Neural speech tracking but not alpha lateralization supports listening success in a large cohort of aging individuals

Sarah Tune[1,2]*,
Mohsen Alavash[1,2],
Lorenz Fiedler[1,2] §,
& Jonas Obleser[1,2] *

1 – Department of Psychology, University of Lübeck, 23562 Lübeck, Germany
2 – Center for Brain, Behavior, and Metabolism, University of Lübeck, 23562 Lübeck, Germany

* *Author correspondence*:
Sarah Tune, Jonas Obleser
Department of Psychology
University of Lübeck
Maria-Goeppert-Str. 9a
23562 Lübeck
Email: sarah.tune@uni-luebeck.de; jonas.obleser@uni-luebeck.de

§ current address: Eriksholm Research Centre, Snekkersten, Denmark

Running title: Modelling neural filters and behavioral outcome in attentive listening

Number of words in Abstract: 218
Number of words in Introduction: 796
Number of words in Discussion: 1785

33 pages, 6 Figures, 0 Tables, Includes Supplemental Information

Keywords: EEG, alpha power lateralization, neural speech tracking, speech processing, auditory selective attention

Conflict of Interest: The authors declare no competing financial interests.

Author contributions: S.T., M.A., and J.O. designed the experiment; S.T. and M.A. oversaw data collection and preprocessing of the data; S.T., M.A. and L.F. analyzed the data; S.T., L.F., M.A., and J.O. wrote the paper.

Acknowledgments: Research was funded by the European Research Council (grant no. ERC-CoG-2014-646696 "Audadapt" awarded to J.O.). The authors are grateful for the help of Franziska Scharata in acquiring the data.

Abstract

Successful speech comprehension requires the listener to differentiate relevant from irrelevant sounds. Most studies on this problem address one of two candidate neural auditory filter solutions: Selective neural tracking of attended versus ignored speech in auditory cortex ("neural speech tracking"), and the lateralized modulation of alpha power in wider temporo–parietal cortex. However, how the two neural filter strategies relate to one another and to single-trial behavioral outcome still lacks understanding. Using an unprecedentedly large, age-varying sample (N=155; age=39–80 years) in a demanding dual-talker task, we observed attentional–cue-driven modulation of both neural filters. Trial-to-trial neural filter states varied independently of one another and speak to a functional trade-off between neurobiological attentional filter mechanisms. Single-trial models support the functional relevance of neural speech tracking, but not alpha lateralization, to overt behavior: Stronger neural speech tracking at the state- and trait-level led to more accurate and faster responses, respectively. Notably, increased chronological age and hearing loss hindered listening success but did not systematically affect the fidelity of neural filter strategies nor their impact on behavior. Our results functionally integrate behavior with trial-by-trial changes in neural state and highlight the translational potential of neural speech tracking in audiological settings. Such large-scale analyses, however, caution against simplified accounts of neural filtering strategies for behavior often derived from younger, smaller samples.

Significance statement

Successful listening crucially depends on intact attentional filters that separate relevant from irrelevant information. Research into their neurobiological implementation has focused on one of two filter strategies: the lateralization of alpha power and selective neural speech tracking. Closing the gap between hitherto separate lines of research, we used electroencephalography and a dual-talker task in a large sample of aging listeners to directly relate state- and trait-level changes in neural filter strategies and to probe their functional relevance to listening success. We demonstrate the co-existence of largely independent neural filters that establish alternating regimes of strong alpha lateralization and neural speech tracking. Additionally, our results support neural speech tracking but not alpha lateralization as a potential neural marker of listening success.

Introduction

Real-life listening is characterized by the concurrence of sound sources that compete for our attention (1). Successful speech comprehension thus relies on the selective filtering of relevant and irrelevant inputs. How does the brain instantiate attentional filter mechanisms that allow for the controlled inhibition and amplification of speech? Recent neuroscientific approaches to this question have focused on two potential neural filter strategies originating from distinct research traditions:

From the visual domain stems an influential line of research that support a role of alpha-band (~8–12 Hz) oscillatory activity in the implementation of controlled, top-down suppression of behaviorally-irrelevant information (2-5). Importantly, across modalities, it was shown that spatial-attention tasks are neurally supported by a hemispheric lateralization of alpha power over occipital, parietal but also the respective sensory cortices (6-15). This suggests that asymmetric alpha modulation could act as a filter mechanism by modulating sensory gain already in early processing stages.

In addition, a prominent line of research focuses on the role of low-frequency (1–8 Hz) neural activity in auditory and, broadly speaking, perisylvian cortex in the selective representation of speech input ("neural speech tracking"). Slow cortical dynamics temporally align with (or "track'') auditory input signals to prioritize the neural representation of behaviorally-relevant sensory information (16-19). In human speech comprehension, a key finding is the preferential neural tracking of attended compared to ignored speech in superior temporal brain areas close to auditory cortex (20-24).

However, with few exceptions (6), these two proposed neural auditory filter strategies have been studied independently of one another (but see 25, 26 for recent results on visual attention). Also, they have often been studied using tasks that are difficult to relate to natural, conversation-related listening situations (27, 28).

We thus lack understanding whether or how modulations in lateralized alpha power and the neural tracking of attended versus ignored speech in wider auditory cortex interact in the service of successful listening behavior. Moreover, few studies using more real-life listening and speech-tracking measures were able to explicitly address the functional relevance of the discussed neural filter strategies, that is, their potency to explain behavioral listening success (22).

In the present EEG study, we aim at closing these gaps by leveraging the statistical power and representativeness of a large, age-varying participant sample. We use a novel dichotic listening paradigm to enable a synoptic look at concurrent changes in auditory alpha power and neural speech tracking at the single-trial level. More specifically, our linguistic variant of a classic Posner paradigm (29) emulates a challenging dual-talker listening situation in which speech comprehension is supported by two different listening cues (30). These cues encourage the use of two complementary cognitive strategies to improve comprehension: A spatial-attention cue guides auditory attention in space, whereas a semantic cue affords more specific semantic predictions of upcoming speech. Previous research has shown that the sensory analysis of speech and, to a lesser degree, the modulation of alpha power are influenced by the availability of higher-order linguistic information (31-35).

Varying from trial to trial, both cues were presented either in an informative or uninformative version. This manipulation allowed us to understand how concurrent changes in the neural dynamics of selective attention and the resulting listening behavior are connected.

We focus on four main research questions (see Fig. 1). Note that in addressing these, we model additional known influences on listening success: age, hearing loss, as well as hemispheric asymmetries in speech processing due to the well-known right-ear advantage (36, 37).

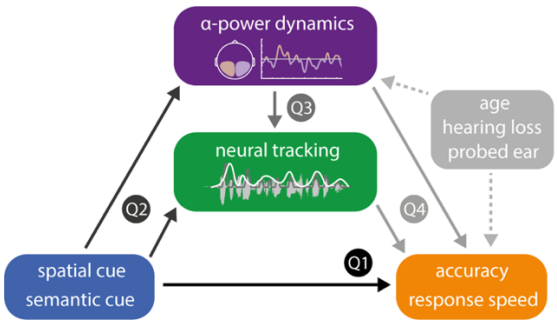

Figure 1. Schematic illustration of the research questions addressed in the present study. The dichotic listening task manipulated the attentional focus and semantic predictability of upcoming input using two separate visual cues. We investigated whether informative cues would enhance behavioral performance (Q1). In line with (Q2), we also examined the degree to which listening cues modulated the two auditory neural measures of interest: neural speech tracking and lateralization of auditory alpha power. Finally, we assessed (Q3) the co-variation of neural measures, and (Q4) their potency in explaining behavioral performance. Furthermore, we investigated the impact of age, hearing loss, and probed ear on listening success and its underlying neural strategies.

First, informative listening cues should increase listening success: These cues allow the listener to deploy auditory selective attention (compared to divided attention), and to generate more specific (compared to only general) semantic predictions, respectively.

Second, we asked how the different cue–cue combinations would modulate the two key neurobiological measures of selective attention—alpha power lateralization and neural speech tracking. We aimed to replicate previous findings of increased alpha power lateralization and a preferential tracking of the target compared to the distractor speech signal under selective (compared to divided) spatial attention. At the same time, we capitalized on the large, age-varying sample to quantify the hitherto contested dependence of these neurobiological filters on participants' chronological age and hearing loss (11, 38-40).

Third, an important and often neglected research question pertains to a direct, trial-by-trial relationship of these two candidate neural measures: Do changes in alpha power lateralization impact the degree to which attended and ignored speech signals are neurally tracked by low-frequency cortical responses?

Our final research question is arguably the most relevant one for all translational aspects of auditory attention; it has thus far only been answered indirectly when deeming these neurobiological filter mechanisms "attentional": To what extent do alpha power and neural speech tracking allow us to explain behavioral outcome—i.e., listening success—at the individual level in a single trial?

Results

We recorded and source-localized EEG signals in a large, age-varying sample (N=155) of healthy middle-aged and older adults (age=39–80 years, see also Fig. 2–supplement 1) who performed a challenging dichotic listening task (30). In this linguistic variant of a classic Posner paradigm, participants listened to two concurrent five-word sentences spoken by the same female talker, and were asked to identify the final word in one of the two sentences. Sentence pairs were temporally aligned by the onset of these task-relevant final words which led to slightly asynchronous sentence onsets.

Importantly, sentence presentation was preceded by two visual cues. First, a spatial-attention cue encouraged the use of either selective or divided attention by providing informative or uninformative instructions about the to-be-attended, and thus later probed, ear. The second cue indicated the semantic category that applied to both final target words. The provided category could represent a general or specific level, thus allowing for more or less precise prediction of the upcoming speech signal (Fig. 2a, b). While this listening task does not tap into the most naturalistic forms of speech comprehension, it still approximated a dual-talker listening situation to probe the neural underpinnings of successful selective listening (30).

Using generalized linear mixed-effects models on single-trial data, we focus on two key neurobiological measures of auditory attention: the lateralization of 8–12 Hz alpha power, emerging from auditory as well as parietal cortex, and the differential neural tracking of attended versus ignored speech by slow (1–8 Hz) auditory cortical responses. We investigate the how spatial cues, age, and hearing status modulate behavior and neural filters, whether neural filters operate independently, and to which extent they influence selective listening success.

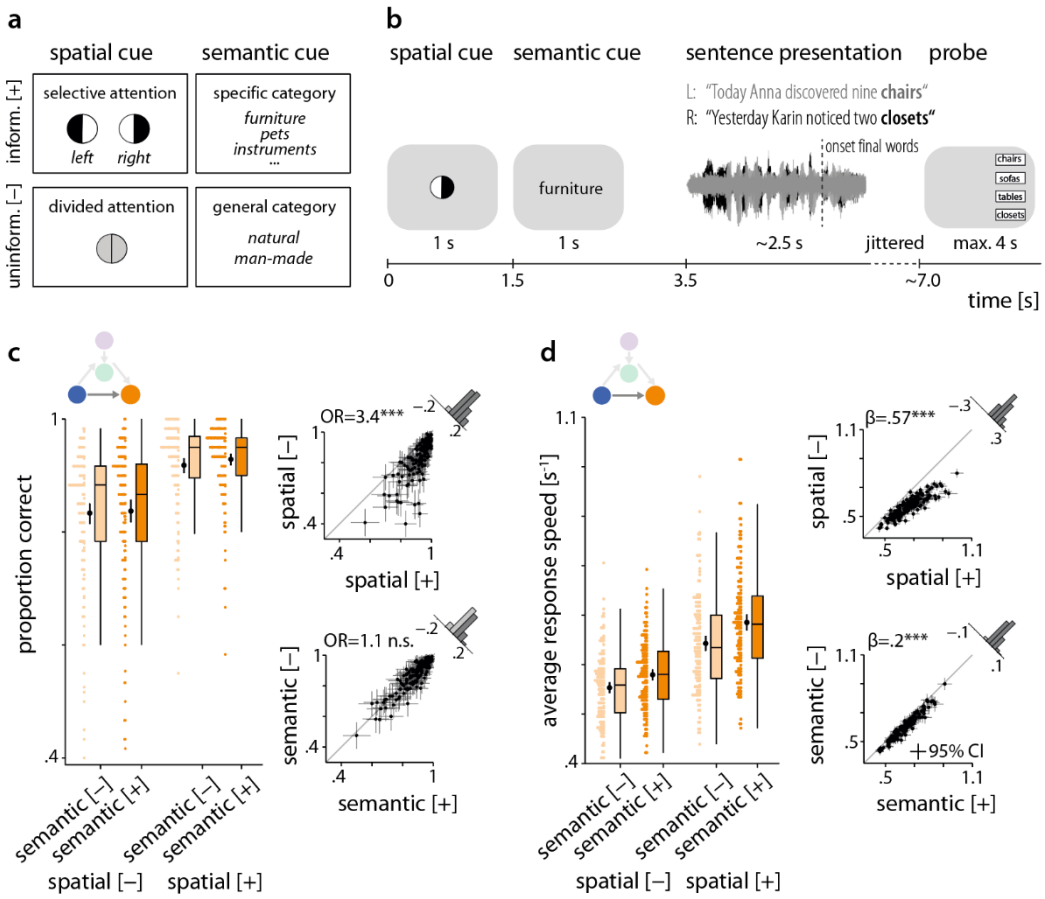

Figure 2. Experimental design and behavioral benefit from informative cues.
(a) Visualization of the employed 2×2 design (30). The level of the spatial and semantic cue differed on a trial-by-trial basis. Top row shows the informative [+] cue levels, bottom row the uninformative [–] cue levels.
(b) Schematic representation of the temporal order of events for a given trial. Successive display of the two visual cues precedes the dichotic presentation of two sentences spoken by the same female talker. After sentence presentation, participants had to select the final word from four alternative words.
(c) Grand average and individual results of accuracy shown per cue-cue combination. Colored dots are individual (N=155) trial-averages, black dots and vertical lines show group means ± bootstrapped 95% confidence intervals (left side). Individual cue benefits displayed separately for the two cues (top: spatial cue, bottom: semantic cue). Black dots indicate individual trial-averages ± bootstrapped 95 % confidence intervals. Histograms show the distribution of the difference for informative vs. uninformative levels. OR: Odds ratio parameter estimate from generalized linear mixed-effects models (right side); *** = $p < .001$.
(d) As in (c) but shown for response speed; β: slope parameter estimate from general linear mixed-effects models.

The following figure supplements are available for figure 2:
Figure supplement 1. (a) Histogram showing age distribution of N=155 participants across two-year age bins. (b) Individual (N=155) and mean air-conduction thresholds (PTA) averaged across the left and right ear, and four age groups.
Figure supplement 2. Analysis of response types split into correct responses, spatial stream confusions and random errors.

Informative spatial cues improve listening success

For behavioral performance, we tested the impact of informative versus uninformative listening cues on listening success. Overall, participants achieved a mean accuracy of

87.8 % ± sd 9.1 % with a mean reaction time of 1742 ms ± sd 525 ms; as response speed: 0.62 s^{-1} ± sd 0.17 s^{-1}.

As expected, behavior depended on the different combinations of listening cues (Fig. 2c,d). Informative compared to uninformative spatial-attention cues yielded a strong behavioral benefit. In selective-attention trials, participants responded more accurately and faster (accuracy: generalized linear mixed-effects model (GLMM); *odds ratio (OR)*=3.5, *std. error (SE)* =.12, $p<.001$; response speed: linear mixed-effects model (LMM); β=.57, SE=.04, $p<.001$). That is, when cued to one of the two sides, participants responded on average 261 ms faster and their probability of giving a correct answer increased by 6 %.

Also, participants responded generally faster in trials in which they were given a specific, more informative semantic cue (LMM; β=.2, *SE*=.03, $p<.001$), most likely reflecting a semantic priming effect that led to faster word recognition.

As in a previous fMRI implementation of this task (30), we did not find evidence for any interactive effects of the two listening cues on either accuracy (GLMM; *OR*=1.3, *SE*=.21, p=.36) or response speed (LMM; β=.09, *SE*=.06, p=.31). Moreover, the breakdown of error trials revealed a significantly higher proportion of spatial stream confusions (6 % ± sd 8.3 %) compared to random errors (3 % ± sd 3.4 %; paired t-test on logit-transformed proportions: t_{155} = 6.53, $p<.001$; see also Fig. 2–supplement 2). The increased rate of spatial stream confusions (i.e., responses in which the last word of the to-be-ignored sentence was chosen) attests to the distracting nature of dichotic sentence presentation and thus heightened task difficulty.

Spatial attention modulates both alpha power lateralization and neural speech tracking in auditory cortex

In line with our second research question, following source projection of EEG data, we probed whether the presence of an informative spatial attention cue would lead to reliable modulation of both alpha power and neural speech tracking within an *a priori* defined auditory region of interest (ROI; see Fig. 3–supplement 1 and Methods for details).

For alpha power, we expected an attention-induced lateralization due to a decrease in power contralateral and concomitant increase in power ipsilateral to the focus of attention. For neural speech tracking, we expected stronger neural tracking of attended compared to ignored speech under selective attention but no such systematic difference in the neural tracking of probed and unprobed sentences in divided-attention trials. Accordingly, our analyses of alpha power and neural tracking focused on attentional modulation index measures that contrast the relative strength of neural responses to target versus distractor stimuli.

We compared alpha power changes ipsi- and contralateral to the probed ear to derive a temporally-resolved single-trial measure of alpha power lateralization [alpha lateralization index (ALI) = ($\alpha\text{-power}_{ipsi} - \alpha\text{-power}_{contra}$) / ($\alpha\text{-power}_{ipsi} + \alpha\text{-power}_{contra}$)] (12).

As shown in Figure 3a, a spatial cue—that is, the instruction to pay attention to a given side—elicited a pronounced lateralization of 8–12 Hz alpha power within the auditory ROI. Lateralization of alpha power was evident following the spatial cue itself and during dichotic sentence presentation with its strongest peak around final word presentation.

As expected, the statistical analysis of alpha power lateralization during sentence presentation (time window: 3.5–6.5 s) revealed a significant modulation by attention that was additionally influenced by the probed ear (LMM; spatial cue x probed ear: β=.13, SE=.02, p<.001; Fig. 3b). Follow-up analysis showed a significant difference in alpha power lateralization between selective- and divided-attention trials when the right ear but not when the left was probed (LMM, right ear probed: β=.12, SE=.01, p<.001; LMM, left ear probed: β=–.016, SE=.013, p=.55; see Table S3 for full model details). This pattern suggests that when given an uninformative spatial cue, participants presumably payed overall more attention to the left-ear stimulus leading to an increase in alpha lateralization for probed-left compared to probed-right trials.

Notably, we did not find any evidence for a modulation by the semantic cue nor any joint influence of the spatial and semantic cue on alpha lateralization during sentence presentation (LMM; semantic cue main effect: β=–.01, SE=.01, p=.53, spatial x semantic cue: β=–.02, SE=.02, p=.53).

Not least, the extent of overall as well as attention-specific alpha power lateralization was unaffected by participants' chronological age and hearing loss (p-values > .27 for main effects of age, PTA, and their respective interactions with the spatial-attention cue; see also Table S6 for a corresponding analysis of alpha power during the interval of the final word).

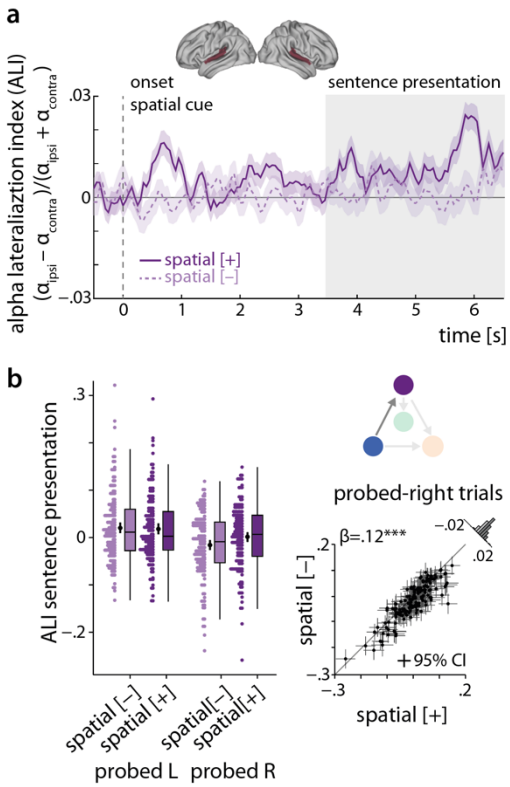

Figure 3. Informative spatial cue elicits increased alpha power lateralization before and during speech presentation.
(a) Attentional modulation of 8–12 Hz auditory alpha power shown throughout the trial for all N=155 participants. Brain models indicate the spatial extent of the auditory region of interest (shown in red). Purple traces show the temporally resolved alpha lateralization index (ALI) for the informative (solid dark purple line) and the uninformative spatial cue (dashed light purple line), each collapsed across the semantic cue levels. Positive values indicate relatively higher alpha power in the hemisphere ipsilateral to the attended/-probed sentence compared to the contralateral hemisphere. Shaded grey area show the temporal window of sentence presentation.
(b) Alpha lateralization index (ALI) during sentence presentation (3.5–6.5 s) shown separately per spatial cue condition and probed ear (left plot). Colored dots represent trial-averaged individual results, black dots and error bars indicate the grand-average and bootstrapped 95% confidence intervals. For attend-right/ignore-left trials we observed a significant difference in alpha power lateralization between selective- and divided-attention trials (right plot). Black dots represent individual mean ALI values ± bootstrapped 95 % confidence intervals. Histogram shows distribution of differences in ALI for informative vs. uninformative spatial-cue levels. β: slope parameter estimate from the corresponding general linear mixed-effects model; *** = p<.001.

In close correspondence to the alpha-power analysis, we investigated whether changes in attention or semantic predictability would modulate the neural tracking of attended versus ignored speech. We used linear backward ('decoding') models to reconstruct the onset envelopes of the to-be-attended and ignored sentences (for simplicity hereafter referred to as *attended* and *ignored*) from neural activity in the auditory ROI. Reconstruction models were trained on selective-attention trials only, but then utilized to reconstruct attended (probed) and ignored (unprobed) envelopes for both attention conditions (see Methods and Fig. 4a along with supplement 1 for details).

In line with previous studies (21, 23, 41), the forward transformed temporal response functions (TRFs) show increased encoding of attended compared to ignored speech in the time window covering the $N1_{TRF}$ and $P2_{TRF}$ component (see Fig. 4b, left panel). Here, however, this was observed particularly for right-ear inputs processed in the left auditory ROI.

Further attesting to the validity of our reconstruction models, reconstructed attended envelopes were overall more similar to the envelope of the to-be-attended sentence than to that of the to-be-ignored sentence, and vice versa for the reconstructed ignored envelopes (see Fig. 4b, right panel).

As shown in Figure 4c, the differential neural tracking of attended and ignored envelopes (probed and unprobed envelopes under divided attention) was modulated

by attention. Following an informative spatial cue, the neural tracking index becomes increasingly positive during the second half of sentence presentation with its highest peaks around final word onset.

The statistical analysis of single-trial index values averaged for the time interval of final word presentation confirmed this pattern: The difference in the neural tracking of attended and ignored sentence was generally more pronounced under selective compared to divided attention. However, this effect was also modulated by differences in sentence onset: The difference in neural tracking between the two attention conditions was reduced when the attended/probed sentence started ahead of the distractor sentence. This effect was driven by an increase in differential neural tracking for divided attention in such trials: in absence of an informative spatial cue, participants' attention was captured by the sentence with the earlier onset. Consequently, we observed overall more positive index values when the earlier sentence was probed compared to when it was not probed (LMM, earlier onset x spatial cue: β=–.05, SE=.02, p=.049, see Fig. 4d and Table S4 for full model details).

We also found a neural correlate of the known right-ear advantage for verbal materials, that is, an overall stronger tracking of left-ear inputs. This effect was independent of spatial-attention cueing (LMM; probed ear main effect: β=–.03, SE=.01, p=.023; spatial cue x probed ear: β=.02, SE=.02, p=.54). As for alpha power, we did not observe any modulation of neural tracking by the semantic cue, nor any joint influence of the spatial and semantic cue (LMM; semantic cue main effect: β=–.01, SE=.01, p=.53, interaction spatial x semantic cue: β=–.02, SE=.02, p=.53).

Again, participants' age and hearing status did not prove significant predictors of neural tracking (p-values > .54 for main effects of age, PTA, and their respective interactions with the spatial-attention cue, see also Table S7 for a corresponding analysis of neural tracking during the entire sentence presentation).

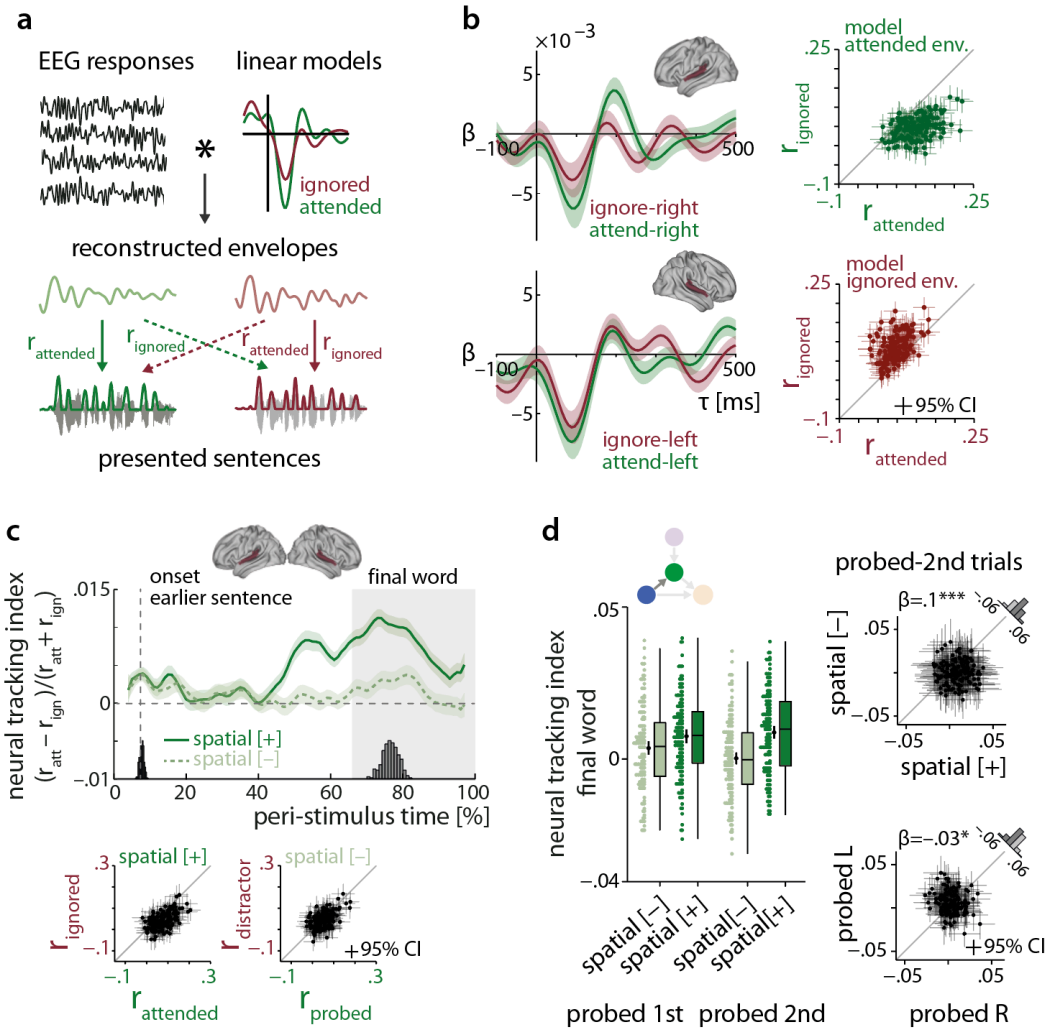

Figure 4. Neural speech tracking of attended and ignored sentences.
(a) Schematic representation of the linear backward model approach used to reconstruct onset envelopes on the basis of recorded EEG responses. Following the estimation of linear backward models using only selective-attention trials (see Fig. 4–supplement 1 for details), envelopes of attended and ignored sentences were reconstructed via convolution of EEG responses with the estimated linear backward models in all functional parcels within the auditory ROI. Reconstructed onset envelopes were compared to the onset envelopes of presented sentences to assess neural tracking strength and decoding accuracy (see Supplements).
(b) Left: forward-transformed temporal response functions (TRFs) for attended (green) and ignored (red) speech trained on selective-attention trials. Models are averaged across all N=155 participants, and shown separately for the left and right auditory ROI. Error bands indicate 95 % confidence intervals. Right: Single-subject average Pearson correlation of reconstructed envelopes with the envelopes of the presented sentences. Top panel shows results for the reconstruction of the attended envelope, bottom panel for the ignored envelope along with 95% CI. Shown for all N=155 participants.
(c) Top: Grand average peri-stimulus time course (see Supplements for details) of neural tracking index shown separately for selective (solid dark green curve) and divided attention (dashed light green curve) along with error band reflecting ± 1 SE. Histograms indicate the distribution of sentence and final word onsets, respectively. Grey shaded area reflect the time interval of final word presentation that was used for statistical analysis. Attentional modulation in the neural tracking of attended (under divided attention: probed) speech. Bottom: 45° plot contrasting the neural tracking of attended (probed) and ignored (distractor) envelope during final word presentation for selective and divided attention, respectively.
(d) Left: Neural tracking index shown separately for spatial-attention conditions and for trials in which cued/probed sentences started ahead of ('probed first') or after the distractor ('probed second'). Colored dots represent trial-averaged individual results, black dots and error bars indicate the grand-average and bootstrapped 95 % confidence intervals. Right: In probed-second trials there was a significant difference in neural tracking between selective- and divided-attention trials (top plot), as well as overall more selective tracking in cued/probed-left trials. Black dots in 45° plot represent individual mean neural tracking index values ± bootstrapped 95 % confidence intervals. Histogram shows distribution of differences in neural

tracking in informative vs. uninformative spatial-cue trials, and probed-left vs. probed-right trials, respectively. β: slope parameter estimate from the corresponding general linear mixed-effects model; *** = p<.001, *= p<.5.

The following figure supplements are available for figure 4:
Figure supplement 1. Training and testing of envelope reconstruction models.
Figure supplement 2. Decoding accuracy of attended and ignored speech shown per attention condition (N=155).

Trial-to-trial neural speech tracking is independent of synchronous alpha power lateralization

Our third major goal was to investigate whether the modulation of alpha power and neural speech tracking reflect two dependent neural strategies of auditory attention at all. We asked whether neural tracking could be explained by auditory alpha power lateralization either at the state level (i.e., within an individual from trial to trial) or at the trait level (i.e., between individual mean levels; see Statistical analysis for details). If modulations of alpha power over auditory cortices indeed act as a neural filter mechanism to selectively gate processing during early stages of sensory analysis then heightened levels of alpha power lateralization should lead to a more differentiated neural tracking of the attended vs. ignored speech and thus more a positive neural tracking index (cf. Fig. 5a).

However, in the analysis of the task-relevant final word period, we did not find evidence for an effect of alpha power lateralization on neural tracking at neither the state nor the trait level (Fig. 5b, LMM; ALI within-subject effect: β=–.008, *SE*=.005, *p*=.35; ALI between-subject effect: β=–.0007, *SE*=.007, *p*=.98; see Table S5 for full model details). This notable absence of an alpha lateralization–speech tracking relationship was independent of the respective spatial-attention condition and probed ear (all *p-values*>.35).

To complement our fine-grained single-trial level investigation into the neural-neural relationship with a coarser, yet time-resolved analysis, we related the temporal dynamics of both neural measures in an exploratory between-subjects cross-correlation analysis. As shown in Figure 5c, under selective attention, neural tracking and alpha power lateralization follow different temporal trajectories with neural tracking peaking earlier than alpha lateralization around final word presentation. The average cross-correlation of the two neural time courses during sentence presentation confirms a systematic temporal delay with fluctuations in neural tracking leading those in alpha power by about 520 ms (see Fig. 5d).

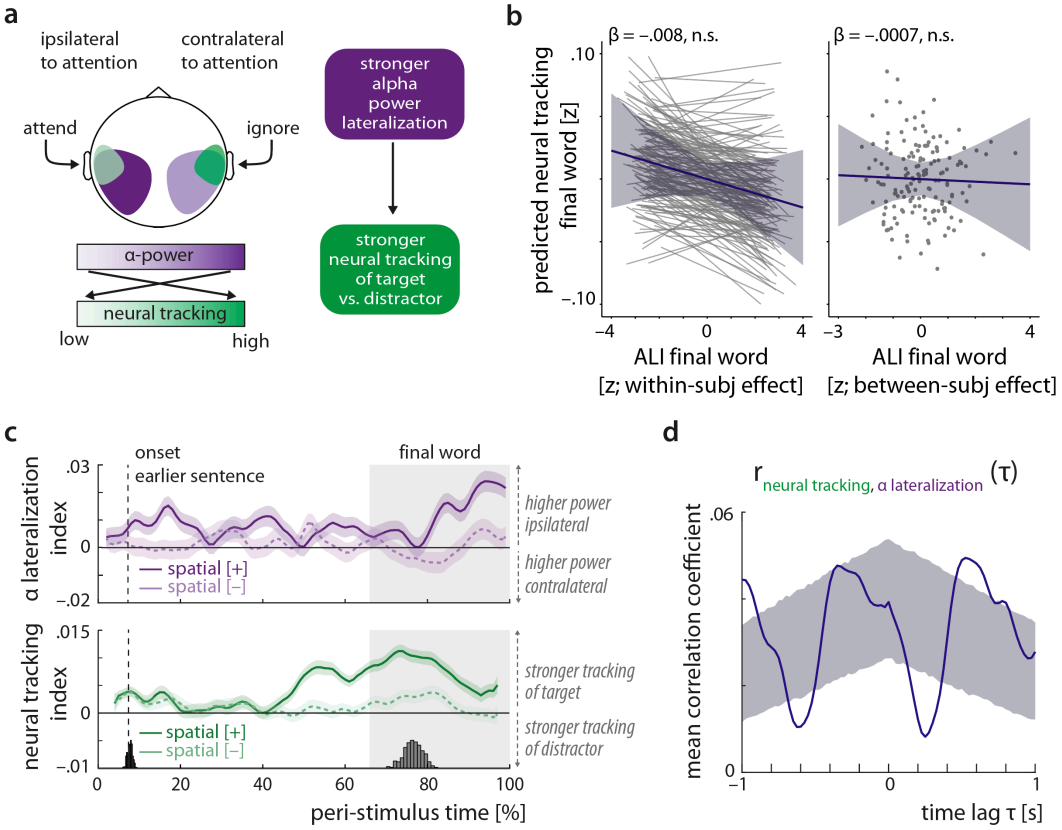

Figure 5. Relationship of alpha power lateralization and neural tracking.
(a) Hypothesized relationship of alpha power and neural speech tracking within the auditory region of interest. Changes in alpha lateralization are assumed to drive changes in neural tracking. Schematic representation for an attend-left trial.
(b) Independence of neural tracking and alpha power lateralization during the final word presentation as shown by the predictions from the same linear mixed effect model. Plots show the predicted, non-significant effect of within- and between-subject variations in alpha lateralization on selective neural tracking, respectively. Blue lines indicates the respective predicted fixed effects ± 95 % CI, grey thin lines in the left plot show N=155 subject-specific random slopes (included for illustrative purposes only), and grey dots show average predictions per subject. β: slope parameter estimate from the corresponding general linear mixed-effects model; *** = p<.001.
(c) Grand-average time courses of alpha lateralization and neural tracking during sentence presentation mapped to the same peri-stimulus time axis. Shown separately for selective-attention (darker, solid curves) and divided-attention trials (lighter, dashed curves). Error bands reflect ± 1 SE. Note how the peak in neural tracking under selective attention precedes the peak in alpha power lateralization.
(d) Mean normalized cross-correlation of trial-averaged neural tracking and alpha power lateralization time courses. Upper and lower bound of shaded area reflect the 97.5th and 2.5th percentile of surrogate data derived from 5000 independently permuted time courses of alpha power and neural tracking.

Neural speech tracking but not alpha lower lateralization explains listening behavior

Having established the functional independence of alpha power lateralization and speech tracking at the single-trial level, the final piece of our investigation is to probe their relative functional importance for behavior outcome.

Using the same (generalized) linear mixed-effects models as in testing our first research question (Q1 in Fig.1), we investigated whether changes in task performance could be explained by the independent (i.e., as main effects) or joint influence (i.e., as an interaction) of neural measures. Again, we modelled the influence of the two neural filter strategies on behavior at the state and trait level (42).

For response accuracy, our most important indicator of listening success, we observed an effect of trial-by-trial variation in neural tracking both during the presentation of the final word and across the entire sentence: participants had a higher chance of responding correctly in trials in which they neurally tracked the cued/probed sentence more strongly than the distractor sentence (see Fig. 6a, left panel). For changes in neural tracking extracted from the entire sentence presentation, this effect occurred independently of other modelled influences (GLMM; main effect neural tracking (within-subject effect): *OR*=1.06, *SE*=.02, *p*=.03; see Table S12 for full model details) while it was generally less pronounced and additionally modulated by the probed ear for the period of the task-relevant final word (GLMM; probed ear x neural tracking (within-subject effect): *OR*=1.1, *SE*=.04, *p*=.03; see Table S1 for full model details).

The data held no evidence for any direct effects of trial-to-trial or participant-to-participant variation in alpha lateralization during sentence or final word presentation on accuracy (all *p*-values > .18; see Tables S1 and S12 for full model details). We also did not find evidence for any joint effects of alpha power and neural tracking extracted from either of the two time windows (all *p*-values >.33). Importantly, the absence of an effect did not hinge on differences in neural measures across spatial-cue, or probed-ear levels as relevant interactions of neural measures with these predictors were included in the model (all *p*-values > .55).

The observed effects of neural filters on response speed depended on the analyzed time window: while participants with relatively higher average levels of neural tracking during sentence presentation responded overall faster (LMM, neural tracking (between-subject effect): β=0.08, SE=.03, *p*=.01; see Fig 6a, right panel and Table S13 for full model details), we found a joint effect of neural dynamics measured during final word presentation. Under selective but not divided attention, response speed was jointly predicted by trial-to-trial variation in both alpha lateralization and neural tracking (LMM; spatial cue x ALI (within-subject effect) x neural tracking (within-subject effect): β=0.08, SE=.03, *p*=.01; see Table S13 for full model details). In short, responses were fastest in trials where relatively elevated levels in one neural measure were paired with relatively reduced levels in the other.

In line with the literature on listening behavior in aging adults (e.g. 43, 44), the behavioral outcome was further reliably predicted by age, hearing loss, and probed ear. We observed that participants' performance varied in line with the well-attested *right-ear advantage* (REA, also referred to as *left-ear disadvantage*) in the processing of linguistic materials (37). More specifically, participants responded both faster and more accurately when they were probed on the last word presented to the right compared to the left ear (response speed: LMM; β=.08, *SE*=.013, *p*<.001; accuracy: GLMM; *OR*=1.25, *SE*=.07, *p*=.006; see also Fig. 6–supplement 1) .

Increased age led to less accurate and slower response (accuracy: GLMM; *OR*=.80, *SE*=.08, *p*=.025; response speed: LMM; β=–.15, *SE*=.03, *p*<.001). In contrast,

increased hearing loss led to less accurate (GMM; *OR*=.75, *SE*=.08, *p*=.002) but not slower responses (LMM; *β*=–.05, *SE*=.03, *p*=.21, see Table S1, S2, and Fig.6–supplement 2).

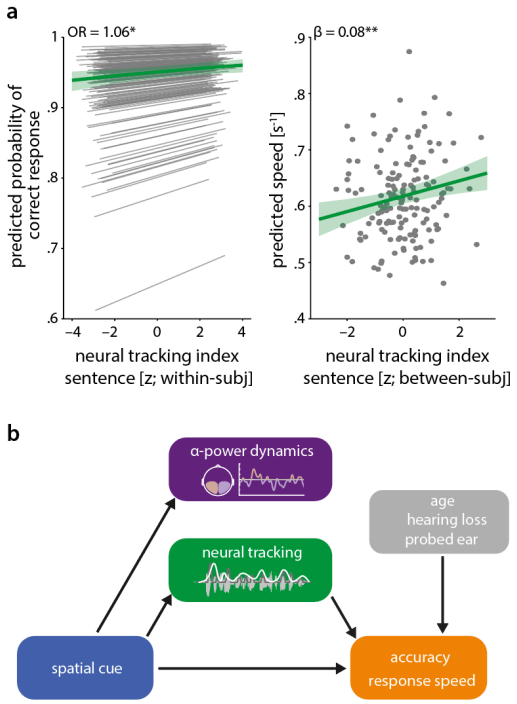

Figure 6. Neural speech tracking predicts listening behavior.
(a) Model predictions for the effect of neural tracking on behavior for N=155 participants. Left panel shows the predicted group-level fixed effect (green line ± 95 % CI) of trial-to-trial variation in neural tracking on accuracy. Grey thin lines indicate estimated subject-specific slopes. Right panel shows the predicted group-level fixed effect of neural tracking at the between-subject level on response speed (green line ± 95 % CI). Grey dots indicate subject-level model predictions. OR = odds ratio, *β* = β: slope parameter estimate from the corresponding general linear mixed-effects model; ** = p<.01, * = p<.05.
(b) Summary of results. Black arrows highlight statistically significant effects from (generalized) single-trial linear mixed-effects modelling. Notably, changes in age and hearing loss did not modulate the fidelity of the two key neural measures.

Control analyses

We ran additional control analyses to validate our main set of results. We asked whether the observed independence of alpha lateralization and neural tracking hinged on the precise time window and cortical site from neural measures were extracted. One set of control models thus included alpha lateralization during spatial-cue rather than sentence presentation as a predictor of neural tracking during sentence and final word presentation. Next, we related the two neural filters during the entire sentence presentation rather than only during the final word. Finally, we tested the hypothesis that neural speech tracking might be driven not primarily by alpha-power modulations emerging in auditory cortices, but rather by those generated in domain-general attention networks in parietal cortex (45). We therefore ran control models including alpha lateralization within the inferior parietal lobule. However, none of these additional analyses found evidence for an effect of alpha lateralization on neural speech tracking (see Tables S8–S11).

With respect to the functional relevance of neural measures to overt behavior we tested whether changes in age or hearing loss would modulate the relationship of neural tracking and alpha lateralization with listening behavior. However, the inclusion of the respective interaction terms did not further improve the statistical models of accuracy and response speed (all *p*-values > .44).

Discussion

The present study has utilized the power of a large, representative sample of middle-aged and older listeners to explicitly address the question of how the two most eminent neurobiological implementations of attentional filtering, typically studied in isolation, relate to one another, and how they jointly shape listening success.

In our dichotic listening task, we source-localized the electroencephalogram, and observed systematic cue-driven changes within auditory cortex in alpha power lateralization and the differential neural tracking of attended versus ignored speech.

These results do provide unprecedented large-sample support for their suggested roles as neural signatures of selective attention. Additionally, they address important but still insufficiently answered questions: how do the two neural filters relate to one another, and how do they influence listening success in a demanding real-life listening situation? In addition, we leveraged our large age-varying sample to ask how chronological age and hearing loss affect the fidelity of neural filter strategies and their influence on behavior.

First, an informative spatial-attention cue not only boosted both neural measures but also consistently boosted behavioral performance. Listening behavior was additionally influenced by both trial-to-trial and individual-to-individual variation in neural speech tracking with relatively stronger tracking leading to better performance.

Second, when related at the single-trial, single-subject level, we found the modulation of the two neural attentional filter mechanisms to be statistically independent, underlining their functional segregation and speaking to two distinct neurobiological implementations. Yet, when related in a coarser, between-subjects analysis across time, peaks in selective neural tracking systematically preceded those in alpha power lateralization.

Importantly, while chronological age and hearing loss reliably decreased behavioral performance they did not systematically affect the fidelity of neural filter strategies nor their influence on behavior.

Increased neural tracking but not alpha lateralization important for listening success?

The present study aimed at connecting the instantiation of selective attention at the neurobiological and behavioral level: we have explicitly addressed the often overlooked question of how neural filter states (i.e., fluctuations from trial-to-trial) impact behavior, here single-trial listening success (46-49).

Notably, using a sophisticated linear-model approach that probed the impact of both state- and trait-level modulation of neural filters on behavior, we only found evidence for a direct influence of neural speech tracking but not alpha lateralization on behavioral performance (see Fig. 6).

At first glance, the presence of a brain–behavior relationship for neural tracking but not alpha power may appear unexpected given the observed robust spatial-cue

effect on both neural filters and behavior. However, previous studies have provided only tentative support for the behavioral relevance of the two neural filter solutions, and the observed brain–behavior relations do not always converge.

To date, the available evidence for the behavioral relevance of selective neural speech tracking is generally sparse given the emphasis on more naturalistic, yet more complex language stimuli (50, 51). While these stimuli provide a window onto the most natural forms of speech comprehension, they are not easily paired with fine-grained measures of listening behavior. This makes it particularly challenging to establish a direct link between differential neural speech tracking and listening success (20, 21, 23, 52-55). Our current results thus provide important additional fine-grained support to the previously observed beneficial role of a more pronounced cortical representation of behaviorally relevant speech signals (22).

The link between alpha power lateralization and performance has been probed in various spatial-attention tasks. Collectively, the results suggest that the presence of a stable brain-behavior relationship hinges on a number of different factors.

First of all, the establishment of a systematic brain-behavior relationship depends on the age groups tested: most of the evidence associating increased levels of alpha power lateralization with better task performance in spatial-attention tasks stems from smaller samples of young adults (9, 12, 56, 57). By contrast, the observation of such a systematic relationship in middle-age and older adults is obscured by considerable variability in age-related changes at the neural (and to some extent also behavioral) level (11, 58-61; see discussion below).

Second, previous findings and their potential interpretation differ along (at least) two dimensions: (i) whether the functional role of alpha power lateralization is studied during attention cueing, stimulus anticipation, or stimulus presentation (6, 58, 62), and (ii) whether behavior is related to the overall strength of power lateralization or its stimulus-driven rhythmic modulation (cf. 9, 11). Depending on these factors, the observed brain–behavior relations may relate to different top-down and bottom-up processes of selective auditory attention.

Third, as shown in a recent study by Wöstmann et al. (62), the neural processing of target and distractor are supported by two uncorrelated lateralized alpha responses emerging from different neural networks. Notably, their results provide initial evidence for the differential behavioral relevance of neural responses related to target selection and distractor suppression.

The present study deliberately aimed at establishing a brain–behavior relationship at the most fine-grained trial-by-trial, single-subject level as this type of analysis provides the most compelling evidence (63-65). Importantly, the absence of a significant relationship of alpha lateralization and behavior did not hinge on this particular focus as we probed the relationship at the neural trait level (i.e. variation between individuals) as well.

575 ***Are fluctuations in lateralized alpha power and selective neural speech tracking***
576 ***functionally connected?***

577 We investigated attention-related changes in two neural filter strategies that (i) involve
578 neurophysiological signals operating at different frequency regimes, (ii) are assumed to
579 support auditory attention by different neural mechanisms, and (iii) are typically studied
580 in isolation (3, 19). Here, we found both neural filter strategies to be impacted by the
581 same spatial-attention cue which afforded new insights into their neurobiological
582 dependence.

583 We asked whether changes in alpha power, assumed to reflect a neural
584 mechanisms of controlled inhibition, and changes in neural speech tracking, assumed
585 to reflect a sensory-gain mechanism, are functionally connected signals (16, 17). Indeed,
586 there is preliminary evidence suggesting that these two neural filter strategies may
587 exhibit a systematic relationship (6, 11, 27, 28, 66).

588 However, our in-depth trial-by-trial analysis revealed independent modulation
589 of alpha power and neural speech tracking. Therefore, the present results speak against
590 a consistent, linear relationship of momentary neural filter states. We observed instead
591 the coexistence of two complementary but seemingly independent neurobiological
592 solutions to the implementation of auditory selective attention.

593 Our single-trial results are well in line with recent reports of independent
594 variation in alpha band activity and steady-state (SSR) or frequency following responses
595 (FFR) in studies of visual spatial attention (25, 26, 67). Additionally, the inclusion of
596 single-trial alpha power lateralization as an additional training feature in a recent
597 speech-tracking study failed to improve the decoding of attention (68, but see 69 for
598 the joint influence of alpha power and neural tracking on speech intelligibility).

599 Taken together with findings from previous electrophysiological studies (6, 27,
600 28), our results provide new evidence in favor of a proposed functional trade-off
601 between neurobiological attentional filter mechanisms (70). The assumption of
602 mutually exclusive states of strong neural tracking and alpha power lateralization is
603 receiving additional support from our exploratory between-subjects cross-correlation
604 analysis: we observed a systematic temporal delay with peaks in neural tracking leading
605 those in alpha lateralization. The direction and duration of this delay were closely in line
606 with previous findings (9, 11) and, at this coarser level of analysis, speak against a
607 hierarchy of neural processing in which lateralized alpha responses govern the
608 differential tracking of attended versus ignored speech (71). The differences in temporal
609 dynamics of the two neural filters, here observed between individuals, highlight an
610 additional challenge in the establishment of a systematic single-trial brain-brain
611 relationship within a given time window.

612 In sum, we suggest that the intricacies of how neural auditory filter strategies
613 interact hinge on several factors including the particular task demands, and potentially
614 the level of temporal and spatial resolution at which the two neural signatures are
615 examined (for comparison see ref. 28).
616

Do age and hearing loss affect neural filter strategies?

The detrimental effect of increasing age and associated hearing loss on speech comprehension in noisy listening situations is well attested (72) and borne out by the current results. The extent to which the neural implementation of attentional filters is affected by age and hearing loss, and in how far they may constitute neural markers of age-related speech comprehension problems, on the other hand, remains poorly understood.

As in a previous study on a subset of the current sample, we found the fidelity of alpha power lateralization unchanged with age (11). Other studies on auditory attention, however, have observed diminished and less sustained alpha power lateralization for older compared to younger adults that were to some extend predictive of behavior (58, 73, 74).

Our observation of preserved neural speech tracking across age and hearing levels only partially agree with earlier findings. They are in line with previous reports of differential neural tracking of attended and ignored speech for hearing impaired older adults that mimicked the attentional modulation observed for younger or older normal-hearing adults (38-40, 75). As revealed by follow-up analysis (data not shown), however, our data do not provide evidence for a differential impact of hearing loss on the neural tracking of attended or ignored speech as found in some of these studies. We also did not find evidence for overall increased levels of cortical neural tracking with age as observed in earlier studies (76, 77).

The discrepancy in results may be explained by differences in (i) the studied populations (i.e., whether groups of younger and older participants were contrasted compared to the modelling of continuous changes in age and hearing loss), (ii) whether natural stories or short Matrix sentence speech materials were used (78), or (iii) by differences in task details. In sum, the results suggest that the commonly observed adverse effects of age and hearing loss on speech-in-noise processing are not readily paired with concomitant changes at the neural level.

Conclusion

In a large, representative sample of aging listeners, we have provided support for the roles of alpha power lateralization and neural speech tracking as neural signatures of spatial attention. Only for neural speech tracking, we have been able to establish a mechanistic link from trial-to-trial neural filtering during the concurrent sound input to the ensuing behavioral outcome. This link is independent from age and hearing status, which points to the potency of neural speech tracking to serve as a separate marker of comprehension problems in clinical settings, and a basis for translational neurotechnological advances.

This key advance notwithstanding, our results also highlight the level of complexity associated with establishing statistically robust brain–behavior relationships in the deployment of auditory attention. To understand how the brain enables successful selective listening it is necessary that studies go beyond

characterization of neurobiological filter strategies alone, and further jointly account for the variability in both neural states and behavioral outcomes.

662 ## Materials and Methods

663 ### *Data collection*

664 The analyzed data are part of an ongoing large-scale study on the neural and cognitive
665 mechanisms supporting adaptive listening behavior in healthy middle-aged and older
666 adults ("The listening challenge: How ageing brains adapt (AUDADAPT)";
667 https://cordis.europa.eu/project/rcn/197855_en.html). This project encompasses the
668 collection of different demographic, behavioral, and neurophysiological measures
669 across two time points. The analyses carried out on the data aim at relating adaptive
670 listening behavior to changes in different neural dynamics (30).
671

672 ### *Participants and procedure*

673 A total of $N = 155$ right-handed German native speakers (median age 61 years; age range
674 39–80 years; 62 males; see Fig. 2–supplement 1 for age distribution) were included in
675 the analysis. Handedness was assessed using a translated version of the Edinburgh
676 Handedness Inventory(79). All participants had normal or corrected-to-normal vision,
677 did not report any neurological, psychiatric, or other disorders and were screened for
678 mild cognitive impairment using the German version of the 6-Item Cognitive
679 Impairment Test (6CIT; 80).
680 During the EEG measurement, participants performed six blocks of a
681 demanding dichotic listening task (see Fig. 2 and Supplements for details on task and
682 sentence materials).
683 As part of our large-scale study on adaptive listening behavior in healthy aging
684 adults, prior to the EEG session, participants also underwent a session consisting of a
685 general screening procedure, detailed audiometric measurements, and a battery of
686 cognitive tests and personality profiling (see ref. 11 for details). Only participants with
687 normal hearing or age-adequate mild-to-moderate hearing loss were included (see Fig.
688 2–supplement 1 for individual audiograms). As part of this screening procedure an
689 additional 17 participants were excluded prior to EEG recording due to non-age-related
690 hearing loss or a medical history. Three participants dropped out of the study prior to
691 EEG recording and an additional 9 participants were excluded from analyses after EEG
692 recording: three due to incidental findings after structural MR acquisition, and six due
693 to technical problems during EEG recording or overall poor EEG data quality.
694 Participants gave written informed consent and received financial compensation (8€
695 per hour). Procedures were approved by the ethics committee of the University of
696 Lübeck and were in accordance with the Declaration of Helsinki.
697
698
699

700 ## Behavioral data analysis

701 We evaluated participants' behavioral performance in the listening task with respect to
702 accuracy and response speed. For the binary measure of accuracy, we excluded trials in
703 which participants failed to answer within the given 4-s response window ('timeouts').
704 Spatial stream confusions, that is trials in which the sentence-final word of the to-be-
705 ignored speech stream were selected, and random errors were jointly classified as
706 incorrect answers. The analysis of response speed, defined as the inverse of reaction
707 time, was based on correct trials only. Single-trial behavioral measures were subjected
708 to (generalized) linear mixed-effects analysis and regularized regression (see Statistical
709 analysis).
710

711 ## EEG data analysis

712 The preprocessed continuous EEG data (see Supplements for details on data collection
713 and preprocessing) were high-pass-filtered at 0.3 Hz (finite impulse response (FIR) filter,
714 zero-phase lag, order 5574, Hann window) and low-pass-filtered at 180 Hz (FIR filter,
715 zero-phase lag, order 100, Hamming window). The EEG was cut into epochs of –2 to 8 s
716 relative to the onset of the spatial-attention cue to capture cue presentation as well as
717 the entire auditory stimulation interval.
718 For the analysis of changes in alpha power, EEG data were down-sampled to
719 f_s=250 Hz. Spectro-temporal estimates of single-trial data were then obtained for a time
720 window of –0.5 to 6.5 s (relative to the onset of the spatial-attention cue) at frequencies
721 ranging from 8 to 12 Hz (Morlet's wavelets; number of cycles = 6).
722 For the analysis of the neural encoding of speech by low-frequency activity, the
723 continuous preprocessed EEG were down-sampled to f_s=125 Hz and filtered between
724 f_c=1 and 8 Hz (FIR filters, zero-phase lag, order: $8f_s/f_c$ and $2f_s/f_c$, Hamming window). The
725 EEG was cut to yield individual epochs covering the presentation of auditory stimuli,
726 beginning at noise onset until the end of auditory presentation.
727 Following EEG source and forward model construction (see Supplements for
728 details), sensor-level single-trial data in each of our two analysis routines were projected
729 to source space by matrix-multiplication of the spatial filter weights. To increase signal-
730 to-noise ratio in source estimates and computationally facilitate source-level analyses,
731 source-projected data were averaged across grid-points per cortical area defined
732 according to the HCP functional parcellation template (81, 82). This parcellation
733 provides a symmetrical delineation of each hemisphere into 180 parcels for a total of
734 360 parcels. We constrained the analysis of neural measures to an a priori defined,
735 source-localized auditory region of interest (ROI) as well as one control ROI in the inferior
736 parietal lobule (see Supplements for details).
737

738 ### *Attentional modulation of alpha power*

739 Absolute source power was calculated as the square-amplitude of the spectro-temporal
740 estimates. Since oscillatory power values typically follow a highly skewed, non-normal
741 distribution, we applied a nonlinear transformation of the Box-Cox family ($power_{trans}$ =

(power^p −1)/p with p=0.5) to minimize skewness and to satisfy the assumption of normality for parametric statistical tests involving oscillatory power values (83). To quantify attention-related changes in 8–12 Hz alpha power, per ROI, we calculated the single-trial, temporally resolved alpha lateralization index as follows (9, 11, 12):

$$ALI = (\alpha\text{-power}_{ipsi} - \alpha\text{-power}_{contra})/(\alpha\text{-power}_{ipsi} + \alpha\text{-power}_{contra}) \quad (1)$$

To account for overall hemispheric power differences that were independent of attention modulation, we first normalized single-trial power by calculating per parcel and frequency the whole-trial (–0.5–6.5 s) power averaged across all trials and subtracted it from single trials. We then used a robust variant of the index that applies the inverse logit transform [(1 / (1+ exp(–x))] to both inputs to scale them into a common, positive-only [0;1]-bound space prior to index calculation.

For visualization and statistical analysis of cue-driven neural modulation, we then averaged the ALI across all parcels within the auditory ROI and extracted single-trial mean values for the time window of sentence presentation (3.5–6.5 s), and treated them as the dependent measure in linear mixed-effects analysis (see Statistical Analysis below). They also served as continuous predictors in the statistical analysis of brain-behavior and brain-brain relationships. We performed additional analyses that focused on the ALI in auditory cortex during presentation of the sentence-final word and spatial-attention cue, respectively. Further control analyses included single-trial ALI during sentence and final-word presentation that were extracted from the inferior parietal ROI.

Estimation of envelope reconstruction models

To investigate how low-frequency (i.e., < 8 Hz) fluctuations in EEG activity relate to the encoding of attended and ignored speech, we trained stimulus reconstruction models (also termed decoding or backward models) to predict the onset envelope (see Supplements for details) of the attended and ignored speech stream from EEG (84, 85). In this analysis framework, a linear reconstruction model *g* is assumed to represent the linear mapping from the recorded EEG signal, *r(t,n),* to the stimulus features, *s(t)*:

$$\hat{s}(t) = \sum_n \sum_\tau g(\tau, n) r(t + \tau, n) \quad (2)$$

where $\hat{s}(t)$ is the reconstructed onset envelope at time point *t*. We used all parcels within the bilateral auditory ROI and time lags τ in the range of –100 ms to 500 ms to compute envelope reconstruction models using ridge regression (86):

$$g = (R^T R + \lambda m I)^{-1} R^T s \quad (3)$$

where R is a matrix containing the sample-wise time-lagged replication of the neural response matrix r, λ is the ridge parameter for regularization, m is a scalar representing the mean of the trace of RTR (87), and I is the identity matrix. We followed the procedures described in ref. (88) to estimate the optimal ridge parameter.

Compared to linear forward ('encoding') models that derive temporal response functions (TRFs) independently for each EEG channel or source, stimulus reconstruction models represent multivariate impulse response functions that exploit information from all time lags and EEG channels/sources simultaneously. To allow for a neurophysiological interpretation of backward model coefficients, we additionally transformed them into linear forward model coefficients following the inversion procedure described in ref. (89). All analyses were performed using the multivariate temporal response function (mTRF) toolbox (84, version 1.5) for Matlab.

Prior to model estimation, we split the data based on the two spatial attention conditions (selective vs. divided), resulting in 120 trials per condition. Envelope reconstruction models were trained on concatenated data from selective-attention trials, only. Prior to concatenation, single trials were zero-padded for 600ms to reduce discontinuity artefacts, and one trial was left out for subsequent testing. On each iteration, two different backward models were estimated, an envelope reconstruction model for the-be-attended speech stream (short: attended reconstruction model), and one for the to-be-ignored speech stream (short: ignored reconstruction model. Reconstruction models for attended and ignored speech signals were trained separately for attend-left and attend-right trials which yielded 120 decoders (60 attended, 60 ignored) per attentional setting. For illustrative purposes, we averaged the forward transformed models of attended and ignored speech per hemisphere across all participants (Fig. 4B).

Evaluation of neural tracking strength

We analyzed how strongly the attended compared to ignored sentences were tracked by slow cortical dynamics by quantifying the envelope reconstruction accuracy for individual trials. To this end, we reconstructed the attended and ignored envelope of a given trial using a leave-one-out cross-validation procedure. The two envelopes of a given trial were reconstructed using the models trained on all but the current trial from the same attention condition. The reconstructed onset envelope obtained from each model was then compared to onset envelopes of the actually presented speech signals using a 250-ms sliding window. The resulting time courses of Pearson-correlation coefficients, $r_{attended}$ and $r_{ignored}$, reflect a temporally resolved measure of single-trial neural tracking strength or reconstruction accuracy (23) (see Fig. 4 and supplement 1).

We proceeded in a similar fashion for divided-attention trials. Since these trials could not be categorized based on the to-be-attended and -ignored side, we split them based on the ear that was probed at the end the trial. Given that even in the absence of a valid attention cue, participants might still (randomly) focus their attention to one of the two streams, we wanted to quantify how strongly the probed and unprobed envelopes were tracked neurally. We used the reconstruction models trained on selective-attention trials to reconstruct the onset envelopes of divided-attention trials. Sentences presented in probed-left/unprobed-right trials were reconstructed using the attend-left/ignore-right reconstruction models while probed-right/unprobed-left trials used the reconstruction models trained on attend-right/ignore-left trials.

Attentional modulation of neural tracking

In close correspondence to the alpha lateralization index, we calculated a neural tracking index throughout sentence presentation. The index expresses the difference in neural tracking of the to-be-attended and ignored sentence (in divided attention: probed and unprobed, respectively) (cf. 22):

$$\text{Neural tracking index} = (r_{attended} - r_{ignored})/(r_{attended} + r_{ignored}) \quad (4)$$

Positive values of the resulting index indicate that the attended envelope was tracked more strongly than the ignored envelope, and vice versa for negative values. Since individual sentences differed in length, for visualization and statistical analysis, we mapped their resulting neural tracking time courses onto a common time axis expressed in relative (percent) increments between the start and end of a given stimulus. We converted the original time course using a sliding window covering 3% (1% overlap) of the original sentence length. The same procedure was applied to the time course of alpha power lateralization following up-sampling to 125 Hz. Single-trial measures for the interval of final word presentation were averaged across the final 35% of sentence presentation as this interval covered final word onset across all 240 sentence pairs. We used the single-trial neural tracking index as (in-)dependent variables in our linear mixed-effects analyses (see below).

Statistical analysis

We used (generalized) linear mixed-effect models to answer the research questions outlined in Figure 1. This approach allowed us to jointly model the impact of listening cues, neural filter strategies and various additional covariates known to influence behavior. These included the probed ear (left/right), whether the later-probed sentence had the earlier onset (yes/no), as well as participants' age and hearing acuity (pure-tone average across both ears).

To arbitrate between state-level (i.e., within-subject) and trait-level (i.e., between-subject) effects, our models included two separate regressors for each of the key neural measures. Between-subject effect regressors consisted of neural measures that were averaged across all trials at the single-subject level, whereas the within-subject effect was modelled by the trial-by-trial deviation from the subject-level mean (cf. 42).

Deviation coding was used for categorical predictors. All continuous variables were z-scored. For the dependent measure of accuracy, we used generalized linear mixed-effects model (binomial distribution, logit link function). For response speed, we used general linear mixed-effects model (gaussian distribution, identity link function). Given the large sample size, *p*-values for individual model terms are based on Wald *t-as-z*-values for linear models (90) and on z-values and asymptotic Wald tests in generalized linear models. All reported p-values are corrected to control for the false discovery rate (91).

In lieu of a standardized measure of effect size for mixed-effects models, we report odds ratios (OR) for generalized linear models and standardized regression coefficients (β) for linear models along with their respective standard errors (SE).

All analyses were performed in R (92 version 3.6.1) using the packages lme4 (93), and sjPlot (94).

Model selection

To avoid known problems associated with a largely data-driven stepwise model selection that include the overestimation of coefficients (95) or the selection of irrelevant predictors (96), the inclusion of fixed effects was largely constrained by our *a priori* defined hypotheses. The influence of visual cues and of neural measures were tested in same brain-behavior model. The brain-behavior model of accuracy and response speed included random intercepts by subject and item. In a data-driven manner, we then tested whether model fit could be further improved by the inclusion of subject-specific random slopes for the effects of the spatial-attention cue, semantic cue, or probed ear. The change in model fit was assessed using likelihood ratio tests on nested models.

Data availability

The complete dataset associated with this work including raw data, EEG data analysis results, as well as corresponding code will be publicly available under https://osf.io/28r57/.

References

1. Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25(5):975–979.

2. Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186.

3. Foxe JJ, Snyder AC (2011) The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Front Psychology* 2:154.

4. Händel BF, Haarmeier T, Jensen O (2011) Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J Cogn Neurosci* 23(9):2494–2502.

5. Rihs TA, Michel CM, Thut G (2007) Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. *Eur J Neurosci* 25(2):603–610.

6. Kerlin JR, Shahin AJ, Miller LM (2010) Attentional gain control of ongoing cortical speech representations in a "cocktail party." *J Neurosci* 30(2):620–628.

7. Ahveninen J, Huang S, Belliveau JW, Chang W-T, Hämäläinen M (2013) Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. *J Cogn Neurosci* 25(11):1926–1943.

8. Müller N, Weisz N (2011) Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. *Cereb Cortex* 22(7):1604–1613.

9. Wöstmann M, Herrmann B, Maess B, Obleser J (2016) Spatiotemporal dynamics of auditory attention synchronize with speech. *Proceedings of the National Academy of Sciences* 113(14):3873–3878.

10. Wöstmann M, Vosskuhl J, Obleser J, Herrmann CS (2018) Opposite effects of lateralised transcranial alpha versus gamma stimulation on auditory spatial attention. *Brain Stimulation* 11(4):752–758.

11. Tune S, Wöstmann M, Obleser J (2018) Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *Eur J Neurosci* 48(7):2537–2550.

12. Haegens S, Handel BF, Jensen O (2011) Top-Down Controlled Alpha Band Activity in Somatosensory Areas Determines Behavioral Performance in a Discrimination Task. *J Neurosci* 31(14):5197–5204.

13. Bauer M, Kennett S, Driver J (2012) Attentional selection of location and modality in vision and touch modulates low-frequency activity in associated sensory cortices. *J Neurophysiol* 107(9):2342–2351.

14. Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci* (20):1–6.

931 15. Kelly SP, Lalor EC, Reilly RB, Foxe JJ (2006) Increases in alpha oscillatory power
932 reflect an active retinotopic mechanism for distracter suppression during
933 sustained visuospatial attention. *J Neurophysiol* 95(6):3844–3851.

934 16. Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as
935 instruments of sensory selection. *Trends Neurosci* 32(1):9–18.

936 17. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of
937 Active Sensing and perceptual selection. *Current Opinion in Neurobiology*
938 20(2):172–176.

939 18. Henry MJ, Obleser J (2012) Frequency modulation entrains slow neural
940 oscillations and optimizes human listening behavior. *Proc Natl Acad Sci USA*
941 109(49):20095–20100.

942 19. Obleser J, Kayser C (2019) Neural Entrainment and Attentional Selection in the
943 Listening Brain. *Trends Cogn Sci* 23(11):913–926.

944 20. Zion Golumbic EMZ, et al. (2013) Mechanisms Underlying Selective Neuronal
945 Tracking of Attended Speech at a ''Cocktail Party''. *Neuron* 77(5):980–991.

946 21. Ding N, Simon JZ (2012) Emergence of neural encoding of auditory objects
947 while listening to competing speakers. *Proc Natl Acad Sci USA* 109(29):11854–
948 11859.

949 22. Mesgarani N, Chang EF (2012) Selective cortical representation of attended
950 speaker in multi-talker speech perception. *Nature* 485(7397):233–236.

951 23. O'Sullivan JA, et al. (2014) Attentional Selection in a Cocktail Party Environment
952 Can Be Decoded from Single-Trial EEG. *Cereb Cortex* 25(7):1697–1706.

953 24. Horton C, D'Zmura M, Srinivasan R (2013) Suppression of competing speech
954 through entrainment of cortical oscillations. *J Neurophysiol* 109(12):3082–3093.

955 25. Keitel C, et al. (2019) Stimulus-Driven Brain Rhythms within the Alpha Band: The
956 Attentional-Modulation Conundrum. *J Neurosci* 39(16):3119–3129.

957 26. Gundlach C, Moratti S, Forschack N, Müller MM (2020) Spatial Attentional
958 Selection Modulates Early Visual Stimulus Processing Independently of Visual
959 Alpha Modulations. *Cereb Cortex* 16:637–18.

960 27. Henry MJ, Herrmann B, Kunke D, Obleser J (2017) Aging affects the balance of
961 neural entrainment and top-down neural modulation in the listening brain. *Nat*
962 *Commun* 8:15801.

963 28. Lakatos P, et al. (2016) Global dynamics of selective attention and its lapses in
964 primary auditory cortex. *Nat Neurosci* 19(12):1707–1717.

965 29. Posner MI (1980) Orienting of attention. *Quarterly Journal of Experimental*
966 *Psychology* 32(1):3–25.

967 30. Alavash M, Tune S, Obleser J (2019) Modular reconfiguration of an auditory
968 control brain network supports adaptive listening behavior. *Proc Natl Acad Sci*
969 *USA* 116(2):660–669.

31. Presacco A, Simon JZ, Anderson S (2016) Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol* 116(5):2356–2367.

32. Sohoglu E, Peelle JE, Carlyon RP, Davis MH (2012) Predictive Top-Down Integration of Prior Knowledge during Speech Perception. *J Neurosci* 32(25):8443–8453.

33. Peelle JE, Gross J, Davis MH (2013) Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cereb Cortex* 23(6):1378–1387.

34. Obleser J, Weisz N (2012) Suppressed Alpha Oscillations Predict Intelligibility of Speech and its Acoustic Details. *Cereb Cortex* 22(11):2466–2477.

35. Wöstmann M, Lim S-J, Obleser J (2017) The human neural alpha response to speech is a proxy of attentional control. *Cereb Cortex* 27(6):3307–3317.

36. Broadbent DE, Gregory M (1964) Accuracy of recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology* 16(4):359–360.

37. Kimura D (1961) Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology/Revue canadienne de psychologie* 15(3):166–171.

38. Decruy L, Vanthornhout J, Francart T (2020) Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hearing Research* 393:107961.

39. Presacco A, Simon JZ, Anderson S (2019) Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS ONE* 14(3):e0213899–26.

40. Petersen EB, Wöstmann M, Obleser J, Lunner T (2017) Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J Neurophysiol* 117(1):18–27.

41. Fiedler L, Wöstmann M, Herbst SK, Obleser J (2019) Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *NeuroImage* 186:33–42.

42. Bell A, Fairbrother M, Jones K (2018) Fixed and random effects models: making an informed choice. *Qual Quant* 53(2):1051–1074.

43. Passow S, et al. (2012) Human aging compromises attentional control of auditory perception. *Psychol Aging* 27(1):99–105.

44. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013) A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing Research* 300:18–32.

45. Banerjee S, Snyder AC, Molholm S, Foxe JJ (2011) Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms? *J Neurosci* 31(27):9923–9932.

1010 46. Waschke L, Tune S, Obleser J (2019) Local cortical desynchronization and pupil-
1011 linked arousal differentially shape brain states for optimal sensory performance.
1012 *eLife Sciences* 8:1868–27.

1013 47. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017)
1014 Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93(3):480–
1015 490.

1016 48. van Ede F, Köster M, Maris E (2012) Beyond establishing involvement:
1017 quantifying the contribution of anticipatory α- and β-band suppression to
1018 perceptual improvement with attention. *J Neurophysiol* 108(9):2352–2362.

1019 49. Ding N, Simon JZ (2014) Cortical entrainment to continuous speech: functional
1020 roles and interpretations. *Front Hum Neurosci* 8(364.):13367.

1021 50. Hamilton LS, Huth AG (2018) The revolution will not be controlled: natural
1022 stimuli in speech neuroscience. *Language, Cognition and Neuroscience* 27(9):1–
1023 10.

1024 51. Sassenhagen J (2018) How to analyse electrophysiological responses to
1025 naturalistic language with time-resolved multiple regression. *Language,*
1026 *Cognition and Neuroscience* 0(0):1–17.

1027 52. Ding N, Simon JZ (2012) Neural coding of continuous speech in auditory cortex
1028 during monaural and dichotic listening. *J Neurophysiol* 107(1):78–89.

1029 53. Broderick MP, Anderson AJ, Di Liberto GM, Crosse MJ, Lalor EC (2018)
1030 Electrophysiological Correlates of Semantic Dissimilarity Reflect the
1031 Comprehension of Natural, Narrative Speech. *Curr Biol* 28(5):1–11.

1032 54. Brodbeck C, Hong LE, Simon JZ (2018) Rapid Transformation from Auditory to
1033 Linguistic Representations of Continuous Speech. *Curr Biol* 28(24):3976–
1034 3982.e6.

1035 55. Fiedler L, Woestmann M, Herbst SK, Obleser J (2018) Late cortical tracking of
1036 ignored speech facilitates neural selectivity in acoustically challenging
1037 conditions. *NeuroImage*:1–37.

1038 56. Thut G (2006) **α**-band electroencephalographic activity over occipital cortex
1039 indexes visuospatial attention bias and predicts visual target detection. *J*
1040 *Neurosci* 26(37):9494–9502.

1041 57. Bengson JJ, Mangun GR, Mazaheri A (2012) The neural markers of an imminent
1042 failure of response inhibition. *NeuroImage* 59(2):1534–1539.

1043 58. Dahl MJ, Ilg L, Li S-C, Passow S, Werkle-Bergner M (2019) Diminished pre-
1044 stimulus alpha-lateralization suggests compromised self-initiated attentional
1045 control of auditory processing in old age. *NeuroImage* 197:414–424.

1046 59. Hong X, Sun J, Bengson JJ, Mangun GR, Tong S (2015) Normal aging selectively
1047 diminishes alpha lateralization in visual spatial attention. *NeuroImage*
1048 106(C):353–363.

60. Leenders MP, Lozano-Soldevilla D, Roberts MJ, Jensen O, De Weerd P (2018) Diminished Alpha Lateralization During Working Memory but Not During Attentional Cueing in Older Adults. *Cereb Cortex* 28(1):21–32.

61. Mok RM, Myers NE, Wallis G, Nobre AC (2016) Behavioral and neural markers of flexible attention over working memory in aging. *Cereb Cortex* 26(4):1831–1842.

62. Wöstmann M, Alavash M, Obleser J (2019) Alpha Oscillations in the Human Brain Implement Distractor Suppression Independent of Target Selection. *J Neurosci* 39(49):9797–9805.

63. Pernet CR, Sajda P, Rousselet GA (2011) Single-trial analyses: why bother? *Front Psychology* 2:322.

64. Riecke L, Formisano E, Sorger B, Başkent D, Gaudrain E (2018) Neural Entrainment to Speech Modulates Speech Intelligibility. *Curr Biol* 28:1–9.

65. Zoefel B, Archer-Boyd A, Davis MH (2018) Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr Biol* 28(3):401–408.e5.

66. Wöstmann M, Obleser J (2016) Acoustic detail but not predictability of task-irrelevant speech disrupts working memory. *Front Hum Neurosci* 10(781):201–9.

67. Zhigalov A, Herring JD, Herpers J, Bergmann TO, Jensen O (2019) Probing cortical excitability using rapid frequency tagging. *NeuroImage* 195:59–66.

68. Teoh ES, Lalor EC (2019) EEG decoding of the target speaker in a cocktail party scenario: considerations regarding dynamic switching of talker location. *J Neural Eng* 16(3):036017–30.

69. Hauswald A, Keitel A, Chen Y-P, Roesch S, Weisz N (2020) Degradation levels of continuous speech affect neural speech tracking and alpha power differently. 1–33.

70. Zoefel B, VanRullen R (2017) Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. *Front Neurosci* 11:225–13.

71. Antonov PA, Chakravarthi R, Andersen SK (2020) Too little, too late, and in the wrong place: Alpha band activity does not reflect an active mechanism of selective attention. *NeuroImage* 219:117006.

72. Pichora-Fuller MK (2017) Using the Brain when the Ears are Challenged Helps Healthy Older Listeners Compensate and Preserve Communication Function. 1–14.

73. Rogers CS, Payne L, Maharjan S, Wingfield A, Sekuler R (2018) Older adults show impaired modulation of attentional alpha oscillations: Evidence from dichotic listening. *Psychol Aging* 33(2):246–258.

74. Getzmann S, Klatt L-I, Schneider D, Begau A, Wascher E (2020) EEG correlates of spatial shifts of attention in a dynamic multi-talker speech perception scenario in younger and older adults. *Hearing Research* 398:108077–46.

75. Fuglsang SA, Märcher-Rørsted J, Dau T, Hjortkjær J (2020) Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. *J Neurosci* 40(12):2562–2572.

76. Decruy L, Vanthornhout J, Francart T (2019) Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *J Neurophysiol* 122(2):601–615.

77. Presacco A, Simon JZ, Anderson S (2016) Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol* 116(5):2346–2355.

78. 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(6):1586–1597.

79. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. 9(1):97–113.

80. Jefferies K, Gale TM (2013) 6-CIT: Six-Item Cognitive Impairment Test. *Cognitive Screening Instruments* (Springer, London, London), pp 209–218.

81. Glasser MF, et al. (2016) A multi-modal parcellation of human cerebral cortex. *Nature* 536(7615):171–178.

82. Keitel A, Gross J (2016) Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLoS Biol* 14(6):e1002498–22.

83. Smulders FTY, Oever ten S, Donkers FCL, Quaedflieg CWEM, van de Ven V (2018) Single-trial log transformation is optimal in frequency analysis of resting EEG alpha. *Eur J Neurosci* 44:94–14.

84. Crosse MJ, Di Liberto GM, Bednar A, Lalor EC (2016) The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front Hum Neurosci* 10:604.

85. Lalor EC, Foxe JJ (2010) Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur J Neurosci* 31(1):189–193.

86. Hoerl AE, Kennard RW (1970) Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12(1):55–67.

87. 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 Trans Neural Syst Rehabil Eng* 25(5):402–412.

88. Fiedler L, et al. (2017) Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. *J Neural Eng* 14(3):036020.

89. Haufe S, et al. (2014) On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87(C):96–110.

90. Luke SG (2017) Evaluating significance in linear mixed-effects models in R. *Behav Res* 49(4):1494–1502.

91. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57(1):289–300.

92. R Core Team (2019) R: A language and environment for statistical computing. doi:https://www.R-project.org/.

93. Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *J Stat Soft* 67(1):1–48.

94. Lüdecke D, Lüdecke D (2018) Data Visualization for Statistics in Social Science [R package sjPlot version 2.6.1]. doi:10.5281/zenodo.1308157.

95. Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* 56(1):267–288.

96. Derksen S, Keselman HJ (1992) Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology* 45(2):265–282.