linguaggio e Sordità" – CIRCLeS; University of Trento, Italy;

4. IRCCS Fondazione Stella Maris, Pisa, Italy;

5. University of Milano-Bicocca, Milano, Italy;

6. IRCCS Materno Infantile Burlo Garofolo, Trieste, Italy;

7. IRCCS Meyer, Azienda Ospedaliero-Universitaria Meyer, Firenze, Italy.

Corresponding author:

Davide Bottari

davide.bottari@imtlucca.it

Keywords: Neural Tracking of Speech, Cochlear Implants, EEG, Sensitive Periods, Experience-Dependent Plasticity

SUMMARY

Infants are born with biologically constrained biases that favor language acquisition. One is the auditory system's ability to track the envelope of continuous speech. However, to what extent the synchronization between brain activity and this pivotal speech feature relies on postnatal auditory experience remains unknown. To uncover this, we studied individuals with or without access to functional hearing in the first year of life after they received cochlear implants (CIs) for hearing restoration. We measured the neural synchronization with continuous speech envelope in children with congenital bilateral profound deafness (CD; minimum auditory deprivation 11 months) or who acquired profound deafness later in development (AD; minimum auditory experience after birth 12 months), as well as in hearing controls (HC). Speech envelope tracking was unaffected by the absence of auditory experience in the first year of life. At short timescales, neural tracking had a similar magnitude in CI users and HC. However, in CI users, it was substantially delayed, and its timing depended on the age of hearing restoration. Conversely, we observed alterations at longer timescales, possibly accounting for the comprehension deficits observed in children with CI. These findings highlight (i) the resilience of sensory components of speech envelope tracking to the lack of hearing in the first year of life, supporting its strong biological bias, (ii) the crucial role of when functional hearing restoration takes place in mitigating the impact of atypical auditory development, (iii) the vulnerability of higher hierarchical levels of speech-envelope tracking in CI users. Neural tracking of continuous speech could provide biomarkers along the processing hierarchy between sensory and core linguistic operations, even after cochlear implantation.

INTRODUCTION

Slow brain activity temporally aligns with the regularities of speech. Through this mechanism, the auditory cortex tracks speech features, such as signal amplitude modulations. These fluctuations, or speech envelope, have energy peaks around the syllabic rate and have been posited as a pivotal feature for speech comprehension. Evidence shows that, in quiet conditions, adults can understand heavily degraded speech, provided the envelope is preserved (Shannon et al., 1995). On the contrary, suppressing the speech envelope impairs comprehension (Drullman et al. 1994a; b). Importantly, intelligibility is markedly impaired in case of poor neural tracking of speech envelope (Vanthornhout et al., 2018).

The primary role of the neural tracking of speech envelope is moreover substantiated by the early development of this brain mechanism. Neural speech tracking has been documented in newborns and young infants (Kalashnikova et al., 2018; Jessen et al., 2019; Ortiz Barajas et al., 2021; Attaheri et al., 2022), suggesting a strong biological predisposition. Nonetheless, recent evidence suggested that linguistic experience in the first year of life modulates speech envelope tracking (Ortiz Barajas et al., 2021). Yet, it is unknown to what extent the development of speech envelope tracking relies on postnatal hearing experience and, thus, which would be the impact of auditory deprivation on this pivotal brain function.

Typical brain development requires temporal overlapping between neural system readiness and appropriate environmental statistics (Reh et al., 2020; Werker & Hensch, 2015; Bottari & Berto, 2021). Individuals facing a period of sensory deprivation provide the unique opportunity for causally assessing whether, in the absence of specific input within a specific phase of life, neural functions typically develop or not (Kral et al., 2019; Röder & Kekunnaya, 2021; Ricciardi et al., 2020). That is, whether neural functions have sensitive phases in which specific experience must be provided for shaping neural circuitries subtending certain computations (Hensch et al., 2005) or whether their development is mainly guided by biological predispositions instead.

Following a period of profound bilateral sensorineural hearing loss (profound deafness from here onward), in which sounds cannot reach the auditory system, the cochlear implant (CI) provides the possibility of partial auditory restoration (Winn & Nelson, 2021; Pavani & Bottari, 2022; Gates et al., 1995). The importance of early access to sounds for the development of auditory functions is supported by studies assessing language acquisition and neurophysiological responses in CI individuals. Seminal studies uncovered that the latency of auditory responses falls within typical developmental trajectory only when children are implanted before 3.5 years of age (i.e., P1 and N1 waves of the ERPs, e.g., Sharma et al., 2002a, 2002b, 2005, 2015; Eggermont & Ponton, 2003). However, the activity of the temporal cortex in CI individuals has been measured in reaction to simple and short-lived sounds, such as syllables. Thus, two key questions remain unanswered: (i) to what extent does the CI provide the possibility to develop hearing-like neural tracking of continuous speech? (ii) is there a perinatal sensitive period in which auditory input must be available for the development of this function?

To fill these gaps, we measured the degree of synchronization between brain activity and continuous speech in hearing and CI children with different onsets of bilateral profound

deafness (congenital or acquired). Children born with congenital deafness (CD) have been auditory deprived at least during the first 11 months of life before cochlear implantation. Conversely, all children with acquired deafness (AD) were born with some degree of functional hearing, and profound deafness emerged only after 12 months of age (see Figure 1). The children's age range (3 – 18 years old) was chosen to measure developmental trajectories of the synchronization between brain activity and continuous speech in all groups.

Results revealed that neural tracking was not influenced by the presence or absence of auditory experience in the first year of life, indicating a biological predisposition of the human brain's ability to synchronize with the speech envelope despite temporary auditory deprivation from birth. However, the neural tracking dynamic at a short timescale, typically capturing the sensory component of speech processing, was delayed in CI children, and the implantation age (i.e., access to functional hearing) modulated the effect. Nevertheless, at this timescale, the neural tracking magnitude was unimpaired in CI children. Finally, results suggested that alterations at higher-level speech processing, yet captured by the envelope tracking dynamic at a longer timescale, could account for CI children's speech comprehension deficits. This latter finding revealed that the neural tracking of speech envelope could be employed to extract biomarkers of sensory and core linguistic operations in developing populations with hearing sensory substitution devices.

RESULTS

The sample comprised 69 children (age range 3 – 18 years old): thirty-two children with profound sensorineural hearing loss and using CI. Half of which had bilateral profound congenital deafness (CD group; mean age at cochlear implantation 27.7 months, range: 11 – 132 months), and half acquired bilateral profound deafness after the first year of life (AD group; mean age at cochlear implantation 53.6 months, range: 17 – 120 months, see Figure 1 and Table S1 reporting CI users' clinical characteristics). Thirty-seven children with typical hearing were also recruited for the study (hearing controls, HC). The CD, AD, and HC groups were age and gender-matched. To ensure consolidated auditory experience, all CI children were tested at least six months after cochlear implant activation (Sharma et al., 2002b). Participants were asked to listen to four stories (~ three minutes each), consisting of continuous speech, whose content was chosen according to the participant's age. At the

end of each story, children were asked to respond to a two-alternative forced-choice comprehension questionnaire (see Task and experimental procedure).

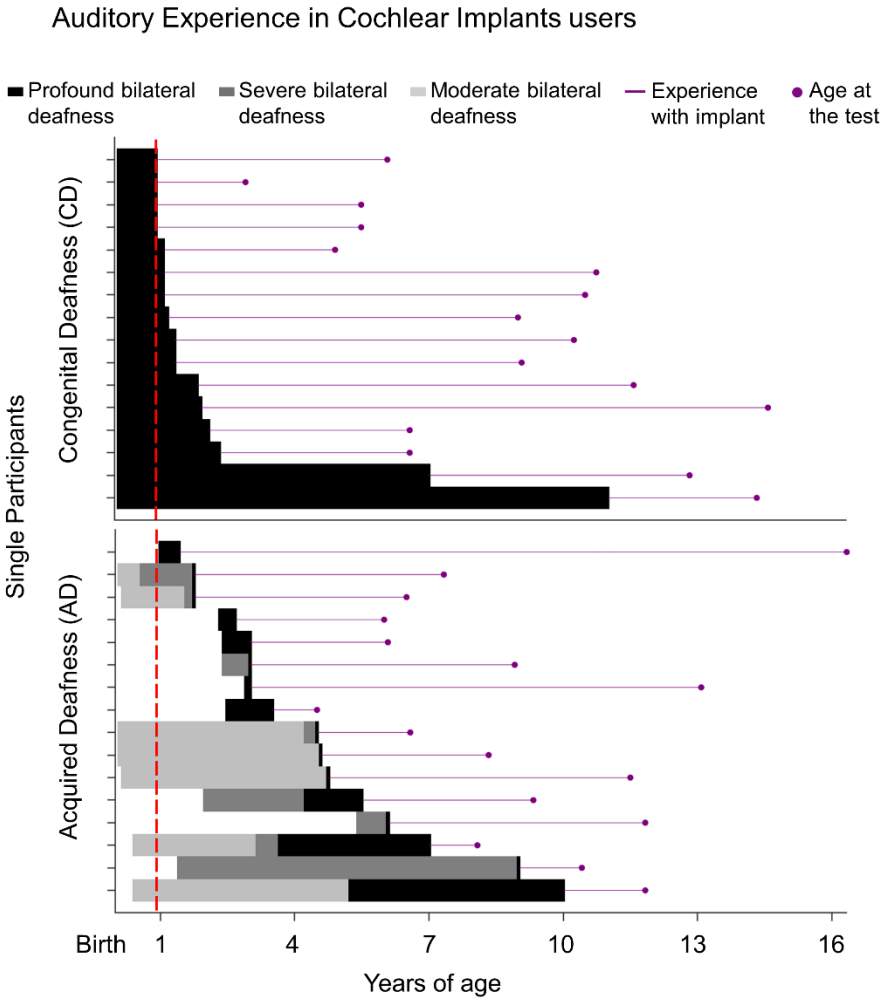

Figure 1. Auditory experience in CI users. The plot graphically represents the auditory experience that characterizes each CI participant. Periods of profound bilateral deafness are depicted in black, whereas severe and moderate bilateral deafness are rendered with different shades of gray. The red dashed line highlights the different hearing experiences between congenital deafness (CD) and acquired deafness (AD) groups: only the CD group faced auditory deprivation throughout the first year of life. The magenta lines illustrate experiences with implants until the day of testing, represented by the dot.

Neural tracking of speech in hearing and cochlear-implanted children

First, we assessed whether the neural tracking of speech envelope could be measured in HC and CI children irrespective of deafness onset. The temporal response function (TRF) was computed employing an encoding model to predict slow brain oscillations (2 - 8 Hz) from speech envelope (below 8 Hz, Crosse et al., 2016, 2021). We estimated the auditory response function within a frontocentral cluster of sensors (Cz, Fz, FC1, and FC2) that typically capture auditory response functions in children and adults (e.g., Jessen et al., 2019; Fiedler et al., 2019) and in CI users (Paul et al., 2020). The TRF model was compared with a set of null-TRF models fitted on randomly mismatched pairs of speech envelopes-EEG response trials (Combrisson & Jerbi, 2015). As a result, in both HC and CI children, we observed significant speech envelope neural tracking (see Figure 2A). Specifically, in HC children, auditory responses were characterized by a prominent positivity between 0 and 110 ms time lags (all $p_{FDR} < 0.05$; peak TRF = 0.065, SE = 0.008; d = 1.24, 95th confidence interval (CI95) = 0.78 – 1.64), and second negativity between 200 and 320 ms time lags (all $p_{FDR} < 0.05$; peak TRF = -0.052, SE = 0.009; d = -0.86, CI95 = -1.21 – -0.39). In the CI group, the first positive response emerged at time lags between 20 and 270 ms (all $p_{FDR} < 0.05$; peak TRF = 0.090, SE = 0.010; d = 1.92, CI95 = 1.43 – 2.49), and then the negativity occurred after 390 ms time lag (all $p_{FDR} < 0.05$; peak TRF = -0.054, SE = 0.007; d = -0.93, CI95 = -1.44 – -0.48). Results highlighted that in both groups of children (i.e., HC and CI), the neural tracking could be robustly measured with twelve minutes of natural speech; the activity was characterized by two main phases of short and long timescales of brain-speech tracking occurring within 600 ms.

Once the existence of the auditory TRF was verified in both groups, we investigated whether it was possible to measure a developmental trajectory of the neural tracking of speech envelope. We reasoned that, with age, neural tracking would become more efficient. Therefore, the TRF, representing the synchronization between the neural signals and the continuous speech, would become less spread over time as typically observed in developmental ERP and TRF studies (e.g., Barriga-Paulino et al., 2017; Jessen et al., 2019) and as suggested by energy landscape analysis methods applied to EEG signals (e.g., Watanabe, 2021; Watanabe et al., 2019). In turn, we hypothesized that with increasing age, the amplitude of speech envelope TRF would be condensed in fewer time lags. That is, TRF would have a higher variance of values over time, as there would be more time lags with no substantial neural tracking (thus, the signal's sparsity would increase with age). To test this, we computed at the single participant level, the Global Field Power (GFP; Lehmann &

Skrandies, 1980) of the TRFs (i.e., GFP-TRF) for a better estimate of the dynamic of the signal and to avoid a space-dependent index (Michel et al., 2009). For each GPF-TRF between -100 and 600 ms, we estimated the marginal moments, i.e., variance, mean, kurtosis, and skew, to characterize the data distributions. Finally, we tested with a linear model whether the estimated marginal moments of the GFP-TRF were associated with children's age (see the Encoding model (TRF) section in the Quantification and statistical analysis). In the HC group a clear developmental pattern emerged, highlighting an association between neural tracking signal and age, with an increase of variance (signal sparsity) and a decrease of mean (adjusted $R^2 = 0.62$, $F_{(4,32)} = 15.7$, $p < 0.001$; variance: $\beta = 3.27$, SE = 0.95, $p = 0.002$; mean: $\beta = -5.89$, SE = 0.91, $p < 0.001$, see Figure 2C upper panel). A similar pattern was observed in CI children (adjusted $R^2 = 0.42$, $F_{(4,27)} = 6.6$, $p < 0.001$; variance: $\beta = 2.08$, SE = 0.68, $p = 0.005$; mean: $\beta = -3.11$, SE = 0.65, $p < 0.001$, see Figure 2C lower panel).

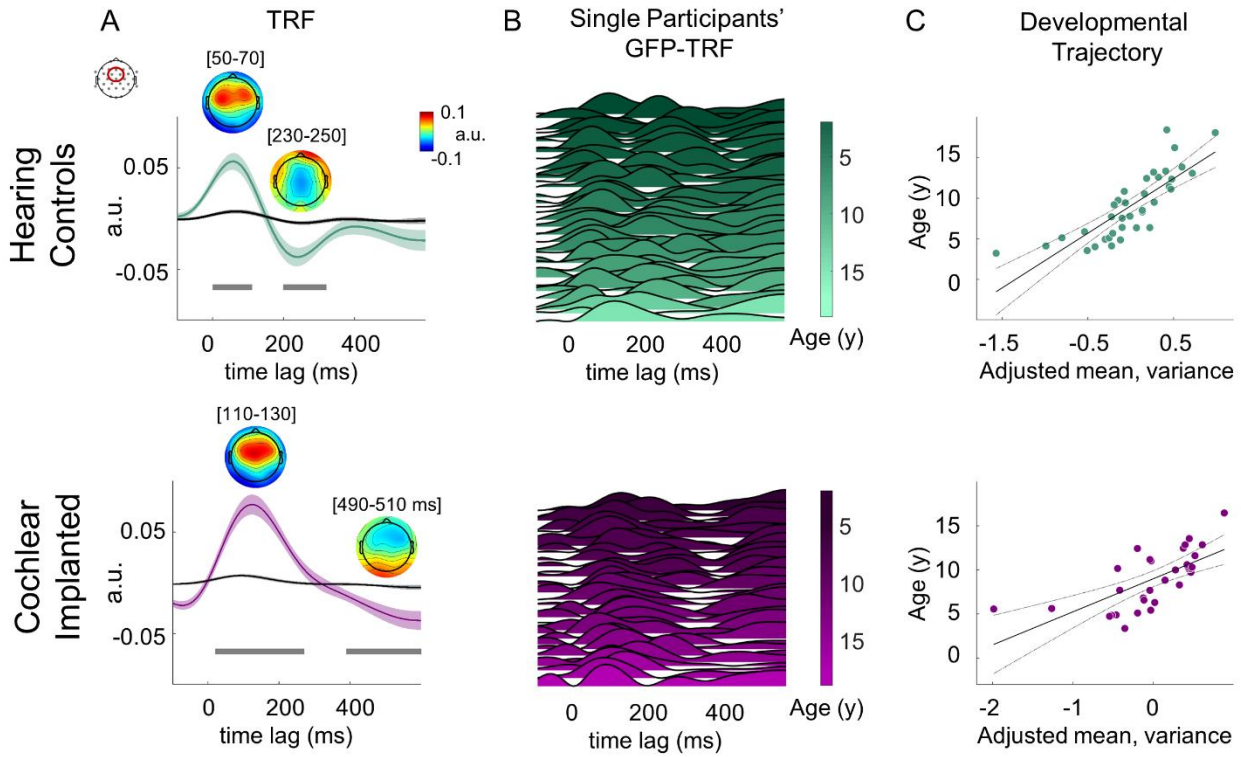

Figure 2. Neural tracking of speech in HC and CI groups. (A) Grand average speech envelope TRFs (olive color HC, magenta color CI) with topographies to represent peak distribution over the scalp and grand average null-TRFs (grey color) at frontocentral electrodes (Cz, Fz, FC1, and FC2) between -100 and 600 ms time lags. Shaded areas represent SE of the mean. Grey horizontal bars indicate time lags (between 0 and 600 ms)

at which speech envelope TRFs significantly differed from the null-TRF (t-test, FDR corrected $p_{FDR} < 0.05$). (B) Single participants' Global Field Power of their speech envelope TRF (GFP-TRF, normalization was performed for visualization purposes) sorted by age; HC group in the upper panel, and CI group in the lower panel. (C) A partial regression plot of the linear regression model shows the combined contribution of variance and mean of the GFP-TRF in predicting children's age in both HC (upper panel) and CI participants (lower panel).

How postnatal auditory experience affects neural tracking of speech

Once the existence and development of the auditory response function were assessed in both HC and CI children, we compared the TRFs across groups.

First, spatiotemporal profiles of TRFs were compared between the CI and the HC groups with a cluster-based permutation test performed at the whole brain level (across all sensors) and comprising the TRFs at every time lag between 0 and 600 ms. Results revealed a significant difference between CI and HC groups ($p_{clust} < 0.05$), with the largest significant effect at time lags between 110 and 290 ms in a large frontocentral cluster of sensors (d = -1.69, CI95 = -2.14 – -1.25, see Figure 3A). The latency of the first peak of the neural tracking (i.e., GFP-TRF P1) in the CI group was substantially delayed compared to the HC group ($t_{(67)} = -2.99$; p = 0.004, HC mean = 86.5 ms, SE = 8.7; CI mean = 119.1 ms, SE = 6.0; d = -0.71, CI95 = -1.33 – -0.11, see Figure 3B). This result suggested that speech neural tracking at a short timescale is delayed in cochlear implanted children.

Next, we investigated the specific role of auditory experience in the first year of life. The two groups of CI children, with congenital deafness (CD) and acquired deafness (AD), were contrasted to investigate the existence of a perinatal sensitive period in which auditory input must be present for the development of neural tracking of speech envelope (note in both CD and AD groups the TRF exceeded the null-TRF at frontocentral sensors, and for both the neural tracking was delayed compared to HC; see Supplementary materials and Figure S3A). The data revealed a clear overlap between the TRFs of the two groups (see Figure 3C). No difference emerged between CD and AD (cluster-based permutation test performed across all sensors between 0 and 600 ms, no clusters were found at $p_{clust} < 0.05$), unveiling that neural tracking of speech envelope was not affected by the lack of hearing input within the first year of life. A direct comparison between the first peak (GFP-TRF P1) latency of CD and AD groups revealed no difference (CD vs. AD: $t_{(30)} = 0.36$, p = 0.722; CD mean = 121.3 ms, SE = 10.4; AD mean = 116.9 ms, SE = 6.3, see also Figure S3B).

We then investigated whether the age at which auditory input was restored with cochlear implantation could account for the delayed neural tracking observed in CI users. We tested the association between the latencies of the neural tracking first peak (GFP-TRF P1) and the age at which CI children received their first implant (given the lack of difference between CD and AD groups, the correlation was run across all CI users irrespective of their deafness onset). Coherently with previous ERP findings (e.g., Sharma et al., 2005), results indicated that the later the implantation occurred, the more delayed the neural tracking (R^2 = 0.115, $F_{(1,30)}$ = 5.04, p = 0.032, β = 0.38, CI95 = 0.04 – 0.72, see Supplementary materials and Figure S4A). However, given the high variability of the neural tracking first peak (GFP-TRF P1) for the earliest implanted children (see Figure 3D), we further investigated this association by fitting multiple linear functions to identify discontinuities (the multivariate adaptive regression splines method, Friedman, 1991). We found a significant regression model with two basis functions and a discontinuity knot at 21 months (R^2 = 0.173; cross-validated R^2_{GCV}: 0.06, p = 0.036, hinge function max (0, x1 -21) β = 0.45, CI95= 0.44 – 1.05, see Quantification and statistical analysis for more details in the section Encoding model (TRF)), revealing that the positive relationship between implantation age and the latency of the neural tracking clearly emerged from 21 months onward.

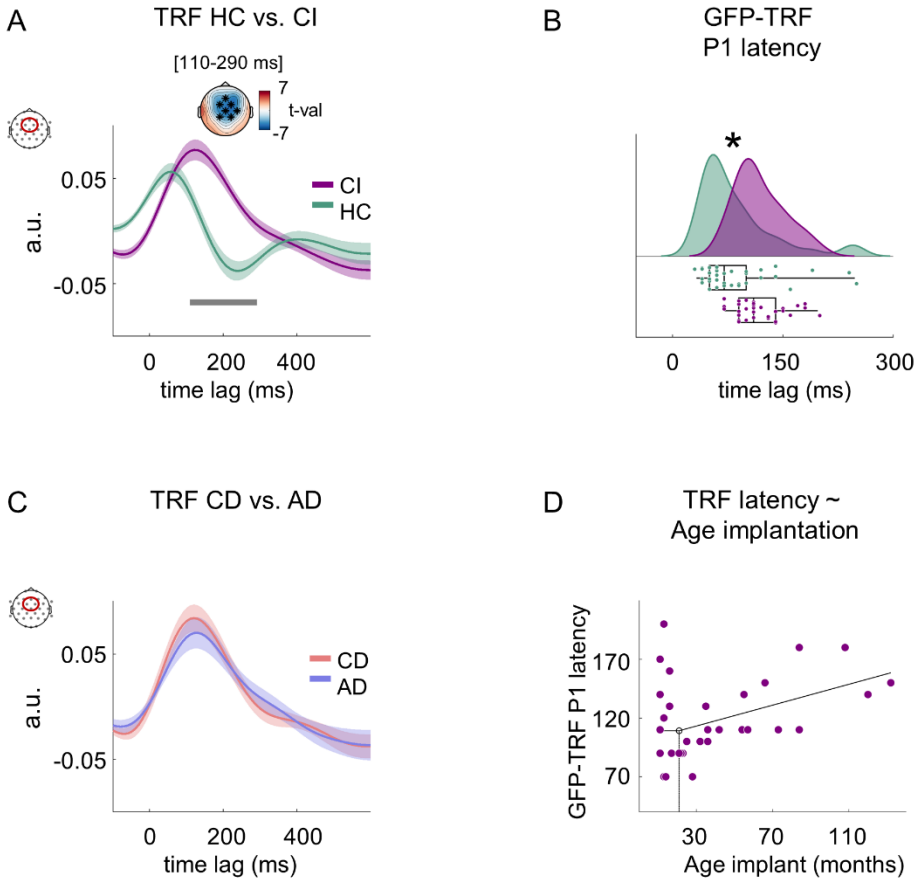

Figure 3. The impact of auditory deprivation on neural tracking of speech. (A) The speech envelope TRF at the frontocentral sensors between -100 and 600 ms time lags for HC (olive) and CI (magenta) groups. The topography shows the statistical difference between TRFs in HC and CI children between 110 and 290 ms; significant sensors at $p_{clust} < 0.05$ are highlighted with black asterisks. (B) The first peak latency (GFP-TRF P1) was delayed in CI compared to HC ($t_{(67)} = -2.99$; $p = 0.004$). (C) Speech envelope TRF at the central sensors between -100 and 600 ms time lags for CD (red) and AD (blue) groups. For all TRFs, the continuous line represents the group mean and the shaded area the SE. The data of CD and AD overlapped (no difference in the cluster-based test), suggesting that the auditory experience in the first year of life does not affect the speech envelope neural tracking in CI children. (D) Piecewise-linear regression between the age at which CI children received the first implant (auditory restoration) and the neural tracking first peak (GFP-TRF P1), highlighting that a positive relationship starts from 21 months of age.

Neural tracking of speech and comprehension

To assess speech comprehension, we investigated the outcome of the behavioral questionnaire, which comprised questions concerning story details (see Behavioral section in the Quantification and statistical analysis). Results revealed impaired scores in CI children

compared to HC (one-way ANOVA: $F_{(2,66)} = 20.72$, $p < 0.001$). Both CD and AD groups showed significantly lower accuracy with respect to HC children (HC: mean 85.70 % accuracy ± SE 2.50; CD: mean 56.77 % accuracy ± SE 5.72; AD: mean 61.98 % accuracy ± SE 3.61). Post-hoc Bonferroni corrected t-tests confirmed significant differences between HC and both CI groups (HC vs. CD $p_{Bonf} < 0.001$, d = 1.60, CI95 = 0.93 – 2.31; HC vs. AD $p_{Bonf} < 0.001$, d = 1.56, CI95 = 0.83 – 2.34; see Figure 4A), while no difference emerged between the two CI groups ($p_{Bonf} = 1.00$). Given the observed neural tracking delay in the CI group, we investigated whether the behavioral performance was associated with the GFP-TRF P1 latency. No significant correlation emerged in CI ($r_{(30)} = 0.020$, $p = 0.912$), nor in HC ($r_{(35)} = 0.045$, $p = 0.791$). This suggested that the delayed neutral tracking alone could not be responsible for lower speech comprehension scores in CI participants.

While the TRF delay was not crucial for comprehension, a difference between the two groups' speech envelope TRF dynamic was evident even when accounting for the neural tracking delay (see Figure 3A and 4B). To test this hypothesis, we shifted the CI's data, aligning the first TRF peak (P1, at the frontocentral cluster) of the CI group to the homolog peak of the HC group (see Figure 4B; see Supplementary materials). The cluster-based permutation test (across all sensors and time lags) between HC's TRFs and CI's temporally shifted TRFs revealed reduced neural tracking in CI individuals compared to HC at time lags between 130 and 260 ms ($p_{clust} < 0.05$; d = -1.11, CI95 = -1.50 – -0.69). At this time range, the second major phase of the auditory temporal response function clearly emerged only in the HC group (see the significant difference with the null-TRF in Figure 2A). Conversely, the first TRF phase (P1) did not differ between HC and CI, suggesting unaltered TRF magnitude at a short timescale. Following these observations, we explored whether the degree of neural tracking (the TRF magnitude) would be associated with children's comprehension scores (Vanthrnhout et al., 2018) by performing a series of partial correlations (at all sensors) between point-by-point TRF values and response accuracy accounting for the role of age (see Figure 4C). In HC, the association was significant between 150 and 250 ms (averaging across all sensors, $p_{FDR} < 0.05$, mean $r^2 = 0.065$, SE = 0.010). Instead, in the CI group, no significant association emerged for the same analysis (see Figure 4D). Overall, these findings revealed that in HC, the second phase magnitude (between 150 and 250 ms), indicating longer timescale neural tracking, was associated with their comprehension scores. Precisely at this latency (note, once accounting for their TRF delay), the CI group had reduced neural tracking. This altered dynamic could thus account for the comprehension deficit in the CI group.

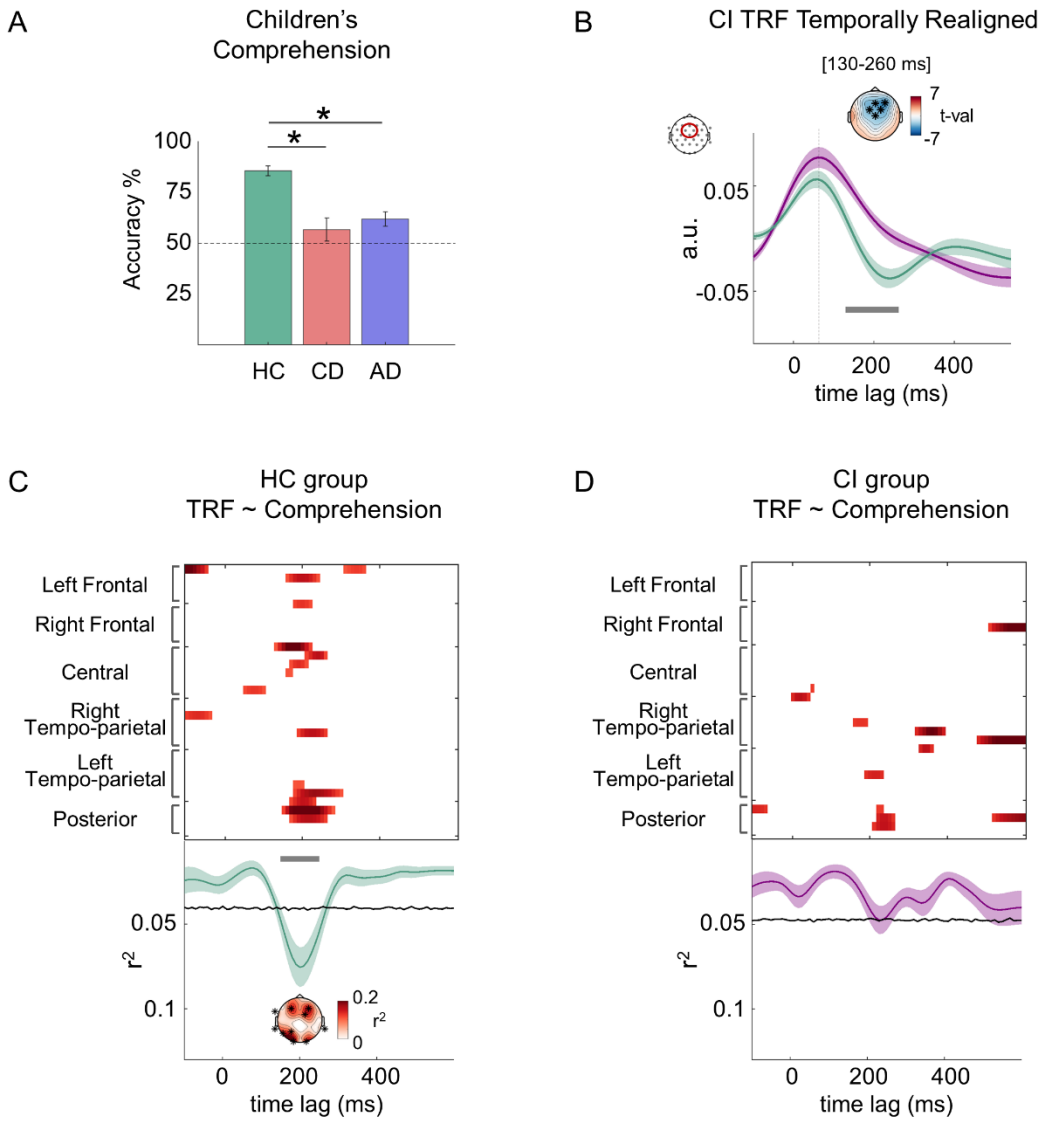

Figure 4. The relationship between neural tracking and speech comprehension. (A) Comprehension scores in HC, CD, and AD groups. Significant differences emerged between HC and both CD and AD groups. (B) The plot shows the TRF for HC and CI at frontocentral sensors. The first peak of the CI's TRF was temporally realigned to the first peak of the HC group to account for the CI's neural tracking delay. The topography shows that, despite accounting for the delay, significant differences between the two groups emerged selectively in the time window [130 – 260 ms]; significant sensors are highlighted with black asterisks. (C) The plot shows the r^2 values with $p < 0.05$ for the correlation between the TRF magnitude and children's comprehension scores for all sensors and time lags for the HC group. Below, the r^2 averaged across sensors is plotted as a function of time, and the shadow represents the SE. The black line represents the boundary of the null effect, that is, the 95th percentile of the r^2 null distribution (note that the y-axis is reversed; higher averaged r^2 values emerged below the black line). A significant difference emerged between 150 and 250 ms ($p_{FDR} < 0.05$). The topography shows the r^2 at the peak (200 ms), and black asterisks highlight significant sensors. (D) The plot shows the r^2 values with $p < 0.05$ of the correlation between the TRF amplitude and children's comprehension score for all sensors and time lags for the

CI group. Below, the r^2 averaged across sensors is plotted as a function of time, and the shadow represents the SE. No significant results emerged in the CI group ($p_{FDR} > 0.05$).

DISCUSSION

Despite the period of auditory deprivation and cochlear implants conveying only partial input to the brain (Winn & Nelson, 2021; Pavani & Bottari, 2022), our data showed the existence of neural tracking of speech envelope in both CI groups (i.e., CD and AD). Neural tracking was unaffected by the presence or absence of auditory experience in the first year of life.

No perinatal sensitive period for the neural tracking of the speech envelope

The TRF's latency and magnitude remarkably overlapped in individuals with congenital deafness (CD) and acquired deafness (AD). CI children of both groups were homogenous in terms of age, gender, and experience with the implants (see Figure 1) but differed for their bilateral profound deafness onset, which was ensured by objective examinations (screening at birth, auditory brainstem measurements or auditory thresholds measured in free field). Contrasting these two groups of children provided the unique opportunity to test whether functional acoustic input within the first year of life is necessary to develop neural tracking of speech envelope. The data clearly revealed the absence of a perinatal sensitive period for the development of this brain function.

The first phase of life after birth is crucial for developing several acoustic functions (see Sanes & Woolley, 2011). The processing of basic properties of language, such as phonemes and syntax, appears to rely upon, partially sequential, critical periods occurring within the first years of life (Werker & Hensch, 2015; Friedmann & Rusou, 2015). However, certain computations are already available at birth. In newborns, neural tracking in delta and theta oscillations have been found to be modulated by the presentation of different languages, suggesting that slow neural oscillations represent a pivotal mechanism for speech processing since the earliest phases of life (Ortiz Barajas et al., 2023). Our findings expand these observations by providing evidence that the presence or lack of auditory experience within the first year of life does not affect the neural tracking of speech envelope below 8 Hz. Ultimately, this evidence advocates that at birth the brain is endowed with strong biological constraints for the tracking of speech that are strikingly resilient to sensory deprivation. Notwithstanding, a developmental trajectory of the TRF properties (i.e.,

variance and mean amplitude) emerged in both hearing control and children using CIs. This further supports the sensitivity of neural tracking to capture changes along typical and atypical development (see also Ortiz Barajas et al., 2021; Pérez-Navarro et al., 2023).

Delayed neural tracking in children with CIs

The neural tracking of speech envelope was markedly delayed in the CI children compared to HC individuals (of about 60 ms). This is coherent with previous observations employing simple speech units (e.g., Eggermont & Ponton, 2003; Sharma et al., 2005). Moreover, in adults, the severity of hearing impairment is positively associated with neural tracking delay (Gillis et al., 2021). Coherently, difficult acoustic listening conditions, such as when speech is presented in noise, are known to delay auditory responses to speech stimuli measured with ERPs or neural tracking in both children and adults (Billings et al., 2011; Ding & Simon, 2013; Gustafson et al., 2019; Yasmin et al., 2023). Taken together, these findings suggest that delayed neural tracking observed here in both CD and AD groups could be influenced by reduced efficiency of continuous speech processing.

Notably, the TRF delay measured here was associated with the age of cochlear implantation. Coherently with seminal electrophysiological studies employing short-lived speech sounds (e.g., “ba”; Sharma et al., 2002a; 2002b; 2005; 2015), our results highlighted the pivotal role of when auditory restoration takes place for the development of speech tracking. The later the child receives the implant, the more delayed the first TRF peak (P1), representing auditory processing of language at a short timescale (within the first 150 ms of brain-speech tracking). Longitudinal studies on cochlear implanted children established a sensitive period for the development of basic auditory responses to syllables within the first 3.5 – 4.0 years of age (Kral & Sharma, 2012; Sharma & Campbell, 2011), strongly advocating for early implantation. In the case of children implanted at an early age (< 3.5 years), the P1 latency of the ERPs consistently fell within the 95% confidence interval of typical development. Conversely, children who underwent implantation after the age of 7 years never reached the typical latency range of early auditory responses (Sharma et al., 2005; 2007). Here, we found further support for these observations, showing experience-dependent effects associated with the timing of implantation, and revealed that they also seem to emerge for early, sensory-based components of the neural tracking of continuous speech envelope. However, in our data, the benefit of early implantation was evident only from 21 months of age. The neural tracking latency (GFP-TRF P1) in children implanted

before this age was highly variable (from 70 ms to 200 ms), and it was not explained by the age of implantation, nor by other factors like chronological age, experience with the implant and age at which hearing aids were provided before implantation (see Supplementary materials). This result represents a clear limitation for understanding relationships between speech neural tracking and clinical characteristics of children implanted in earliest development. The etiology of children implanted before 21 months of age comprised different profiles (e.g., GJB2 connexin 26, congenital CMV, Waanderburg, and perinatal complications associated with prematurity). Most of them had bilateral profound congenital deafness onset (ten out of eleven, see Supplementary materials and Figure S4B). Noteworthy, in the case of congenital deafness onset, the clinical practice does not assess whether auditory input was available in the intrauterine life since auditory screening is performed only after birth (Lieu et al., 2020). Intriguingly, recent evidence suggests that fetal linguistic experience shapes neural synchronization with language measured at birth (Mariani et al., 2023).

Neural tracking dynamics uncover higher-order deficits of speech processing in children with CIs

In HC, the neural tracking magnitude occurring at time lags between 150 and 250 ms was associated with higher comprehension scores. This result is consistent with recent evidence suggesting that the magnitude of neural tracking with a similar timescale is associated with comprehension of continuous speech (Etard & Reichenbach, 2019). Coherently, studies using noise-vocoded speech demonstrated that neural tracking occurring at about 200 ms was strongly reduced when the speech was degraded and comprehension impaired (Chen et al., 2023). Even accounting for the TRF delay, by time aligning CI's neural data with hearing control one, CI individuals had hampered tracking around this longer timescale (see Figure 4B). Noteworthy, after accounting for this delay, the sensory component of the speech neural tracking (P1) did not differ between HC and CI, suggesting unaltered neural tracking magnitude at a short timescale. These observations suggested selective higher-order deficits of speech processing in cochlear implanted children. Accordingly, the behavioral performance of both CD and AD groups was markedly impaired. These findings provide the first evidence for the identification of a possible biomarker associated with natural speech comprehension in CI individuals. TRF magnitude at this latency range could be employed to verify their speech understanding when behavioral measures are difficult to

acquire, as in the case of infants, and to estimate the development of speech processing after implantation.

In conclusion, the data clearly highlighted that speech envelope tracking leverages a robust biological predisposition, resilient to a period of auditory deprivation from birth. However, results pointed toward the crucial role of early auditory restoration in mitigating atypical auditory development and possibly ameliorating speech processing efficiency. Finally, we substantiated in CI children a clear vulnerability of higher hierarchical levels of speech-envelope tracking, which are associated with speech comprehension. Overall, by employing speech envelope tracking and investigating individuals with atypical access to sensory input, the present findings support a model of hierarchical and interdependent processing levels for continuous speech elaboration, which are endowed with strong biological constraints. Despite its simple nature, neural tracking of speech envelope is a promising method to assess auditory and speech functions in developing cochlear implanted individuals and could help explain the variability in outcomes that typically characterize this population.

ACKNOWLEDGMENTS

Fundings: PRIN 20177894ZH

AUTHOR CONTRIBUTIONS: Conceptualization D.B., A.F.; Formal Analysis A.F., D.B., G.H., E.B.; Investigation A.F., M.F., E.N., A.M., D.B.; Data Curation A.F., D.B.; Writing – Original Draft D.B., A.F.; Writing – Review & Editing D.B., A.F., G.H., F.P., E.N., E.O., E.R., M.F., M.B., B.B., A.M., E.B.; Visualization D.B., A.F., G.H.; Supervision D.B.; Project Administration D.B.; Funding Acquisition D.B., E.R., F.P.

CONFLICT OF INTEREST STATEMENT: The authors declare no competing financial interests.

METHODS

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Sample size estimation

We estimated the minimum sample size needed to measure a clear P1 in the auditory response function. We expected the P1 to emerge between 0 and 150 ms lags (Fiedler et al., 2019) and in the frontocentral sensors (Fz, Cz, FC1, FC2), and it should be higher than in the null-TRF. Using the data of 10 pilot HC subjects, we computed the mean and SD acoustic TRF (mean = 0.044, SD = 0.036) and null-TRF (mean = 0.005, SD = 0.007). Through simulations suited for cluster-based permutation tests (500 randomizations; Wang & Zhang, 2021), we estimated a minimum sample size of 16 subjects to reach a power of 0.95 (lower threshold).

Participants

A total of 81 children participated in the study. They were categorized according to their hearing status: cochlear implanted children (CI) or hearing control children (HC).

All CI children received cochlear implantation at least six months before the EEG acquisition (Sharma et al., 2002b) to ensure a stable implant functioning and that the auditory system had accumulated some auditory experience (see Table S1 for more detailed CI participants' information). A total of 44 children with cochlear implants were recruited at the Meyer Hospital of Florence (Italy) and the IRCCS Materno Infantile Burlo Garofolo of Trieste (Italy). A few CI participants were excluded due to different reasons: a two-year-old child was discarded because they could not comply with the experimental session; two children were excluded because their IQ was below the age standard; iii) one child was reimplanted after many years from the first implantation following an ear infection; iv) one child was discarded due to the bad quality of the EEG signal (15 electrodes were detected as bad channels). The remaining CI participants (N = 39) were classified according to their deafness onset and thus their access to auditory input in early development: Children with profound bilateral congenital deafness (congenital deafness CD) and children who acquired profound bilateral deafness during the development (acquired deafness (AD).

Congenital deafness was ensured by the following criteria: (a) having failed to pass the neonatal screening for otoacoustic emissions, which in Italy is performed before the hospital discharge (typically < 1 week after birth); (b) receiving a diagnosis of profound bilateral deafness (hearing thresholds ≥ 90 dB in both ears) following the objective evaluation of auditory brain-stem responses (ABR) within two months of age (mean age: 37.87 days; range 21 – 60 days).

In contrast, AD children had at least some auditory experiences in early development (i.e., a minimum of 12 months). To ensure the presence of such auditory experience, we combined the following clinical information: (a) whether they passed otoacoustic emissions neonatal screening at least with one ear; (b) an ABR indicating normal hearing before the diagnosis of deafness or a diagnosis of deafness that was not profound bilaterally (e.g., moderate deafness at least in one ear) made by ABR or behavioral test (hearing thresholds < 90 dB in at least one ear); (c) family report indicating residual hearing for the first period of life. All AD patients received a diagnosis of profound bilateral deafness before cochlear implantation (hearing thresholds ≥ 90 dB in both ears for children under 2 years old and > 75 dB for children older than 2 years old; Berrettini et al., 2011, age range: 12 – 107 months. In case the exact date of this test was not available, we estimated the onset of profound bilateral deafness one month before the date of the first cochlear implantation; N=6).

It was not possible to ensure whether profound bilateral deafness was congenital or acquired in seven children, and thus, their data were excluded. The final sample of cochlear implanted participants comprised thirty-two children: sixteen CD children (mean age = 8.81 years; SD = 3.52, eight females and eight males) and sixteen AD children (mean age = 9.17 years; SD = 3.15, nine females and seven males). As expected, the age of the diagnosis of profound bilateral deafness differed significantly between the two groups ($t_{(15.02)} = -7.35$, $p < 0.001$, $d = -2.53$, CI95 = -3.58 – -1.53). Importantly, no difference emerged between CD and AD in their experience with the implant ($t_{(30)} = 1.53$; $p = 0.136$). The mean age at cochlear implantation was 27.7 months (range: 11 – 132) for the CD group and 53.6 months (range: 17 – 120) for the AD group.

A group of age- and gender-matched hearing controls children (HC) was recruited as control group (N = 37; mean age = 9.04 years; SD = 4.10, seventeen females and twenty males). No significant difference emerged between the three groups neither for age ($F_{(2,66)} = 0.037$; $p = 0.964$) nor for gender ($\chi^2_{(2)} = 0.479$; $p = 0.787$). HC children were recruited among public schools in Lucca (Italy) and at the MultiLab of Milano-Bicocca University (Italy). None of the children who participated in the study had any additional sensory deficits or neurological disorders (medical records and/or family reports). All participants were oralists; their first language (L1) was Italian (one CD, two AD, and three HC participants were bilingual).

The study was approved by the local Ethical Committee (Comitato Etico Regionale per la sperimentazione clinica della Regione Toscana: Numero registro 34/2020, and Comitato etico congiunto per la ricerca espressione di parere delibera n. 17/2020). Before

participating in the experiment, written informed consent was signed by the participants' parents and by the children themselves if they were older than seven years of age. The experimental protocol adhered to the principles of the Declaration of Helsinki (2013).

METHOD DETAILS

Stimuli and experimental procedure

Speech stimuli

The speech stimuli were 3-minute length stories read by a native Italian speaker. We chose different stories according to the children's age in order to provide each participant with speech materials suitable for their age. Three different age ranges [3 to 6], [7 to 10], and [11 to 15] years old – were defined according to Italian school cycles. For each age group, we selected ten stories from popular Italian books suitable for that age range. Stories were read by a person whose diction had been formally trained and were recorded in a sound attenuated chamber (BOXY, B-Beng s.r.l., Italy) with an iPhone 7 (camera with 12MP, video resolution in HD, 720p with 30fps, at a sampling frequency of 48000 kHz) and an external condenser microphone (YC-LM10 II, Yichuang). All audio recordings were imported in iMovie (version 10.3.1), *the noise reduction* at 100% was applied, and each file was cut to have 2 seconds of silence before the story's title and a few seconds of silence at the end of the story. Then, the audio was imported into Audacity® (version 2.4.2, https://www.audacityteam.org/ using *ffmpeg* and *lame* functions to isolate the audio from the video). The audio files imported in Audacity were preprocessed with the following steps: they were converted from stereo to mono, amplified (default value in Audacity and avoid clipping were selected), down-sampled to 44100 Hz, and set to a 32-bit sample. Finally, we imported all audio files in Matlab to perform RMS equalization to achieve an equal loudness for all the stimuli (RMS value = 0.03).

Speech stimuli were presented to participants using Psychopy® software (PsychoPy3, v2020.1.3), and the sound was delivered by a single front-facing loudspeaker (Bose Companion® Series III multimedia speaker system, country, USA) placed in front of the participants behind the computer screen, approximately 70 cm distance from their heads. Stimuli were delivered at ~80 dB, measured at the place of the loudspeaker (Meterk MK09 Sound Level Meter).

We reasoned that to validly study continuous speech processing, participants should receive acoustic stimulation that represented their typical everyday input. Thus, we decided not to alter the sounds (e.g., with vocoding procedures) presented to hearing control children to simulate what the cochlear implanted children could hear.

Task and experimental procedure

Participants were asked to listen carefully to the stories while looking at a cross displayed at the center of the screen. At the beginning of each story, a white cross was displayed, and after two seconds of silence, the story's title was presented, and then the story began. The cross was always presented in the middle of the screen and its color was randomly generated and changed every 1 to 20 seconds to keep the children's gaze attracted throughout the story. One experimenter was always sitting beside the child and checked that they maintained eye contact with the screen throughout the stimuli presentation. At the end of each story, children were asked to answer an ad-hoc questionnaire. Each questionnaire comprised two comprehension questions, 2-alternative-forced-choice; most of them were yes-no questions (e.g., "Did Lorenza like to travel?") and few alternative answer questions for the young children (e.g., "What did the kitten fairy give to Lorenza?", possible answer in the picture "ball of wool" or "doll"). For younger children (age range between 3 – 6 years), questions were performed verbally by the narrator's voice, and the alternative answers were supported by drawings representing the content. For older children (> 7 years old), questions were presented via text on the screen and read by the experimenter. Each participant was presented with four stories, randomly drawn among the ten selected for their age range to obtain more generalizable results (eleven HC participants and ten CI listened to only three stories). Narrated stories were unknown to most of the participants (three HC children and two CI children had previously heard one story each). During the whole duration of the experiment, their EEG activity was recorded.

EEG recording and preprocessing

EEG data were collected continuously during the entire experimental session, using a Brain Products system (ActiCHampPlus) with elastic caps (Easy cap Standard 32Ch actiCAP snap) suited for children and having 32 active channels (500 Hz sampling rate). Note that for CI participants, electrodes placed very close to the magneto of the cochlear implants

were disconnected (mean number of disconnected electrodes = 3.50, SD = 1.44; range 1 – 7). Continuous EEG data acquired during each story presentation were concatenated and offline preprocessed using the EEGLAB toolbox (Delorme & Makeig, 2004), implementing a validated preprocessing pipeline (Stropahl et al., 2018; Bottari et al., 2020).

Prototypical artifact cleaning. Continuous EEG recordings were low-pass filtered (cut-off = 40Hz; window type = Hanning; filter order = 50), downsampled to 250 Hz to reduce the computational time, and high-pass filtered (cut-off =1Hz; window type = Hanning; filter order = 500). The filtered downsampled data were segmented into consecutive 1-second epochs. Noisy segments were removed using joint probability (threshold across all channels = 3 SD; Delorme et al., 2007). To remove stereotypical artifacts (e.g., blink and eye movement) data were submitted to Independent Component Analysis (ICA, based on the extended Infomax, Bell and Sejnowski, 1995; Jung et al., 2000b, 2000a). The computed ICA weights were applied to the continuous *raw* (unfiltered) data (Stropahl et al., 2018; Bottari et al., 2020). Components associated with blinks and eye movement artifacts were identified using CORRMAP, a semiautomatic procedure in which a prototypical topography for each type of artifact (i.e., eye movement and blink) is selected. All the components that correlate more than 80% with the template were removed. (Viola et al., 2009). For CI participants, the mean number of removed components was 2.09 ± 0.39 SD, and for HC 2.00 ± 0.00 SD.

CI artifact cleaning. EEG studies involving CI users have to deal with electrical artifacts from the CI. Due to the specific way we model the EEG data as a function of continuous stimulus feature, we could not use previous approaches (e.g., CIAC plug-in for EEGLAB). Thus, we developed a novel method to clean electrical artifacts from EEG data caused by the implant in CI participants. This study employed temporal response functions (TRF) to investigate how continuous speech processing develops as a function of the auditory experience (typical or congenital or acquired deafness) and hearing status (hearing controls and cochlear implanted children). This analysis involves associating dynamic changes in speech features (e.g., envelope) to changes in the EEG data, at specific time lags. Due to their function, we expected CI artifacts to occur at around 0 ms time lag (Deprez et al., 2017; Somers et al., 2018). To search for such activity in the data, we performed the following steps. Briefly, we decompose the EEG recordings in components with the purpose of separating physiological and noise sources. Then, we applied the TRF approach to each component to obtain a set of component-TRFs. Finally, with the aim of identifying artifacts with about zero lag, and by using a minimal set of parameters extracted from TRFs in HC,

we identified and discarded the artifactual components and reconstructed back EEG recordings.

In detail, data cleaned by their stereotypical artifacts (blinks and eye movements) were reprocessed using the second order blind identification (SOBI) algorithm to identify independent components based on second-order statistics, making it suitable to separate temporally correlated signals, and maximizing activity related to CI artifacts (Paul et al., 2020). The same procedure explained above for ICA (filtering and rejection of noisy segments) was applied prior to SOBI estimation. Specifically, continuous cleaned EEG recordings were low-pass filtered (cut-off = 40 Hz, window type = Hanning, filter order = 50), downsampled to 250 Hz to reduce the computational time, and high-pass filtered (cut-off = 1 Hz, window type = Hanning, filter order = 500). The filtered downsampled data were segmented into consecutive 1-second epochs. Noisy segments were removed using joint probability (threshold across all channels = 3 SD; Delorme et al., 2007). Then, SOBI was computed, and the SOBI weights were applied to the original (unfiltered) data cleaned by their stereotypical artifacts. Data were downsampled to 250 Hz, filtered between 2 – 8 Hz, epoched from 6 seconds to 2.5 minutes for each story, downsampled again to 100 Hz in order to match the sampling rate of the acoustic features for TRF estimation, and segmented into 50-second trials (see the following paragraph “Filtering, Removing Bad Channels and Epoching” for a more detailed explanation of these steps). TRF model was applied to each SOBI component across -100 – 600 ms time lags (same parameters used for the envelope TRF encoding model; see below). Thus, we obtained a TRF of each SOBI component for each subject.

Subsequently, we implemented an algorithm to remove the components classified as containing mainly CIs artifact signals. We started defining the criteria to reject components using the data of the HC participants (who have no CI) to identify parameters values at a 5% false positive rate. First, we normalized SOBI components by extracting the absolute value of each timepoint and scaling it with the maximum intensity of the series. Then, we modeled normalized SOBI components by fitting in each time series a set of Gaussian responses: (i) a single gaussian was restricted to peak between -100 and 0 ms (that should contain mainly artifactual activity); (ii) up to five gaussians were limited to peak between 50 and 500 ms (that should include mainly neural activity). Two criteria were used to decide whether a SOBI contained artifacts: (a) the *ratio* between R^2 of the gaussian fitted before zero and R^2 gaussians fitted after zero; (b) the *beta* of the gaussian fitted before zero. The rationale was that we expected to find most implant activity around zero, while the rest (after

a physiological delay) should be considered brain activity. We mapped in HC our R^2 *ratio* and *beta* onto a plane to identify a decision boundary that isolates portions of the parameters space with both high R^2 *ratio* and *beta* and retains a false positive rate of 5%. This decision boundary was applied to CI data using the same procedure described above. Artefactual components with high R^2 *ratio* and *beta* were removed (number of components removed per participant mean ± SD: CI 2.47 ± 2.19, CD: 2.69 ± 2.39, and AD: 2.25 ± 2.02).

Filtering, removing bad channels, and epoching. After the removal of SOBI components associated with CI artifacts, (unfiltered) data cleaned from artifacts of CI and HC groups were then low-pass filtered (cut-off = 40 Hz; window type = Hanning; filter order = 50), downsampled to 250 Hz, and high-pass filtered (cut-off = 0.1Hz, window type= Hanning; filter order = 5000). Noisy channels were identified based on the automatic bad channel detection algorithm (*clean_channels* function of *clean_data* plugin of EEGLAB; correlation threshold = 0.8 and sample size =1; all the other parameters were kept as default). Noisy channels were then interpolated using spherical spline interpolation (mean interpolated electrodes per subject ± SD, in CI participants: 1.88 ± 1.60, in HC: 2.30 ± 1.15). Disconnected channels near the magneto of the cochlear implant were also interpolated. Following interpolation, data were re-referenced to the average reference. EEG data were then filtered according to the envelope frequency of interest: between 2 and 8 Hz (high-pass filter: cut-off = 2 Hz, window type = Hanning, filter order = 250, and low-pass filter: cut-off= 8 Hz, window type = Hanning, filter order = 126) as previously performed (Mirkovic et al., 2015; O'Sullivan et al., 2015). The timing of each epoch was adjusted to +99 ms onset delay measured by the AV device (EGI). Then, preprocessed EEG data of each listened story were epoched starting from 6 seconds and lasting 2.5 minutes. The first 6 seconds of each story, including 2 seconds of silence, the story title, and the beginning of the story, were removed to avoid the stimulus onset response as much as possible (Crosse et al., 2021). Finally, epochs were downsampled to 100 Hz, concatenated, and segmented into 50-second trials, resulting in 12 trials per subject (or nine for the children in which we collected three instead of four stories). Trials were created in order to perform a cross-validation procedure in the analysis. Data were z-scored to optimize the cross-validation procedure while estimating the regularization parameter (Crosse et al., 2016).

Extraction of the speech envelope

The audio file of each listened story was loaded for each subject, and the first 6 seconds were discarded, as for the EEG signal. Then, for each story, the acoustic envelope was extracted, taking the absolute value of the Hilbert transform of the original piece of the story and applying a low-pass filter with an 8 Hz cut-off (3rd-order Butterworth filter, *filtfilt* MATLAB function).

For each subject, the speech envelope of each story was then concatenated in the same order they were presented to each participant and segmented into corresponding 50-second trials, resulting in twelve trials per subject (or nine trials for subjects who have listened to only three stories). The envelopes were downsampled to 100 Hz to match the EEG data (e.g., Mirkovic et al., 2015; O'Sullivan et al., 2015) and normalized by dividing each amplitude value by the maximum one to optimize the estimation of the regularization parameter (Crosse et al., 2016).

Estimation of TRF

The forward model. To investigate how acoustic speech envelope is encoded in the children's brain, we used a linear forward model known as temporal response function (TRF, incorporated in *mTRF* toolbox, Crosse et al., 2016). TRF can be seen as a filter that describes the mapping between ongoing stimulus features (here, envelope) and the ongoing neural response. This approach allows the prediction of previously unseen EEG responses from the stimulus feature and has been extensively used to model the neural tracking of continuous speech envelope. Mathematically, the encoding model is described by the following function:

$$r(t,n) = \sum_{\tau} \quad w(\tau,n)s(t-\tau) + \varepsilon(t,n),$$

Where t = 0, 1, … T is time, $r(t,n)$ is the neural EEG response from an individual channel n at time t, s is the stimulus feature(s) at each moment ($t - \tau$), τ is the range of time lags between s and r, $w(\tau,n)$ are the calculated regression weights over time lags (TRF), and $\varepsilon(t,n)$ is a residual response at each channel n at time t not explained by the TRF model (Crosse et al., 2016). Specifically, the TRF at each time lag (τ) represents how the unit change in the amplitude of the speech envelope would affect the EEG response τ ms later (Lalor et al., 2009).

We fitted separate TRF models at the single subject level to predict response in each of the 32 EEG channels from the acoustic feature (i.e., the envelope) using time lags from -100 to 600 ms in steps of 0.01. The TRF at -100 ms time lag represents how the amplitude change of the speech envelope would affect the EEG response 100 ms earlier, while the TRF at 600 ms time lag represents how the amplitude change of the speech envelope would affect the EEG response 600 ms later (the same is for all the time lags comprised in the time lags window). Importantly, since the TRF model is conducted separately for each channel, their interpolation during preprocessing (in case of bad channels or vicinity with the cochlear implant) did not affect the model results. To train the model, a leave-one-out cross-validation procedure was used. All trials except one were used to train the model to predict the neural response from the speech envelope, and the left-out trial was used to test the model. This procedure was performed for each trial; the prediction model for every trial was computed and then averaged together to obtain the TRF model for each channel.

Regularization parameter estimation. Importantly, the regularization parameter was estimated to avoid overfitting in the regression model obtained by the training data. Overfitting consists of fitting the noise in the data unrelated to the stimulus, thus preventing generalization to different datasets. Regularization is achieved by selecting the TRF models' optimal regularization parameter (λ). A set of ridge values ($\lambda = 10^{-3}, 10^{-2}, 10^{-1}, 1, 10, \ldots, 10^{9}, 10^{10}$) is used to compute the model for time lags from -100 to 600 ms through a leave-one-out cross-validation procedure. To determine the optimal regularization parameter (λ) for each participant, we used the mean squared error (MSE) value – averaged across trials and electrodes – between the actual and the predicted EEG responses; the λ value reaching the lowest MSE value was selected. The identified λ value for the envelope model was 10^{4}. These values emerged for most participants, and in order to generalize results, we decided to keep λ constant across all channels and participants.

Estimation of the null effect. To verify that neural tracking was greater than a null effect in all groups, we computed a null-TRF model for each participant (Combrisson & Jerbi, 2015). We permuted the 50-second pairs of trials to obtain mismatched envelope and EEG response pairs, and then, the TRFs were fitted on these randomly mismatched trials of speech envelopes-EEG responses (*mTRFpermute* function with 100 iterations; Crosse et al., 2016, 2021). Then, all these null-TRF models (one for each channel) computed across the iterations were averaged to obtain a null-TRF model that served as a control. This procedure was done separately for each participant and each channel.

QUANTIFICATION AND STATISTICAL ANALYSIS

For all analyses, the threshold level for statistical significance was set at 95% (alpha = 0.05, two tails). We referred to p_{FDR} when we performed FDR correction for multiple comparisons and to p_{clust} when we performed a cluster-based permutation test. To report the magnitude of the effect, we computed the unbiased estimate of Cohen's d; we acknowledge that the effect can be inflated since it is computed on the same sample. Its confidence interval was computed with a bootstrap method with 1000 permutations. When multiple time points and/or channels were significant, Cohen's d was computed on the mean of the time window in which the significant effect emerged and across the significant channels.

Behavioral

To assess any difference in children's comprehension, we computed the accuracy percentage (correct answers) for each participant, and we ran a univariate ANOVA with group (HC, CD, AD) as between-participant factors.

Encoding model (TRF)

Assessing the existence of the speech neural tracking (TRF) within each group. To test whether we could measure an auditory temporal response function (TRF model) within each group, we first selected a cluster of four frontocentral channels (Cz, Fz, FC1, and FC2; note that none of these channels were disconnected in CI children) typically capturing at the scalp level auditory responses with evoked potentials (for review see Steinschneider et al., 2011) and auditory response functions in children and adults (e.g., Jessen et al., 2019; Paul et al., 2020). Notably, the electrodes of this frontocentral cluster are far from cochlear implants. Then, we performed comparisons between the frontocentral TRF model and the frontocentral null-TRF by running paired t-tests within time lags [0 – 600 ms] (i.e., every 10 ms) separately for HC and CI groups and also separately for CD and AD subgroups (two-tailed, q=0.05, FRD correction, Benjamini & Yekutieli, 2001). We also performed cluster-based permutation tests (Maris & Oostenveld, 2007) in the FieldTrip toolbox (Oostenveld et al., 2011) between the TRFs and the null-TRFs within each group to confirm that the same results emerged testing the frontocentral cluster are stable also testing across all electrodes

(see Supplementary materials and figure S1). A cluster was defined along electrodes × time lags dimensions. Cluster-based permutation tests were performed at the whole brain level (across all electrodes) and time lags between 0 and 600 ms, using the Monte-Carlo method with 1000 permutations. Cluster-level statistics were calculated by taking the sum of the t-values within every cluster (minimum neighbor channel = 2; cluster alpha was set to 0.05, which was used for thresholding the sample-specific t-statistics). Identified clusters were considered significant for the permutation test at $p > 0.025$ (the probability of falsely rejecting the null hypothesis). The alpha level 0.05 was thus divided by 2 ($p = 0.025$) to account for a two-sided test (positive and negative clusters).

Investigating the developmental trajectory separately for the HC and CI groups. We also assessed whether the neural tracking of speech would follow a developmental trajectory in HC and CI groups. We reasoned that the development of speech processing would affect the sparsity of the TRF signal. Namely, we expected the TRF signal to become more sparse with age. In this case, sparsity would indicate selectivity (in the time domain) of natural activity. This would result in high-variance amplitude distributions. To this aim, we estimated marginal moments of the TRF signal as they allow us to describe the sparsity of a signal (e.g., marginal moments play a crucial role in envelope discrimination; see Lorenzi et al., 1999; Strickland & Viemeister, 1996).
We performed a linear regression model with all the z-scored marginal moments (mean, variance, kurtosis, and skewness) of the normalized Global Field Power of the TRFs (GFP-TRF) as independent variables and the children's age as dependent variable. The reason for choosing the GFP instead of selecting specific channels of interest is that this approach allows for a more objective and reference-free characterization of temporal dynamics of the global electric field (see Michel et al., 2009).

Testing neural tracking differences between groups. To test differences in the spatiotemporal profile of TRFs between groups, we performed cluster-based permutation tests with the same parameters defined above. First, we performed a cluster-based permutation using independent-sample t-statistics between HC and CI groups. Then, we performed the same cluster-based permutation test between CD and AD.

Correlation between neural tracking delay and implantation age. In order to test whether the neural tracking delay measured in the CI group was affected by age at cochlear implantation, we performed a simple linear regression between the age at which children received the first implant and the individual latency of the GFP-TRF first peak (P1). Once

more, we employed GFP to extract a more reliable peak latency (see Michel et al., 2009). Given the high variability in the neural tracking latency (GFP-TRF P1) of the earliest implanted children, we explored whether a piecewise-linear regression model can significantly explain the data using the adaptive regression splines toolbox (Friedman, 1991). We defined the best number of linear basis functions for our model using the pruning procedures. Thus, we tested the final model with the identified number of basis functions (i.e., 2, including the intercept).

Correlation between neural tracking and speech comprehension. Finally, to explore a possible correlation between the neural responses and the behavioral performance, a point-by-point partial correlation -accounting for the effect of age- between the TRF value at each time point across the whole time window [0 – 600 ms] at each electrode and the percentage of accuracy was performed separately in HC and CI groups. We computed the averaged r^2 across channels for each time-point, and we constructed a null distribution of r^2 with 1000 permutations by shuffling r^2 across time. The actual r^2 effect was compared to the null r^2 distribution, and the empirical p-values obtained were corrected in time with FDR.

SUPPLEMENTARY MATERIALS

Table Characteristics of CI participants

PARTICIPANTS INFORMATION				RESIDUAL HEARING ABILITY BEFORE COCHLEAR IMPLANT SURGERY							TEST INDICATING PROFOUND BILATERAL DEAFNESS BEFORE IMPLANTATION			OTHER CLINICAL INFORMATION								
ID	GROUP	SEX	AGE y.	Newborn hearing screening with otoacoustic emissions	Hearing threshold <70 dB HL at least in one ear assessed via ABR or free field behavioural tone audiometry			Severe hearing loss (mean hearing threshold 71-90 dB HL) at least in one ear assessed via ABR or free field behavioural tone audiometry			Diagnosis of profound (≥90 dB HL) bilateral sensorineural hearing loss before implantation assessed via ABR or behaviorally in free in field			Hearing loss etiology	FAMILIARITY	Implant information				Age of first hearing aid fitting before CI months	BILIGUAL	Gestation / type of birth
					Age	LEFT ear	RIGHT ear	Age	LEFT ear	RIGHT ear	Age	LEFT ear	RIGHT ear			Side	Age first implant months	Age second implant months	Experience with CI months			
CI 1	CD	F	6	fail bilateral							38 days	95	100	GJB2 (Connexin 26)	yes	Bilateral	11	36	61	6	no	
CI 2	CD	F	14	fail bilateral							25 days	100	90	GJB2 (Connexin 26)	no	Bilateral	132	160	39	1	no	
CI 3	CD	M	13	fail bilateral							49 days	90	90	GJB2 (Connexin 26)	no	Right	84		69	6	no	
CI 4	CD	F	10	fail bilateral							25 days	90	90	GJB2 (Connexin 26)	no	Left	16		106	6	no	NICU
CI 5	CD	M	7	fail bilateral							21 days	100	100	GJB2 (Connexin 26) congenital CMV	yes	Left	28		50	5.5	no	twin pregnancy without complications
CI 6	CD	M	7	fail bilateral							21 days	100	100	GJB2 (Connexin 26) congenital CMV	yes	Left	25		53	5.5	no	twin pregnancy without complications
CI 7	CD	M	15	fail bilateral							31 days	100	100	congenital CMV	no	Right	23		151	never used	no	
CI 8	CD	F	3	fail bilateral							57 days	100	100	GJB2 (Connexin 26)	yes	Bilateral	11	11	23	6	no	
CI 9	CD	M	6	fail bilateral							60 days	>100	>100	GJB2 (Connexin 26)	yes	Bilateral	11	12	54	25	no	
CI 10	CD	F	9	fail bilateral							30 days	>90	>90	Prematurity and perinatal complications	no	Bilateral	16	31	92	5	no	premature birth (25 weeks) ceasarean delivary. NICU
CI 11	CD	F	6	fail bilateral							33 days	> 90	> 90	GJB2 (Connexin 26)	no	Bilateral	11	12	54	3	yes	

CI 12	CD	F	5	fail bilateral							34 days	100	100	congenital CMV	yes	Bilateral	13	13	45	7	no	
CI 13	CD	M	12	fail bilateral							49 days	> 90	> 90	Unknown (not syndromic)	no	Bilateral	22	72	116	4	no	
CI 14	CD	M	11	fail bilateral							40 days	> 90	> 90	Unknown (not syndromic)	no	Bilateral	13	27	115	5	no	
CI 15	CD	M	9	fail bilateral							40 days	>100	>100	Waanderburg sindrome	yes	Bilateral	14	35	93	1	no	
CI 16	CD	F	11	fail bilateral							53 days	100	100	congenital CMV	no	Bilateral	13	36	112	3.5	no	
CI 17	AD	F	9	fail bilateral	ABR showing not severe bilateral deafness			24 month	85	90	50 month	100	100	Unknown (not syndromic)	yes	Bilateral	66	66	45	24	no	birth at 31 weeks; ceasarean delivary. NICU
CI 18	AD	M	10	pass bilateral				17 month	80	90	107 month	Bilateral profound sensorineural hearing loss assessed in free field		GJB2 (Connexin 26) and congenital CMV	no	Left	108		16	18	yes	
CI 19	AD	F	12	not done	5 month	60	60				62 month	Bilateral profound sensorineural hearing loss assessed in free field		Renal tubular acidosis and enlarged vestibular acqueduct	yes	Right	120		21	6	no	
CI 20	AD	M	12	pass bilateral				65 month	Bilateral severe sensorineural hearing loss assessed in free field		72 month	Bilateral profound sensorineural hearing loss assessed in free field		GJB2 (Connexin 26) and congenital CMV	no	Bilateral	73	122	68	72	no	
CI 21	AD	F	6	fail bilateral	screening ABR pass						29 month	>100	>100	TBCID24	yes	Right	35		37	30	no	birth at 41 weeks; vaginal delivary
CI 22	AD	F	7	left pass	1 month	15	55	50 month	100	80	53 month	Bilateral profound sensorineural hearing loss assessed in free field		congenital CMV	no	Left	54		24	38	no	premature birth at 8 months
CI 23	AD	M	6	pass bilateral							28 month	>100	>100	WFS1 and USH2A	no	Bilateral	32	50	39	30	no	
CI 24	AD	M	8	fail bilateral	5 month	Bilateral moderate to severe sensorineural hearing loss		37 month	Bilateral severe sensorineural hearing loss assessed		43 month	Bilateral profound sensorineural hearing loss assessed in free field		perinatal suffering, NICU, diabetes	no	Right	84		12	6	no	perinatal suffering, NICU, diabetes

						assessed in free field			in free field											
CI 25	AD	F	9	fail bilateral	1°ABR normal			29 month		Bilateral severe sensorineural hearing loss assessed in free field	35 month	Bilateral profound sensorineural hearing loss assessed in free field		Unknown	yes	Bilateral	36	36	70	30	no	
CI 26	AD	M	12	fail bilateral	2 month	75	55				56 month	Bilateral profound sensorineural hearing loss assessed in free field		enlarged vestibular aqueduct	no	Bilateral	57	99	80	6	no	
CI 27	AD	F	16	pass bilateral							12 month	Bilateral profound sensorineural hearing loss assessed with ABR		congenital CMV	no	Right	17		178	6	no	
CI 28	AD	F	5	pass bilateral							30 month	>100	90	GJB2 (Connexin 26)	yes	Left	42		11	32	no	
CI 29	AD	F	13	fail bialteral	No clinical evidence of profound sensorineural hearing loss but of progressive deterioration			No clinical evidence of profound sensorineural hearing loss but of progressive deterioration			35 month	>100	>100	GJB2 (Connexin 26)	no	Bilateral	36	36	120	34	no	
CI 30	AD	M	7	fail bialteral	1 month	60	50	6 month	80	80	20 month	>90	>90	congenital CMV	no	Bilateral	21	28	66	4	no	
CI 31	AD	F	7	fail bialteral	2 month	30	30	18 month	95	85	20 month	Bilateral profound sensorineural hearing loss assessed in free field		GJB2 (Connexin 26)	no	Bilateral	21	26	56	18	yes	
CI 32	AD	M	8	pass bilateral	at birth	30	30				54 month	Before implantation tested for profound bialteral deafness in free field		enlarged vestibular aqueduct	no	Bilateral	55	66	44	30	no	31 week caesarean delivery due to interruption intrauterine development

Table S1. Characteristics of each CI participant included in the final sample. Degrees of hearing loss were defined following American Speech-Language-Hearing Association (https://www.asha.org/public/hearing/degree-of-hearing-loss/, see also Lieu et al., 2020).

The existence of TRF with respect to the null-TRF across all electrodes

Within each group, a cluster-based permutation test (same parameters as in the main analysis reported above) was performed to assess the difference between TRFs and the null-TRFs across all electrodes and all time lags between 0 and 600 ms. In the HC group, four positive and two significant clusters emerged ($p_{clust} < 0.05$, see Figure S1A). In the CI group, two positive and two negative significant clusters emerged (all $p_{clust} < 0.05$, Figure S1B). Also, when we tested the two subgroups of CI, CD and AD separately, significant differences between the TRF and the null-TRF emerged. Within CD, one positive and two negative significant clusters emerged ($p_{clust} < 0.05$, Figure S1C), and within the AD group, one positive and one negative significant cluster emerged ($p_{clust} < 0.05$, Figure S1D).

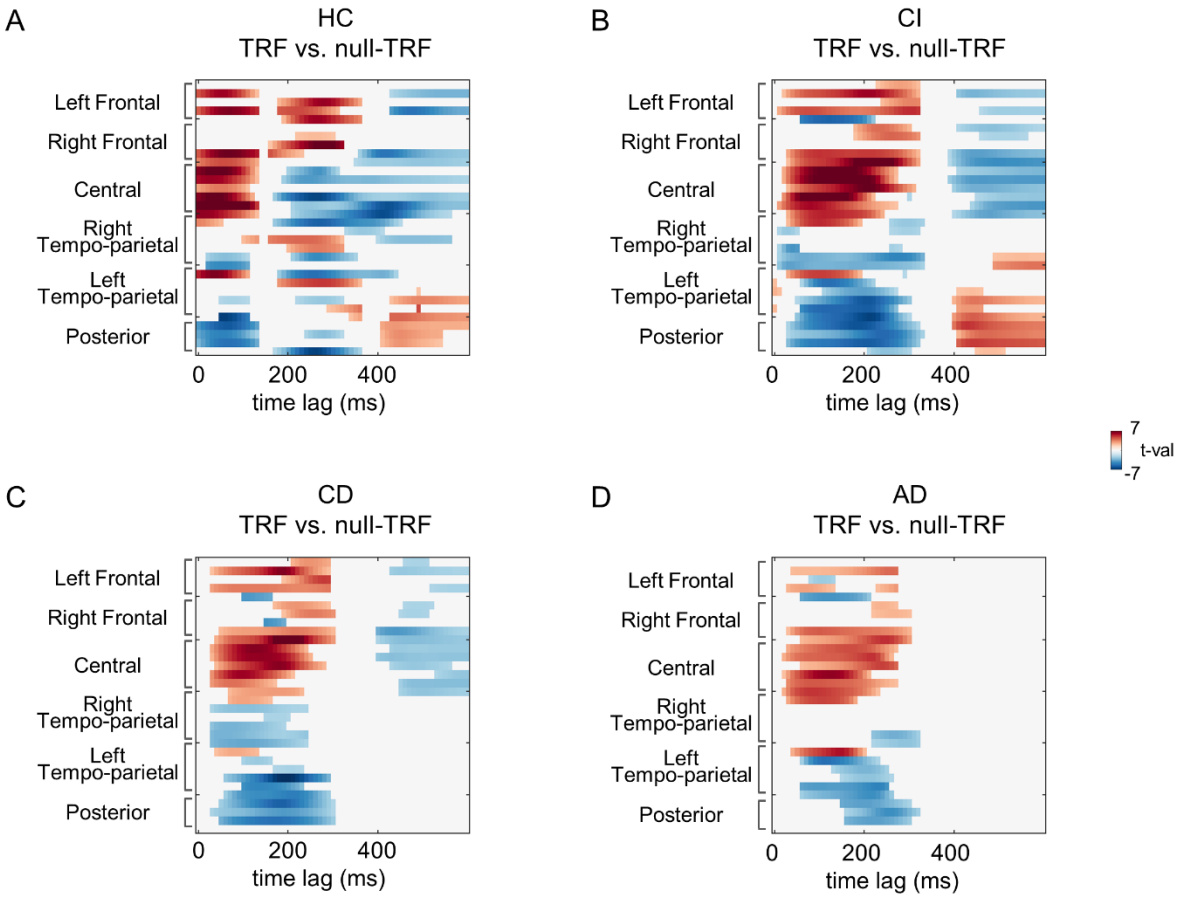

Figure S1. Existence of neural tracking of speech in all groups. Plots representing the results of cluster-based permutation tests contrasting the TRF and the null-TRF for HC, CI, CD, and AD groups (A, B, C, and D panels respectively).

Spatiotemporal dynamics of neural tracking within HC and CI groups

The topographies reported in Figure S2 show the TRF temporal dynamics in the whole time window of interest for both HC and CI children.

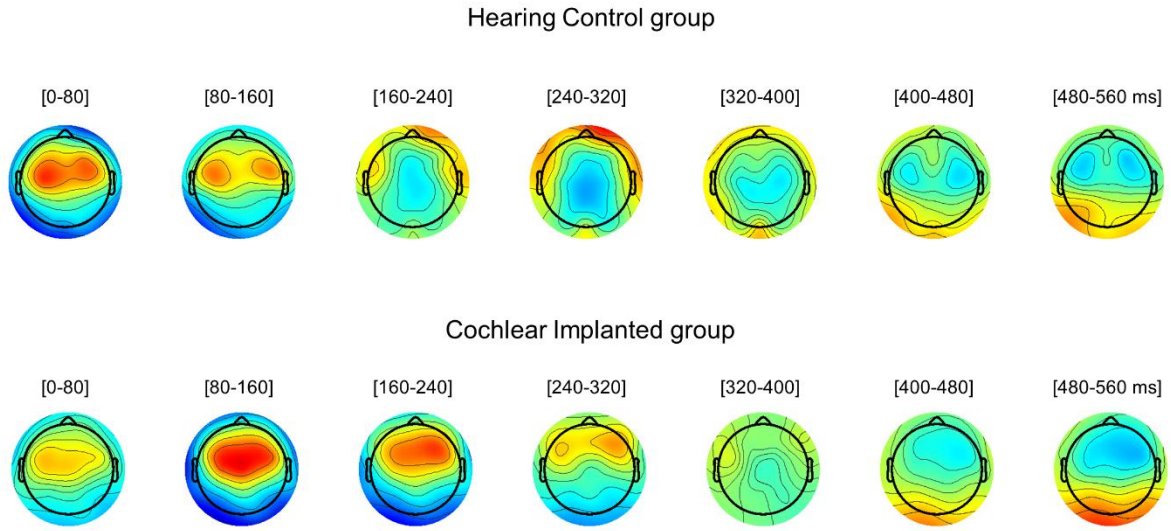

Figure S2. The spatiotemporal dynamics of HC's and CI's TRFs. Topography showing the TRFs across time (successive time window of 80 ms), separately for hearing control (HC) and cochlear implanted (CI) groups.

Frontocentral TRF in CD and AD

For both CD and AD, the frontocentral cluster TRF was significantly different from the null-TRF, highlighting a clear P1 (CD, all $p_{FDR} < 0.05$; AD, all $p_{FDR} < 0.05$, see Figure S3A).

Delayed neural tracking in CD and AD

We investigated whether the delay that emerged in CI's TRF compared to HC's TRF was significant in both CD and AD subgroups. Two separate independent t-tests performed on the latency first peak (GFP-TRF P1) contrasting CD vs. HC and AD vs. HC confirmed that neural tracking of both CD and AD group was significantly delayed (HC vs. CD: $t_{(51)} = -2.32$, $p = 0.024$, $d = -0.69$, CI95 = -1.57 – -0.12; HC vs. AD: $t_{(51)} = -2.18$, $p = 0.034$, $d = -0.64$, CI95

= -1.35 – -0.12). Note that the distributions of CD's and AD's GFP-TRF P1 latencies are superimposed, and no outliers were present (see Figure S3B).

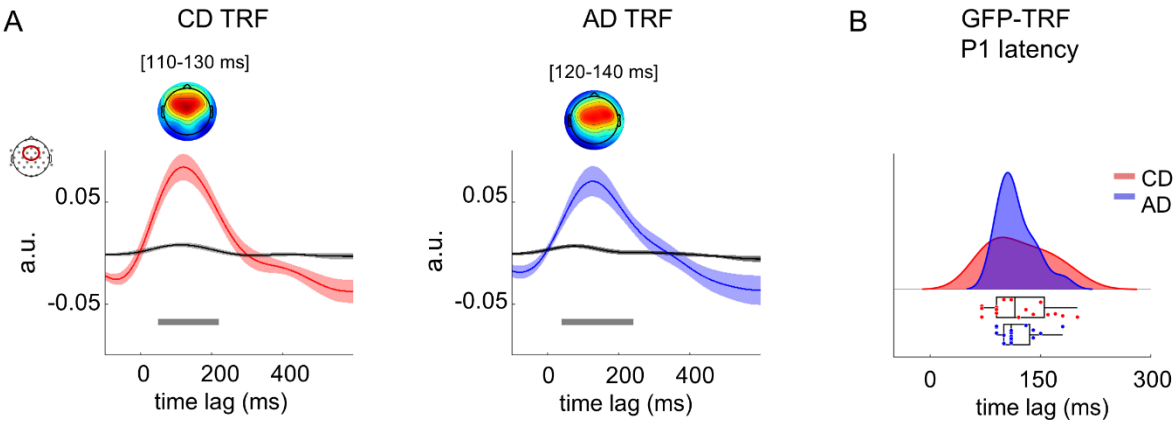

Figure S3. Frontocentral TRF and P1 latency in the two CI groups. (A) Grand average TRFs (blue CD, red CI) and grand average null-TRFs (grey color) were measured at frontocentral electrodes (Cz, Fz, FC1, and FC2) between -100 and 600 ms time lags. Shaded areas represent SE of the mean. Grey horizontal bars indicate time lags (between 0 and 600 ms) at which TRFs differed significantly from the null-TRF (running t-tests, FDR corrected $p_{FDR} < 0.05$). (B) Distributions of the first peak (GFP-TRF P1) latency for CD and AD participants.

Relationship between GFP-TRF P1 latency and age at cochlear implantation

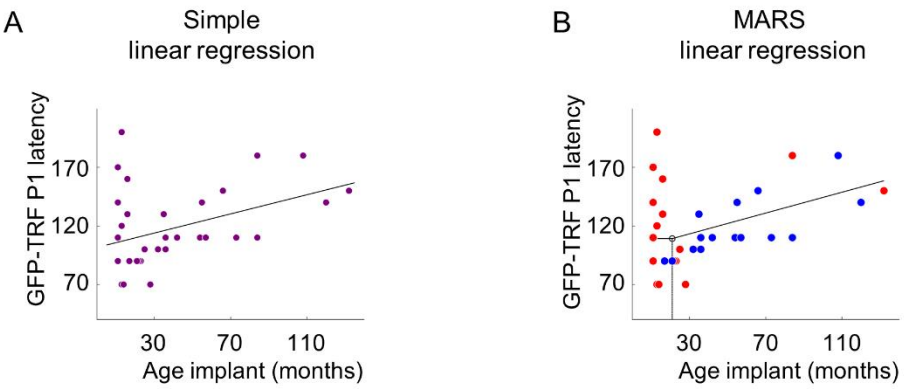

Figure S4. Relationship between neural tracking latency and implantation age. (A) The plot shows the simple linear regression between when the child received the implant and the latency of the first peak of the neural tracking (GFP-TRF P1). (B) The plot shows the piecewise linear regression highlighting that most children implanted earlier than 21 months

were diagnosed with bilateral profound congenital deafness (CD children are represented in red, while AD in blue).

Given the high variability in the latency of the first peak neural tracking (GFP-TRF P1) for the children implanted before 21 months of age, we assessed whether some other clinical characteristics could explain this variance. We ran three separate simple linear regressions within this subgroup of CI participants who were implanted before 21 months of age to explore the impact on the P1 latency of chronological age, experience with the implant, and age at which hearing aids were provided before implantation, respectively. None of these factors explain the variance of the P1 latency in the earliest implanted children (all $p > 0.05$).

CI TRF temporally realigned

To account for the delay in the CI group, we realigned their first peak (P1) to the first peak of the HC group. The shift amount was 60 ms, equal to the difference between the mean of HC's first peak (60 ms) and CI's first peak (120 ms) computed on the frontocentral TRF cluster.

Results without cleaning of the artifacts

Comparable results emerged even without the artifact-cleaning procedure.

When we tested the existence of the auditory response function (between 0 and 600 ms), the frontocentral cluster TRF emerged to be significantly different from the null-TRF at two time windows between 20 and 260 ms and between 400 and 600 ms ($p_{FDR} < 0.05$).

Cluster-based permutation test between CI and HC groups performed at the whole brain level (across all channels) and comprising the TRF at every time lag between 0 to 600 ms revealed a significant difference between CI and HC groups (all $p_{clust} < 0.05$).

No difference emerged between CD and AD (no clusters were found at $p_{clust} < 0.05$).

REFERENCES

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., Rocha, S., Brusini, P., Mead, N., Olawole-Scott, H., Boutris, P., Gibbon, S., Williams, I., Grey, C., Flanagan, S., &

Goswami, U. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. Neuroimage 247:118698 Available at: https://doi.org/10.1016/j.neuroimage.2021.118698.

Barriga-Paulino, C. I., Rodríguez-Martínez, E. I., Arjona, A., Morales, M., & Gómez, C. M. (2017). Developmental trajectories of event related potentials related to working memory. Neuropsychologia, 95, 215-226.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1129–1159.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Annals of statistics, 1165-1188.

Berrettini, S., Arslan, E., Baggiani, A., Burdo, S., Cassandro, E., Cuda, D., ... & Forli, F. (2011). Analysis of the impact of professional involvement in evidence generation for the HTA Process, subproject “cochlear implants”: methodology, results and recommendations. Acta Otorhinolaryngologica Italica, 31(5), 273.

Billings, C. J., Bennett, K. O., Molis, M. R., & Leek, M. R. (2011). Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. Ear and hearing, 32(1), 53.

Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., & Röder, B. (2020). EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. Neuroimage 223:117315.

Bottari, D., & Berto, M. (2021). Three factors to characterize plastic potential transitions in the visual system. Neurosci Biobehav Rev 126:444–446 Available at: https://doi.org/10.1016/j.neubiorev.2021.03.035.

Chen, Y. P., Schmidt, F., Keitel, A., Rösch, S., Hauswald, A., & Weisz, N. (2023). Speech intelligibility changes the temporal evolution of neural speech tracking. NeuroImage, 268, 119894.

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. J Neurosci Methods 250:126–136 Available at: http://dx.doi.org/10.1016/j.jneumeth.2015.01.010.

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. Front Hum Neurosci 10:1–14.

Crosse, M. J., Zuk, N. J., Di Liberto, G. M., Nidiffer, A., Molholm, S., & Lalor, E. C. (2021). Linear modeling of neurophysiological responses to naturalistic stimuli: Methodological considerations for applied research.

Delorme, A., & Makeig, S. (2004) EEGLAB: an open sorce toolbox for analysis of single-trail EEG dynamics including independent component anlaysis. J Neurosci Methods 134:9–21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage, 34(4), 1443-1449.

Deprez, H., Gransier, R., Hofmann, M., van Wieringen, A., Wouters, J., & Moonen, M. (2017). Characterization of cochlear implant artifacts in electrically evoked auditory steady-state responses. Biomedical Signal Processing and Control, 31, 127-138.

Ding, N., & Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. Journal of Neuroscience, 33(13), 5728-5735.

Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. The Journal of the Acoustical Society of America, 95(5), 2670-2680.

Drullman, R., Festen, J. M., & Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. The Journal of the Acoustical Society of America, 95(2), 1053-1064.

Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: Correlations with changes in structure and speech perception. Acta Otolaryngol 123:249–252.

Etard, O., & Reichenbach, T. (2019). Neural speech tracking in the theta and in the delta frequency band differentially encode clarity and comprehension of speech in noise. Journal of Neuroscience, 39(29), 5750-5759.

Fiedler, L., Wöstmann, M., Herbst, S. K., & Obleser, J. (2019). Late cortical tracking of

ignored speech facilitates neural selectivity in acoustically challenging conditions. Neuroimage, 186, 33-42.

Friedman, J. H. (1991). Multivariate adaptive regression splines. The annals of statistics, 19(1), 1-67.

Friedmann, N., & Rusou, D. (2015). Critical period for first language: the crucial role of language input during the first year of life. Current opinion in neurobiology, 35, 27-34.

Gates, G. A., Daly, K., Dichtel, W. J., Dooling, R. J., Gulya, A. J., Hall, J. W., ... & Hall, Z. W. (1995). Cochlear implants in adults and children. JAMA, 274(24), 1955-1961.

Gillis, M., Decruy, L., Vanthornhout, J., & Francart, T. (2021). Hearing loss is associated with delayed neural responses to continuous speech Running title : Hearing loss delays neural responses to speech : KU Leuven , Department of Neurosciences , ExpORL , 3000 Leuven , Belgium : Institute for Systems Research, Univ.

Gustafson, S. J., Billings, C. J., Hornsby, B. W. Y., & Key, A. P. (2019). Effect of competing noise on cortical auditory evoked potentials elicited by speech sounds in 7- to 25-year-old listeners. Hear Res 373:103–112.

Jessen, S., Fiedler, L., Münte, T. F., & Obleser, J. (2019). Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. Neuroimage 202:116060 Available at: https://doi.org/10.1016/j.neuroimage.2019.116060.

Jung, T-P., Makeig, S., Humphries, C., Lee, T-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000a). Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37:163–178.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000b). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin Neurophysiol 111:1745–1758.

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. Sci Rep 8:1–8.

Kral, A., Dorman, M. F., & Wilson, B. S. (2019). Neuronal Development of Hearing and Language: Cochlear Implants and Critical Periods. Annu Rev Neurosci 42:47–65.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. Trends in neurosciences, 35(2), 111-122.

Lalor, E. C., Power, A. J., Reilly, R. B., & Foxe, J. J. (2009). Resolving precise temporal processing properties of the auditory system using continuous stimuli. J Neurophysiol 102:349–359.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr Clin Neurophysiol 48:609–621.

Lieu, J. E., Kenna, M., Anne, S., & Davidson, L. (2020). Hearing loss in children: a review. Jama, 324(21), 2195-2205.

Lorenzi, C., Berthommier, F., & Demany, L. (1999). Discrimination of amplitude-

modulation phase spectrum. J. Acoust. Soc. Am. 105, 2987–2990.

Mariani, B., Nicoletti, G., Barzon, G., Ortiz Barajas, M. C., Shukla, M., Guevara, R., ... & Gervain, J. (2023). Prenatal experience with language shapes the brain. Science Advances, 9(47), eadj3524.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Michel, C. M. (Ed.). (2009). Electrical neuroimaging. Cambridge University Press.Mirkovic, B., Debener, S., Jaeger, M., & De Vos, M. (2015). Decoding the attended speech stream with multi-channel EEG: Implications for online, daily-life applications. J Neural Eng 12.

Ortiz Barajas, M. C., Guevara, R., & Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. Developmental cognitive neuroscience, 48, 100915.

Ortiz Barajas, M. C., Guevara, R., & Gervain, J. (2023). Neural oscillations and speech processing at birth. Iscience, 26(11).

O’Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. Cereb Cortex 25:1697–1706.

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. Trends Cogn Sci 23:913–926.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011.

Paul, B. T., Uzelac, M., Chan, E., & Dimitrijevic, A. (2020). Poor early cortical differentiation of speech predicts perceptual difficulties of severely hearing-impaired listeners in multi-talker environments. Sci Rep 10:1–12.

Pavani, F., & Bottari, D. (2022). Neuroplasticity following cochlear implants. In Handbook of Clinical Neurology (Vol. 187, pp. 89-108). Elsevier.

Pérez-Navarro, J., Klimovich-Gray, A., Lizarazu, M., Piazza, G., Molinaro, N., & Lallier, M. (2023). The contribution of early language exposure to the cortical tracking of speech. bioRxiv, 2023-09.

Reh, R. K., Dias, B. G., Nelson III, C. A., Kaufer, D., Werker, J. F., Kolb, B., ... & Hensch, T. K. (2020). Critical period regulation across multiple timescales. Proceedings of the National Academy of Sciences, 117(38), 23242-23251.

Ricciardi, E., Bottari, D., Ptito, M., Röder, B., & Pietrini, P. (2020). The sensory-deprived brain as a unique tool to understand brain development and function. Neurosci Biobehav Rev 108:78–82.

Röder, B., & Kekunnaya, R. (2021). Visual experience dependent plasticity in humans. Curr Opin Neurobiol 67:155–162.

Sanes, D. H., & Woolley, S. M. N. (2011). A behavioral framework to guide research on central auditory development and plasticity. Neuron 72:912–929.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270(5234), 303-304.

Sharma, A., Campbell, J., & Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. Int J Psychophysiol 95:135–144 Available at: https://pubmed.ncbi.nlm.nih.gov/24780192/ [Accessed October 3, 2022].

Sharma, A., & Campbell, J. (2011). A sensitive period for cochlear implantation in deaf

children. The Journal of Maternal-Fetal & Neonatal Medicine, 24(sup1), 151-153.

Sharma, A., Dorman, M., & Spahr, A. (2002a). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. Ear Hear 23:532–539.

Sharma, A., Dorman, M., Spahr, A., & Todd, N. W. (2002b). Early cochlear implantation in children allows normal development of central auditory pathways. Ann Otol Rhinol Laryngol 111:38–41.

Sharma, A., Dorman, M. F., & Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. Hear Res 203:134–143.

Sharma, A., Gilley, P. M., Dorman, M. F., & Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. International journal of audiology, 46(9), 494-499.

Somers, B., Verschueren, E., & Francart, T. (2018). Neural tracking of the speech envelope in cochlear implant users. Journal of neural engineering, 16(1), 016003.

Steinschneider, M., Liégeois-Chauvel, C., & Brugge, J. F. (2011). Auditory Evoked Potentials and Their Utility in the Assessment of Complex Sound Processing. In: The Auditory Cortex (Springer, Boston M, ed), pp 535–559.

Strickland, E. A., & Viemeister, N. F. (1996). Cues for discrimination of envelopes. The Journal of the Acoustical Society of America, 99(6), 3638-3646.

Stropahl, M., Bauer, A. K. R., Debener, S., & Bleichner, M. G. (2018). Source-Modeling auditory processes of EEG data using EEGLAB and brainstorm. Front Neurosci 12:309 Available at: www.frontiersin.org [Accessed July 3, 2021].

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., & Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. Journal of the Association for Research in Otolaryngology, 19, 181-191.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., & Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. Clin Neurophysiol 120:868–877

Wang, C., & Zhang, Q. (2021). Word frequency effect in written production: Evidence from

ERPs and neural oscillations. Psychophysiology, 58(5), e13775.

Watanabe, T. (2021). Causal roles of prefrontal cortex during spontaneous perceptual switching are determined by brain state dynamics. Elife, 10, e69079.

Watanabe, T., Rees, G., & Masuda, N. (2019). Atypical intrinsic neural timescale in autism. Elife, 8, e42256.

Werker, J. F., & Hensch, T. K. (2015). Critical Periods in Speech Perception: New Directions. Annu Rev Psychol 66:173–196 Available at: http://www.annualreviews.org/doi/10.1146/annurev-psych-010814-015104.

Winn, M. B., & Nelson, P. B. (2021). Cochlear implants. In Oxford Research Encyclopedia of Linguistics.

Yasmin, S., Irsik, V. C., Johnsrude, I. S., & Herrmann, B. (2023). The effects of speech masking on neural tracking of acoustic and semantic features of natural speech. Neuropsychologia, 186, 108584.