Delta-band neural envelope tracking predicts speech intelligibility in noise in preschoolers

Tilde Van Hirtum [a,*], **Ben Somers, Eline Verschueren, Benjamin Dieudonné, Tom Francart**

[a] *KU Leuven – University of Leuven, Department of Neurosciences, Experimental Oto-rhino-laryngology, Herestraat 49 bus 721, 3000 Leuven, Belgium.*

Correspondence*:
Tilde Van Hirtum
Herestraat 49, bus 721, 3000 Leuven, Belgium
tilde.vanhirtum@kuleuven.be

ABSTRACT

Behavioral tests are currently the gold standard in measuring speech intelligibility. However, these tests can be difficult to administer in young children due to factors such as motivation, linguistic knowledge and cognitive skills. It has been shown that measures of neural envelope tracking can be used to predict speech intelligibility and overcome these issues. However, its potential as an objective measure for speech intelligibility in noise remains to be investigated in preschool children. Here, we evaluated neural envelope tracking as a function of signal-to-noise ratio (SNR) in 14 5-year-old children. We examined EEG responses to natural, continuous speech presented at different SNRs ranging from -8 (very difficult) to 8 dB SNR (very easy). As expected delta band (0.5-4 Hz) tracking increased with increasing stimulus SNR. However, this increase was not strictly monotonic as neural tracking reached a plateau between 0 and 4 dB SNR, similarly to the behavioral speech intelligibility outcomes. These findings indicate that neural tracking in the delta band remains stable, as long as the acoustical degradation of the speech signal does not reflect significant changes in speech intelligibility. Theta band tracking (4-8 Hz), on the other hand, was found to be drastically reduced and more easily affected by noise in children, making it less reliable as a measure of speech intelligibility. By contrast, neural envelope tracking in the delta band was directly associated with behavioral measures of speech intelligibility. This suggests that neural envelope tracking in the delta band is a valuable tool for evaluating speech-in-noise intelligibility in preschoolers, highlighting its potential as an objective measure of speech in difficult-to-test populations.

Keywords: Neural tracking, EEG, Children, Speech in noise, Speech intelligibility

1 INTRODUCTION

Globally, 34 million children (∼0.2% of all children aged 0 to 18 years) have disabling hearing loss, i.e., hearing loss greater than 30 decibels (dB) in the better ear (report of World Health Organization, 2021). The prevalence rises to 5% when mild and unilateral hearing losses are also considered (Wang et al., 2019). Unaddressed hearing loss has been proven to affect children's speech and language development, educational attainments and social skills. Through early detection and interventions many of these impacts can be mitigated, highlighting the importance of accurate hearing diagnostics (O'Donoghue, 2013). Evaluation of speech intelligibility is a fundamental component of hearing loss assessment and rehabilitation. It can determine speech intelligibility and discrimination of speech features, and provides insight into the perceptual abilities of an individual (Eggermont, 2017). The current gold standard in measuring speech intelligibility relies heavily on behavioral tests. While these tests are reliable and fast in healthy adults, it can be difficult to assess speech intelligibility in children. A child's abilities and limitations in the language domain as well as their cognitive abilities strongly affect the results of a behavioral test. Moreover, the type of response task, the speech material, and children's motivation and involvement should also be considered (Mendel, 2008). In addition, measures of speech intelligibility are often restricted to testing in quiet, due to a lack of appropriate and reliable speech in noise tasks (van Wieringen & Wouters, 2022). However, the high prevalence of noise in children's natural listening environments, especially in settings where children learn and play (e.g., clamorous classrooms and rowdy playgrounds), poses serious challenges for a still immature auditory system (e.g., Ambrose et al., 2014; Neuman et al., 2010). Therefore it is essential that we can measure the ability to understand speech in these noisy, everyday environments.

An objective approach, using EEG to measure neural tracking in response to natural running speech, could overcome the current challenges in pediatric hearing assessment and provide a more reliable measure of a child's speech intelligibility (for a review see

Gillis et al., 2022b). Neural tracking refers to the process by which neural responses in a listener's brain time-lock to dynamic patterns of the presented speech, such as the speech envelope. The speech envelope contains acoustical information (Rosen, 1992) and reflects phoneme, syllable and word boundaries (Peelle et al., 2013), which are critical for speech intelligibility (Shannon et al., 1995).

Indeed, a rich literature has found that during speech perception, auditory neural activity tracks the temporal fluctuations of the speech envelope in frequency bands matching the important occurrences of speech information, i.e., phrases/sentences (below 2 Hz) and syllables (2-8 Hz) (Vander Ghinst et al., 2019; Ding et al., 2016; Bourguignon et al., 2013; Ahissar et al., 2001; Gross et al., 2013; Meyer et al., 2017; Molinaro & Lizarazu, 2017). More importantly regarding diagnostic purposes, many researchers have shown that neural envelope tracking is affected by the intelligibility of the presented speech (Peelle et al., 2013; Gross et al., 2013; Di Liberto et al., 2018) and is significantly correlated with behavioral measures of speech intelligibility (Ding et al., 2014; Vanthornhout et al., 2018; Lesenfants et al., 2019a; Verschueren et al., 2021).

Most of the above-mentioned studies have been conducted in adults. Research on neural envelope tracking in children is scarce, despite the functional relevance that neural tracking might have for objectively measuring children's speech intelligibility. Moreover, cross-sectional evidence shows that auditory neural activity changes drastically during childhood (e.g., Cragg et al., 2011; Vander Ghinst et al., 2019; Schneider & Maguire, 2019; Panda et al., 2020), which makes it difficult to extrapolate outcomes from neural tracking research in adults to children. Furthermore, previous research in children mostly involved older children ($>$ 10 years) with dyslexia, focusing on the relation between neural tracking and reading development/experience. These studies show consistent coherence in the delta frequency range (0.5-4 Hz) (Molinaro et al., 2016) and theta (4-8 Hz) (Abrams et al., 2009; Molinaro et al., 2016) using natural speech. In addition, they found that dyslexic children have impaired tracking at low frequencies ($<$ 2 Hz) compared to typically developing

children (Molinaro et al., 2016; Power et al., 2013, 2016; Di Liberto et al., 2018; Destoky et al., 2022).

Although evidence of the neural tracking mechanisms in typically developing children is limited, more recent studies have been able to show successful neural tracking of continuous natural speech in both infants (<1 year) (Kalashnikova et al., 2018; Tan et al., 2022; Attaheri et al., 2022) and young children between 4 and 9 years old (Vander Ghinst et al., 2019; Ríos-López et al., 2020; Tan et al., 2022). For example, Ríos-López et al. (2020) showed that speech-brain coupling already occurs at 4 years of age in the delta-band frequency range, but not at theta frequencies. Similarly, Vander Ghinst et al. (2019) found significant speech tracking in children aged 6-9 at $<$ 1 Hz frequencies, while neural envelope tracking was reduced or even absent in the theta band.

Furthermore, neural envelope tracking in children has typically been evaluated under optimal, noiseless listening conditions. Findings from multiple, behavioral studies provided evidence that mature performance on a wide range of speech-in-noise measures is established by about 9-10 years of age, while younger children require a higher signal-to-noise ratio (SNR) to achieve adult-like performance (for a review see Leibold 2017). Thus, children seem more susceptible to the detrimental effects of noise on speech intelligibility. Only one study to our knowledge has studied neural tracking of speech-in-noise in typically developing children using magnetoencephalography (MEG). In accordance to behavioral measures, Vander Ghinst et al. (2019) found that neural tracking differs between typically developing children and adults and, that noise differentially corrupts neural tracking in children. In adults, their results showed a clear effect of noise at delta frequencies (<1 Hz), that is, a decrease in coherence as SNR decreased and speech is less intelligible. However, in children increasing noise decreased coherence more strongly than in adults. Additionally, children's coherence was drastically reduced or even absent in comparison with adults in the theta band regardless of SNR. Generally, these results are in line with previous behavioral studies showing children's poorer speech intelligibility in adverse listening conditions

(Johnson, 2000; Talarico et al., 2007; Neuman et al., 2010; Leibold, 2017). However, we cannot conclude from this study that delta and/or theta coherence is directly related to speech-in-noise intelligibility, since only indirect behavioral measures (e.g., intelligibility rating) were included. In addition, MEG-based recordings were used and various practical aspects of MEG in its current form (e.g., cost) pose a limitation to its large-scale usability in clinical practice, compared to electroencephalography (EEG) (Destoky et al., 2019).

Our recent research provides a framework for investigating neural tracking of different speech features, including the speech envelope (Lesenfants et al., 2019b; Verschueren et al., 2021; Gillis et al., 2022b), using EEG. The approach combines both linear decoding (backward-modelling) and encoding (forward-modelling) models, providing complementary information about neural tracking. The backward model is a model to reconstruct the speech envelope from the associated EEG recording, whereas the forward model predicts the EEG responses to speech and can be used to study the spatio-temporal dynamics of the response similarly to event related potentials (ERPs). This is the first study to use this method in preschoolers and we aim to investigate the validity of a measure that has previously been used only with adults (e.g., Vanthornhout et al., 2018). More specifically, in this study, we investigate (i) the effect of SNR on neural envelope tracking in preschoolers, and (ii) whether neural envelope tracking reflects speech intelligibility by evaluating the correspondence between neural envelope tracking and behavioral measures of speech intelligibility in noise. We have two specific hypotheses. First, we predicted that as SNR increases, neural envelope tracking increases, given that previous research has shown that stronger neural responses are associated with better speech intelligibility (Ahissar et al., 2001; Peelle et al., 2013; Ding et al., 2014; Vanthornhout et al., 2018). Secondly, we hypothesize that, similar to studies in adults (Vanthornhout et al., 2018; Lesenfants et al., 2019a), behaviorally measured speech intelligibility in noise is significantly correlated with our neural, objective measure.

2 MATERIALS AND METHODS

2.1 Participants

Fourteen 5-year-old children, recruited from the third year of kindergarten, participated in the experiment (6 female). All children were native Dutch speakers, had normal peripheral hearing (hearing thresholds $\leq$ 30 dB HL for frequencies from 0.5 to 4 kHz) and had normal or corrected to normal vision. None of the children were at risk for a cognitive or language delay nor had a family history of developmental disorders, as reported by the parents. The study protocol was approved by the local Medical Ethics Committee (reference no. S57102) and all parents provided written informed consent before the experiment. All children received a gift voucher for participating.

2.2 Behavioral measurements

2.2.1 Speech-in-noise intelligibility

Speech intelligibility in noise was assessed using the Leuven Intelligibility Peuter Test (Lilliput; van Wieringen & Wouters, 2022). This test consist of 20 lists of 11 three-phoneme consonant-vowel-consonant words (e.g., “bus”) uttered by a female speaker. All lists were presented in a stationary speech-weighted noise, matching the long-term average spectrum of the speaker. The noise level was fixed at 65 dB SPL, whereas the speech level was adjusted to obtain the targeted SNRs. Each child started with 1 training list at 0 dB SNR. Thereafter, different lists were presented monaurally to the right ear at four fixed SNRs: 0, -3, -6 and -9 dB SNR. If necessary, additional lists at -12 dB SNR (i.e., if scores at -9 dB SNR $>$50%) were presented. Two lists were presented per SNR. Children were instructed to recall each word as accurately as possible. The result of the Lilliput test is a phoneme score, i.e. the number of phonemes correctly repeated by the child. The average phoneme score per SNR was used for further analysis. Words were played through Peltor

H7A headphones using the software platform APEX (Francart et al., 2008) on a Samsung Galaxy Tab A tablet.

2.2.2 Receptive language

Children's receptive language skills were examined using the Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL; Dunn & Dunn, 1997). The PPVT-III-NL was administered in order to confirm the absence of a language impairment. It is a norm-referenced test that consists of selecting 1 of 4 pictures corresponding to a given word, at increasing levels of difficulty. Raw scores were calculated as the total responses correct and converted to age-adjusted standard scores (100 ± 15; mean $\pm$ SD) according to the standard values included in the test manual. The mean of the PPVT-III scores was 114 (SD = 12.82), ranging from 86 to 127 showing that all children performed well within age norms.

2.3 EEG experimental procedure

2.3.1 Speech material and procedure

During the EEG recordings, children listened to four different stories of "Little Polar Bear", the children's series by Hans de Beer, narrated by the same native Flemish speaker as the Lilliput. Each story was 10 to 12 minutes long and segmented in two-minute fragments taking sentence boundaries into account. The stories were presented in a stationary speech-weighted noise, obtained by the long-term average spectrum of all stories together. For all stories the speech level was fixed at 60 dB A, while the noise level was adjusted to obtain five SNR conditions: 8, 4, 0, -4 and -8 dB SNR. Every two-minute segment was presented randomly at a different SNR. In total, every SNR was presented three times (i.e., 6 minutes), with the exception of the easiest SNR condition (8 dB SNR). The 8 dB SNR condition was presented for an additional 6 minutes at the end of the recording session.

2.3.2 Data acquisition

EEG signals were recorded using a Biosemi ActiveTwo System at a sample rate of 8192 Hz. This system uses 64 Ag/AgCl electrodes, distributed over the scalp according to the international 10-20 system. The electrode offsets were kept between -30mV and 30mV to ensure stable recording.

All recordings were carried out in a soundproof and electrically shielded room, using a child-friendly and age-appropriate protocol. Children were comfortably seated in a chair approximately 1 m from a sound-permeable screen. Speech stimuli were presented at a sample rate of 48000 Hz through a GENELEC (8020A) loudspeaker positioned at head-height of the seated child using the software platform APEX (Francart et al., 2008) and an RME Fireface UC soundcard (Haimhausen, Germany). To mimic a storytelling scenario children watched images of the corresponding "Little Polar Bear" books projected on the screen in front of them while listening to the stories.

Children were instructed to attend to the stories and to avoid moving as much as possible. Additionally, children were accompanied by an experienced test leader in the EEG cabin to monitor alertness and movement. During each story, each child was asked 3 multiple-choice comprehension questions to improve attention. Short recesses between each two-minute segment and each different story were also used to motivate the children and, if necessary, allow them to rest or move before resuming the story.

2.4 EEG data analysis

We measured neural tracking of the speech envelope as a function of stimulus SNR using an envelope reconstruction approach (backward-modelling) and temporal response function estimation (forward-modelling). All signal processing was performed offline using Matlab R2016b (The MathWorks Inc, 2016).

2.4.1 Envelope reconstruction

To measure neural envelope tracking, we used the envelope reconstruction approach described in detail by Vanthornhout et al. (2018) and Verschueren et al. (2020). In summary, the speech envelope was extracted according to Biesmans et al. (2017) who showed good reconstruction accuracy using a gammatone filter bank followed by a power law. The acoustic envelope was then downsampled from 48000 to 256 Hz in order to decrease computation time. Next, the speech envelope was band-pass filtered between 0.5-4 Hz (delta band) and 4-8 Hz (theta band) using a Chebyshev filter (with 80 dB attenuation and 10% outside the passband) and further downsampled to 128 Hz. Similarly to the speech envelope, the EEG data was first downsampled from 8192 Hz to 256 Hz. Next, a multi-channel Wiener filter (Somers et al., 2018) was applied to the EEG data to remove common EEG artifacts such as eye blinks and muscle artifacts. Bad EEG channels were interpolated from their 5 neighbouring channels using Fieldtrip (Oostenveld et al., 2011). We then re-referenced each EEG signal to a common-average reference. Finally, the EEG data was band-pass filtered similarly to the speech envelope and further downsampled to 128 Hz.

For each SNR condition, the reconstructed envelope is obtained by applying a condition-specific linear decoder, calculated using ridge regression as implemented in the mTRF toolbox (Lalor et al., 2006). Each decoder is basically a spatiotemporal filter that combines the 64 EEG channels and their time-shifted versions from 0 to 250 ms (the integration window) into a reconstruction of the envelope. After normalization, envelope decoders were trained using a leave-one-out 6-fold cross-validation scheme: for each SNR condition, containing 6 minutes of EEG data, 5 minutes were used to train the decoder, which was then applied on the EEG of the remaining minute to obtain a 1-minute envelope reconstruction. This was repeated 6 times to obtain envelope reconstructions for all folds. All reconstructions within a SNR condition were concatenated and compared to the original envelope using a bootstrapped Spearman correlation, resulting in an envelope tracking

value for that SNR condition. The significance level of this correlation value is calculated by constructing a null distribution by correlating 1000 random permutations of the real and reconstructed envelopes with each other, and taking the 2.5 and 97.5 percentiles to obtain a 95% confidence interval.

2.4.2 Temporal response function (TRF) estimation

The envelope reconstruction approach is a powerful analysis tool, which integrates information of multiple EEG channels and their time-shifted versions to reconstruct the envelope. However, this type of analysis does not allow an interpretation of the spatial pattern of the response (Haufe et al., 2014). We therefore conducted a linear forward modelling approach, in addition to the envelope reconstruction approach. This forward modelling approach predicts EEG given a speech representation and results in a temporal response function (TRF) for each channel. A TRF represents an impulse response function of how the brain responds to the stimulus envelope. The main advantage of estimating TRFs is that their morphology provides valuable insights concerning the neural response latency, amplitude and topology across the scalp. The first signal processing steps are identical to the envelope reconstruction approach. Only the band-pass filtering was carried out within a broader frequency band (0.5-25 Hz), after which the EEG data is normalized. As with the backward-modelling approach, TRFs are computed using 6-fold cross-validation for every SNR condition, where 5 minutes are used for training and 1 minute for validation. TRFs are computed for every channel using the boosting algorithm as described by David et al. (2007), implemented in the python Eelbrain toolbox (Brodbeck, 2020). In short, this TRF estimation algorithm works by starting from an all-zero TRF model and iteratively improving this model by adding a weight at a latency that causes the largest improvement in the response prediction. The iterations terminate when no significant improvements to the model can be made anymore for a given step size by which the weights are changed. The advantage of this method is that the resulting TRF is sparse, i.e., for a well-chosen step size, only the peaks that contribute most to the response prediction are included in the TRF

while other TRF weights are zero. TRFs were computed over a latency range (integration window) between -200 and 500 ms with an adaptively decreasing stepsize. For further analysis and visualisation, the estimated TRFs were convolved with a Gaussian kernel of 9 samples long (SD=2) to smooth over the time lags. Next, smoothed TRFs were averaged across all folds.

2.4.3 Reliability of envelope reconstruction scores

To investigate potential changes in neural tracking over the course of the recording session, we calculated a measure of reliability. In particular, test-retest reliability can be used to reflect the variation in multiple measurements on the same subject under the same conditions. Here we quantified test-rest reliability by means of intraclass correlation (ICC) (Shrout PE, 1979), which is one of the most commonly used reliability measures in the neuroimaging field (Bennett & Miller, 2010). In this study, every SNR was presented three times (i.e., 6 minutes), with the exception of the easiest SNR condition. For the 8 dB SNR condition, an additional 6 minutes were presented at the end of each measurement. Test-retest reliability was estimated between the envelope reconstructions scores for the first and last 6 minutes, under the assumption of a single-measurement, absolute-agreement, two-way mixed model, i.e., ICC(A,1) (Koo & Li, 2016). ICC estimates and their 95% confidence intervals were calculated using the irr package (Gamer et al., 2019).

2.5 Statistical analysis

Statistical analysis were carried out in R (version 4.0.3; R Core Team, 2020). All tests were performed with a significance level of $\alpha = 0.05$ unless otherwise stated.

To investigate behavioral speech intelligibility in noise, the speech reception threshold (SRT; the SNR corresponding to a 50% percentage-correct score) and slope were determined. For each child individually a sigmoid function was fitted to their average phoneme scores

using the following formula:

$$score(SNR) = \frac{1}{1 + e^{-\frac{SNR - SRT}{slope}}} \tag{1}$$

To assess the relation between stimulus SNR and neural envelope tracking, we fitted a linear mixed effect model (LME) per filterband using maximum likelihood criteria with the nlme package (Pinheiro et al., 2022). The following general formula was used:

$$neuralTracking \sim \ SNR + random = participant \tag{2}$$

where "neuralTracking" represents the Spearman correlation between the real and the reconstructed envelope, TRF amplitude or latency depending on the outcome measure being investigated. In addition, a random intercept per participant was included to account for dependencies between measures from the same child. Residual plots of the selected model were analyzed to assess the assumption of normality and did not reveal any violations. All significant effects are discussed in the results section by reporting the β estimates, with the corresponding standard error (SE), degrees of freedom (df) and test statistics (t-value and p-value). Where applicable, post-hoc comparisons were carried out using a non-parametric Wilcoxon signed-rank test (2-tailed, $p < 0.01$), with Bonferroni correction for multiple comparisons.

Finally, the Spearman correlation between the envelope reconstruction scores and behaviorally measured speech intelligibility was computed to examine the relation between neural envelope tracking and speech intelligibility. Since different SNR conditions were used for the behavioral experiment, we estimated the percentage-correct scores corresponding to the SNR conditions (i.e. -8, -4, 0, 4 and 8 dB SNR) used for the EEG measurement from the individually fitted performance intensity functions.

3 RESULTS

3.1 Behavioral speech intelligibility

For each child, we fitted a psychometric curve through their average percentage correct scores per SNR from which the SRT and slope were derived. The mean of the individual behavioral SRTs was -5.52 dB SNR (SD = 0.94 dB), with an average slope of 7.99%/dB (SD = 2.39%/dB). Individual speech intelligibility scores evaluated at the different SNR conditions are shown in in Figure 1 together with the fitted performance intensity function on the group average and individual SRT estimates.

3.2 Envelope reconstruction

3.2.1 Effect of SNR on envelope reconstruction

Figure 2 shows neural tracking of the speech envelope as a function of the SNR using a 0-250 ms integration window in the delta (0.5-4 Hz) and theta band (4-8 Hz). These results were obtained by a leave-one-out cross-validation approach within each 6-minute SNR block.

In the delta band, an increase of neural envelope tracking with increasing SNR was found ($b = 0.005$, SE = 0.0006, $p < 0.001$). However, we observed that neural tracking did not show a strictly monotonic increase with SNR. To investigate in more detail if neural tracking differed significantly between adjacent SNR conditions in the delta band, post-hoc comparisons were carried out using a non-parametric Wilcoxon signed-rank test (2-tailed, $p < 0.01$), with Bonferroni correction for multiple comparisons. Comparisons between increasing SNR conditions demonstrated that neural tracking increased significantly from -8 to -4 dB SNR ($p = 0.003$; r = -0.792) and from -4 to 0 dB SNR ($p = 0.007$; r = -0.724), but not in the two conditions with the lowest noise (0 to 4 dB SNR: $p = 0.216$, 4 to 8 dB SNR: $p = 0.808$). Therefore, a sigmoid function was fitted on the data across all subjects. Its

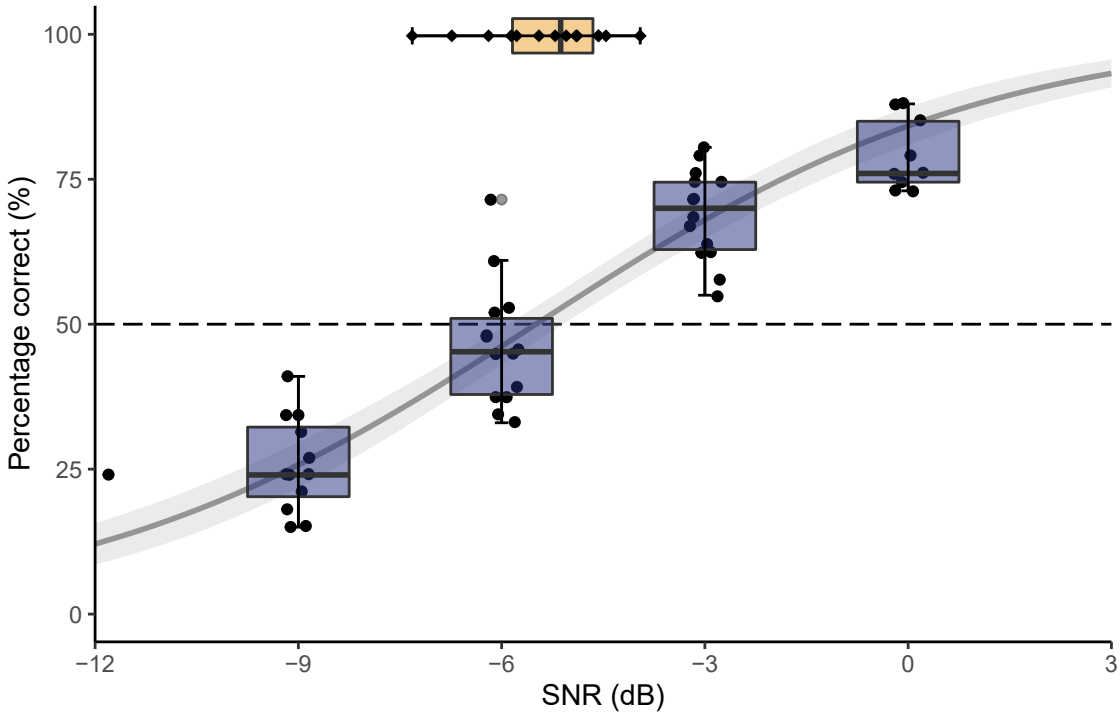


Figure 1. Behavioral speech intelligibility. Individual's speech intelligibility (percentage correct) as a function of stimulus SNR and the fitted performance intensity function of the group average. Black dots represent the individual scores. The grey line shows the sigmoid fitted to the data ± the error of the fit of the model. The orange boxplot on top shows the distribution of the SRT. Individual SRT estimates are presented as black diamonds to show variation.

298 midpoint was determined at -6.98 dB SNR and, convergence of the function (as calculated
299 by the 95th%-value of the fit) was found around 1.5 dB SNR suggesting that neural tracking
300 indeed increases with SNR until reaching a plateau between 0 and 4 dB SNR.

301 In the theta band, neural tracking again increased with SNR (b = 0.002, SE = 0.0006,
302 $p < 0.001$). However, statistically significant responses were limited. Statistically significant
303 neural tracking was observed at 8 dB SNR in 11 out of 14 children, at 4 and 0 dB SNR in 5
304 out of 14 children, and at -4 and -8 dB SNR in only 3 out of 14 children. Whereas in the
305 delta band, statistically significant neural tracking was found in 14 of 14 children down to

0 dB SNR, at 13 of 14 children at -4 dB SNR, and at 9 of 14 children at -8 dB SNR. For further analysis, our frequency range of interest was therefore restricted to the delta band.

3.2.2 Test-retest reliability of envelope reconstruction scores

Test-retest reliability of neural tracking over the course of the recording session was determined by means of ICC(A,1) for the first and last 6 minutes of speech presented at 8 dB SNR. ICC estimates were calculated based on a single-measure, absolute-agreement, 2-way mixed-effects model. We found a high degree of test-retest reliability of the envelope reconstruction values in the delta band. The average measure ICC was 0.876, with a 95% confidence interval from 0.629 to 0.96 ($F(13,10.2) = 18.3$, $p < 0.001$).

3.2.3 Effect of speech intelligibility on envelope reconstruction

In order to assess the relation between speech intelligibility and neural envelope tracking, we calculated the percentage-correct scores for each SNR condition used in the EEG experiment using the individual psychometric curves. Figure 3 shows that neural tracking in the delta band increased with increasing speech intelligibility ($r = 0.498$; $p < 0.001$, Spearman rank correlation), suggesting that the better a child can understand speech, the higher the neural envelope tracking for that child.

3.3 Effect of SNR on TRFs

The envelope reconstruction analysis showed an S-shape tendency of neural tracking as a function of SNR. That is, a monotonic increase in envelope tracking over SNR, until saturating between 0 and 4 dB SNR. To better understand the spatio-temporal properties of this effect, we calculated TRFs of every 6-minute segment (i.e., each SNR condition). Based on visual inspection of the topographies, 12 frontocentral channels were selected (presented by the black dots in Figure 4B) and averaged per subject, resulting in one TRF per SNR for each subject. Figure 4A shows a prominent peak appearing around $\sim$83 ms.

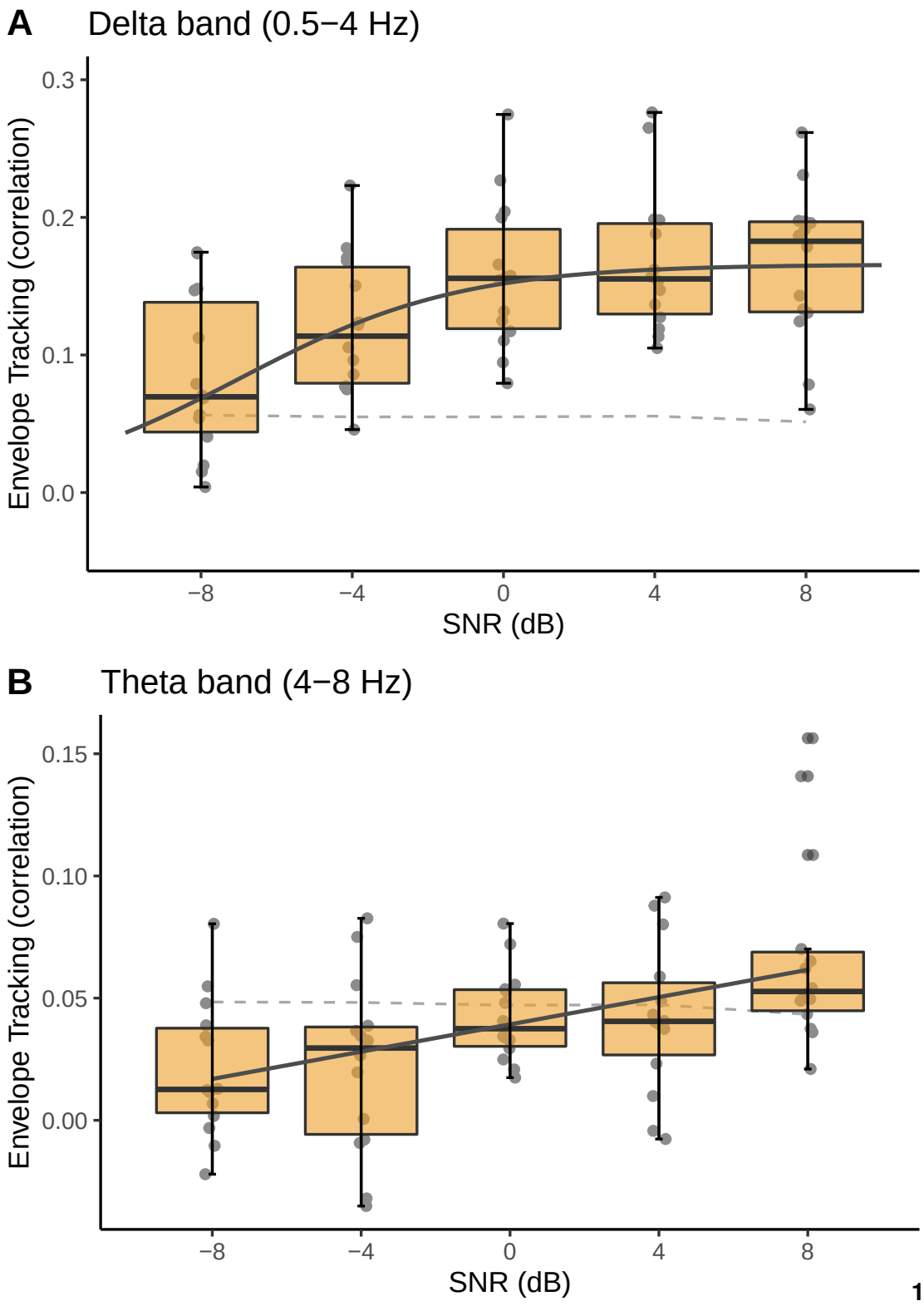


Figure 2. Neural envelope tracking as a function of stimulus SNR. **A** In the delta band, neural envelope tracking increases with increasing SNR until reaching a plateau around 0 dB SNR. The individual data points in each boxplot represent individual envelope reconstruction scores (i.e., correlation) to show variation. The solid line (dark grey) shows the sigmoid fitted to the data. The dashed line (light grey) shows the significance level of the envelope reconstruction scores. **B** In the theta band, neural envelope tracking monotonically increases with increasing SNR. However, limited statistically significant correlations were found below 8 dB SNR.

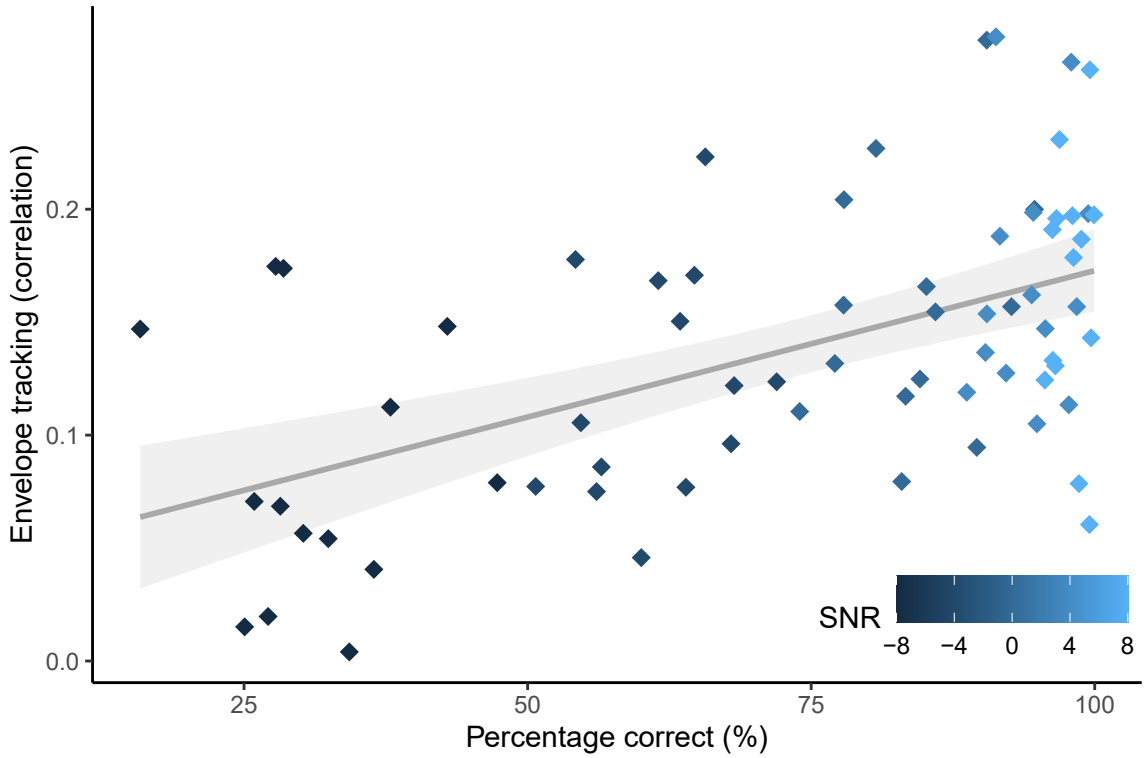


Figure 3. Relation between behavioral speech intelligibility and envelope tracking across all SNRs in the delta band. Envelope tracking increases with increasing speech intelligibility. The color gradient was used for illustrative purposes to mark SNR. The grey line shows the linear model fit. The shaded area represents the 95% confidence interval for the predictions of the linear model.

The topography of this peak (70-90 ms) is shown in Figure 4B. To investigate the influence of stimulus SNR on the peak latency and amplitude in more detail, the maximum TRF amplitude for every subject per SNR was calculated between 0 and 150 ms. As shown in Figure 4C and confirmed by statistical analysis, with increasing SNR, the peak amplitude increases (b = 4.71 $\times$ 10^{-4} , SE = 1.27 $\times$ 10^{-4}, $p < 0.001$). In contrast, the latency of this peak decreases with increasing SNR (b = -1.22, SE = 0.32, $p < 0.001$). Post-hoc comparisons between the adjacent SNR conditions demonstrated that amplitude increased significantly from -4 to 0 dB SNR (p = 0.029; r = -0.582) , whereas latency decreased from -4 to 0 dB SNR (p = 0.002; r = -0.808). We did not find significant differences between the

two noisiest (-8 to -4 dB SNR: $p > 0.05$) nor the two conditions with the lowest noise (0 to 4 dB SNR: $p > 0.05$, 4 to 8 dB SNR: $p > 0.05$)

4 DISCUSSION

We examined the influence of stimulus SNR on neural envelope tracking of natural, continuous speech in young children. We recorded EEG in 14 normal-hearing pre-schoolers while they listened to age-appropriate stories at different stimulus SNRs. Neural tracking was investigated using both envelope reconstruction and TRF analysis. Our findings demonstrated that neural tracking of the speech envelope in the delta band (0.5-4 Hz) increased monotonically with increasing SNR, until stabilizing between 0 and 4 dB SNR, suggesting that neural envelope tracking remains stable as long as speech is intelligible. By contrast, neural tracking in the theta band (4-8 Hz) appeared to be more sensitive to noise. As SNR increased, neural tracking increased. However it is important to note that our results in the theta band did not fully indicate the presence of significant envelope reconstruction scores below 8 dB SNR. Lastly, we found that neural envelope tracking in the delta band was related to behavioral speech-in-noise performance, suggesting its potential for a realistic and objective measure of speech intelligibility in clinical practice.

We hypothesized that neural tracking of the speech envelope would increase as SNR, and thus speech intelligibility, increases. We demonstrated that this is indeed the case. However, in the delta band, we did not find a strictly monotonic increase of neural envelope tracking with SNR. As the SNR decreased, envelope reconstruction scores remained stable until $<$0 dB SNR, showing a S-shape correlation over SNRs (i.e., flat, followed by a decrease as SNR further decreases). Our finding aligns with previous results reported by Ding & Simon (2013) in which the effect of noise on neural entrainment to slow temporal modulations ($<$4 Hz) was investigated. Based on MEG recordings, they found that neural responses in the delta band remained stable as the SNR decreased, until the noise background is more than twice as strong as the speech. In accordance, a more recent MEG study including children aged 6-9 (Vander Ghinst et al., 2019) found no difference in coherence to speech in

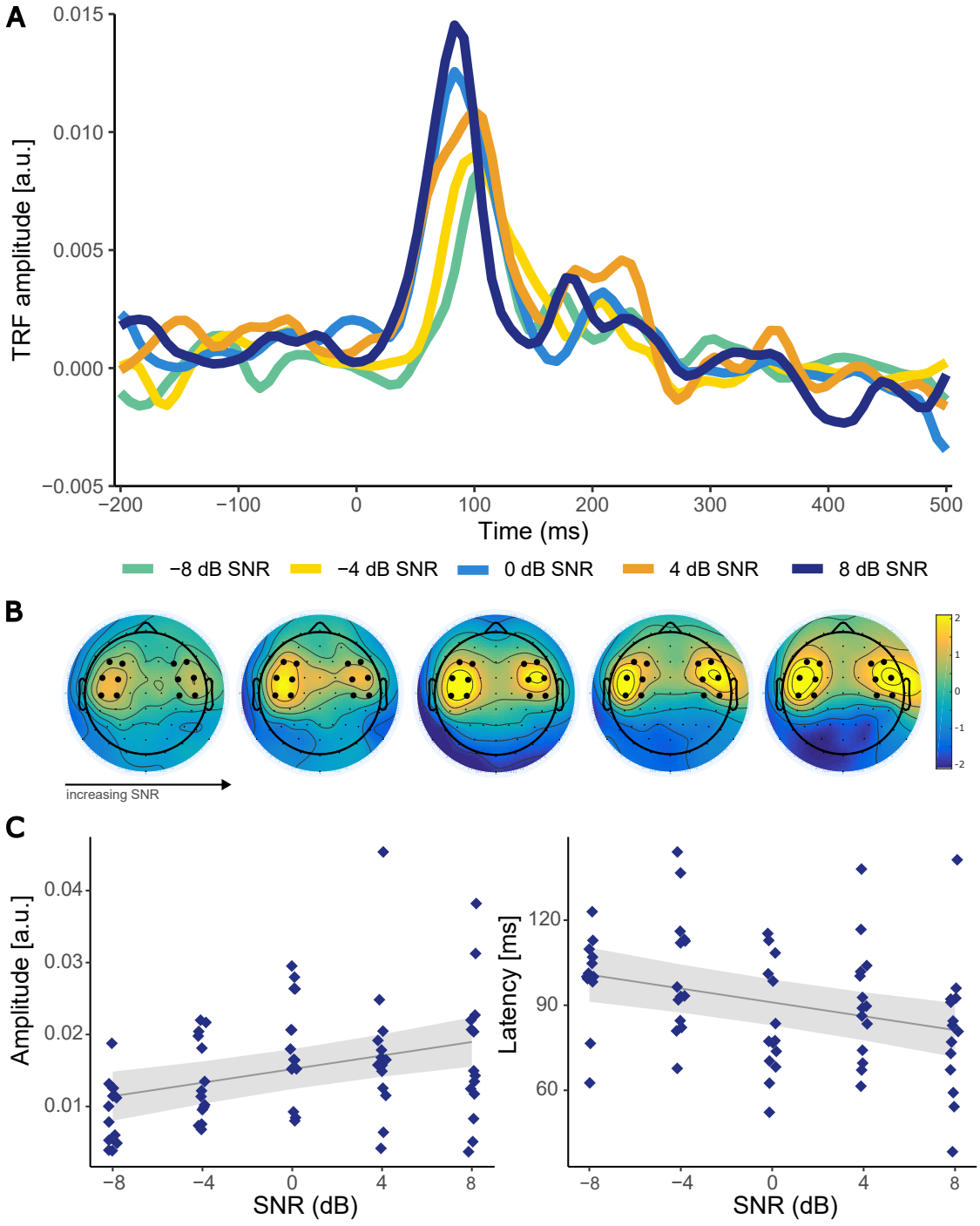


Figure 4. TRFs as a function of stimulus SNR. **A** Mean TRF activity over participants per SNR. **B** Topographies showing the associated peak topographies in the mean TRF (∼70-90 ms) per SNR. Black dots represent the centro-frontal channels selected to calculate the TRFs in the time-domain in panel A. **C** The effect of SNR on peak amplitude (left) and latency (right) of the TRF. Every diamond (blue) represents an individual participant. The grey line shows the linear model fit. The shaded area represents the 95% confidence interval for the predictions of the linear model.

different levels of background noise (SNR ranging from +5 to -5 dB) at 1-4 Hz. Additionally, these results are consistent with the previous EEG work (Iotzov & Parra, 2019), including our own (Vanthornhout et al., 2018; Lesenfants et al., 2019a; Decruy et al., 2019), yielding an increase in neural tracking with SNR. However, Lesenfants et al. (2019a) reported a monotonic increase in the delta band, with neural tracking further increasing from -1 dB SNR to noiseless conditions, thus only reaching a maximum in quiet. The discrepancy with our current results could be attributed to the differences in experimental paradigm. Previous studies (e.g., Vanthornhout et al., 2018) used a 15-minute-long recording of a story presented without noise to train a linear decoder. Then, this decoder was applied to shorter 2-minute-long segments presented at different SNRs. A recent study (Verschueren et al., 2021), however, demonstrated that reconstruction accuracy changes when using different stimulus intensities to train and test the decoder on. This suggests that a decoder is optimized to decode brain responses presented at the intensity it was trained on. Thus, when training and testing on the same stimulus parameters, neural tracking seems robust to stimulus intensity.

Taken together, we hypothesize that acoustically degrading the speech signal, such as by adding background noise, does not influence neural tracking provided that speech remains intelligible. It is well known that speech intelligibility and SNR are highly correlated. Therefore, it can be challenging to disentangle to what extent a difference in neural tracking arises from a change in the acoustics, such as different levels of background noise, or changes in speech intelligibility itself. However, our results revealed that changes in auditory stimulus characteristics (i.e., SNR) did not necessarily alter neural tracking, supporting this hypothesis. Similarly to our electrophysiological measures, behaviorally measured speech intelligibility starts to decrease significantly below 0 dB SNR, whereas average scores at 0 dB SNR exceeded 75%, reflecting good speech intelligibility for these young children (van Wieringen & Wouters, 2022). Thus, when acoustical degradation of the speech signal does not reflect a significant change in speech intelligibility, neural tracking in the delta band seems insensitive to stimulus SNR.

In contrast, neural envelope tracking in the theta band was influenced by stimulus SNR even though speech intelligibility remained high, replicating studies performed in adults (Ding & Simon, 2013; Verschueren et al., 2021). Furthermore, the number of children with significant tracking decreased with increasing noise level, suggesting that theta band tracking was more sensitive to SNR compared to the delta band. With limited noise present (i.e., 8 dB SNR), 11 of 14 children (delta band: 14 of 14 children) showed significant neural tracking of the speech envelope, whereas only 3 of 14 children (delta band: 9 of 14 children) still had significant neural responses when more noise was added (i.e., -4 dB SNR). This finding seems inconsistent with previous studies reporting significant theta tracking in adults (Bourguignon et al., 2013; Ding & Simon, 2013; Molinaro et al., 2016; Vanthornhout et al., 2018; Vander Ghinst et al., 2019; Lesenfants et al., 2019a). Yet, more recent results in young children also failed to demonstrate significant responses in the theta range (Vander Ghinst et al., 2019; Ríos-López et al., 2020).

The discrepancy between frequency bands could be rooted in developmental reasons such as the language acquisition stage of pre-school children. Previous behavioral results show that children's sensitivity to phonological units progresses from larger to smaller units, thus slow to fast information in the speech signal, as linguistic and reading skills improve (Ziegler & Goswami, 2005). However, several studies found evidence of an early-developed ability to discriminate syllables already in preterm infants (Mahmoudzadeh et al., 2013; Nittrouer, 2006). Therefore, this hypothesis is at odds with the view that tracking at theta frequencies correspond to the processing of syllabic units (e.g. Ding & Simon, 2012b; Gross et al., 2013; Ding et al., 2016). Additionally, theta responses are generally lower compared to delta tracking (e.g. Vander Ghinst et al., 2019; Verschueren et al., 2021), supporting an alternative hypothesis that delta and theta tracking reflect different speech perception mechanisms. That is, delta band neural tracking involves more higher-order speech specific processing, whereas theta band activity is more sensitive to acoustic processing of the speech signal (Molinaro & Lizarazu, 2017; Ding et al., 2016)

This explanation can be further strengthened by the fact that we found a significant correlation between speech intelligibility and neural envelope tracking in the delta band. This is in line with previous studies (Meyer et al., 2017; Vanthornhout et al., 2018; Verschueren et al., 2021, 2019; Etard & Reichenbach, 2019) consistently reporting that increased neural envelope tracking reflects better speech intelligibility. This is especially promising for clinical applications in hearing loss assessment and rehabilitation.

In addition to the envelope reconstruction analysis, our TRF analysis revealed one prominent positive TRF peak around 70–90 ms. In contrast, the TRF in adults typically contains two distinct peaks (e.g., Ding & Simon, 2012a, 2013; Vanthornhout et al., 2019; Verschueren et al., 2021, 2022). The first peak, here called the P1, occurs relatively fast with latencies around ∼50 ms and is thought to reflect mainly acoustic processing, whereas the second peak occurs at longer latencies (∼100–150 ms; N1), and might be influenced by top–down processing related to speech intelligibility and attention (Brodbeck & Simon, 2020). In line with observations of cortical auditory evoked potentials (CAEP), we believe the positive peak in this study could be related to the P1 as reported in adults. In general, CAEP traces in young children can be characterized by a large positive peak (P1), emerging between 100–150 ms, whereas components N1 and P2 emerge more gradually with maturation and might not be reliably evoked until the age of 7–9 years old Albrecht et al. (2000); Ponton et al. (2000); Ceponien et al. (1998). Besides the change in temporal pattern of the CAEP (e.g., peaks emerging with increasing age), peak latency decreases with age (Wunderlich et al., 2006). Therefore, age–related changes reflecting maturation of auditory neural processing might explain why our results did not reveal a clear N1 or P2 component.

To further investigate how SNR affects the TRFs, we compared the individual peak amplitudes and latencies at the different stimulus SNRs. Our findings demonstrated that with increasing SNR, the amplitude of P1 increased. Furthermore, the latency of P1 decreased as the SNR increased. These results dovetail with previous research with adult

participants showing that the amplitude of the early TRF peak (∼50 ms) decreases, while the latency increases continuously with SNR (Ding & Simon, 2013; Jan et al., 2022). A delay in neural responses has been hypothesized to reflect a decrease in neural processing efficiency (Bidelman et al., 2019; Gillis et al., 2022a). In addition, increased latencies have previously been related to increasing task demand such as lower stimulus intensity, increasing background noise or vocoded speech both for neural processing of continuous speech (Mirkovic et al., 2019; Verschueren et al., 2021; Kraus et al., 2021) as well as simple sounds (Billings et al., 2015; Van Dun et al., 2016; Maamor & Billings, 2017; McClannahan et al., 2019). In the current study, we observed the same effect of background noise in pre-school children. However, the results in this study only showed significant differences in peak amplitude and/or latency from -4 to 0 dB SNR and not between the lower noise conditions, similar to the previously discussed envelope reconstruction results. This, again, suggests that decreasing the stimulus SNRs only influences the amplitude and latency of the neural response if speech intelligibility is affected. This aligns with the finding of Verschueren et al. (2021) that the P1 latency increased when audibility of the stimulus decreased, but only when audibility affected speech intelligibility and listening effort.

4.1 Limitations and future directions

It is important to note that, in this study particular emphasis has been placed on neural tracking of the speech envelope. Even though the envelope is essential for speech intelligibility (Shannon et al., 1995), only taking into account basic acoustic properties as a measure of speech intelligibility might be an oversimplification. As reviewed in detail by Brodbeck & Simon (2020), neural responses can be predicted more accurately when not only considering other acoustic features, but also higher-level linguistic speech representations. Therefore, neural envelope tracking should be considered a measure to assess whether the pre-conditions for speech intelligibility are met, rather than a direct

472 measurement of speech intelligibility. Nevertheless, neural envelope tracking is shown to
473 be quite robust, therefore remaining a good candidate for use in a clinical setting.

474 A second caveat of this study is that we included normal-hearing children as a first
475 cohort, since speech intelligibility in noise had never been evaluated in such young children.
476 However, for clinical applications of our measure, future research should study the impact
477 of parameters such as hearing loss and hearing aid fitting. A number of studies have already
478 demonstrated the use of CAEPs for validation of hearing aid fitting in hearing-impaired
479 children (e.g., Chang et al., 2012; Glista et al., 2012; Baydan et al., 2019). Overall, their
480 findings showed that aided behavioral thresholds were strongly correlated with CAEP
481 responses to short non-speech as well as speech sounds. However, we believe that the use
482 of continuous speech is more ecologically relevant compared to speech sounds and could
483 potentially provide more information concerning the functionality of the hearing aid in
484 daily life.

5 CONCLUSION

485 To summarize, the results of the present study showed that neural envelope tracking
486 increases with increasing SNR. However, this increase is not strictly monotic as delta
487 band tracking converges between 0 and 4 dB SNR, similarly to the behavioral outcomes.
488 Moreover, these results were confirmed by the TRF analysis, suggesting that neural tracking
489 correlates with speech intelligibility rather than stimulus SNR. In contrast, the ability of
490 children's brain to track the speech envelope in the theta band was drastically reduced and
491 more easily corrupted by increasing noise in comparison to the delta band. Lastly, we found
492 that neural envelope tracking in the delta band was directly associated with behavioral
493 measures of speech intelligibility. Altogether, our findings provide a unique basis for a
494 behavior-free evaluation of speech intelligibility in difficult-to-test populations, such as
495 young children.

AUTHOR STATEMENT

Tilde Van Hirtum: Formal analysis, investigation, data curation, writing-original draft, visualization and project administration. **Ben Somers**: Investigation, writing-review and editing, software. **Eline Verschueren**: Investigation, writing-review and editing and project administration. **Benjamin Dieudonné**: Writing-review and editing, software. **Tom Francart**: Conceptualization, methodology, writing-review and editing, supervision and funding acquisition.

DECLARATION OF INTEREST

None.

ACKNOWLEDGEMENTS

The authors are grateful to all the children and their parents who made time to participate in this study. The authors would also like to thank Mathilde Van Hecke for her help in data acquisition. Financial support was provided by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme [grant number 637424, Tom Francart]; the KU Leuven research fund [grant number C3/20/045] and VLAIO innovation mandate (IM) [grant number HBC.2021.0203, Ben Somers].

REFERENCES

Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience*, *29*, 7686–7693. doi:10.1523/JNEUROSCI.5242-08.2009.

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, *98*, 13367–13372. doi:10.1073/pnas.201400998.

Albrecht, R., Suchodoletz, W., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clinical Neurophysiology*, *111*, 2268–2276. doi:https://doi.org/10.1016/S1388-2457(00)00464-8.

Ambrose, S., Van Dam, M., & Moeller, M. P. (2014). Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear and Hearing*, *35*, 139–147. doi:https://doi.org/10.1097/AUD.0b013e3182a76768.

Attaheri, A., Áine Ní Choisdealbha, 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. doi:https://doi.org/10.1016/j.neuroimage.2021.118698.

Baydan, M., Batuk, M., & G., S. (2019). Relationship between aided cortical auditory evoked responses and aided behavioral thresholds. *International Journal of Pediatric Otorhinolaryngology*, *125*, 98 – 102. doi:10.1016/j.ijporl.2019.05.015.

Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, *1191*, 133–155. doi:https://doi.org/10.1111/j.1749-6632.2010.05446.x.

Bidelman, G. M., Price, C. N., Shen, D., Arnott, S. R., & Alain, C. (2019). Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. *Hearing research*, *382*, 107795.

Biesmans, W., Das, N., Francart, T., & Bertrand, A. (2017). Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*, 402–412. doi:10.1109/TNSRE.2016.2571900.

Billings, C. J., Penman, T. M., McMillan, G. P., & Ellis, E. (2015). Electrophysiology and perception of speech in noise in older listeners: effects of hearing impairment and age. *Ear and hearing*, *36*, 710.

543 Bourguignon, M., De Tiège, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert,
544 P., Goldman, S., Hari, R., & Jousmäki, V. (2013). The pace of prosodic phrasing
545 couples the listener's cortex to the reader's voice. *Human Brain Mapping*, *34*, 314–326.
546 doi:https://doi.org/10.1002/hbm.21442.
547 Brodbeck, C. (2020). Eelbrain 0.34. http://doi.org/10.5281/zenodo.
548 3923991.
549 Brodbeck, C., & Simon, J. Z. (2020). Continuous speech processing. *Current Opinion in*
550 *Physiology*, *18*, 25–31. doi:10.1016/j.cophys.2020.07.014.
551 Ceponien, R., Cheour, M., & Näätänen, R. (1998). Interstimulus interval and auditory event-
552 related potentials in children: evidence for multiple generators. *Electroencephalography*
553 *and Clinical Neurophysiology/Evoked Potentials Section*, *108*, 345–354. doi:https:
554 //doi.org/10.1016/S0168-5597(97)00081-6.
555 Chang, H.-W., Dillon, H., Carter, L., van Dun, B., & Young, S.-T. (2012). The relationship
556 between cortical auditory evoked potential (caep) detection and estimated audibility in
557 infants with sensorineural hearing loss. *International Journal of Audiology*, *51*, 663–670.
558 doi:10.3109/14992027.2012.690076.
559 Cragg, L., Kovacevic, N., McIntosh, A. R., Poulsen, C., Martinu, K., Leonard, G., &
560 Paus, T. (2011). Maturation of EEG power spectra in early adolescence: a longitudinal
561 study. *Developmental science*, *14(5)*, 935–943. doi:10.1111/j.1467-7687.2010.
562 01031.x.
563 David, S. V., Mesgarani, N., & Shamma, S. A. (2007). Estimating sparse spectro-temporal
564 receptive fields with natural stimuli. *Network: Computation in Neural Systems*, *18*,
565 191–212. doi:10.1080/09548980701609235.
566 Decruy, L., Vanthornhout, J., & Francart, T. (2019). Evidence for enhanced neural tracking
567 of the speech envelope underlying age- related speech-in-noise difficulties. *Journal of*
568 *neurophysiology*, *122*, 601–615. doi:10.1152/jn.00687.2018.
569 Destoky, F., Bertels, J., Niesen, M., Wens, V., Vander Ghinst, M., Rovai, A., Trotta, N.,
570 Lallier, M., De Tiège, X., & Bourguignon, M. (2022). The role of reading experience in

atypical cortical tracking of speech and speech-in-noise in dyslexia. *NeuroImage*, *253*, 119061. doi:https://doi.org/10.1016/j.neuroimage.2022.119061.

Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiège, X., & Bourguignon, M. (2019). Comparing the potential of meg and eeg to uncover brain tracking of speech temporal envelope. *NeuroImage*, *184*, 201–213. doi:https://doi.org/10.1016/j.neuroimage.2018.09.006.

Di Liberto, G. M., Lalor, E. C., & Millman, R. E. (2018). Causal cortical dynamics of a predictive enhancement of speech intelligibility. *NeuroImage*, *166*, 247–258. doi:10.1016/j.neuroimage.2017.10.066.

Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage*, *88*, 41–46. doi:10.1016/j.neuroimage.2013.10.054.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*, 158–64. doi:10.1038/nn.4186.

Ding, N., & Simon, J. Z. (2012a). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11854–9. doi:10.1073/pnas.1205381109.

Ding, N., & Simon, J. Z. (2012b). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*, 78–89. doi:10.1152/jn.00297.2011.

Ding, N., & Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *Journal of Neuroscience*, *33*, 5728–5735. doi:10.1523/JNEUROSCI.5297-12.2013.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test–Third Edition (PPVT-III)*. (3rd ed.).

Eggermont, J. (2017). Hearing aids. (pp. 263–288). doi:10.1016/B978-0-12-805398-0.00009-8.

599 Etard, O., & Reichenbach, T. (2019). Neural speech tracking in the theta and in the
600 delta frequency band differentially encode clarity and comprehension of speech in noise.
601 *Journal of Neuroscience*, *39*, 5750–5759.
602 Francart, T., van Wieringen, A., & Wouters, J. (2008). APEX 3: a multi-purpose test
603 platform for auditory psychophysical experiments. *Journal of Neuroscience Methods*,
604 *172*, 283–293. doi:10.1016/j.jneumeth.2008.04.020.
605 Gamer, M., Lemon, J., & ¡puspendra.pusp22@gmail.com¿, I. F. P. S. (2019). *irr:*
606 *Various Coefficients of Interrater Reliability and Agreement*. URL: https://CRAN.
607 R-project.org/package=irr r package version 0.84.1.
608 Gillis, M., Decruy, L., Vanthornhout, J., & Francart, T. (2022a). Hearing loss is associated
609 with delayed neural responses to continuous speech. *European Journal of Neuroscience*,
610 *55*, 1671–1690. doi:https://doi.org/10.1111/ejn.15644.
611 Gillis, M., Van Canneyt, J., Francart, T., & Vanthornhout, J. (2022b). Neural tracking
612 as a diagnostic tool to assess the auditory pathway. *Hearing Research*, *426*, 108607.
613 doi:https://doi.org/10.1016/j.heares.2022.108607.
614 Glista, D., Easwar, V., Purcell, D. W., & Scollie, S. (2012). A pilot study on cortical
615 auditory evoked potentials in children: Aided caeps reflect improved high-frequency
616 audibility with frequency compression hearing aid technology. *International Journal of*
617 *Otolaryngology*, *2012*. doi:10.1155/2012/982894.
618 Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013).
619 Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS*
620 *biology*, *11*, e1001752. doi:10.1371/journal.pbio.1001752.
621 Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-d., Blankertz, B., & Bießmann,
622 F. (2014). NeuroImage On the interpretation of weight vectors of linear models in
623 multivariate neuroimaging. *NeuroImage*, *87*, 96–110. doi:10.1016/j.neuroimage.
624 2013.10.067.
625 Iotzov, I., & Parra, L. C. (2019). Eeg can predict speech intelligibility. *Journal of Neural*
626 *Engineering*, *16*, 036008. doi:10.1088/1741-2552/ab07fe.

627 Jan, M., Kuruvila, I., & Hoppe, U. (2022). Prediction of speech intelligibility by means of
628 eeg responses to sentences in noise. *Frontiers in Neuroscience*, *16*, 876421. doi:doi.
629 org/10.3389/fnins.2022.876421.
630 Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise.
631 *Journal of Speech, Language, and Hearing Research*, *43*, 144–157. doi:10.1044/
632 jslhr.4301.144.
633 Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018).
634 Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech.
635 *Scientific reports*, *8(1)*, 13745. doi:10.1038/s41598-018-32150-6.
636 Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation
637 coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163.
638 doi:https://doi.org/10.1016/j.jcm.2016.02.012.
639 Kraus, F., Tune, S., Ruhe, A., Obleser, J., & Wostmann, M. (2021). Unilateral acoustic
640 degradation delays attentional separation of competing speech. *Trends in Hearing*, *25*.
641 doi:10.1177/23312165211013242.
642 Lalor, E. C., Pearlmutter, B. A., Reilly, R. B., McDarby, G., & Foxe, J. J. (2006). The
643 VESPA: A method for the rapid estimation of a visual evoked potential. *NeuroImage*, *32*,
644 1549–1561. doi:10.1016/j.neuroimage.2006.05.054.
645 Leibold, L. J. (2017). Speech Perception in Complex Acoustic Environments:
646 Developmental Effects. *Journal of Speech, Language, and Hearing Research*, *60*,
647 3001–3008.
648 Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., & Francart, T. (2019a).
649 Predicting individual speech intelligibility from the neural tracking of acoustic- and
650 phonetic-level speech representations. *Hearing Research*, *380*, 1–9. doi:10.1016/j.
651 heares.2019.05.006.
652 Lesenfants, D., Vanthornhout, J., Verschueren, E., & Francart, T. (2019b). Data-driven
653 spatial filtering for improved measurement of cortical tracking of multiple representations
654 of speech. *Journal of Neural Engineering*, *16*, 066017. doi:10.1088/1741-2552/

655 ab3c92.
656 Maamor, N., & Billings, C. J. (2017). Cortical signal-in-noise coding varies by noise
657 type, signal-to-noise ratio, age, and hearing status. *Neuroscience Letters*, *636*, 258–264.
658 doi:https://doi.org/10.1016/j.neulet.2016.11.020.
659 Mahmoudzadeh, M., Ghislaine, D.-L., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J.,
660 Grebe, R., & Wallois, F. (2013). Syllabic discrimination in premature human infants
661 prior to complete formation of cortical layers. *PNAS*, *110*, 4846–4851. doi:10.1073/
662 pnas.1212220110.
663 McClannahan, K. S., Backer, K., & Tremblay, K. L. (2019). Auditory evoked responses in
664 older adults with normal hearing, untreated, and treated age-related hearing loss. *Ear*
665 *and hearing*, *40*, 1106–1116.
666 Mendel, L. L. (2008). Current considerations in pediatric speech audiometry. *International*
667 *Journal of Audiology*, *47*, 546–553. doi:10.1080/14992020802252261.
668 Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2017). Linguistic
669 Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral*
670 *Cortex*, *27*, 4293–4302. doi:10.1093/cercor/bhw228.
671 Mirkovic, B., Debener, S., Schmidt, J., Jaeger, M., & Neher, T. (2019). Effects of
672 directional sound processing and listener ' s motivation on EEG responses to continuous
673 noisy speech : Do normal-hearing and aided hearing-impaired listeners differ? *Hearing*
674 *Research*, *377*, 260–270. doi:10.1016/j.heares.2019.04.005.
675 Molinaro, N., & Lizarazu, M. (2017). Delta(but not theta)-band cortical entrainment
676 involves speech-specific processing. *European Journal of Neuroscience*, *9*, 1–9. doi:10.
677 1111/ejn.13811.
678 Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-
679 synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, *37*,
680 2767–2783. doi:https://doi.org/10.1002/hbm.23206.
681 Neuman, A., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010). Combined effects of
682 noise and reverberation on speech recognition performance of normal-hearing children

683 and adults. *Ear and Hearing*, *31*, 336–344. doi:https://doi.org/10.1097/AUD.
684 0b013e3181d3d514.
685 Nittrouer, S. (2006). Children hear the forest. *The Journal of the Acoustical Society of*
686 *America*, *120*, 1799–1802. doi:10.1121/1.2335273.
687 O'Donoghue, G. (2013). Cochlear implants — science, serendipity, and success. *New*
688 *England Journal of Medicine*, *369*, 1190–1193. doi:10.1056/NEJMp1310111.
689 Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source
690 Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data.
691 *Computational Intelligence and Neuroscience*, . doi:10.1155/2011/156869.
692 Panda, E., Emami, Z., Valiante, T., & EW, P. (2020). EEG phase synchronization during
693 semantic unification relates to individual differences in children's vocabulary skill.
694 *Developmental science*, *24(1)*, e12984. doi:10.1111/desc.12984.
695 Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in
696 Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, *23*,
697 1378–1387. doi:10.1093/cercor/bhs118.
698 Pinheiro, J., Bates, D., & R Core Team (2022). *nlme: Linear and Nonlinear Mixed Effects*
699 *Models*. URL: https://CRAN.R-project.org/package=nlme r package
700 version 3.1-161.
701 Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation
702 of human central auditory system activity: evidence from multi-channel evoked
703 potentials. *Clinical Neurophysiology*, *111*, 220–236. doi:https://doi.org/10.
704 1016/S1388-2457(99)00236-9.
705 Power, A., Mead, N., Barnes, L., & Goswami, U. (2013). Neural entrainment to rhythmic
706 speech in children with developmental dyslexia. *Frontiers in Human Neuroscience*, *7*,
707 777. doi:10.3389/fnhum.2013.00777.
708 Power, A. J., Colling, L. J., Mead, N., Barnes, L., & Goswami, U. (2016). Neural encoding
709 of the speech envelope by children with developmental dyslexia. *Brain and Language*,
710 *160*, 1–10. doi:https://doi.org/10.1016/j.bandl.2016.06.006.

R Core Team (2020). R: A language and environment for statistical computing. URL: https://www.r-project.org/.

Ríos-López, P., Molinaro, N., Bourguignon, M., & Lallier, M. (2020). Development of neural oscillatory activity in response to speech in children from 4 to 6 years old. *Developmental science*, *23(6)*, e12947. doi:10.1111/desc.12947.

Rosen, S. (1992). Tem poral information in speech : acoustic , auditory and linguistic aspects. *Philosophical Transactions of the Royal Society*, *336*, 367–373.

Schneider, J. M., & Maguire, M. J. (2019). Developmental differences in the neural correlates supporting semantics and syntax during sentence processing. *Developmental Science*, *22*, e12782. doi:https://doi.org/10.1111/desc.12782.

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

Shrout PE, F. J. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428. doi:10.1037//0033-2909.86.2.420.

Somers, B., Francart, T., & Bertrand, A. (2018). A generic EEG artifact removal algorithm based on the multi-channel Wiener filter. *Journal of neural engineering*, *15*, 036007. doi:10.1088/1741-2552/aaac92.

Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., Foenander, K., Grayden, D., & Paolini, A. (2007). Effect of Age and Cognition on Childhood Speech in Noise Perception Abilities. *Audiology and Neurotology*, *12*, 1420–3030.

Tan, J. S. H., Kalashnikova, M., Di Liberto, G. M., Crosse, M. J., & Burnham, D. (2022). Seeing a talking face matters: The relationship between cortical tracking of continuous auditory-visual speech and gaze behaviour in infants, children and adults. *NeuroImage*, *256*, 119217. doi:10.1016/j.neuroimage.2022.119217.

The MathWorks Inc (2016). MATLAB and Statistics Toolbox Release 2013a.

Van Dun, B., Kania, A., & Dillon, H. (2016). Cortical auditory evoked potentials in (un) aided normal-hearing and hearing-impaired adults. *Seminars in hearing*, *37*, 9.

van Wieringen, A., & Wouters, J. (2022). Lilliput: age-appropriate words in speech-weighted noise and in quiet for young Dutch speaking children. *International Journal of Audiology*, . doi:10.1080/14992027.2022.2086491.

Vander Ghinst, M., Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Jousmäki, V., Hari, R., Goldman, S., & De Tiège, X. (2019). Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. *The Journal of neuroscience*, *39(15)*, 2938–2950. doi:10.1523/JNEUROSCI.1732-18.2019.

Vanthornhout, J., Decruy, L., & Francart, T. (2019). Effect of task and attention on neural tracking of speech. *Frontiers in Neuroscience*, *13*. doi:10.3389/fpsyg.2019.00449.

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.

Verschueren, E., Gillis, M., Decruy, L., Vanthornhout, J., & Francart, T. (2022). Speech understanding oppositely affects acoustic and linguistic neural tracking in a speech rate manipulation paradigm. *Journal of Neuroscience*, *42*, 7442–7453. doi:10.1523/JNEUROSCI.0259-22.2022.

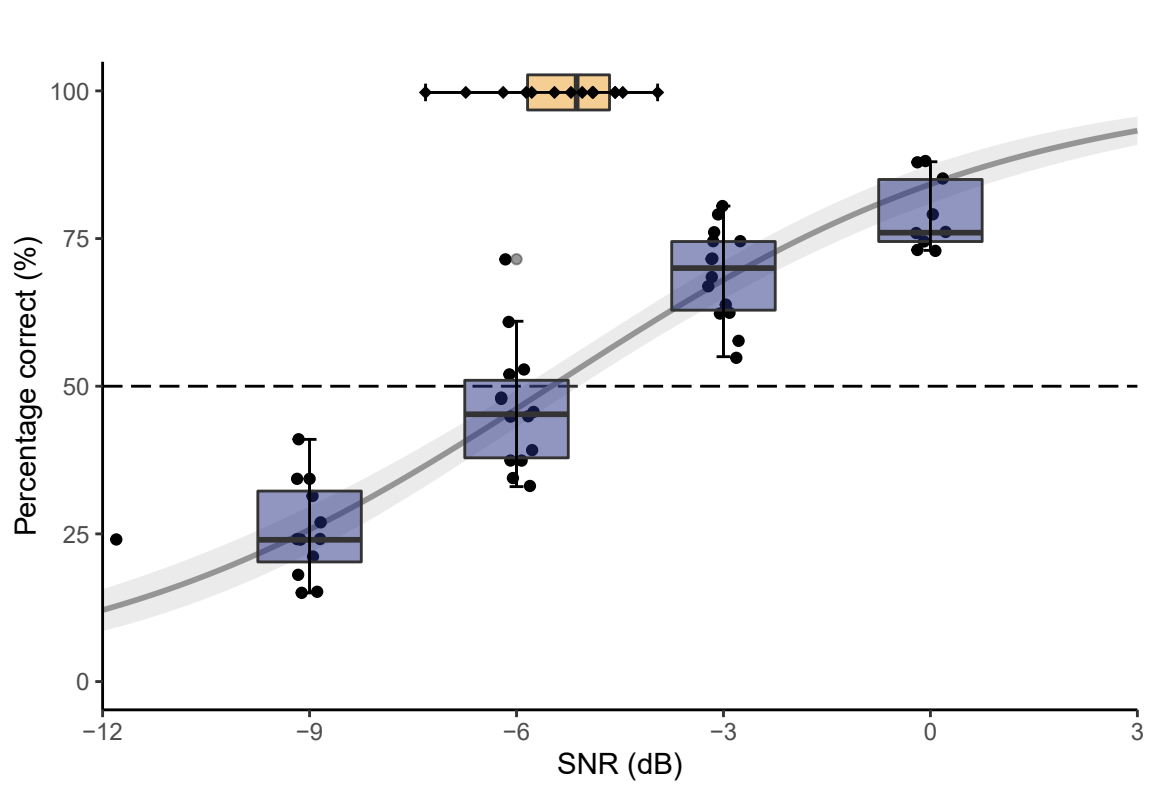
Verschueren, E., Somers, B., & Francart, T. (2019). Neural envelope tracking as a measure of speech understanding in cochlear implant users. *Hearing research*, *373*, 23–31. doi:10.1016/j.heares.2018.12.004.

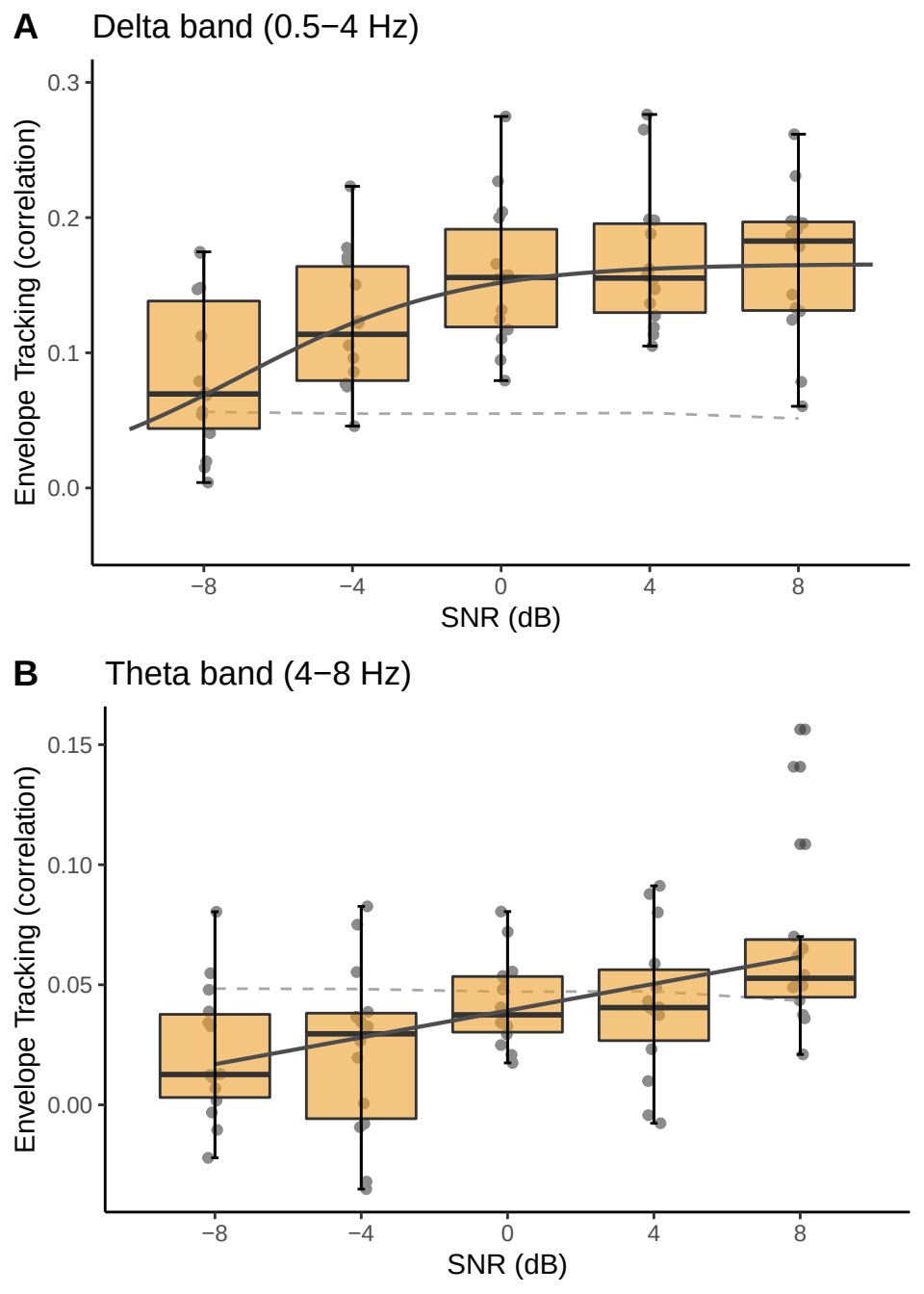
Verschueren, E., Vanthornhout, J., & Francart, T. (2020). The effect of stimulus choice on an eeg-based objective measure of speech intelligibility. *Ear and Hearing*, *41*, 1586–1597.

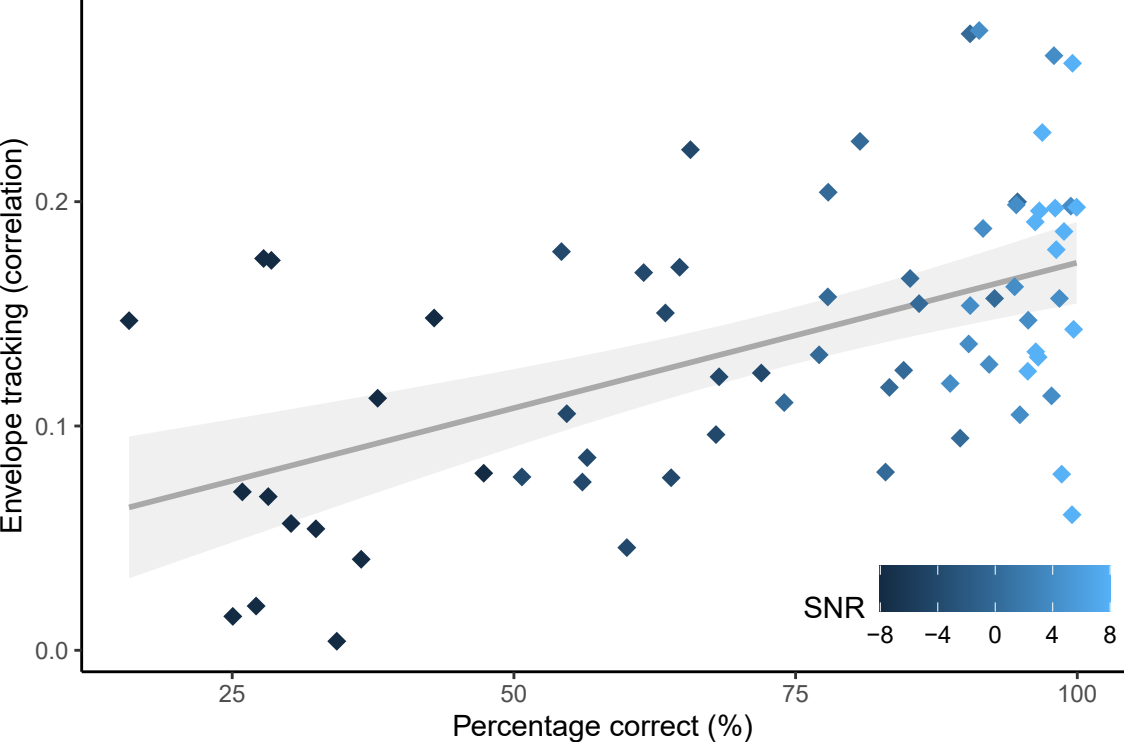
Verschueren, E., Vanthornhout, J., & Francart, T. (2021). The effect of stimulus intensity on neural envelope tracking. *Hearing Research*, *403*, 108175.

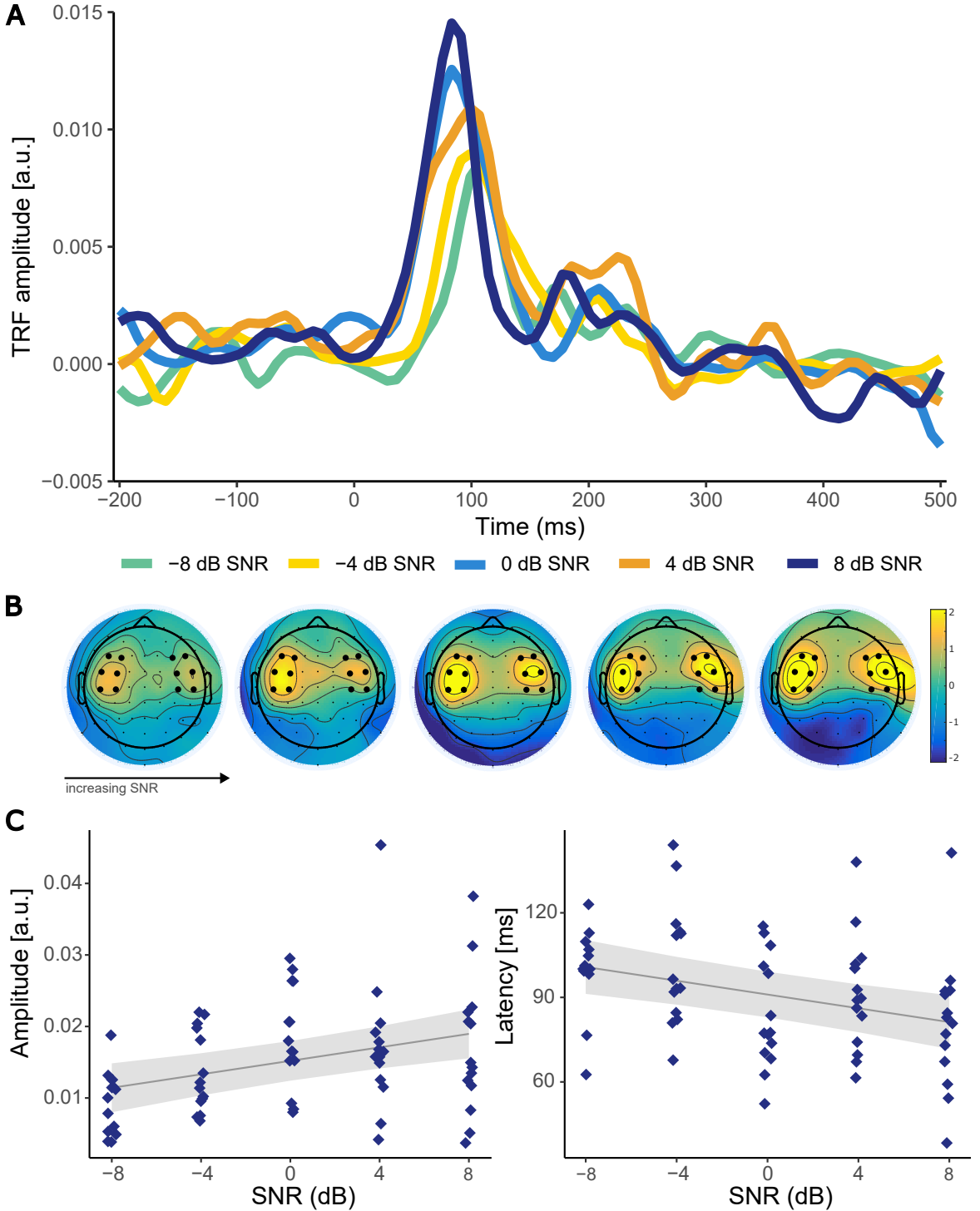
Wang, J., Sung, V., Carew, P., Burt, R. A., Liu, M., Wang, Y., Afandi, A., & Wake, M. (2019). Prevalence of childhood hearing loss and secular trends: A systematic review and meta-analysis. *Academic Pediatrics*, *19*, 504–514. doi:https://doi.org/10.1016/j.acap.2019.01.010.

766 World Health Organization (2021). Deafness and hearing loss.
767 https://www.who.int/news-room/fact-sheets/detail/
768 deafness-and-hearing-loss. Accessed: 2021-04-01.
769 Wunderlich, J. L., Cone-Wesson, B. K., & Shepherd, R. (2006). Maturation of the cortical
770 auditory evoked potential in infants and young children. *Hearing Research*, *212*, 185–202.
771 doi:https://doi.org/10.1016/j.heares.2005.11.010.
772 Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and
773 skilled reading across languages: a psycholinguistic grain size theory. *Psychological*
774 *Bulletin*, *131*, 3–29. doi:10.1037/0033-2909.131.1.3.


100
75
50
25
0
Percentage correct (%)
−12
−9
−6
−3
0
3
SNR (dB)


A
Delta band (0.5−4 Hz)
Envelope Tracking (correlation)
0.3
0.2
0.1
0.0
−8
−4
0
4
8
SNR (dB)
B
Theta band (4−8 Hz)
Envelope Tracking (correlation)
0.15
0.10
0.05
0.00
−8
−4
0
4
8
SNR (dB)


Envelope tracking (correlation)
0.0
0.1
0.2
25
50
75
100
Percentage correct (%)
SNR
−8
−4
0
4
8


A
TRF amplitude [a.u.]
0.015
0.010
0.005
0.000
−0.005
−200
−100
0
100
200
300
400
500
Time (ms)
−8 dB SNR
−4 dB SNR
0 dB SNR
4 dB SNR
8 dB SNR
B
increasing SNR
2
1
0
-1
-2
C
Amplitude [a.u.]
0.04
0.03
0.02
0.01
−8
−4
0
4
8
SNR (dB)
Latency [ms]
120
90
60
−8
−4
0
4
8
SNR (dB)