Neural entrainment to both target and distracting speech causally contributes to speech perception in a multi-speaker scenario

Authors: Jules Erkens[a,b], Mathieu Marx[a,b], Benedikt Zoefel[a,b]

Affiliations:

[a] Centre de Recherche Cerveau et Cognition (CerCo), CNRS UMR 5549, Toulouse, France

[b] Université de Toulouse III Paul Sabatier, Toulouse, France

Corresponding authors:

Jules Erkens (jules.erkens@cnrs.fr), Benedikt Zoefel (benedikt.zoefel@cnrs.fr)

The authors report no conflict of interest.

Acknowledgements: This work was supported by the Fondation pour l'Audition (grant number FPA-RD-2021-10).

Abstract

Segregating important stimuli from distractors is crucial for successful speech perception. Neural activity aligned to speech, also termed “neural entrainment”, is thought to be instrumental for this purpose. However, the relative contribution of neural entrainment to target and distractors in speech perception remained unclear. In this study, we used transcranial alternating current stimulation (tACS) to manipulate entrainment to two simultaneously presented sequences of rhythmic speech while participants attended to one of them. A random temporal relationship between speech streams allowed us to disentangle effects of tACS on target and distractor processing, and to examine their combined effect on a behavioural measure of speech perception. We found that the phase relation between tACS and both target and distracting speech modulated word report accuracy, and to a similar degree. The phasic modulation of target processing correlated with that of the distractor across subjects, and their combined effect on speech perception was stronger than each of the two alone. In contrast to our expectation, the tACS phases leading to most accurate perception were uncorrelated between target and distracting speech modulations. Together, our results suggest that entrainment to target and distracting speech jointly and causally contributes to speech perception. They also demonstrate how effect sizes might be increased in future work and for technological or clinical applications.

Significance statement

Synchronising neural activity to moments of important information in speech is beneficial for its perception. Successful speech perception is a consequence of both target processing and the suppression of potential distractors; so far, it has remained unclear whether brain-speech synchronisation plays a more important role for either of these two tasks. We used transcranial brain stimulation (tACS) in a multi-speaker scenario to independently manipulate synchronisation to target- or distracting speech, respectively. We found that both manipulations led to changes in speech perception that were of comparable magnitude, and their combined effect was stronger than either of them alone. These results give an insight into the cortical processing of competing speech streams and how brain-speech synchronisation causally contributes to their perception. Our findings also have implications for practical applicability of brain stimulation protocols, showing how they could be used and made more effective in improving speech perception. This approach could have potential for medical application such as additional support in hearing aids.

Introduction

Neural entrainment is the process of synchronizing ongoing neural activity to rhythmic stimuli (Lakatos et al., 2008; Schroeder & Lakatos, 2009). Through the combined effort of brain imaging (Henry & Obleser, 2012; Kösem et al., 2018; Petersen et al., 2017) and brain stimulation studies (Heimrath et al., 2016; Helfrich et al., 2014; Riecke et al., 2018; Thut et al., 2011; Zoefel, Allard, et al., 2019), evidence has accumulated that neural entrainment structures human perception. Presumably, this is achieved by aligning neural resources with relevant information in the stimulus, or by aligning moments of reduced excitability to distractors, leading to their suppression (Schroeder & Lakatos, 2009). Neural entrainment also plays an important role for the perception of speech (Peelle et al., 2013; Peelle & Davis, 2012). In a multi-speaker setting, Horton et al. (2013) showed entrainment to both target and distracting speech, but in an anti-phase relation. Although these results were correlative by nature, they are consistent with the notion that entrained neural activity simultaneously enhances target speech (alignment of high excitability phase) and suppresses distractors (alignment of low excitability phase).

Despite recent advances, the causal role of neural entrainment in multi-speaker scenarios remains poorly understood. Brain stimulation methods allow us to make the step from correlation to causation that would not be possible with brain imaging alone (Herrmann et al., 2016; Vosskuhl et al., 2018; Zoefel & Davis, 2017). Transcranial alternating current stimulation (tACS; Figure 1A) is a non-invasive stimulation method that can modulate entrainment by subthreshold polarization of the resting membrane potential (Antal & Paulus, 2013; Bland & Sale, 2019). TACS has already been successfully used to modulate speech processing (van Bree et al., 2021; Wilsch et al., 2018; Zoefel et al., 2018, 2020). This modulation is often achieved by varying the tACS waveform relative to the speech stimulus, and testing for consequences in perception. Few studies have applied tACS to modulate neural entrainment in multi-speaker scenarios. Riecke et al. (2018) presented participants with two competing speakers, one assigned as the target and one as the distractor. The perception of target speech depended on the delay between tACS and the two speech streams. However, as the delay between the two was fixed, it was not possible to disentangle the causal role that entrainment plays for the processing of target and distracting speech, respectively.

Keshavarzi et al. (2021) independently varied the timing of tACS relative to target and distracting speech. In separate experimental conditions, they either applied tACS as a theta-filtered version (4-8 Hz) of the envelope of the target speech, or as that of the ignored speech, filtered likewise. They showed that the perception of target speech is disrupted when its envelope was conveyed through tACS at a phase shift other than 0° (i.e. if it differed from the acoustically presented envelope). The tACS waveform that conveyed the distractor envelope also modulated target perception, with a 180° shift (relative to target) leading to most accurate perception. These results imply that entrainment to both target- and distracting speech causally modulates speech perception. However, as entrainment to the target- and distractor stimuli were modulated in separate conditions, their combined effects could not be investigated.

In the current study, we used rhythmically spoken speech stimuli to simultaneously vary the phase of ongoing tACS relative to both target and distracting speech. A random temporal relationship between targets and distractors allowed us to disentangle their effects in the same trials (Figure 1B): For a given phase relation between tACS and target speech, the distractor was presented at all possible tACS phases, and any phasic effect the latter produces would cancel out on average (and vice versa). In this way, we were we able to quantify causal effects of neural entrainment on target- and distracting speech, both combined and separately from each other, as well as any phasic relation between the two (Figure 1C).

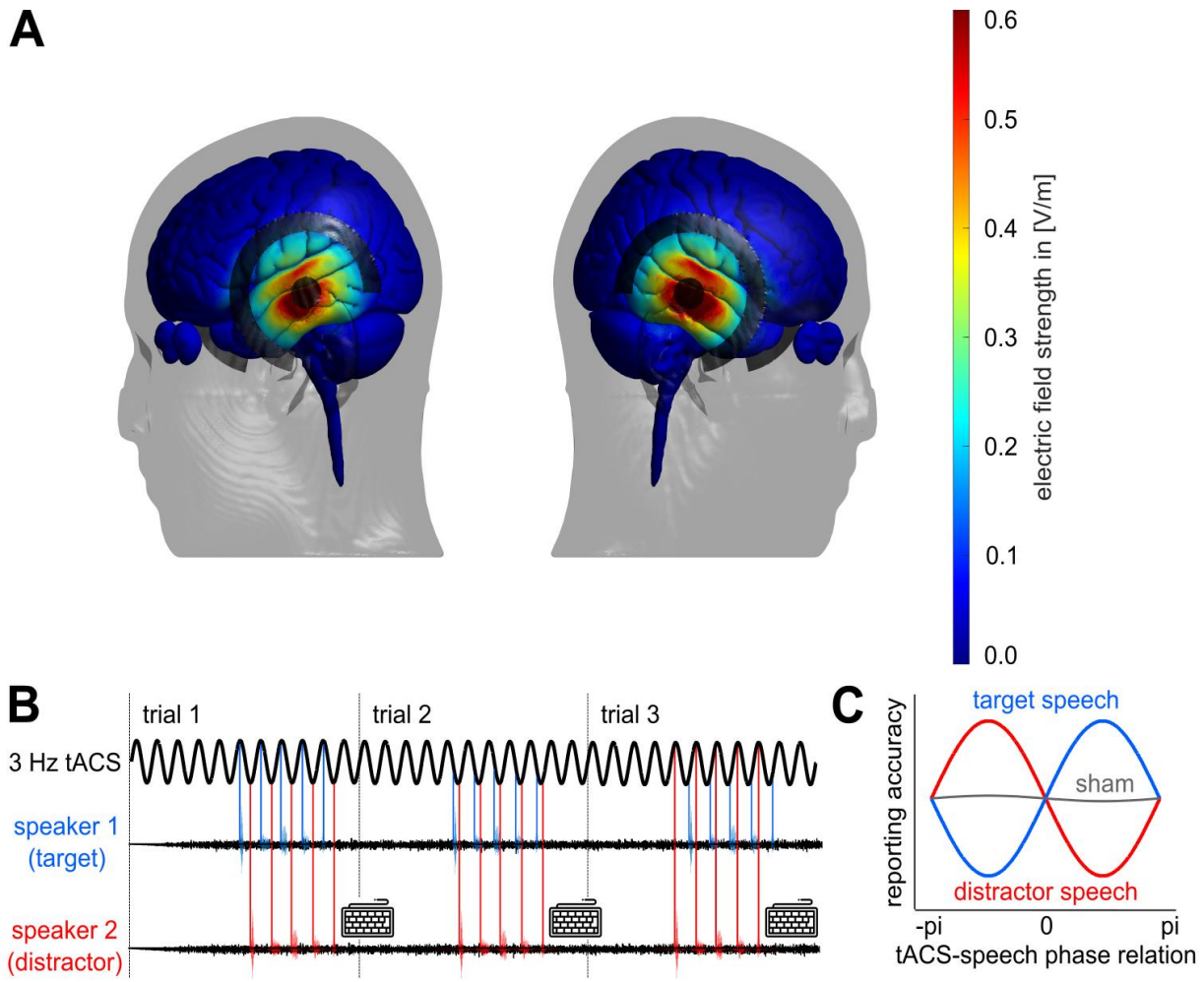

Figure 1. A*: TACS electrode locations used during the experiment, together with an estimate of the electric field that was induced in a template brain and for the average stimulation intensity (2.1 mA peak to peak). By placing the center electrode at positions T7 and T8 of the conventional EEG 10-10 system, we stimulated auditory brain regions, including the superior and middle temporal gyrus.* ***B****: Diagram of the experimental setup. Continuous 3-Hz tACS was applied during the word report task, and speech stimuli were presented at a different phase of the ongoing tACS in each trial. The relative delay between distractor and target speech was random and varied between trials. Consequently, the tACS phase for target speech was uncorrelated with that for distracting speech, and their effects could be disentangled. After each five-word sequence, participants were asked to type in the middle three words from the target speech (keyboard symbol).* ***C****: Hypothetical effect of tACS induced entrainment on task performance. We tested whether the accuracy of reporting target words depends on tACS phase relative to target speech (red) or distracting speech (blue). In the example shown, enhancing/disrupting target, and suppressing/boosting distracting speech was equally effective (relative to sham, grey). We hypothesized that, although the specific phase relation between tACS and speech that leads to enhanced perception might differ across individual participants, it would be consistently opposite between the two conditions (target vs distracting speech).*

Methods

Participants

Twenty-three participants (Mean age 25.52, St. Dev. 3.87, 11 women) were recruited for the experiment. They provided informed consent under a process approved by the Comité de Protection des Personnes (CPP) Nord Ouest IV (protocol 2020-A02707-32). All participants were native French speakers, reported having normal hearing, and had no history of neurological disease or any other exclusion criteria for tACS.

Stimuli and task

Stimuli consisted of sequences of one-syllable French words presented at a comfortable level using in-ear headphones (ER2 Earphones, Etymotic Research Inc., USA). Words were randomly selected from a pool of ~500 words that were spoken individually to a metronome at ~2 Hz (inaudible to participants). Each trial consisted of two sequences of five words, embedded in background noise and time-compressed to 3 Hz using the change tempo function of Audacity (version 3.2.0, Audacity Team, 2021). The two sequences were presented simultaneously, but with a variable lag relative to each other (Figure 1B), described in detail below. The level of background noise was determined in pre-tests to achieve ~50 % task accuracy and kept identical across participants in the main experiment. One of the two-word sequences was spoken by a male speaker, and the other one by a female speaker. The first and last words were ‘pause’ in each sequence and irrelevant for participants’ task. This was done as the first and last words had reduced overlap with the other speech stream, and would thus be easier to perceive. Participants were instructed to attend to one of the two speech streams (male or female), and ignore the other. After each sequence, participants were asked to report the middle three words from the five-word sequence they had attended, using a standard computer keyboard. Participants performed ten blocks of this task; the first two blocks familiarized the participant with the task (performance was not included in the analysis), followed by eight experimental blocks. Each block consisted of 36 trials of approximately 9.3 seconds each. The identity of the target speaker alternated between male and female between blocks. Participants were encouraged to guess or fill in partial words if they were uncertain. In between block, participants were allowed to take breaks at their leisure. The task was

presented using Matlab (The MathWorks Inc., 2022. MATLAB version: 9.7 (R2019b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com) on a Windows 10 PC.

Electric stimulation

During some experimental blocks, tACS was administered using two battery-driven stimulators (DC-Stimulator MR, Neuroconn GmbH, Ilmenau, Germany), one for each hemisphere. To assure synchronization between current and sound, stimulators and participants' headphones were driven by the output of the same high-quality sound card (Fireface, UCX, RME, Germany). The tACS electrode configuration consisted of a bilateral ring setup that produced reliable modulation of speech perception and neural oscillatory activity in previous works (van Bree et al., 2021; Zoefel, Allard, et al., 2019). Electrodes were attached with adhesive, conductive ten20 paste (Weaver and Company, Aurora, CO, USA). The center electrodes with a diameter of 20 mm were attached at positions T7 and T8 of the conventional EEG 10-10 system, placed to target auditory cortical regions bilaterally. The outer ring electrodes, with a band diameter of 13 mm, were placed around these center electrodes, with 25 mm distance on all sides. A quarter circle was cut out of the outer rings as to not overlap with the ear (Figure 1A). The electrode montage was prepared in such a way that the skin impedances on both sides of the head were comparable and never exceeded 10 kOhm. Current intensity was adjusted to a level comfortable to the participant, with an average intensity of 2.1 mA (peak to peak; std: 0.8 mA). Figure 1A shows the estimated electric field (for a template brain) for this electrode montage and average stimulation intensity.

During six experimental blocks, tACS was applied for the entire block. At the start and end of the block, tACS was faded in and out using the first and second half of a 200-ms Hanning window, respectively. The two remaining blocks served as sham blocks, for which current was faded in and out immediately (i.e. within 200 ms), without stimulation for the remainder of the block. This was done to simulate the typical sensations caused by tACS, which are most pronounced in the first few seconds of stimulation. The order of stimulation and sham blocks was randomized, and the two sham blocks always consisted of one male and one female target speaker block.

Experimental design

For both speech streams separately, the five words were presented so that the perceptual center (Scott, 1998) of each word aligned at one out of six different phases of ongoing 3-Hz tACS (Figure

1B). These six phases covered one complete 3-Hz cycle, i.e., they varied from 0° to 300° in steps of 60°, or from 0 ms to 278 ms in steps of 56 ms, respectively. The tACS phase for each trial was selected pseudo-randomly, and separately for both speech streams. Consequently, the phase (and temporal) delay between target- and distracting speech was also random, ranging from complete overlap (in-phase) to no overlap (anti-phase). Our design created 36 different combinations of tACS-target and tACS-distractor phase relations. Within each block, each of these combinations was presented once. For the sham blocks, the same 36 target-distractor phase combinations were used to maintain the relative timing of the two speech streams. This resulted in a total of 6 (blocks) x 36 = 216 tACS trials (36 per phase and attended/distractor condition) and 2 (blocks) x 36 = 72 sham stimulation trials (36 per attended/distractor condition as tACS phase is not defined for sham).

Analysis: Word report accuracy

To evaluate participants' accuracy in reporting words from the target sequence, their written responses and the target words were first converted to a phonological representation using the phonemizer Python plugin (Bernard & Titeux, 2021). Responses and targets were then compared using Levenshtein distance, i.e. the minimum number of substitutions, deletions and insertions necessary to turn one word into the other (Levenshtein, 1966). The Levenshtein distance is typically normalized by dividing its absolute value by the length of target or response word (de la Higuera & Mico, 2008). In this way, however, the same absolute Levenshtein distance can produce multiple normalized outcomes, depending on the length of these words. To avoid this issue, we used an alternative approach that was validated in simulations during pre-tests. This approach "corrected" participants' responses by regressing single-trial Levenshtein distances on the length of corresponding target words, and subtracting the explained variance. All subsequent analyses were then performed using the residuals from this regression analysis.

As each trial consisted of three target words, participants entered up to three words after each trial, and these might not follow the original order. We therefore calculated Levenshtein distances for each combination of target and response words within a trial and selected the combination with the highest accuracy. Note (1) that the same procedure was applied to the sham condition which was later compared with tACS (see next section) and (2) that it does not bias accuracy specifically for some tACS phases and therefore cannot have produced spurious phase effects.

Analysis: Comparison with sham

The random phase relationship between targets and distractors allowed us to disentangle their causal contributions to word report accuracy: For a given phase relation between tACS and target speech, the distractor was presented at all possible tACS phases, and any phasic effect the latter produces would cancel out on average (and vice versa). However, due to the nature of the design, the relative timing between the two speech streams as a whole was not identical across tACS phases: Whereas target speech preceded the distractor for some tACS phases, the opposite was true for others. Note that these differences in *global* timing would only affect the context of the first and last words of the sequence (as these were either preceded or followed by a word from the other speech stream or not, depending on the tACS phase), but should not have affected the middle three words that were used to quantify accuracy. We nevertheless controlled for these timing-related effects by making use of the fact that they should be equally prevalent during tACS and sham stimulation. For each relative timing between speech streams (corresponding to the 36 target-distractor phase combinations described above) and for single participants, we first calculated the trial-averaged accuracy (corrected Levenshtein distance as described in the previous section) in the sham condition. From the accuracy in each tACS trial, we then subtracted the average accuracy in the sham condition obtained for the respective target-distractor combination. This approach allowed us to isolate effects of tACS phase (which should only be present in the tACS condition) from generic effects of presentation timing. It also allowed us to express performance relative to a baseline (sham) condition.

Analysis: Phasic modulation of word report accuracy

To quantify whether tACS had a phasic effect on word report accuracy (corrected for target word length and compared with sham), we applied a statistical approach that was shown to be particularly sensitive at detecting such effects (Zoefel et al., 2019). We used a regression model to test how strongly tACS phase (sine- and cosine-transformed) predicts single-trial measures of performance. This model yields two regression coefficients for individual participants (one for sine, one for cosine predictor) that together reflect the magnitude of phasic modulation. The two coefficients were extracted for tACS phase effects on target and distracting speech, respectively. Separately for each of the two conditions, they were then combined to a single coefficient using a root-mean square operation.

To test for reliable phase effects, we compared the coefficients from the individual regression models that included tACS phase as a predictor with those obtained from an intercept-only regression model, using an F-test. This comparison yielded one p-value per participant. Individual p-values were combined to a group level p-value using Fisher's method. We performed this analysis separately for the two conditions (target and distractor tACS), as well as for a combined regression model which uses the tACS phase from both conditions (i.e. two sine and two cosine-transformed phases) as predictor for behavioural outcomes.

As an alternative approach to confirm phasic modulation, we used a permutation-based method that has been shown to have similar statistical sensitivity (Zoefel, Davis, et al., 2019). Here, we compared regression coefficients from the model using circular (tACS phase) predictors with those from a simulated null distribution. This null distribution was obtained by randomly assigning tACS phases to single trials before re-computing regression coefficients, separately for each participant. This procedure was then repeated 1000 times, yielding 1000 average regression coefficients that would have been obtained in the absence of a tACS effect (which was abolished by the randomization of phases). The average regression coefficient *d* from the data was then compared with the average regression coefficient μ and its standard deviation σ of the null distribution, according to:

$$z = (d-\mu) / \sigma$$

where a phasic effect was considered reliable if the z-score exceeded a critical value (e.g., $z = 1.645$, corresponding to a significant threshold of $\alpha = 0.05$, one-tailed).

To contrast tACS effects on target and distractor entrainment, we compared regression coefficients between the two conditions by means of a repeated-measures t-test. To test whether some participants are more susceptible to tACS than others, we correlated regression coefficients between conditions using linear (Pearson's) correlation.

Participants typically differ in their "best" or "worst" tACS phases, i.e., those leading to most or least accurate perception, respectively (Riecke et al., 2015; van Bree et al., 2021; Zoefel, Davis, et al., 2019). We extracted participants' best phases from the individual regression models. These phases correspond to the peak of a sine function fitted to word report accuracy as a function of tACS phase. As this method relies on data from all tACS phases rather than defining best phase

exclusively based on maximum performance (Erkens et al., 2020, 2021; Riecke et al., 2018), this method is less affected by task-unrelated noise in the data. We used Rayleigh's test for non-uniformity to test whether participants' best tACS phases were distributed non-uniformly ("circ_rtest" from Circular Statistics Toolbox for MATLAB; Berens, 2009).

We hypothesized opposite best tACS phases when applied to modulate target or distracting speech, respectively (Figure 1C). As an optimal (amplifying) phase for target speech should enhance speech perception, the same phase should disrupt speech perception for distracting speech, as distracting information is amplified. Such a result was already reported by Keshavarzhi et al. (2021). We examined this hypothesis by comparing the circular difference between individual best tACS phases obtained for target and distracting speech, respectively, against 0° ("circ_mtest" from the Circular Statistics Toolbox; Berens, 2009), and by testing whether its 95% confidence interval includes 180°.

The methods described above quantify phasic modulation of speech perception, produced by tACS, remain ignorant about the direction of the effect. It is thus unclear whether speech perception was improved, impaired, or both. As word report accuracy during tACS was expressed relative to that obtained during sham stimulation, positive and negative accuracy values reflect enhanced and impaired speech perception, respectively. However, as participants differed in their best (and worst) tACS phases, group-level statistics required the selection of individual tACS phase bins that led to most (and least) accurate perception, respectively. Such a selection leads statistical bias (Asamoah et al., 2019); for instance, accuracy at the individual "best" phase is per definition a maximum, and likely to be positive (i.e. better than sham). However, this bias does not affect the *difference* between sham and best and worst tACS phases (corresponding to peak and trough of the fitted sine function), respectively. Instead of testing for enhancement or disruption of speech perception, we therefore tested whether the tACS-induced change in accuracy, relative to the sham condition, was larger for one or the other. This was done by contrasting change in accuracy (tACS minus sham) at individual best (enhanced) and worst (disrupted; multiplied by -1 for comparability) tACS phases in the two conditions by means of a repeated-measures ANOVA,

We also repeated the selection procedure in the permutated datasets, described above, by extracting their best and worst tACS phases. This allowed us to quantify changes in accuracy that are due to

the selection of behavioural extrema rather than reflecting a genuine tACS effect. For visualization purposes, we also re-aligned participants' accuracy at their best phase to a common phase bin, and phase-wrapped the remaining data (for both original and permuted datasets; cf. Figure 3A).

We used a similar procedure to explore combined effects of target and distractor modulation. We first extracted word report accuracy separately for each of the 6x6 combinations of target-tACS and distractor-tACS phase relations (described above). We then re-aligned the resulting matrix based on individual best tACS phases for both target (first dimension) and distracting speech (second dimension). Consequently, maximal accuracy was aligned to the center of the matrix (cf. Figure 3C). We repeated the re-alignment for the permuted dataset and contrasted the outcomes with the original re-aligned data.

Results

There was no difference in overall word report accuracy between tACS and sham stimulation ($t(22) = .350$, $p = .730$, $d = .073$; repeated-measures t-test), suggesting no generic effect of tACS (independent of phase) on speech perception.

We defined participants' individual "best" phase for speech perception as the peak of a sine function fitted to their mean word report accuracy per phase (dotted red line in Figure 2A; see Materials and Methods). The distribution of these phases showed no significant deviation from uniformity, neither for tACS applied to modulate target speech ($p = .181$, $z = 1.72$; Rayleigh's test for non-uniformity) nor applied to distracting speech ($p = .156$, $z = 1.86$). This result implies that there was no consistent best phase across participants, as described previously (Henry & Obleser, 2012; Neuling et al., 2012; Riecke et al., 2015). Note that our principal approach to detect tACS phase effects does not rely on such consistent best tACS phases (see Materials and Methods).

We used linear regression with circular predictors (sine and cosine-transformed tACS phase) to quantify phasic modulation for both conditions (target and distracting speech), as well as their combination, for individual participants (Figure 2A). Figure 2 shows individual regression coefficients (root-mean squared sine and cosine coefficients) from this analysis and measures of their distribution, for both conditions. These coefficients reflect the magnitude of tACS-induced

phasic modulation in the respective condition. We found that tACS reliably modulated word report accuracy, irrespective of whether it was applied to alter entrainment to target speech (p = .016; Fisher's method to combine individual p-values) or entrainment to distracting speech (Fisher's p = .0008). This result was confirmed by an alternative statistical approach, contrasting regression coefficients with a simulated null distribution, shown in Figure 2B (target: p = .011, z = 2.29, Cohen's D = .411; distractor: p = .0005, z = 3.25, Cohen's D = .476). There was no difference between the two conditions, indicating that neural entrainment to target and distractor equally contributed to speech perception ($t(22)$ = -.731, p = .472). Finally, the phasic effect increased further when tACS phase relations to both target and distracting speech were used combinedly to predict behavioural outcomes (Fisher's p = .00005). As degrees of freedom are considered in the underlying F-test (see Materials and Methods), this result cannot simply be explained by the addition of more predictors in the regression model, but indicates that entrainment to both target and distracting speech jointly influence speech perception.

We furthermore found that the phasic modulation in the target condition reliably correlated with that in the distractor condition (Figure 2C; r = .638, p = .001). In other words, participants with a strong tACS-induced modulation of target entrainment also showed one for distractor entrainment, possibly due to individual differences in sensitivity to tACS.

When comparing best tACS phases for target and distracting speech, the distribution of their difference showed no significant deviance from uniformity (Figure 2D, Rayleigh's test: p = .656, z = .429). Given the uniform distribution, testing for specific phase relationships (including the hypothesized phase opposition) was not meaningful. This finding contradicts our hypothesis that optimal tACS phases for target and distractor modulation are in anti-phase (Figure 1B).

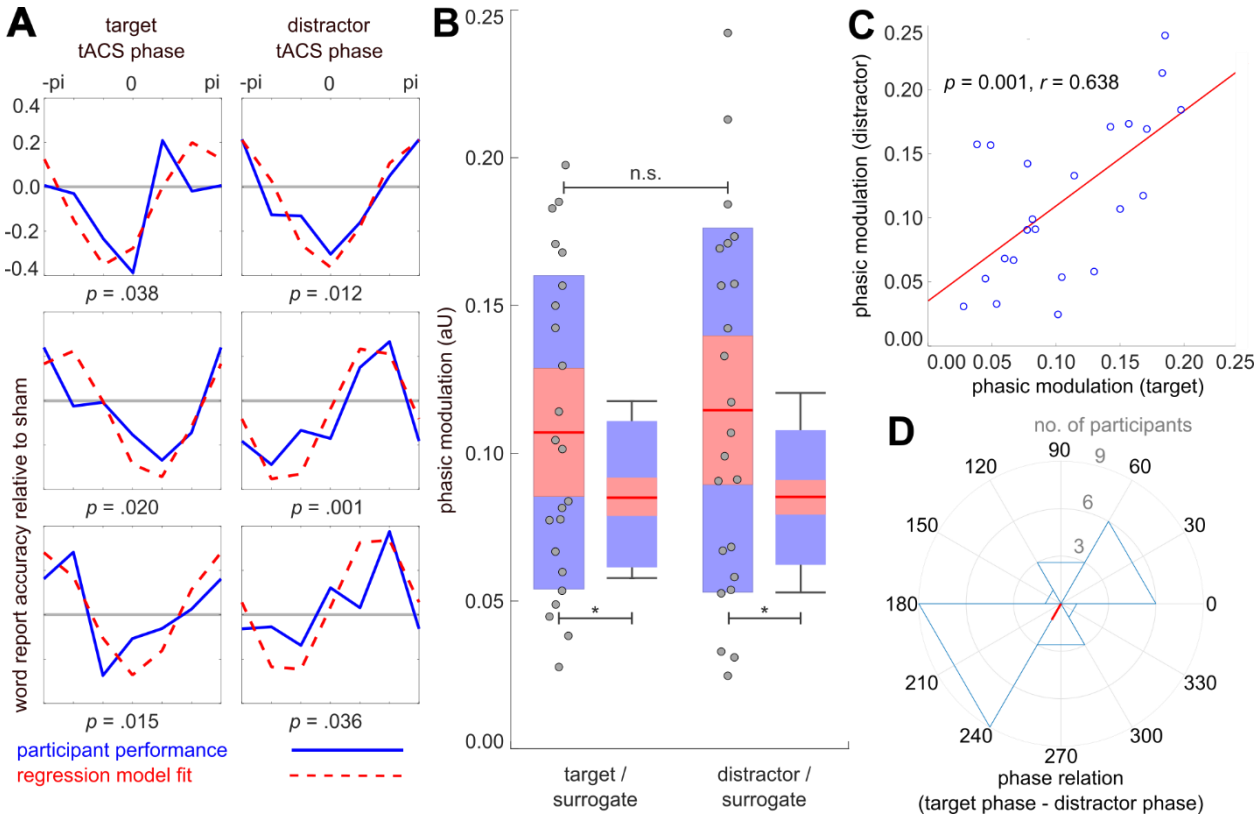

Figure 2. A*: Word report accuracy as a function of tACS phase for six example participants (blue line). Accuracy is expressed relative to sham stimulation which included the same relative delays between target and distracting speech but no tACS (see Materials and Methods). The red dotted line shows the fit of a sine function to the data. Its amplitude is equivalent to coefficients obtained from regressing accuracy on tACS phase (sine and cosine), shown in B. These coefficients were contrasted with an intercept-only regression model to obtain individual p-values indicated below each panel. Shown are three participants with lowest p-values for target (left) and distractor (right) modulation, respectively. The peak of the sine fit was defined as individual participants' best phase. Note that, for visualization of the phasic effect only, the +/-pi bin was plotted twice.* ***B****: Individual regression coefficients (root-mean square of sine and cosine coefficients) for target and distracting speech and their distribution, as well as that obtained from a permutation procedure. Points, red lines, red areas and blue areas correspond to individual participants, their average, standard deviation, and 95% confidence interval, respectively. Asterisks indicate a reliable difference (p < 0.05) between original data and surrogate distribution* ***C****: Between-subject correlation between coefficients obtained for modulation of target speech and distracting speech. Points correspond to individual participants.* ***D****: Distribution of circular differences between individual best tACS phases for the modulation of target speech and distracting speech. The direction and length of the red line illustrate the average difference and its consistency across participants, respectively.*

Finally, we examined whether the phasic effects we observed corresponded to enhanced or disrupted speech perception (or both), respectively. Figure 3A shows average word report accuracy (relative to sham) after being re-aligned to individual best tACS phases in each condition (see Materials and Methods). Note that this analysis was not designed to assess phasic effects (this was already done in analyses described above), but to contrast enhancing and disrupting effects on perception. Due to the alignment, the behavioural data necessarily follows a sinusoidal shape, and behavioural extrema are located at the aligned best phase and its anti-phase, respectively. For illustration purposes, we tested how much of this is due to the re-alignment itself by repeating the same procedure in a permuted dataset (grey line in Figure 3A). Note that the actual data diverges from the simulated null effect in both directions (positive and negative).

To contrast enhancing and disrupting tACS effects, we compared tACS-induced changes in accuracy, relative to the sham condition, between individual best and worst tACS phases. Note that, although the re-alignment can make these changes look larger than they are, this bias does not affect their *difference* (difference to sham at the best vs worst phase). Individual changes in accuracy and measures of their distribution are shown in Figure 3B (tACS minus sham, multiplied by -1 for the worst phase). We found no difference between tACS-induced enhancement and disruption in either of the two conditions (repeated-measures ANOVA, main effect target-distractor: $F(22) = .076$, $p = .786$, main effect enhancement-disruption: $F(22) = .106$, $p = .748$), nor an interaction ($F(22) = .158$, $p = .695$). Given the significant phasic effects reported above, this result indicates that tACS led to *both* an enhancement and disruption of speech perception, depending on the phase relation it was applied to the stimulus.

Figure 3C shows results when data was re-aligned according to the combination of individual best tACS phases for both target and distracting speech (see Materials and Methods). Again, the enhancement (relative to the simulated null distribution) and disruption of performance at best and worst phase combinations are of similar magnitude, respectively. Critically, however, the change in accuracy at both phase combinations is larger for this combined version than the change observed in individual conditions. Again, this result implies that entrainment to targets and entrainment to distractors jointly and causally influence speech perception.

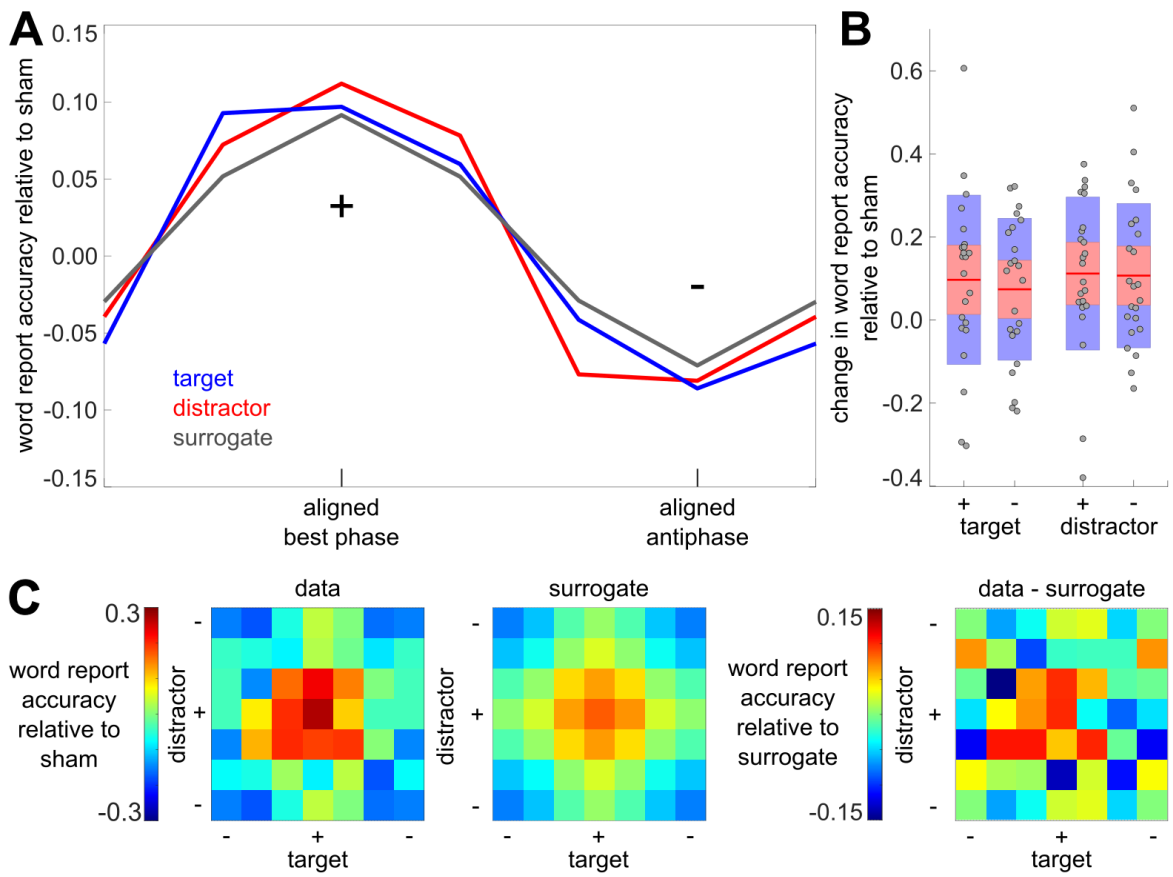

Figure 3. A*: Averaged word report accuracy (relative to sham) that was re-aligned according to individual best tACS phases (peak of fitted sine function), separately for the two tACS conditions (blue and red) and the permuted dataset (grey). For the latter, the average across 1000 permutations is shown.* ***B****: tACS-induced changes in accuracy (tACS minus sham) at the best (+; peak of fitted sine) and worst (-; trough of fitted sine; multiplied by -1 for comparability) tACS phase, respectively. Points, red lines, red areas and blue areas correspond to individual participants, their average, standard deviation, and 95% confidence interval, respectively.* ***C****: Word report accuracy, re-aligned individually based on best tACS phases from both target and distractor conditions. The left panel shows the mean of the original data, the middle panel shows the mean of the surrogate data (averaged across subjects and permutations), and the right panel shows the difference between the two. For all three plots, the combination of best phases was aligned to the center of the square, while worst phases are furthest away from the center. Note, that for visualization only, the combination of worst phases is plotted twice.*

Discussion

Using tACS to independently manipulate neural activity aligned to target and distracting speech, presented simultaneously, we here show that (1) neural entrainment to both of these speech streams causally contributes to speech perception; (2) these effects are of similar magnitude and (3) correlated across subjects; and (4) their combined effect goes beyond each of the two individual modulations alone. In the following, we embed our findings into the literature and discuss differences and commonalities with previous work.

In any complex hearing scenario, including the often-dubbed ‘cocktail party’, we are tasked with separating relevant sounds from distractors. Recent research has examined neural dynamics (including “neural entrainment”) in cocktail-party scenarios, but findings have mostly been restricted to correlative evidence. In one of these studies, Orf et al. (2022) compared EEG “tracking” of target and distracting speech with that of a neutral stimulus. Although all speech stimuli evoked tracking responses, only the target but not distracting speech produced clear differences to the neutral speech. Although common measures of “tracking” are different from the brain-speech phase coherence targeted in the current work and definitions of “tracking” and “entrainment” can be inconsistent (Lakatos et al., 2019; Obleser & Kayser, 2019), this finding nevertheless seems to imply that neural following of attended speech is functionally more relevant than that of distractors. It might also contradict our result (as well as that by Keshavarzi et al, 2021) that the causal contributions of target and distractor entrainment to speech perception are of similar magnitude. This difference can be explained by several factors. A fundamental element of the study by Orf et al. (2022) is the comparison between a sometimes task-relevant distractor and a neutral, never task-relevant stimulus. It could be that the neutral stimulus acted as a distractor as well, so that the manipulation of its processing with tACS would have benefited perception of target speech if used in an experimental scenario like ours. It is also possible that entrainment to distractors is particularly relevant in challenging situations. In a similar study, Fiedler et al. (2019) found that distracting speech evokes a late tracking component that is of opposite polarity to that seen for target speech, but only when it is difficult to understand, implying an additional, active suppression of distractors. In line with this notion, speakers were presented at different locations in the task used by Orf et al (2022), whereas in our study both speakers were presented binaurally through in-ear headphones. Spatial separation cues are known to be important for speech

segregation (Shinn-Cunningham, 2005; Shinn-Cunningham & Best, 2008), and cortical tracking of distracting stimuli might have become more relevant due to the increased difficulty in absence of these cues.

Other studies have shown that distracting stimuli are represented in cortical activity (Evans et al., 2016), and that this representation dissipates further up in the auditory hierarchy (Zion Golumbic et al., 2013). Furthermore, this representation depends on the structure of the distractor, with more speech-like stimuli being represented for longer (Har-shai Yahav & Zion Golumbic, 2021). Entrainment to distractor stimuli would allow the listener to separate sound sources not only based on the attributes of the target alone, but also based on those of the distractor (Ding & Simon, 2012, 2014). At earlier stages of processing, target and distractor streams might therefore need to be processed separately instead of competing for neural resources. In addition, a representation of distractors might facilitate a switch between target and distracting stimuli as it is often necessary in natural scenarios. Together, although the underlying mechanisms might not be identical, there is converging evidence for processing of both targets and distractors in challenging listening situations.

Brain stimulation work, including the present study, has confirmed these previous findings and demonstrated a causal role of entrained neural dynamics for speech processing, not only for attended (e.g. Keshavarzi et al., 2020; Riecke et al., 2018; Wilsch et al., 2018; Zoefel et al., 2018) but also distracting speech (Keshavarzi et al., 2021; Riecke et al., 2018). Our study is closely related to that by Keshavarzi et al. (2021), who applied tACS to independently modulate entrainment to target and distracting speech and found effects on speech perception in both cases. Although our results confirm these findings, we did not replicate the anti-phase relationship between target and distractor entrainment that was apparent in their work. This difference might be due to changes in the experimental design that allowed us to investigate combined causal effects of target and distractor entrainment. In the study by Keshavarzi and colleagues (2021), a filtered envelope of naturally spoken speech was apply via tACS and delayed relative to the acoustic stimulus. However, natural speech contains local variations in rate and phase, and applying tACS with a fixed delay might lead to different effects, depending on instantaneous variations in these parameters. Global phase shifts, induced by tACS, can therefore be hard to interpret. For instance, delaying the theta-filtered envelope with tACS might not only have altered theta-neural activity

relative to the acoustic stimulus, but also de-synchronized theta activity with that at slower time scales (e.g., delta), which also plays a role for speech perception (Ding & Simon, 2014; Etard & Reichenbach, 2019; Keitel et al., 2018; Meyer et al., 2016; Molinaro & Lizarazu, 2018). In our study, we used rhythmically spoken and unrelated mono-syllabic words instead of natural speech. This manipulation avoided the concomitant fluctuations in information at multiple times scales that natural speech has (Doelling et al., 2014; Ghitza, 2012; Giraud & Poeppel, 2012). It therefore allowed us to confirm previous results by Keshavarzi et al (2021) in a setting with reduced local changes in frequency and phase in the acoustic stimulus. An important, novel element in the experimental design is the manipulation of target and distractor entrainment in the same trials. This allowed us to quantify combined effects of target and distractor entrainment, which was not possible in previous designs. Indeed, we found that the phase relations between tACS and target and distracting speech, respectively, provide complementary predictive value for behavioural measures of speech perception.

These differences in experimental design might also explain why we did not find the hypothesized anti-phase relation between best tACS phases for target and distractor modulation. As stated above, there is evidence that distracting speech is processed and detectable in the auditory hierarchy until it is suppressed at later stages (Zion Golumbic et al., 2013). Indeed, we often perceive ongoing distracting speech, but we do not process their lexical information until it demands our attention (Arons, 1992; Power et al., 2012). It is therefore possible that the role of entrainment for distractor processing depends on the hierarchical stage it operates at. Whilst our tACS protocol manipulated entrainment at the word/syllable level, stimuli used by Keshavarzi et al. (2021) included grammatical and structural information. It is possible that (only) the latter induced competition between target and distracting speech at a higher hierarchical level, and led to opposite best tACS phases for speech perception depending on whether tACS was used to manipulate target or distractor processing. In our study, best tACS phases for target and distracting speech were uncorrelated, which implies that their processing involves distinct neural mechanisms, at least at lower hierarchical levels.

It has been previously proposed that tACS could have medical applications, specifically when integrated into conventional hearing aids (Erkens et al., 2021; Keshavarzi & Reichenbach, 2020; Riecke et al., 2015). Ageing and hearing-impaired listeners do not only struggle to understand

speech in noise, but also show changes in neural entrainment to speech, implying a functional role for speech perception in noisy scenarios (Cabeza et al., 2002; Henry et al., 2017; Petersen et al., 2017). It is a common complaint amongst users that hearing aid can be inefficient in complex hearing scenarios (Fischer et al., 2020; Kochkin, 2000); moreover, speech stream segregation has been shown to worsen with age (Getzmann et al., 2017; Getzmann & Näätänen, 2015; Pichora-Fuller et al., 1995; Pichora-Fuller & Singh, 2006). TACS could therefore provide a way to assist struggling listeners in the processing of speech by enhancing entrainment in such situations. It remained an open question, however, if tACS should be applied to enhance entrainment to target speakers or disrupt that to distractors. Our study suggests that both versions might work, and to a similar degree. More so, we found that their combined effects go beyond what can be achieved with either of these two possibilities alone. This finding is important as, just as in our study, reported tACS effects are typically small (Bland & Sale, 2019; Riecke, 2016; Riecke & Zoefel, 2018), and approaches to enhance them to clinically relevant levels are urgently needed. Our design required rhythmic speech to examine these combined effects in the same trials. An important step would therefore be the development of a stimulation protocol to simultaneously modulate target and distracting speech in natural scenarios. Another relevant finding from our study is the correlation between target and distractor effects across participants. This result indicates that some participants were more susceptible to tACS, possibly due to anatomical factors (Kasten et al., 2019). If tACS were to be used as a medical device, these individual differences in effectiveness should be explored.

Conclusion

We here demonstrate that neural entrainment to target and distracting speech is causally relevant for speech perception in a multi-speaker scenario. The fact that best tACS phases for target and distractor modulations were uncorrelated in our, but correlated in others' work, illustrates the importance of experimental design in tACS research. Although speculatively, it might also reflect distinct and partly independent neural representations of the two speech streams in the auditory pathway. Together, these findings illustrate the functional role of entrainment in challenging listening situations, and bring us closer to incorporating tACS in future medical or technological applications.

Data and Code Availability Statement

Data and code will be made available upon request to the authors.

References

Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience*, *7*(June), 1–4. https://doi.org/10.3389/fnhum.2013.00317

Arons, B. (1992). A Review of The Cocktail Party Effect. *Journal of the American Voice I/O Society*, *12*(7), 35–50.

Asamoah, B., Khatoun, A., & Mc Laughlin, M. (2019). Analytical bias accounts for some of the reported effects of tACS on auditory perception. *Brain Stimulation*, *12*(4), 1001–1009. https://doi.org/10.1016/j.brs.2019.03.011

Berens, P. (2009). **CircStat**: A *MATLAB* Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*(10). https://doi.org/10.18637/jss.v031.i10

Bernard, M., & Titeux, H. (2021). Phonemizer: Text to Phones Transcription for Multiple Languages in Python. *Journal of Open Source Software*, *6*(68), 3958. https://doi.org/10.21105/joss.03958

Bland, N. S., & Sale, M. V. (2019). Current challenges: The ups and downs of tACS. *Experimental Brain Research*, *237*(12), 3071–3088. https://doi.org/10.1007/s00221-019-05666-0

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage*, *17*(3), 1394–1402. https://doi.org/10.1006/nimg.2002.1280

de la Higuera, C., & Mico, L. (2008). A contextual normalised edit distance. *2008 IEEE 24th International Conference on Data Engineering Workshop*, 354–361. https://doi.org/10.1109/ICDEW.2008.4498345

Ding, N., & Simon, J. Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*(1), 78–89. https://doi.org/10.1152/jn.00297.2011

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*(May), 1–7. https://doi.org/10.3389/fnhum.2014.00311

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035

Erkens, J., Schulte, M., Vormann, M., & Herrmann, C. S. (2020). Lacking Effects of Envelope Transcranial Alternating Current Stimulation Indicate the Need to Revise Envelope Transcranial Alternating Current Stimulation Methods. *Neuroscience Insights*, *15*, 263310552093662. https://doi.org/10.1177/2633105520936623

Erkens, J., Schulte, M., Vormann, M., Wilsch, A., & Herrmann, C. S. (2021). Hearing Impaired Participants Improve More Under Envelope-Transcranial Alternating Current Stimulation When Signal to Noise Ratio Is High. *Neuroscience Insights*, *16*, 263310552098885. https://doi.org/10.1177/2633105520988854

Etard, O., & Reichenbach, T. (2019). Neural Speech Tracking in the Theta and in the Delta Frequency Band Differentially Encode Clarity and Comprehension of Speech in Noise. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *39*(29), 5750–5759. https://doi.org/10.1523/JNEUROSCI.1828-18.2019

Evans, S., McGettigan, C., Agnew, Z. K., Rosen, S., & Scott, S. K. (2016). Getting the Cocktail Party Started: Masking Effects in Speech Perception. *Journal of Cognitive Neuroscience*, *28*(3), 483–500. https://doi.org/10.1162/jocn_a_00913

Fiedler, L., Wöstmann, M., Herbst, S. K., & Obleser, J. (2019). Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *NeuroImage*, *186*, 33–42. https://doi.org/10.1016/j.neuroimage.2018.10.057

Fischer, T., Caversaccio, M., & Wimmer, W. (2020). Multichannel acoustic source and image dataset for the cocktail party effect in hearing aid and implant users. *Scientific Data*, *7*(1), 440. https://doi.org/10.1038/s41597-020-00777-8

Getzmann, S., Jasny, J., & Falkenstein, M. (2017). Switching of auditory attention in “cocktail-party” listening: ERP evidence of cueing effects in younger and older adults. *Brain and Cognition*, *111*, 1–12. https://doi.org/10.1016/j.bandc.2016.09.006

Getzmann, S., & Näätänen, R. (2015). The mismatch negativity as a measure of auditory stream segregation in a simulated “cocktail-party” scenario: Effect of age. *Neurobiology of Aging*, *36*(11), 3029–3037. https://doi.org/10.1016/j.neurobiolaging.2015.07.017

Ghitza, O. (2012). On the Role of Theta-Driven Syllabic Parsing in Decoding Speech: Intelligibility of Speech with a Manipulated Modulation Spectrum. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00238

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Publishing Group*, *15*(4). https://doi.org/10.1038/nn.3063

Har-shai Yahav, P., & Zion Golumbic, E. (2021). Linguistic processing of task-irrelevant speech at a cocktail party. *eLife*, *10*, e65096. https://doi.org/10.7554/eLife.65096

Heimrath, K., Fiene, M., Rufener, K. S., & Zaehle, T. (2016). Modulating Human Auditory Processing by Transcranial Electrical Stimulation. *Frontiers in Cellular Neuroscience*, *10*(March). https://doi.org/10.3389/fncel.2016.00053

Helfrich, R. F., Schneider, T. R., Rach, S., Trautmann-Lengsfeld, S. A., Engel, A. K., & Herrmann, C. S. (2014). Entrainment of brain oscillations by transcranial alternating current stimulation. *Current Biology*, *24*(3), 333–339. https://doi.org/10.1016/j.cub.2013.12.041

Henry, M. J., Herrmann, B., Kunke, D., & Obleser, J. (2017). Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nature Communications*, *8*(May). https://doi.org/10.1038/ncomms15801

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Herrmann, C. S., Strüber, D., Helfrich, R. F., & Engel, A. K. (2016). EEG oscillations: From correlation to causality. *International Journal of Psychophysiology*, *103*, 12–21. https://doi.org/10.1016/j.ijpsycho.2015.02.003

Horton, C., D'Zmura, M., & Srinivasan, R. (2013). Suppression of competing speech through entrainment of cortical oscillations. *Journal of Neurophysiology*, *109*(12), 3082–3093. https://doi.org/10.1152/jn.01026.2012

Kasten, F. H., Duecker, K., Maack, M. C., Meiser, A., & Herrmann, C. S. (2019). Integrating electric field modeling and neuroimaging to explain inter-individual variability of tACS effects. *Nature Communications*, *10*(5427), 1–11. https://doi.org/10.1038/s41467-019-13417-6

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biology*, *16*(3), e2004473. https://doi.org/10.1371/journal.pbio.2004473

Keshavarzi, M., Kegler, M., Kadir, S., & Reichenbach, T. (2020). Transcranial alternating current stimulation in the theta band but not in the delta band modulates the comprehension of naturalistic speech in noise. *NeuroImage*, *210*(October 2019), 116557. https://doi.org/10.1016/j.neuroimage.2020.116557

Keshavarzi, M., & Reichenbach, T. (2020). Transcranial Alternating Current Stimulation With the Theta-Band Portion of the Temporally-Aligned Speech Envelope Improves Speech-in-Noise Comprehension. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00187

Keshavarzi, M., Varano, E., & Reichenbach, T. (2021). Cortical Tracking of a Background Speaker Modulates the Comprehension of a Foreground Speech Signal. *The Journal of Neuroscience*, *41*(23), 5093–5101. https://doi.org/10.1523/JNEUROSCI.3200-20.2021

Kochkin, S. (2000). MarkeTrak V: 'Why my hearing aids are in the drawer': The consumers' perspective. *The Hearing Journal*, *53*(2), 34–41. https://doi.org/10.1097/00025572-200002000-00004

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural Entrainment Determines the Words We Hear. *Current Biology*, *28*(18), 2867-2875.e3. https://doi.org/10.1016/j.cub.2018.07.023

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology*, *29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, *10*, 707.

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2016). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral Cortex*, cercor;bhw228v1. https://doi.org/10.1093/cercor/bhw228

Molinaro, N., & Lizarazu, M. (2018). Delta(but not theta)-band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience*, *48*(7), 2642–2650. https://doi.org/10.1111/ejn.13811

Neuling, T., Rach, S., Wagner, S., Wolters, C. H., & Herrmann, C. S. (2012). Good vibrations: Oscillatory phase shapes perception. *NeuroImage*, *63*(2), 771–778. https://doi.org/10.1016/j.neuroimage.2012.07.024

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004

Orf, M., Wöstmann, M., Hannemann, R., & Obleser, J. (2022). *Auditory neural tracking reflects target enhancement but not distractor suppression in a psychophysically augmented continuous-speech paradigm* (p. 2022.06.18.496558). bioRxiv. https://doi.org/10.1101/2022.06.18.496558

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*(SEP), 1–17. https://doi.org/10.3389/fpsyg.2012.00320

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of Neurophysiology*, *117*(1), 18–27. https://doi.org/10.1152/jn.00527.2016

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, *97*(1), 593–608. https://doi.org/10.1121/1.412282

Pichora-Fuller, M. K., & Singh, G. (2006). Effects of Age on Auditory and Cognitive Processing: Implications for Hearing Aid Fitting and Audiologic Rehabilitation. *Trends in Amplification*, *10*(1), 29–59. https://doi.org/10.1177/108471380601000103

Power, A. J., Foxe, J. J., Forde, E. J., Reilly, R. B., & Lalor, E. C. (2012). At what time is the cocktail party? A late locus of selective attention to natural speech. *European Journal of Neuroscience*, *35*(9), 1497–1503. https://doi.org/10.1111/j.1460-9568.2012.08060.x

Riecke, L. (2016). Studying Effects of Transcranial Alternating Current Stimulation on Hearing and Auditory Scene Analysis. In P. Van Dijk, D. Başkent, E. Gaudrain, E. De Kleine, A. Wagner, & C. Lanting (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (Vol. 894, pp. 371–379). Springer International Publishing. https://doi.org/10.1007/978-3-319-25474-6_39

Riecke, L., Formisano, E., Herrmann, C. S., & Sack, A. T. (2015). 4-Hz transcranial alternating current stimulation phase modulates hearing. *Brain Stimulation*, *8*(4), 777–783. https://doi.org/10.1016/j.brs.2015.04.004

Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2018). Neural Entrainment to Speech Modulates Speech Intelligibility. *Current Biology*, *28*(2), 161-169.e5. https://doi.org/10.1016/j.cub.2017.11.033

Riecke, L., & Zoefel, B. (2018). Conveying temporal information to the auditory system via transcranial current stimulation. *Acta Acustica United with Acustica*, *104*(5), 883–886. https://doi.org/10.3813/AAA.919235

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Scott, S. K. (1998). The point of P-centres. *Psychological Research Psychologische Forschung*, *61*(1), 4–11. https://doi.org/10.1007/PL00008162

Shinn-Cunningham, B. G. (2005). *Influences of spatial cues on grouping and understanding sound*.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective Attention in Normal and Impaired Hearing. *Trends in Amplification*, *12*(4), 283–299. https://doi.org/10.1177/1084713808325306

Thut, G., Schyns, P. G., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology*, *2*(JUL), 1–10. https://doi.org/10.3389/fpsyg.2011.00170

van Bree, S., Sohoglu, E., Davis, M. H., & Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biology*, *19*(2). https://doi.org/10.1371/JOURNAL.PBIO.3001142

Vosskuhl, J., Strüber, D., & Herrmann, C. S. (2018). Non-invasive Brain Stimulation: A Paradigm Shift in Understanding Brain Oscillations. *Frontiers in Human Neuroscience*, *12*(May), 1–19. https://doi.org/10.3389/fnhum.2018.00211

Wilsch, A., Neuling, T., Obleser, J., & Herrmann, C. S. (2018). Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage*, *172*(July 2017), 766–774. https://doi.org/10.1016/j.neuroimage.2018.01.038

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., & Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a 'cocktail party'. *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

Zoefel, B., Allard, I., Anil, M., & Davis, M. H. (2019). Perception of rhythmic speech is modulated by focal bilateral transcranial alternating current stimulation. *Journal of Cognitive Neuroscience*, *32*(2), 226–240. https://doi.org/10.1162/jocn_a_01490

Zoefel, B., Allard, I., Anil, M., & Davis, M. H. (2020). Perception of Rhythmic Speech Is Modulated by Focal Bilateral Transcranial Alternating Current Stimulation. *Journal of Cognitive Neuroscience*, *32*(2), 226–240. https://doi.org/10.1162/jocn_a_01490

Zoefel, B., Archer-Boyd, A., & Davis, M. H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Current Biology*, *28*(3), 401-408.e5. https://doi.org/10.1016/j.cub.2017.11.071

Zoefel, B., & Davis, M. H. (2017). Transcranial electric stimulation for the investigation of speech perception and comprehension. *Language, Cognition and Neuroscience*, *32*(7), 910–923. https://doi.org/10.1080/23273798.2016.1247970

Zoefel, B., Davis, M. H., Valente, G., & Riecke, L. (2019). How to test for phasic modulation of neural and behavioural responses. In *bioRxiv*. https://doi.org/10.1101/517243