

<https://doi.org/10.1038/s42003-025-07788-4>

Resilience and vulnerability of neural speech tracking after hearing restoration



Alessandra Federici¹, Marta Fantoni¹, Francesco Pavani^{2,3}, Giacomo Handjaras¹, Evgenia Bednaya¹, Alice Martinelli^{1,4}, Martina Berto¹, Franco Tralbalzini⁵, Emiliano Ricciardi¹, Elena Nava⁶, Eva Orzan⁷, Benedetta Bianchi⁵ & Davide Bottari¹ ✉

The role of early auditory experience in the development of neural speech tracking remains an open question. To address this issue, we measured neural speech tracking in children with or without functional hearing during their first year of life after their hearing was restored with cochlear implants (CIs), as well as in hearing controls (HC). Neural tracking in children with CIs is unaffected by the absence of perinatal auditory experience. CI users and HC exhibit a similar neural tracking magnitude at short timescales of brain activity. However, neural tracking is delayed in CI users, and its timing depends on the age of hearing restoration. Conversely, at longer timescales, speech tracking is dampened in participants using CIs, thereby accounting for their speech comprehension deficits. These findings highlight the resilience of sensory processing in speech tracking while also demonstrating the vulnerability of higher-level processing to the lack of early auditory experience.

In the case of typical development, the auditory cortex tracks several speech sound features, such as signal amplitude modulations. These fluctuations, or speech envelope, have energy peaks around the syllabic rate and are a pivotal feature for speech comprehension. Adults can understand heavily degraded speech, provided the envelope is preserved¹. On the contrary, however, its suppression impairs comprehension^{2,3}. The primary role of this form of neural speech tracking is substantiated by studies revealing its presence in newborns and young infants^{4–7}, thus suggesting a strong biological predisposition. Nonetheless, recent evidence suggests that linguistic experience in the first year of life modulates neural speech tracking⁶. Indeed, typical brain development requires temporal overlapping between neural system readiness and appropriate environmental statistics^{8–10}. These statistics include regularities in the auditory environment, which support language acquisition in infants (see ref. 11). Exposure to statistical properties of speech sounds during the first year of life has been demonstrated to play a critical role in native language development, for instance, in phoneme categorizations¹². Yet, the extent to which the development of neural speech tracking relies on postnatal auditory experience remains unknown. Individuals facing a period of sensory deprivation provide a unique opportunity for causally assessing whether neural functions have sensitive phases in which specific sensory input must be provided for shaping the associated neural circuitries¹³ 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^{14–16}. Cochlear implants rely substantially upon the sound envelope to convert continuous speech into electric impulses¹⁷. However, despite efficiently conveying this information to the brain, speech comprehension outcomes are very heterogeneous¹⁸. Crucial for understanding language acquisition variability is the age at which hearing is restored with cochlear implantation; results suggested that the sooner, the better^{19–22}. Seminal studies uncovered that the latency of auditory responses in individuals using CIs falls within typical developmental trajectory only when children are implanted before 3.5 years of age^{23–27}. However, the activity of the auditory 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) Does the first year of life encompass a sensitive period during which auditory input is essential for its development?

To fill these gaps, we measured the degree of synchronization between brain activity and continuous speech (fitted at the single participant level) in hearing children (HC), listening to either original or vocoded-speech, and children using CI with different onsets of bilateral profound deafness (congenital or acquired). The comparison between the neural speech tracking of children who were born with congenital deafness (CD) and

¹MoMiLab, IMT School for Advanced Studies Lucca, Lucca, Italy. ²Centro Interdipartimentale Mente/Cervello — CIMEC, University of Trento, Trento, Italy. ³Centro Interuniversitario di Ricerca “Cognizione Linguaggio e Sordità” — CIRCLoS, University of Trento, Trento, Italy. ⁴IRCCS Fondazione Stella Maris, Pisa, Italy. ⁵IRCCS Meyer, Azienda Ospedaliero-Universitaria Meyer, Firenze, Italy. ⁶University of Milano-Bicocca, Milano, Italy. ⁷IRCCS Materno Infantile Burlo Garofolo, Trieste, Italy.

✉ e-mail: davide.bottari@imtlucca.it

experienced auditory deprivation during the first year of life, with children who acquired deafness (AD), born with some degree of functional hearing and whose profound bilateral deafness emerged only after one year of age, allowed to assess the role of the perinatal auditory experience (see Fig. 1). Moreover, to control for the effects of degraded speech provided by the CI, speech neural tracking of CD and AD children was also compared with a group of hearing children who listened to vocoded-speech (HC-v). Participants' age range (3–18 years old) was chosen to measure developmental trajectories of neural speech tracking in hearing and children using CI.

Results

Neural speech tracking in hearing and cochlear-implemented children

First, we assessed whether neural speech tracking could be measured in HC and children using CI irrespective of deafness onset. We estimated the temporal response function (TRF) within a frontocentral cluster of sensors suitable for measuring auditory response functions in children and adults^{5,28} and CI users²⁹. Contrasting the TRF model with the null-TRF in HC and children using CI, we observed significant neural speech tracking in both groups (see Fig. 2A). Specifically, in HC children, auditory responses were characterized by a prominent positivity between 0 and 110 ms time lags (all $p_{\text{FDR}} < 0.05$; peak TRF = 0.065, SE = 0.008; $d = 1.24$, 95th confidence interval (CI95) = 0.78–1.64, see Supplementary Materials 2.1 for details on how Cohen's d was computed), and by a subsequent negativity between 200 and 320 ms time lags (all $p_{\text{FDR}} < 0.05$; peak TRF = -0.051, SE = 0.009; $d = -0.86$, CI95 = -1.21 to -0.39). In the CI group, the first positive response emerged at time lags between 20 and 270 ms (all $p_{\text{FDR}} < 0.05$; peak TRF = 0.089, SE = 0.010; $d = 1.92$, CI95 = 1.43–2.49) and then the negativity occurred after 390 ms time lag (all $p_{\text{FDR}} < 0.05$; peak TRF = -0.054, SE = 0.007; $d = -0.93$, CI95 = -1.44 to -0.48). Results highlighted that in both groups of children (i.e., HC and CI), the neural tracking could be robustly measured with about ten minutes of natural speech. The activity was characterized by two main phases, one at short ($P1_{\text{TRF}}$) and one at long ($N2_{\text{TRF}}$) timescales of brain-speech tracking occurring within 600 ms of

time lag (Supplementary Materials 2.2 Fig. S1 for visualization of the whole spatiotemporal dynamics). From hereafter, we refer to the short timescale of speech tracking for neural responses within 150 ms of time lag and to the long timescale after 150 ms of time lag. Importantly, in both CD and AD subgroups, the frontocentral TRF exceeded the null-TRF computed at frontocentral sensors, as well as in the HC-v group (Supplementary Materials 2.3 and Fig. S2). These results were confirmed when cluster-based permutation analysis was performed across all electrodes (Supplementary Materials 2.4 Fig. S3).

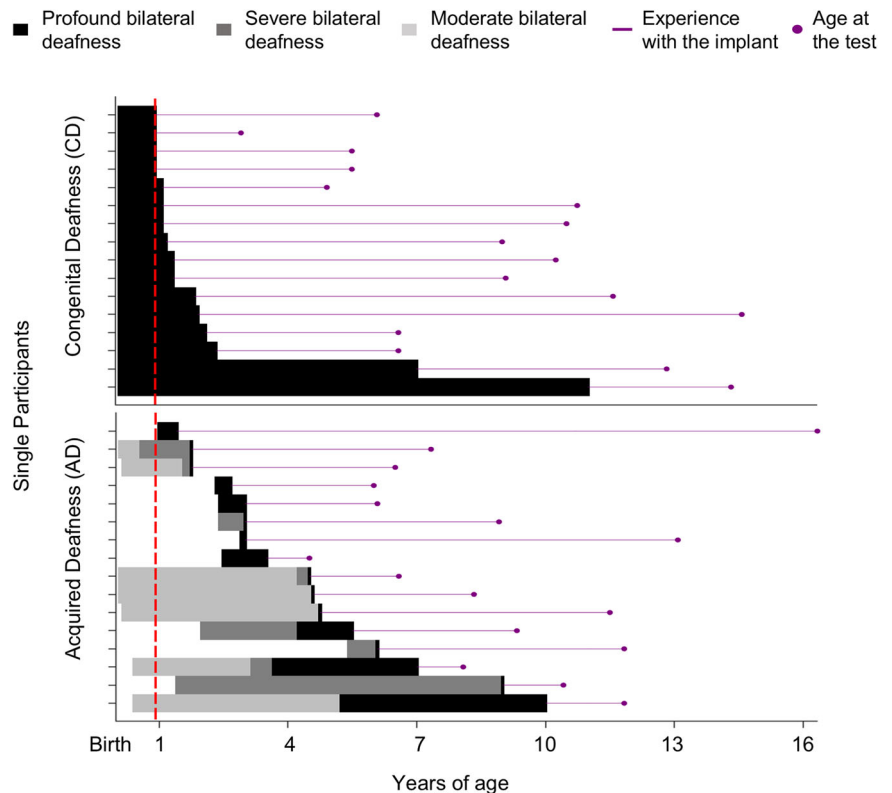
Once the existence of the speech-tracking TRF was verified within each group, we investigated whether it was possible to measure a developmental trajectory of neural tracking in both HC and children using CI. This would strongly advocate for employing neural speech tracking as a reliable measure, even in children who perceive hearing only through cochlear implants. We expected neural tracking to become more efficient (i.e., less spread over time) with age.

In the HC group a clear developmental pattern emerged, highlighting an association between neural tracking signal (GFP-TRF) 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 Fig. 2C upper panel). A similar pattern was observed in children using CI (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 Fig. 2C lower panel).

How temporary auditory deprivation affects neural speech tracking

Once the existence and development of the auditory response function were assessed in both HC and children using CI, we compared the TRFs across groups. First, spatiotemporal profiles of TRFs were compared between the CI and the HC groups using 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 ($p_{\text{clust}} < 0.05$), with the largest

Fig. 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.



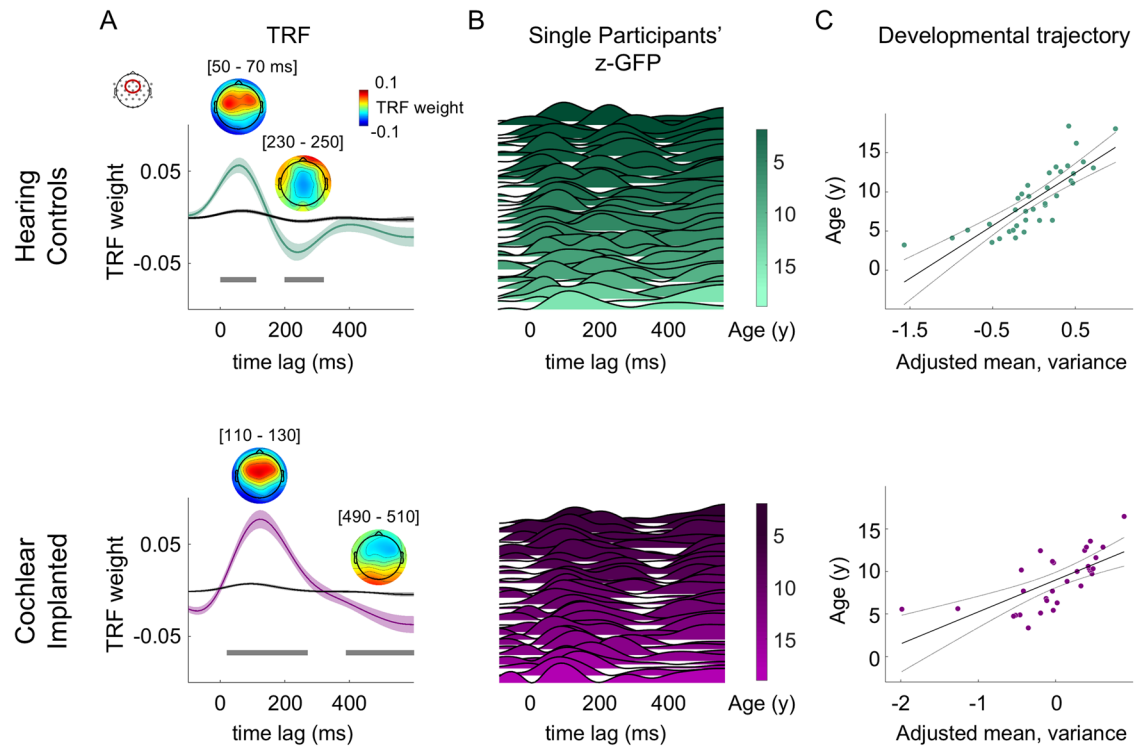


Fig. 2 | Neural speech tracking in HC and children using CI. **A** Grand average speech tracking TRFs (olive color HC, magenta color CI) and grand average null-TRFs (grey color) at frontocentral electrodes (Cz, Fz, FC1, and FC2) between -100 and 600 ms time lags. Shaded areas represent SE of the mean (N HC = 37; N CI = 32). Grey horizontal bars indicate time lags (between 0 and 600 ms) at which speech-tracking TRFs significantly differed from the null-TRF (*t*-test, FDR corrected $p_{FDR} < 0.05$). The topographies show the spatial distribution of the TRF peaks.

B Single participants' Global Field Power of their speech tracking 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** Plots reveal a developmental trajectory in both HC and CI groups, the upper and lower panel, respectively. Partial regression plots of the linear regression models, including only the significant independent variables (i.e., mean and variance as reported in the x-axis label), are shown.

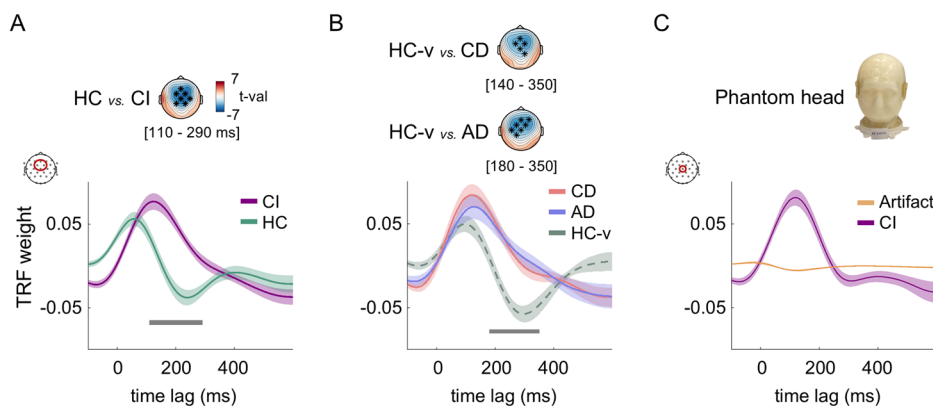


Fig. 3 | The impact of auditory deprivation on neural speech tracking. **A** The speech tracking TRF at the frontocentral sensors between -100 and 600 ms time lags for HC (olive) and CI (magenta) groups. The topography shows the statistical difference between TRFs in HC and children using CI between 110 and 290 ms; significant sensors at $p_{\text{cluster}} < 0.05$ are highlighted with black asterisks. **B** Speech tracking TRF at the frontocentral sensors between -100 and 600 ms time lags for CD (red), AD (blue), and HC-v (dashed greenish) groups. For all TRFs, the continuous line represents the group mean and the shaded area the SE (N HC = 37; N CI = 32; N CD = 16; N AD = 16; N HC-v = 16). 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 neural speech tracking in children using CI. Both CD and AD are significantly different from HC-v groups (significant sensors at $p_{\text{cluster}} < 0.05$ are highlighted with black asterisks), unveiling that the anomalies in CI neural tracking cannot be explained merely by the vocoded speech. **C** Artifact and CI's TRF are shown for the electrode of the frontocentral cluster where the artifact activity is strongest (Cz). Results suggest negligible impact of the electrical activity of the implant on the TRF measured in children using CI.

significant effect at time lags between 110 and 290 ms in a large frontocentral cluster of sensors ($d = -1.69$, $CI95 = -2.14$ to -1.25). This difference revealed that neural tracking in children with CIs was delayed at a short timescale of brain activity ($P1_{\text{TRF}}$), and the subsequent activity was hampered (see Fig. 3A).

Next, we investigated the specific role of auditory experience in the first year of life. The two groups of children using CI, the one with congenital deafness (CD) and acquired deafness (AD), were contrasted to investigate the role of perinatal auditory input in developing neural speech tracking. The data revealed a clear overlap between the TRFs of the two CI groups (see

Fig. 3B). 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 $p_{\text{clust}} < 0.05$). This suggested that the alterations of neural speech tracking observed in CI compared to HC occurred irrespective of whether children experienced or not the lack of hearing input within the first year of life. The cluster-based permutation tests performed between HC and each of the two CI subgroups (i.e., CD and AD) revealed that TRFs of both CD and AD significantly differed from HC's TRF (one positive and one negative cluster, $p_{\text{clust}} < 0.05$).

We further assessed whether the difference between HC and CI's neural tracking could be driven by the degraded stimulation provided by the cochlear implants compared to the natural sounds perceived by HC children. Similarly to the comparison with HC (Fig. 3A), significant differences emerged between hearing children who listen to vocoded-speech (HC-v) and both CD and AD groups ($p_{\text{clust}} < 0.05$), indicating that alterations in CI's neural tracking cannot be explained by the mere degradation of the speech signal (Fig. 3B). Note that listening to a vocoded stimulation slightly delayed the neural tracking in hearing participants (HC-v) with respect to those listening to normal stimuli (HC). Still, this difference did not reach statistical significance (see Fig. S4C in Supplementary Materials 2.5).

Finally, since cochlear implants produce electric impulses, we assessed whether functioning cochlear implants could profoundly affect the recorded and modelled EEG signal. Therefore, we measured the artifact TRF obtained from the electrical activity of cochlear implants during stimuli presentation to a phantom head in the absence of brain activity. Artifact TRF had negligible magnitude and, compared to CI's TRF, had a completely different profile, with the first peak at 0 ms as expected (Fig. 3C). This observation ruled out the possibility that the implant activity merely caused alterations emerging in the neural speech tracking of the CI group.

To sum up, results demonstrated that the difference between CI and HC cannot be explained by the absence of auditory experience during the first year of life, by degradation of the speech signal, or by implant activity. Given that neural tracking of CD and AD children did not differ, we further characterized the alterations of the neural speech tracking of children using CI as a single group and the association of these alterations with clinical profiles, aiming to unravel possible biomarkers of their continuous speech processing.

The first anomaly of CI's neural tracking is the latency. Data revealed that the first peak (i.e., GFP-TRF P1) in the CI group was substantially delayed compared to the HC group ($t_{(67)} = -2.97$; $p = 0.004$, HC mean = 86.8 ms, SE = 8.7; CI mean = 119.1 ms, SE = 6.0; $d = -0.71$, CI95 = -1.24 to -0.12 , see Fig. 4A), suggesting less efficient neural tracking over short time intervals from sound onsets. Given that previous neurophysiological studies²⁵ have highlighted the impact of implantation age on auditory response latency to simple sounds, we assessed whether the age at implantation affects TRF latency at a short timescale. Linear regression indicated that the age at which auditory input was restored with cochlear implantation accounted for the delayed neural tracking in CI users (latency GFP-TRF P1). Data indicated that the later the implantation occurred, the more delayed the neural tracking ($R^2 = 0.115$, $F_{(1,30)} = 5.04$, $p = 0.032$, $\beta = 0.38$, CI 95 = 0.04 – 0.72 , see Fig. 4B). Other clinical variables, such as chronological age, age at bilateral profound deafness onset, age at which hearing aids were provided before implantation, and experience with the implant, did not explain the delay (p -values were greater than 0.22). Given the high variability of the neural tracking first peak for the earliest implanted children (see Fig. 4B), we further investigated the association between implantation age and TRF latency by fitting multiple linear functions to identify discontinuities (Friedman, 1991; see Supplementary Materials 2.6). We found a significant regression model with two basis functions and a discontinuity knot at 21 months ($R^2 = 0.173$; cross-validated $R^2_{\text{GCV}}: 0.06$, $p = 0.036$, hinge function $\max(0, x_1 - 21) \beta = 0.45$, CI95 = 0.44 – 1.05), revealing that the positive relationship between implantation age and the latency of the neural tracking emerged from 21 months onward. The other clinical variables cannot explain the GFP-TRF P1 variability before

21 months of age (see Supplementary Materials Fig. S5 and the limitation section).

A second anomaly of the CI's neural speech tracking emerged at a longer timescale: a difference between the TRF dynamic of CI and HC children was evident even when accounting for the neural tracking delay (see Figs. 3A and 4C). To test this hypothesis, we shifted the CI's data, aligning the first TRF peak ($P1_{\text{TRF}}$ at the frontocentral cluster) of the CI group to the homolog peak of the HC group. We shifted the CI's TRFs (at each sensor) 60 ms earlier to compensate for the delay between HC's first peak (60 ms) and CI's first peak (120 ms) computed at the frontocentral cluster. 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 ($p_{\text{clust}} < 0.05$; $d = -1.11$, CI95 = -1.50 to -0.69). At this long timescale of brain activity, the second major phase of the auditory temporal response function ($N2_{\text{TRF}}$) emerged clearly only in the HC. Conversely, the first TRF phase ($P1_{\text{TRF}}$) did not differ between HC and CI, suggesting unaltered TRF magnitude at a short timescale (Fig. 4C).

Neural tracking and speech comprehension

To assess speech comprehension, we investigated the outcome of the behavioral questionnaire, which comprised questions concerning story details. Results revealed impaired scores in children using CI compared to HC (one-way ANOVA: $F_{(3,81)} = 16.86$, $p < 0.001$). Both CD and AD groups showed significantly lower accuracy with respect to HC and HC-v children (HC: mean 85.70% accuracy $\pm$ SE 2.50; HC-v: mean 86.46% accuracy $\pm$ SE 3.87; CD: mean 56.77% accuracy $\pm$ SE 5.72; AD: mean 61.98% accuracy $\pm$ SE 3.61). Post-hoc Bonferroni corrected t-tests confirmed significant differences between both HC and CI groups (HC vs. CD $p_{\text{Bonf}} < 0.001$, $d = 1.60$, CI95 = 0.93 – 2.31 ; HC vs. AD $p_{\text{Bonf}} < 0.001$, $d = 1.56$, CI95 = 0.83 – 2.34 ; HC-v vs. CD $p_{\text{Bonf}} < 0.001$, $d = 1.48$, CI95 = 0.81 – 2.13 ; HC-v vs. AD $p_{\text{Bonf}} = 0.001$, $d = 1.60$, CI95 = 0.56 – 2.45 , see Fig. 4D), while no difference emerged between the two CI groups nor between the two HC groups ($p_{\text{Bonf}} = 1.00$, statistics reflect the results of both CD vs. AD and HC vs. HC-v comparisons).

Following previous evidence linking neural speech tracking and comprehension^{30–33}, we tested whether the degree of neural tracking (the TRF magnitude) would be associated with children's comprehension scores by performing a series of linear regression models (at all sensors) to predict the point-by-point TRF values from response accuracy, groups, and their interaction, accounting for the impact of participants' age ($\text{TRF}(c,t) = \text{Accuracy} + \text{Group} + \text{Accuracy} \times \text{Group} + \text{Age}$). Note that regarding the CI data, the model was conducted on the CI's TRFs shifted in time because the latency of the $P1_{\text{TRF}}$ was not related to behavioral performance. The linear regression $P1_{\text{TRF}}$ latency with behavioral accuracy revealed a complete absence of a relationship when both groups were considered ($p = 0.13$) and also when it was assessed within HC and CI separately ($p = 0.79$ and $p = 0.91$, respectively). When running regression models to explain TRF magnitudes with behavioral performance, a significant relationship emerged at a longer timescale, between 130 and 330 ms (averaging across all sensors, $p_{\text{FDR}} < 0.05$, mean $R^2 = 0.073$, SE = 0.009, see Fig. 4E and S6A for detailed results in Supplementary Materials 2.7). When we computed the regression model separately for each group, beta accuracy (averaged between 130 and 330 ms) revealed a significant impact of children's comprehension on TRF magnitude at both frontal and posterior sensors in HC children, while in the children using CI, this relationship emerged only at posterior sensors (see Figs. 4F and S6B and C for more details). These findings revealed that the magnitude of the TRF second phase ($N2_{\text{TRF}}$ between 130 and 330 ms), i.e., long timescale neural tracking, was associated with children's comprehension scores. Precisely at this latency (note, once accounting for their neural tracking delay by time-aligning the CI's and HC's TRFs), the CI group had markedly reduced neural speech tracking. Taken together, these findings suggest that this altered dynamic could account for their comprehension deficit.

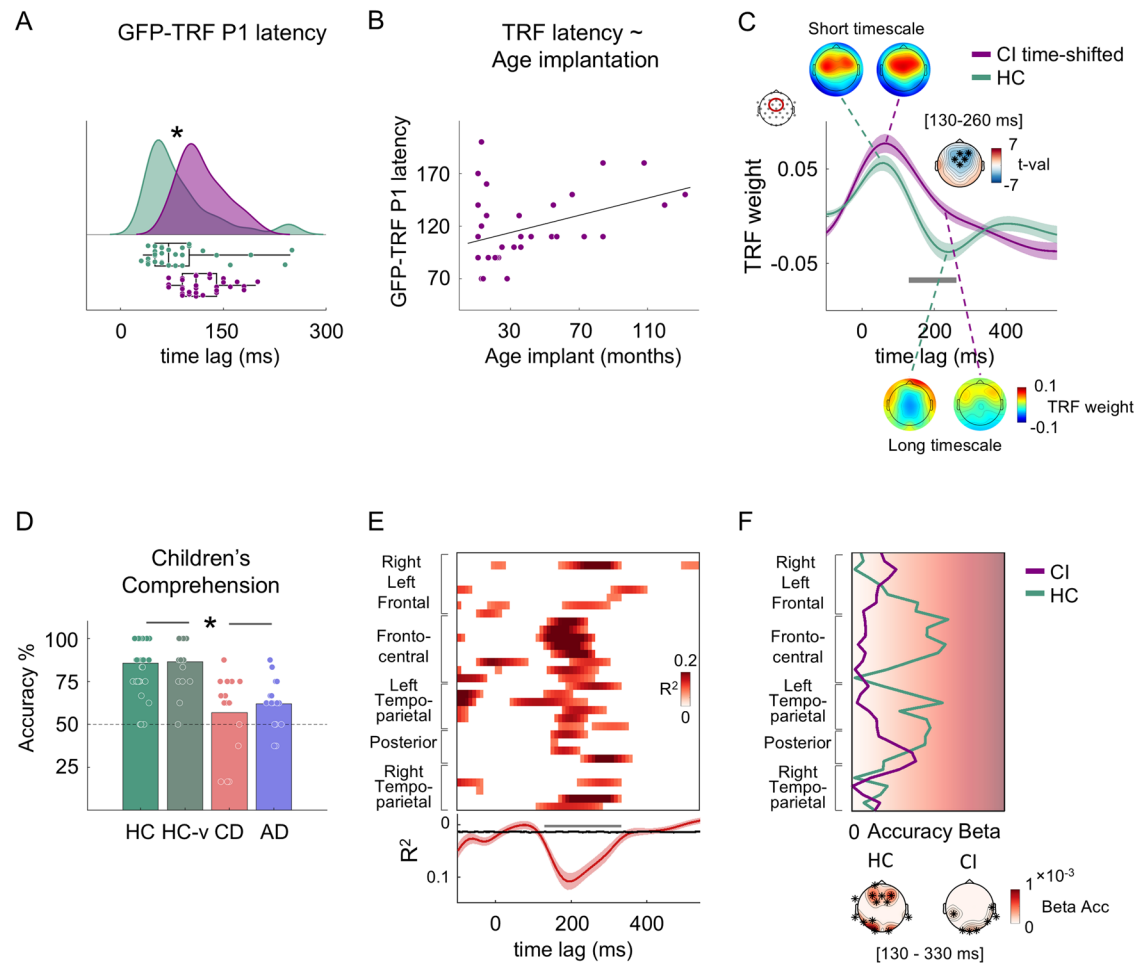


Fig. 4 | Biomarkers of altered neural speech tracking in children with CIs. **A** The shaded areas show the data distributions of the first peak latency (olive color for HC and magenta color for CI); below, in the corresponding boxplots, the central black line represents the median, the box corresponds to the interquartile range, and the whiskers extend to the 2nd and 98th percentiles. The first peak latency (GFP-TRF P1) was delayed in CI compared to HC ($t_{(67)} = -2.99; p = 0.004$). **B** Linear regression predicts the latency of the neural tracking first peak (GFP-TRF P1) from the age at which children using CI received the first implant (auditory restoration). **C** The plot shows the TRF for HC and CI at frontocentral sensors (the continuous line represents the group mean, and the shaded area is the SE, $N_{HC} = 37$, and $N_{CI} = 32$). 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 topographies show the similarity at a short timescale (P1_{TRF}), while a clear difference is evident at a longer timescale (N2_{TRF}). After 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. **D** Mean comprehension scores in HC, HC-v, CD, and AD groups are shown, and error bars represent the SE. Children with CIs (both

CD and AD) had lower scores than the HC and HC-v groups. **E** The plot shows the R^2 values with $p < 0.05$ for the regression model in which the point-by-point TRF values were predicted by behavioral accuracy, group, the interaction between them, and age. 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 130 and 330 ms ($p_{FDR} < 0.05$). **F** The plot shows the absolute values of beta accuracy averaged within the significant time window [130–330 ms] across all sensors (neighboring channels are smoothed for visualization purposes), separately for HC and CI groups. The topographies with black asterisks highlighting significant sensors for each group are shown below. Data revealed the relationship between TRF magnitude between 130 and 330 ms and children's comprehension scores in both groups. CI group had markedly reduced neural speech tracking at this latency and a reduced association between comprehension scores and TRF magnitude.

Discussion

Despite the period of auditory deprivation and the fact that cochlear implants provide only partial input to the brain^{14,15}, our data clearly revealed the possibility of measuring neural speech tracking in children using CI and the informativeness of the associated indices.

First, our data revealed a developmental trajectory in the CI group similar to the HC group, supporting the sensitivity of neural tracking to capture functional changes in both typical and atypical development (see refs. 6,34). The neural tracking became less spread in the time domain, suggesting increased neural tracking efficiency. Most importantly, the TRF remarkably overlapped in individuals with congenital deafness (CD) and acquired deafness (AD). Moreover, at a short timescale, neural speech tracking had a similar magnitude between HC and children with CIs. This

evidence favours a model in which the brain is endowed with neural systems for the early auditory processing of the tracking of speech, which are partially resilient to auditory deprivation during the first year of life. These findings expand previous observations that certain basic computations underpinning language tracking are already available at birth³⁵. However, our data also revealed that neural speech tracking is affected by atypical auditory experiences in children using CI. When we contrasted the neural tracking between CI and HC groups, we observed two main differences associated with different hierarchical levels of speech processing^{32,36}: the CI's neural tracking was characterized by longer latency at a short timescale of neural activity (P1_{TRF}) and dampened magnitude at a longer timescale (N2_{TRF}). Importantly, these anomalies in CI's TRF dynamics were related to biological differences in brain processing: they were not merely caused by

degraded speech stimulations nor by the impact of the CI electrical response artifact. Indeed, neural tracking was similarly affected in CI compared to hearing children, irrespective of whether the latter were exposed to natural speech or vocoded speech. Moreover, the artifact TRF computed from cochlear implants without brain activity could not account for the main results.

Delayed neural tracking in children with CIs

While at a short timescale ($P1_{TRF}$), the neural speech tracking had a similar magnitude, it was markedly delayed in children using CI compared to HC individuals (about 60 ms). This is coherent with previous neurophysiological observations employing simple speech units, such as syllables (longer latency of P1 ERP in children using CI with respect to hearing controls^{25,27}). On the same line, studies measuring neural speech tracking in adult individuals with hearing impairments uncovered that the greater the auditory deficit, the more delayed the first neural tracking response ($P1_{TRF}$ ³⁷), and in children with hearing aids, weaker stimulus intensities delayed the $P1_{TRF}$ ³⁸. Coherently, in hearing adults and children, difficult acoustic conditions, such as when speech is presented with background noise^{37,39,40} or when the syllabic rate increases⁴¹, have been shown to delay the short timescale of neural tracking. Noteworthy, our data on hearing children who listen to vocoded speech showed a tendency for a delayed $P1_{TRF}$ (100 ms, see Figs. S2C, S3C, and S4) with respect to those who listened to normal speech (60 ms). Taken together, this evidence supports the hypothesis that neural tracking delay at a short timescale ($P1_{TRF}$) is a marker for the efficiency with which the brain processes acoustic speech properties during naturalistic continuous stimulation³⁷. Therefore, the delayed $P1_{TRF}$ measured in children using CI could reflect a lower efficiency at the early level of speech processing involved in extracting acoustic features.

Crucially, the TRF delay measured here was associated with the age of cochlear implantation. Coherently with seminal electrophysiological studies employing short-lived speech sounds^{23–26}, our results highlighted the pivotal role of age at which auditory restoration takes place for the development of speech tracking. The later the child receives the implant, the more delayed the first TRF peak ($P1_{TRF}$), which represents the 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 developing basic auditory responses to syllables within the first 3.5–4.0 years of age^{42,43}, strongly advocating for early implantation. In the case of children implanted before 3.5 years of age, 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^{25,44}. Here, we found further support for these observations, showing experience-dependent effects associated with the timing of implantation, revealing that they also seem to emerge for early, sensory-based components of the neural tracking of continuous speech. Importantly, our data shows that the benefit of early implantation was evident only from 21 months of age, suggesting that other variables affecting the development of neural circuits subtending neural speech tracking might play a role.

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

The neural tracking magnitude occurring at a long timescale of brain activity ($N2_{TRF}$, between 130 and 330 ms) was associated with higher comprehension scores. This result is consistent with recent evidence suggesting that the magnitude of neural tracking (at similar timescale) is linked with comprehension of continuous speech³⁰. 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 impaired³¹. Even accounting for the delay, by time-aligning CI's and hearing control's neural data, CI individuals had hampered tracking at this long timescale ($N2_{TRF}$, see Fig. 4C). Moreover, the $N2_{TRF}$ magnitude tended to increase in hearing children listening to vocoded compared to normal speech (see Fig. S4C). A recent study³² demonstrated that at this latency, the

TRF magnitude has a quadratic relationship with speech intelligibility. With medium challenging situations, the TRF magnitude increased, while when comprehension became too difficult, the TRF magnitude decreased. According to this hypothesis, the behavioral performance of both CD and AD groups was markedly impaired, while comprehension was not altered in HC-v. These findings provide the first evidence for a possible identification of a biomarker to assess natural speech comprehension in children using CIs. That is, the TRF magnitude at this latency range ($N2_{TRF}$) could be employed to verify children's speech understanding, especially when behavioral measures are difficult to acquire, as in the case of infants, and to estimate the development of higher-level speech processing after implantation. It remains to be ascertained which processes underly comprehension deficits in children with CI, especially as we only acquired a basic behavioural index of speech understanding. However, neural tracking represents a valid approach to assess the development of speech processing following implantation. The latency of short timescale neural processing ($P1_{TRF}$) could be used to assess the efficiency at the early level of speech processing involved in extracting acoustic features. The magnitude at a longer timescale ($N2_{TRF}$) could be used for estimating comprehension.

It is important to stress that the present findings are the result of the interaction between biological factors and experience: (i) neural tracking in the CI groups followed a developmental trajectory similar to hearing peers, (ii) the latency of CI's neural tracking was affected by the age at which children received the implant, (iii) the dampened neural tracking occurring at long timescale was linked with their speech comprehension deficit and finally, (iv) neural response alterations in children using CI could not be ascribed to the degraded stimulation cochlear implants convey nor to their electrical artifacts. Taken together, these findings revealed that altered CI's TRF dynamics result from their atypical auditory experience. It remains to be clarified to what extent results can be associated with different neural processing strategies implemented by individuals with CI to track continuous speech with respect to hearing children (see, for example refs. 45,46),.

Limitations

Despite results pointing toward the crucial role of early auditory restoration in mitigating atypical auditory development and possibly ameliorating speech processing efficiency, we cannot provide conclusive interpretations as to what causes the high variability of neural tracking latency (GFP-TRF P1) observed across children implanted before 21 months of age (latency from 70 to 200 ms). This certainly represents a relevant issue to address in the future. The etiology of children implanted before 21 months of age comprised different profiles (e.g., GJB2 connexin 26, congenital CMV, Waardenburg syndrome, and perinatal complications associated with prematurity). Most of them had bilateral profound congenital deafness onset (ten out of eleven, see Supplementary Materials Fig. S5). 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 birth⁴⁷. Intriguingly, recent evidence suggests that fetal linguistic experience shapes neural synchronization with language measured at birth⁴⁸. Moreover, even though the other variables that we collected -chronological age, age at which hearing aids were provided before implantation, and experience with the implant- did not significantly explain the variability in the earliest implanted children, other variables, such as the level of performance achieved by children with hearing aids before cochlear implantation, and more general differences in the rehabilitation treatments should be considered. The experiences provided before implantation could certainly play a role in children's brain readiness for processing auditory stimuli after implantation. Larger CI samples are needed to explore the variety of clinical variables that can potentially explain the neural speech tracking outcome.

In conclusion, the data clearly highlighted that speech tracking (of the sound envelope) leverages a robust biological predisposition, being resilient to a period of auditory deprivation in the first year of life, opening up new possibilities for using speech tracking measures in clinical and treatment

applications. However, the data also revealed a specific vulnerability of higher hierarchical levels of neural speech tracking, which are associated with comprehension. Thanks to its simple nature, neural speech tracking 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.

Materials and methods

Participants

A total of 97 children participated in the study. They were categorized according to their hearing status: cochlear implanted or hearing control children. Through simulation, we estimated the minimum sample size needed to measure neural speech tracking within a group of children. Particularly, using the data of ten hearing participants, we estimated that a group of sixteen participants was needed to reliably measure the first phase of the auditory response function ($P1_{TRF}$; see Supplementary Materials 1.1.1).

All children using CI received cochlear implantation at least six months before the experimental test to ensure a consolidated auditory experience²³. A total of forty-four children with cochlear implants were recruited at the Meyer Hospital of Florence (Italy) and the IRCCS Materno Infantile Burlo Garofolo of Trieste (Italy). To characterize CI participants, we collected the following clinical variables: the results of the newborn hearing screening with otoacoustic emissions (pass or fail), the hearing thresholds at each ear to ensure the level of deafness (moderate, severe, or profound), the hearing loss etiology and family history of hearing loss, the age at cochlear implantation, whether the child was provided with hearing aids and when, the child first language, and gestational and birth information. Most importantly, this information allows us to estimate whether the profound bilateral deafness was congenital (CD participant) or started after the first year of life (AD participant). The final sample of CI participants comprised thirty-two children. Sixteen CI were classified as CD children (mean age = 8.81 years; SD = 3.52, eight females and eight males) and sixteen as AD children (mean age = 9.17 years; SD = 3.15, nine females and seven males, see Fig. 1). 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 to -1.53). Importantly, no difference emerged between CD and AD in their experience with the implant ($t_{(30)} = 1.53$; $p = 0.136$). The median age at cochlear implantation was 15 months (IQR = 12, range: 11–132) for the CD group and 48 months (IQR = 36, range: 17–120) for the AD group. See Supplementary Materials 1.1.2 and 1.1.3 for exclusion and inclusion criteria and Supplementary Data 1 for information on the CI participants included in the analysis.

A group of age- and sex-matched hearing children was recruited as the control group (HC, $N = 37$; mean age = 9.04 years; SD = 4.10, seventeen females and twenty males). Moreover, a group of sixteen hearing children was recruited to listen to vocoded-speech stimuli (HC-v, mean age = 8.36 years; SD = 1.49, six females and ten males); we collected an independent group in order to ensure the experimental sessions did not differ across participants (HC, CD, AD, and HC-v). No significant difference emerged between the four groups, neither in age ($F_{(3,81)} = 0.185$; $p = 0.906$) nor in sex ($\chi^2_{(3)} = 1.203$; $p = 0.752$). HC children were recruited in Lucca and Milan (Italy).

None of the children included in the final sample 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, three HC, and one HC-v participants were bilingual). The study was approved by the local Ethical Committee (“Comitato Etico congiunto per la ricerca della Scuola Normale Superiore e della Scuola Superiore Sant’Anna” and “Comitato Etico Regione Toscana – Pediatrico dell’Azienda Ospedaliera Universitaria Meyer” see Supplementary Materials 1.1.4). 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) and all ethical regulations relevant to human research participants were followed.

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 to provide each participant with appropriate speech materials. Three different age ranges, 3–6, 7–10, and 11–15 years old, were defined according to Italian school stages. 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) using an iPhone7 and an external condenser microphone (YC-LM10 II, Yichuang). After a brief preprocessing (see Supplementary Materials 1.2.1), the stimuli were presented to participants using Psychopy® software (PsychoPy3, v2020.1.3).

Speech sound was delivered by a single front-facing loudspeaker (Bose Companion® Series III multimedia speaker system, country, USA) placed in front of the participants at 0° azimuth, approximately 70 cm distant from their heads. Loudness was delivered at ~80 dB SPL (Meterk MK09 Sound Level Meter), and A-weighting (to mimic the sensitivity of human hearing) was applied to sound pressure levels.

The main hearing control group (HC) and CI participants were exposed to normal speech stimuli. However, cochlear implants provided degraded acoustic stimulation to participants with CI. Thus, we also created a vocoded version of the stimuli that were presented to the other group of hearing children (HC-v) to estimate the selective impact of receiving a degraded acoustic stimulation in the absence of auditory deprivation. To simulate acoustically what CI users experience through their devices, we created a vocoded version of our stories with the vocoder function (<https://github.com/egaudrain/vocoder>), applying the frequency spacing as in Cochlear devices and implementing the ACE processing strategy.

Task and experimental procedure

Participants were asked to listen carefully to the stories while looking at a computer screen placed in front of them. A white cross was displayed at the beginning of each story; 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 colour was randomly generated and changed every 1 to 20 s to keep the children’s gazes attracted throughout the story. At the end of each story, children were asked to answer an ad-hoc questionnaire. Each questionnaire comprised two 2-alternative-forced-choice questions (see Supplementary Materials 1.2.2 for more details). The comprehension questions were inserted to induce children to pay attention to the story and assess their comprehension with a basic and quick measure. For each participant, four stories were randomly selected from the pool of ten; some children (eleven HC participants and ten CI) listened to only three stories as they did not want to participate further. Their electrophysiological (EEG) activity was recorded during the entire experimental session.

EEG recording and preprocessing

EEG data were collected using a Brain Products system (ActiCHampPlus) with elastic caps (Easy cap Standard 32Ch actiCAP snap) for children having 32 active channels (500 Hz sampling rate) and recorded with *Brain Vision Recorder*. Note that for CI participants, electrodes placed very close to the magnet 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 preprocessed offline using the EEGLAB toolbox⁴⁹ (Version 14.1.2) with Matlab 2019b, implementing a validated preprocessing pipeline^{50,51}.

Prototypical artifact cleaning. Continuous EEG recordings were low-pass filtered (cut-off = 40 Hz; window type = Hanning; filter order = 50), downsampled to 250 Hz, and high-pass filtered (cut-off = 1 Hz; window type = Hanning; filter order = 500). The filtered downsampled data were segmented into consecutive 1 s epochs. Noisy segments were removed using joint probability (threshold across all channels = 3 SD⁵²). To remove prototypical artifacts (e.g., blink and eye movements), data were submitted to Independent Component Analysis (ICA, based on the extended Infomax^{53–55}). The computed ICA weights were applied to the continuous raw unfiltered data^{50,51}. Components associated with blinks and eye movement artifacts were identified using CORRMAP (version 1.03), 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 removed⁵⁶. 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. We developed a method to clean CI electrical artifacts suited for the measurement of neural speech tracking that involves associating dynamic changes in speech features (e.g., envelope) to changes in the EEG data at specific time lags. We expected CI artifacts to occur at around 0 ms time lag^{57,58}. We combined a data decomposition approach, the Second Order Blind Identification (SOBI, see ref. 29), with an algorithm to remove components classified as containing mainly CI artifact signals. Particularly, to search for CI activity that could generate artifacts in the EEG data, we performed the following steps: (i) we decomposed the EEG recordings into components (using the SOBI algorithm) with the purpose of separating physiological and noise sources; (ii) we applied the Temporal Response Function (TRF) approach (see “Estimation of TRF” section below) to each SOBI component to obtain a set of component-TRFs; (iii) we modeled normalized SOBI by fitting a set of Gaussians (one peaking before zero, and up to five peaking after zero); (iv) we identified artifacts component-TRFs with about zero lag by using a minimal set of parameters (i.e., the ratio between R^2 of the Gaussian fitted before zero and R^2 Gaussians fitted after zero, and the *beta* of the Gaussian fitted before zero) extracted from component-TRFs in HC group 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%; (v) we discarded the identified CI artifact SOBI components and reconstructed back the cleaned EEG recordings (for a more detailed explanation of these steps see Supplementary Materials 1.3.1).

Filtering, removing bad channels, and epoching. After the removal of ICA associated with prototypical artifact and SOBI components associated with CI artifacts, unfiltered data cleaned from artifacts of both 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 = 0.1 Hz, window type = Hanning; filter order = 5000). Noisy channels were identified based on the automatic bad channel detection algorithm (*clean_channels* function of *clean_rawdata2.1* 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 $\pm$ SD, in CI participants: 1.88 ± 1.60 , in HC: 2.30 ± 1.15). Disconnected channels near the magnet 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^{59,60}. Finally, preprocessed EEG data of each story listened to was epoched (2.5 min starting from few seconds after the beginning of the story to avoid stimulus onset response; see Supplementary Materials 1.3.2

for more details), downsampled to 100 Hz, and segmented into 50 s trials, resulting in a total of 12 trials per subject (or nine for the children in which we collected three instead of four stories). Trials were created to perform a cross-validation procedure in the analysis. Data were z-scored to optimize the cross-validation procedure while estimating the regularization parameter⁶¹.

Extraction of the speech envelope

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 in which they were presented to each participant and segmented into corresponding 50 s trials, resulting in twelve trials per subject (or nine trials for subjects who have listened to only three stories). The speech envelopes were downsampled to 100 Hz to match the EEG data and normalized by dividing each amplitude value by the maximum one to optimize the estimation of the regularization parameter⁶¹.

Estimation of TRF

The forward model. To investigate how the children’s brains tracked the speech envelope, we used a linear forward model known as temporal response function (TRF), incorporated in the mTRF toolbox⁶¹ (Version 1.5). TRF can be seen as a filter that describes the mapping between ongoing stimulus features (here, envelope) and ongoing neural activity; in other words, the TRF reveals the neural response associated with a continuous stimulus (see Supplementary Materials 1.4 for details).

We fitted separate TRF models at the single subject level to predict the neural 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 10 ms. The TRF at -100 ms time lag represents how the amplitude change of the speech envelope affects the EEG response 100 ms earlier, while the TRF at 600 ms time lag represents how the amplitude change of the speech envelope affects the EEG response 600 ms later. A leave-one-out cross-validation procedure was used to train the model. 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.

Importantly, the regularization parameter (λ) was estimated to avoid overfitting in the regression model. The identified λ value for the envelope model was 10^4 (see Supplementary Materials 1.4.1 for more details). This value emerged to be the one with the minimum squared error (MSE) for most participants and to generalize results, we kept λ constant across all channels and participants.

Estimation of the null effect (null-TRF). To verify that neural tracking was greater than a null effect within each group, we computed a null-TRF model for each participant⁶². We permuted the 50-second pairs of trials to obtain mismatched envelope and EEG response pairs, and the TRFs were fitted on these randomly mismatched trials of speech envelopes-EEG responses (*mTRFpermute* function with 100 iterations^{61,63}). Then, all these null-TRF models 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.

The artifact TRF. To measure the cochlear implant electrical artifacts, we recorded the EEG activity during stimuli presentation on a phantom head⁶⁴ with cochlear implants inserted below a conductive gel. These data were preprocessed and TRFs were then computed on the electrical signals produced by the CIs in the absence of neural activity (see Supplementary Materials 1.4.2 for detailed information).

Statistics and reproducibility

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.

Behavioral measures

To assess any difference in children's comprehension, we computed the accuracy percentage (correct answers) for each participant, and using IBM SPSS Statistics 20, we ran a univariate ANOVA with Group (HC, CD, AD, HC-v) as a between-participant factor. We performed post-hoc Bonferroni-corrected t -tests to evaluate the difference between groups' comprehension scores.

Encoding model (TRF)

Assessing the existence of neural speech tracking (TRF) within each group. First of all, we assessed the existence of neural speech tracking within HC and CI groups in a frontocentral cluster of electrodes (Cz, Fz, FC1, and FC2) typically capturing auditory responses at the scalp level^{5,29,65} and which was distant from cochlear implants. In each group, we compared the frontocentral TRF and the frontocentral null-TRF by running paired t -tests every 10 ms within time lags [0–600 ms] ($q = 0.05$, FDR correction). The same analyses were also separately performed for the CD and AD subgroups and the HC-v group. We also performed cluster-based permutation tests⁶⁶ in the FieldTrip toolbox⁶⁷ (Version 2019) between the TRFs and the null-TRFs within each group to confirm that results are stable when testing across all electrodes. A cluster was defined along electrodes $\times$ 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 neighboring 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 of 0.05 was thus divided by 2 ($p = 0.025$) to account for a two-sided test (positive and negative clusters).

Developmental trajectory of neural speech tracking. We assessed whether neural speech tracking would follow a developmental trajectory in HC and CI groups. 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 (with an increase of sparsity) as typically observed in developmental ERP and TRF studies^{5,68}. We quantified how the brain organizes neural responses over time, providing a framework to estimate developmental changes in neural tracking and highlighting how the brain becomes more efficient and stable with age and experience. Specifically, we hypothesized that with increasing age, the TRF amplitude would be condensed in fewer time lags, and more time lags would have no substantial neural tracking. In other words, we expected the signal's sparsity to increase with age as the TRF would have a higher variance of values over time. To this aim, first, we computed at the single participant level the Global Field Power (GFP⁶⁹) of the TRFs (i.e., GFP-TRF, see Fig. 2B) for a better estimate of the signal's dynamic and to avoid a space-dependent index⁷⁰. GFP results in a reference-independent descriptor of the potential field. Instead of selecting specific channels of interest, GFP allows for a more objective characterization of the temporal dynamics of the global electric field. For each GFP-TRF between -100 and 600 ms, we estimated the marginal moments (i.e., variance, mean, kurtosis, and skew), which are summary statistics that characterize the data distributions (e.g.^{71,72}). We tested with a linear model whether the z -scores of the estimated marginal moments of the GFP-TRF were associated with children's age.

Neural tracking as a function of auditory input and experience.

To test for differences in the spatiotemporal profile of TRFs between groups, we performed a series of cluster-based permutation tests with the same parameters defined above. First, we performed a test using independent-sample t -statistics between HC and CI groups without any a priori hypothesis on time or space. Then, we performed the same test between CD and AD to investigate the impact of the lack of auditory experience in the first year of life and finally between HC-v vs. CD and HC-v vs. AD to control for the impact of degraded speech provided by the CIs.

Finally, performing linear regression models, we investigated the relationship between neural tracking and CI's clinical profile. We assessed whether the latency of the early phase of neural tracking of continuous speech was associated with the age of hearing restoration (implantation age), as has been consistently found with the latency of early ERP components evoked by syllables^{23–26}. To this aim, we performed a linear regression where the age at which children received the first implant was the independent variable, and the individual latency of the GFP-TRF first peak was the dependent variable. We employed GFP to extract a more reliable peak latency⁷⁰. Then, we explored the relationship between neural tracking and speech comprehension. We expected the TRF magnitude measured across the sensors to be associated with speech comprehension^{30–32}. We performed a linear regression at each TRF time point across the whole-time window [0 – 600 ms] within each electrode (since we were interested in measuring the activity at all sensors, we used all TRFs and not the GFP, which compresses the signal into a single time-series). Particularly, the point-by-point linear regression model was run to predict TRF values from behavioral accuracy, groups (categorical variable), their interaction, and children's age to account for its impact ($\text{TRF}(c,t) = \text{Accuracy} + \text{Group} + \text{Accuracy} \times \text{Group} + \text{Age}$). Then, at each time point, we averaged the R^2 across electrodes to estimate at which latency the model can explain the TRF magnitude. To assess statistical significance, at each time point, a null distribution of R^2 was computed by shuffling the TRF values with 1000 permutations to compare the actual averaged R^2 effect to the null averaged R^2 distribution. The empirical p -values obtained were corrected in time with FDR.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The datasets analysed during the current study (i.e., individual preprocessed data used to compute TRFs) are available in the Mendeley repository "Resilience and vulnerability of neural speech tracking in children with cochlear implants" within the subfolders data in the folder 02_HC_Encoding_model and 03_CI_Encoding_model, <https://data.mendeley.com/datasets/nzg5g2gzrd/273>. The raw data are not publicly available because participants are minors but can be obtained from the corresponding author upon reasonable request.

Code availability

The code to perform all statistical analyses are shared in this Mendeley repository⁷³. The readme file explains all the steps to reproduce the results. EEG data were preprocessed with Matlab 2019b using EEGLAB toolbox (Version 14.1.2). To compute the encoding model, we used mTRF toolbox (Version 1.5), and cluster-based statistic were performed using fieldtrip-20190419. With Matlab 2022b, we computed effect sizes, linear regression model, and a piecewise-linear regression using the Adaptive Regression Splines toolbox (Version 1.13.0). Joyplot and boundedline plugging were used to plot the data.

Received: 8 October 2024; Accepted: 19 February 2025;
Published online: 01 March 2025

References

- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J. & Ekelid, M. Speech Recognition with primarily temporal cues. *Sci.* **270**, 303–304 (1995).
- Drullman, R., Festen, J. M. & Plomp, R. Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* **95**, 1053–1064 (1994).
- Drullman, R., Festen, J. M. & Plomp, R. Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* **95**, 2670–2680 (1994).
- Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C. & Burnham, D. Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Sci. Rep.* **8**, 1–8 (2018).
- Jessen, S., Fiedler, L., Münte, T. F. & Obleser, J. Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. *Neuroimage* **202**, 116060 (2019).
- Ortiz Barajas M. C., Guevara R., Gervain J. The origins and development of speech envelope tracking during the first months of life. *Dev. Cogn. Neurosci.* **48** (2021). <https://doi.org/10.1016/j.dcn.2021.100915>
- Attaheri, A. et al. Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *Neuroimage* **247**, 118698 (2022).
- Reh, R. K. et al. Critical period regulation across multiple timescales. *Proc. Natl. Acad. Sci. USA* **117**, 23242–23251 (2020).
- Werker, J. F. & Hensch, T. K. Critical periods in speech perception: new directions. *Annu. Rev. Psychol.* **66**, 173–196 (2015).
- Bottari, D. & Berto, M. Three factors to characterize plastic potential transitions in the visual system. *Neurosci. Biobehav. Rev.* **126**, 444–446 (2021).
- Dehaene-Lambertz, G. & Spelke, E. S. The infancy of the human brain. *Neuron* **88**, 93–109 (2015).
- Reh, R. K., Hensch, T. K. & Werker, J. F. Distributional learning of speech sound categories is gated by sensitive periods. *Cognition* **213**, 104653 (2021).
- Hensch, T. K. Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* **6**, 877–888 (2005).
- Winn M. B., Nelson P. B. Cochlear Implants. *Oxford Research Encyclopedia of Linguistics*. Published online August 31, 2021. <https://doi.org/10.1093/ACREFORE/9780199384655.013.893>.
- Pavani, F. & Bottari, D. Neuroplasticity following cochlear implants. *Handb. Clin. Neurol.* **187**, 89–108 (2022).
- Gates, G. A. et al. Cochlear Implants in Adults and Children. *JAMA* **274**, 1955–1961 (1995).
- Clark G. *Cochlear Implants: Fundamentals and Applications*. (Springer New York, 2003).
- Tamati, T. N., Pisoni, D. B. & Moberly, A. C. Speech and language outcomes in adults and children with cochlear implants. *Annu. Rev. Linguist.* **8**, 299–319 (2022).
- Houston, D. M., Stewart, J., Moberly, A., Hollich, G. & Miyamoto, R. T. Word learning in deaf children with cochlear implants: effects of early auditory experience. *Dev. Sci.* **15**, 448–461 (2012).
- Niparko, J. K. et al. Spoken language development in children following cochlear implantation. *JAMA* **303**, 1498–1506 (2010).
- Sharma, S. D., Cushing, S. L., Papsin, B. C. & Gordon, K. A. Hearing and speech benefits of cochlear implantation in children: a review of the literature. *Int. J. Pediatr. Otorhinolaryngol.* **133**, 109984 (2020).
- Kral, A., Dorman, M. F. & Wilson, B. S. Neuronal development of hearing and language: cochlear implants and critical periods. *Annu. Rev. Neurosci.* **42**, 47–65 (2019).
- Sharma, A., Dorman, M., Spahr, A. & Todd, N. W. Early cochlear implantation in children allows normal development of central auditory pathways. *Ann. Otol., Rhinol. Laryngol.* **111**, 38–41 (2002).
- Sharma, A., Dorman, M. & Spahr, A. 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 (2002).
- Sharma, A., Dorman, M. F. & Kral, A. The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hear. Res.* **203**, 134–143 (2005).
- Sharma, A., Campbell, J. & Cardon, G. 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 (2015).
- Eggermont, J. J. & Ponton, C. W. 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 (2003).
- Fiedler, L., Wöstmann, M., Herbst, S. K. & Obleser, J. Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *Neuroimage* **186**, 33–42 (2019).
- Paul, B. T., Uzelac, M., Chan, E. & Dimitrijevic, A. Poor early cortical differentiation of speech predicts perceptual difficulties of severely hearing-impaired listeners in multi-talker environments. *Sci. Rep.* **10**, 1–12 (2020).
- Etard, O. & Reichenbach, T. Neural speech tracking in the theta and in the delta frequency band differentially encode clarity and comprehension of speech in noise. *J. Neurosci.* **39**, 5750 (2019).
- Chen, Y. P. et al. Speech intelligibility changes the temporal evolution of neural speech tracking. *NeuroImage* **268**, 119894 (2023).
- Yasmin, S., Irsik, V. C., Johnsrude, I. S. & Herrmann, B. The effects of speech masking on neural tracking of acoustic and semantic features of natural speech. *Neuropsychologia* **186**, 108584 (2023).
- Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z. & Francart, T. Speech intelligibility predicted from neural entrainment of the speech envelope. *J. Assoc. Res. Otolaryngol.* **19**, 181–191 (2018).
- Pérez-Navarro, J. et al. Early language experience modulates the tradeoff between acoustic-temporal and lexico-semantic cortical tracking of speech. *iScience* **27**, 110247 (2024).
- Ortiz-Barajas M. C., Guevara R., Gervain J. Neural oscillations and speech processing at birth. *iScience*. **26** (2023). <https://doi.org/10.1016/j.isci.2023.108187>.
- Brodbeck, C. & Simon, J. Z. Continuous speech processing. *Curr. Opin. Physiol.* **18**, 25–31 (2020).
- Gillis, M., Decruy, L., Vanthornhout, J. & Francart, T. Hearing loss is associated with delayed neural responses to continuous speech. *Eur. J. Neurosci.* **55**, 1671–1690 (2022).
- Van Hirtum, T. et al. Neural envelope tracking predicts speech intelligibility and hearing aid benefit in children with hearing loss. *Hear. Res.* **439**, 108893.a (2023).
- Van Hirtum, T., Somers, B., Verschuere, E., Dieudonné, B. & Francart, T. Delta-band neural envelope tracking predicts speech intelligibility in noise in preschoolers. *Hear. Res.* **434**, 108785 (2023).
- Ding, N. & Simon, J. Z. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J. Neurosci.* **33**, 5728–5735 (2013).
- Verschuere, E., Gillis, M., Decruy, L., Vanthornhout, J. & Francart, T. Speech understanding oppositely affects acoustic and linguistic neural tracking in a speech rate manipulation paradigm. *J. Neurosci.* **42**, 7442–7453 (2022).
- Kral, A. & Sharma, A. Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* **35**, 111–122 (2012).
- Sharma, A. & Campbell, J. A sensitive period for cochlear implantation in deaf children. *J. Matern. Fetal Neonatal Med.* **24**, 151 (2011).
- Sharma, A., Gilley, P. M., Dorman, M. F. & Baldwin, R. Deprivation-induced cortical reorganization in children with cochlear implants. *Int. J. Audio.* **46**, 494–499 (2007).
- Bidelman, G. M. et al. Age-related hearing loss increases full-brain connectivity while reversing directed signaling within the dorsal-

- ventral pathway for speech. *Brain Struct. Funct.* **224**, 2661–2676 (2019).
46. Price, C. N., Alain, C. & Bidelman, G. M. Auditory-frontal channeling in α and β bands is altered by age-related hearing loss and relates to speech perception in noise. *Neuroscience* **423**, 18–28 (2019).
 47. Lieu, J. E. C., Kenna, M., Anne, S. & Davidson, L. Hearing loss in children: a review. *JAMA* **324**, 2195–2205 (2020).
 48. Mariani, B. et al. Prenatal experience with language shapes the brain. *Sci. Adv.* **9**, eadj3524 (2023).
 49. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
 50. Stropahl, M., Bauer, A. K. R., Debener, S. & Bleichner, M. G. Source-Modeling auditory processes of EEG data using EEGLAB and brainstorm. *Front. Neurosci.* **12**, 348523 (2018).
 51. Bottari, D. et al. EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *Neuroimage* **223**, 117315 (2020).
 52. Delorme, A., Sejnowski, T. & Makeig, S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* **34**, 1443–1449 (2007).
 53. Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).
 54. Jung, T. P. et al. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* **37**, 163–178 (2000).
 55. Jung, T. P. et al. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin. Neurophysiol.* **111**, 1745–1758 (2000).
 56. Viola, F. C. et al. Semi-automatic identification of independent components representing EEG artifact. *Clin. Neurophysiol.* **120**, 868–877 (2009).
 57. Deprez, H. et al. Characterization of cochlear implant artifacts in electrically evoked auditory steady-state responses. *Biomed. Signal Process Control* **31**, 127–138 (2017).
 58. Somers, B., Verschuere, E. & Francart, T. Neural tracking of the speech envelope in cochlear implant users. *J. Neural Eng.* **16**, 1–23 (2019).
 59. Mirkovic, B., Debener, S., Jaeger, M. & De Vos, M. Decoding the attended speech stream with multi-channel EEG: Implications for online, daily-life applications. *J. Neural Eng.* **12**, 046007 (2015).
 60. O’Sullivan, J. A. et al. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb. Cortex* **25**, 1697–1706 (2015).
 61. Crosse, M. J., Di Liberto, G. M., Bednar, A. & Lalor, E. C. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. *Front. Hum. Neurosci.* **10**, 1–14 (2016).
 62. Combrisson, E. & Jerbi, K. 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 (2015).
 63. Crosse, M. J. et al. Linear modeling of neurophysiological responses to speech and other continuous stimuli: methodological considerations for applied research. *Front. Neurosci.* **15**, 705621 (2021).
 64. APA Yu A. B., Hairston W. D. *Open EEG Phantom*. Published January 4, 2021. (2024). <https://doi.org/10.17605/OSF.IO/QRKA2>.
 65. Steinschneider M., Liégeois-Chauvel C., Brugge J. F. Auditory evoked potentials and their utility in the assessment of complex sound processing. In: Springer, Boston M, ed. *The Auditory Cortex*. ; 2011:535–559. <https://doi.org/10.1007/978-1-4419-0074-6>.
 66. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
 67. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
 68. Barriga-Paulino, C. I., Rodríguez-Martínez, E. I., Arjona, A., Morales, M. & Gómez, C. M. Developmental trajectories of event related potentials related to working memory. *Neuropsychologia* **95**, 215–226 (2017).
 69. Lehmann, D. & Skrandies, W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* **48**, 609–621 (1980).
 70. Michel C. M. *Electrical neuroimaging*. (Cambridge University Press, 2009).
 71. Lorenzi, C., Berthommier, F. & Demany, L. Discrimination of amplitude-modulation phase spectrum. *J. Acoust. Soc. Am.* **105**, 2987–2990 (1999).
 72. Strickland, E. A. & Viemeister, N. F. Cues for discrimination of envelopes. *J. Acoust. Soc. Am.* **99**, 3638–3646 (1996).
 73. Federici A., et al. *Resilience and vulnerability of neural speech tracking in children with cochlear implants Data set*. 2024; <https://doi.org/10.17632/nzg5g2gzrd.2>.

Acknowledgements

We want to acknowledge all the children and their families participating in this study; Francesca Collese, Irene Sanchez, and Martina Battista for data acquisition; Sara Pintonello and Costanza Fratini for patient recruitment; Edoardo Carini for artifact registration; Heather Formaini for text proofreading. Fundings: Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR) PRIN 20177894ZH to F.P., E.R., D.B. Cochlear implants used for the estimation of the electrical artifact were provided by Cochlear Italia s.r.l. through the IIRG 25-11 to D.B.

Author contributions

Conceptualization: D.B., A.F. Methodology: A.F., D.B., G.H., E.B., M.F., M.B. Investigation: A.F., M.F., E.N., A.M., D.B. Visualization: D.B., A.F., G.H. Funding Acquisition: D.B., E.R., F.P. Project Administration: D.B. Supervision: 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., F.T.

Competing interests

The authors report no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-025-07788-4>.

Correspondence and requests for materials should be addressed to Davide Bottari.

Peer review information *Communications Biology* thanks Brendan Costello, Joaquín Ordóñez, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Jasmine Pan. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025