Neural attentional-filter mechanisms of listening success in middle-aged and older 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 behavioural outcome in attentive listening

Number of words in Abstract: 159
Number of words in Introduction: 868
Number of words in Discussion: 2048

34 pages, 6 Figures, 0 Tables, Includes Supplementary Information

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

Abstract

Successful listening crucially depends on intact attentional filters that separate relevant from irrelevant information. Research into their neurobiological implementation has focused on two potential auditory filter strategies: the lateralization of alpha power and selective neural speech tracking. However, the functional interplay of the two neural filter strategies and their potency to index listening success in an ageing population remains unclear. Using electroencephalography and a dual-talker task in a representative sample of listeners (N=155; age=39–80 years), we here demonstrate an often-missed link from single-trial behavioural outcomes back to trial-by-trial changes in neural attentional filtering. First, we observe preserved attentional–cue-driven modulation of both neural filters across chronological age and hearing levels. Second, neural filter states vary independently of one another, demonstrating complementary neurobiological solutions of spatial selective attention. Stronger neural speech tracking but not alpha lateralization boosts trial-to-trial behavioural performance. Our results highlight the translational potential of neural speech tracking as an individualized neural marker of adaptive listening behaviour.

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 differentiation of relevant and irrelevant inputs. Here, the concept of neural attentional 'filters' serves as an important and pervasive algorithmic metaphor of how auditory attention is implemented at the neural level[2-4]. Neural attentional filters can be instantiated by different mechanistic principles and recent studies have predominantly focused on two potential but nonexclusive neural filter strategies originating from distinct research traditions:

From the visual domain stems an influential line of research that supports a role of alpha-band (~8–12 Hz) oscillatory activity in the implementation of controlled, top-down suppression of behaviourally-irrelevant information[5-8]. 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[9-18]. 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 behaviourally-relevant sensory information[19-22] (see also refs.[23,24] for the neural tracking of contextual semantic information). 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[25-29].

However, with few exceptions[9], these two proposed neural auditory filter strategies have been studied independently of one another (but see refs. [30,31] 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[32,33].

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 behaviour. 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 behavioural listening success[27,28].

As part of an ongoing large-scale project on the neural and cognitive mechanisms supporting adaptive listening behaviour in healthy ageing, the current study aims at closing these gaps by leveraging the statistical power and representativeness of our large, age-varying participant sample. We use a dichotic listening paradigm to enable a synergistic look at concurrent single-trial changes in lateralized alpha power and neural speech tracking.

More specifically, our linguistic variant of a classic Posner paradigm[34] emulates a challenging dual-talker listening situation in which speech comprehension is supported by two different listening cues[35,36]. Of particular interest for the present

scientific endeavour is the spatial-attention cue that guides auditory attention in space. We additionally manipulated the semantic predictability of upcoming speech via a semantic category cue. While the effects of the semantic cue are of secondary importance for the present research questions, its manipulation still allows insights into whether semantic predictability modulates the engagement of neural attentional filter mechanisms, and how it affects listening success in a large cohort of middle-aged and older adults. 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[37-42].

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 behaviour 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[43,44].

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 lateralization and neural speech tracking. We aimed to replicate previous findings of increased alpha 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 our age-varying sample to quantify the hitherto contested dependence of these neurobiological filters on participants' chronological age and hearing loss[14,45-47].

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 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 lateralization and neural speech tracking allow us to explain behavioural outcome—that is, listening success—at the individual level and in a single trial?

Here, we show how an attentional cue modulates neural speech tracking and alpha lateralization independently of age and hearing levels. We demonstrate the co-existence of largely independent neural filters that pose complementary neurobiological implementations of selective attention. Stronger neural speech tracking but not alpha lateralization increases trial-to-trial listening performance. This

emphasizes the potential of neural speech tracking as a diagnostic neural measure of an individual's listening success.

Results

We recorded and source-localized electroencephalography (EEG) signals in an age-varying sample of healthy middle-aged and older adults (N=155; age=39–80 years, see Supplementary Fig. 1) who performed a challenging dichotic listening task. In this linguistic variant of a classic Posner paradigm[35], 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 to 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 approximates a dual-talker listening situation to probe the neural underpinnings of successful selective listening[35].

Using generalized linear mixed-effects models on single-trial data, we focus on two key neurobiological instantiations 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 behaviour and neural filters, whether neural filters operate independently, and to which extent they influence selective listening success.

Informative spatial cues improve listening success

For behavioural performance, we tested the impact of informative versus uninformative 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, behaviour depended on the different combinations of listening cues (Fig. 2c,d). Informative compared to uninformative spatial-attention cues yielded a strong behavioural 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, $p<.001$; see Supplementary Tables 1 and 2). 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. Contrary to our expectations, a more informative semantic cue did not lead to more accurate responses (GLMM; *OR*=1.1, *SE*=.11, p=.69).

As in a previous fMRI implementation of this task[35], 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 Supplementary Fig. 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 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 Supplementary Fig. 3 and Supplementary 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 speech tracking focused on attentional modulation index measures that contrast the relative strength of neural responses to target versus distractor stimuli. In line with previous results, we expected alpha lateralization to be present throughout the auditory sentence presentation but to potentially increase around the task-relevant final word[12,14,48].

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) = $(\alpha\text{-power}_{ipsi} - \alpha\text{-power}_{contra}) / (\alpha\text{-power}_{ipsi} + \alpha\text{-power}_{contra})$][15].

As shown in Figure 3a, an informative 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 lateralization during sentence presentation (time window: 3.5–6.5 s; see control analysis section below for results on the final word period) 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 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 Supplementary Table 3). 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 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 Supplementary Table 6 for a corresponding analysis of alpha power during the interval of the final word).

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, Fig. 4a and Supplementary Fig. 4 for details).

In line with previous studies[26,28,49], 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 (see control analysis section below for results on the entire sentence presentation). However, this effect was also modulated by differences in sentence onset: The difference in neural speech 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 speech 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 Supplementary Table 4 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 speech tracking (p-values > .54 for main effects of age, PTA, and their respective interactions with the spatial-attention cue, see also Supplementary Table 7 for a corresponding analysis of neural tracking during the entire sentence presentation).

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

Our third major goal was to investigate whether neural speech tracking versus the modulation of alpha power reflect two dependent neural mechanisms of auditory attention at all. We asked whether neural speech tracking could be explained by auditory alpha 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 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 lateralization on neural speech 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 Supplementary Table 5). This notable absence of an alpha lateralization–neural speech tracking relationship held irrespective of spatial-attention condition or probed ear (all *p*-values>.35).

To complement our fine-grained single-trial level investigation into the brain-brain 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 speech tracking and alpha lateralization follow different temporal trajectories with neural speech 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 speech tracking leading those in alpha power by about 520 ms (see Fig. 5d).

Neural speech tracking but not alpha lateralization explains listening behaviour

Having established the functional independence of alpha lateralization and neural speech tracking at the single-trial level, the final piece of our investigation was to probe their relative functional importance for behavioural 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 behaviour at the state and trait level[50].

For response accuracy, our most important indicator of listening success, we observed an effect of trial-by-trial variation in neural speech 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 speech 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 Supplementary Table 12) 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 Supplementary Table 1).

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 Supplementary Tables 1 and 12). We also did not find evidence for any joint effects of alpha power and neural speech 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 analysed time window: while participants with relatively higher average levels of neural speech 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 Supplementary Table 13), we found a combined effect of neural dynamics during final word presentation. Under selective but not divided attention, response speed depended on a combination of trial-to-trial variation in both alpha lateralization and neural speech tracking (LMM; spatial cue x ALI (within-subject effect) x neural tracking index (within-subject effect): β=0.08, *SE*=.03, *p*=.01; see Supplementary Table 2). In short, responses were fastest in trials where relatively elevated levels in either neural speech tracking or alpha lateralization were paired with relatively reduced levels in the respective other neural measure thus highlighting the influence of two independent complementary filter solutions (see also Supplementary Fig. 5).

In line with the literature on listening behaviour in ageing adults[51,52], the behavioural 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[44]. 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 Supplementary Fig. 6).

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 Supplementary Tables 1–2, and Supplementary Fig. 7).

Control analyses

We ran additional control analyses to validate our main set of results. First, we asked whether the observed independence of alpha lateralization and neural speech 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 speech 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.

Second, 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[53]. 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 Supplementary Tables 8–11 and Supplementary Fig. 8).

Third, we asked whether our neural speech tracking results were impacted by the range of time lags used for reconstruction, or by the specific decoder model underlying the neural tracking index. Reconstructing envelopes using a shorter time window (50–250 ms) did not significantly change the resulting neural tracking index values (LMM, $\beta = .002$, $SE = .007$, $p = .84$; see also Supplementary Fig. 9). In a separate analysis, we calculated the neural tracking index using only the attended decoder model, and probed its influence on behaviour. The results are overall in line with our main conclusions and particularly underscore the impact of neural speech tracking on response accuracy (see Supplementary Tables 14–17 for details).

Finally, we tested whether changes in age or hearing loss would modulate the relationship of neural tracking and alpha lateralization with listening behaviour. 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

We have utilized the power of a 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 addition, we leveraged our age-varying sample to ask how chronological age and hearing loss affect the fidelity of neural filter strategies and their influence on behaviour.

In our dichotic listening task, we source-localized the electroencephalogram, and primarily focused on systematic spatial-cue-driven changes within auditory cortex in alpha lateralization and the neural tracking of attended versus ignored speech. These results provide large-sample support for their suggested roles as neural instantiations of selective attention.

First, an informative spatial-attention cue not only boosted both neural measures but also consistently boosted behavioural performance. Listening behaviour was additionally influenced by both trial-to-trial and individual-to-individual variation in neural speech tracking, with relatively stronger tracking of the target sentence leading to better performance. An informative semantic cue led to faster responses but did not affect the two neural measures, thus most likely reflecting a priming effect speeding up the analysis of response alternatives rather than the differential processing of the sentences themselves.

Second, when related at the single-trial, single-subject level, the two neural attentional filter mechanisms were found to operate statistically independently of each other. This underlines 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 lateralization.

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

Neural speech tracking but not alpha lateralization predicts listening success

The present study explicitly addressed the often overlooked question of how neural filter states (i.e., fluctuations from trial-to-trial) impact behaviour, here single-trial listening success[54-57]. Using a sophisticated linear-model approach that probed the impact of both state- and trait-level modulation of neural filters on behaviour, we only found evidence for a direct influence of neural speech tracking but not alpha lateralization on behavioural performance even though all three measures were robustly modulated by the presence of a spatial cue (see Fig. 6). What could be the reason for this differential impact of neural measures on behaviour?

To date, the behavioural relevance of selective neural speech tracking is still poorly supported given the emphasis on more naturalistic, yet more complex language stimuli[58,59]. 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 behaviour. This makes it particularly challenging to establish a direct link between differential neural speech tracking and listening success[23,25,26,49,60]. Nevertheless, there is preliminary evidence linking stronger neural tracking to improved comprehension when it is tested at a comparably high level[28] (i.e., content questions on longer speech segments). Our current results thus provide important additional fine-grained and temporally resolved support to the functional relevance of selective neural speech tracking for moment-to-moment listening behaviour[61-63].

Despite a vast number of studies investigating the role of (lateralized) alpha oscillations in attentional tasks, the circumstances under which their top-down modulation may affect the behavioural outcome are still insufficiently understood[31]. Rather, the presence of a stable brain–behaviour relationship hinges on several factors.

First, the link of neural filter state to behaviour seems to be impacted by age: most evidence linking increased alpha lateralization to better task performance in spatial-attention tasks stems from smaller samples of young adults[12,15,64,65]. By contrast, the presence 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 behavioural) level[14,66-69] (see discussion below).

Second, previous findings differ along (at least) two dimensions: (i) whether the functional role of alpha lateralization is studied during attention cueing, stimulus anticipation, or stimulus presentation[9,66,70], and (ii) whether behaviour is related to the overall strength of alpha lateralization or its stimulus-driven rhythmic modulation[12,14]. Depending on these factors, the observed brain–behaviour 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.[70], 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 behavioural relevance of neural responses related to target selection and distractor suppression, respectively.

In summary, it is still a matter of debate by which mechanistic pathway, and at which processing stage the modulation of alpha power will impact behaviour. While it is (often implicitly) assumed that alpha oscillations impact behaviour via a modulation of neural excitability and thus early sensory processing, there is little evidence that shows a direct influence of alpha oscillation on changes in neural excitability and on subsequent behaviour[31,71].

Lastly, the increase in alpha lateralization around final word presentation could at least partially reflect post-perceptual processes associated with response selection rather than the perceptual analysis itself[72]. The observed combined influence of neural tracking and alpha lateralization on response speed but not accuracy would seem compatible with such an interpretation (but see also ref.[73] for the combined influence of non-lateralized alpha power and neural speech tracking on intelligibility in a non-spatial listening task).

Taken together, our results underscore the impact of prioritized sensory encoding of relevant sounds via selective neural speech tracking on listening performance and highlight the difficulty in establishing a comparable link for a neural signature as multifaceted as alpha oscillations[74-76].

Are fluctuations in lateralized alpha power and neural speech tracking functionally connected?

We investigated attention-related changes in two neural filter strategies that (i) involve neurophysiological signals operating at different frequency regimes, (ii) are assumed to support auditory attention by different neural mechanisms, and (iii) are typically studied in isolation[6,22]. Here, we found both neural filter strategies to be impacted by the same spatial-attention cue which afforded insights into their neurobiological dependence.

There is preliminary evidence, mostly from between-subjects analyses, suggesting that the two neural filter strategies may exhibit a systematic relationship[9,14,32,33,77]. How the two neural filter strategies may be connected mechanistically is thus still an open question. We here asked whether concurrent changes in neural filter states would imply a neural hierarchy in which alpha-driven controlled inhibition modulates the amplification of behaviourally relevant sensory information via selective neural speech tracking[78-80].

Our in-depth trial-by-trial analysis revealed independent modulation of alpha power and neural speech tracking. At the same time, in our exploratory between-subjects cross-correlation analysis we observed a systematic temporal delay with peaks in neural speech tracking leading those in alpha lateralization. While the direction and duration of this delay were closely in line with previous findings[12,14], at this coarser level of analysis, they speak against a hierarchy of neural processing in which lateralized alpha responses govern the differential neural tracking of attended versus ignored speech[81].

Our single-trial results are well in line with recent reports of independent variation in alpha band activity and steady-state (SSR) or frequency following responses (FFR) in studies of visual spatial attention[30,31,82]. Additionally, the inclusion of single-trial alpha lateralization as an additional training feature in a recent speech-tracking study failed to improve the decoding of attention[83]. The results from our most fine-grained single-trial level of analysis thus speak against a consistent, linear relationship of momentary neural filter states. We observed instead the co-existence of two complementary but seemingly independent neurobiological solutions to the implementation of auditory selective attention.

How can this finding be reconciled with findings from previous electrophysiological studies[9,32,33] pointing towards a functional trade-off between neurobiological attentional filter mechanisms? And what could be an advantage to independent neural solutions for selective auditory attention?

Our between-subjects cross-correlation analysis appears to provide at least tentative support for a systematic relationship in which peaks in neural speech tracking

precede those in alpha lateralization. A closer inspection of the group-level temporal modulation of neural measures throughout sentence presentation, however, reveals some important differences to previous results. Whereas earlier studies reported an acyclic waxing and waning of neural entrainment and alpha power in response to rhythmic auditory stimulation[12,14,33], in the current study, the two neural measures show different temporal dynamics: neural speech tracking gradually increases leading up to the final word, while alpha lateralization peaks at sentence and final word onset. The temporal dynamics of alpha lateralization, in particular, may point to strategic intermittent engagement of spatial attention in line with task demands[48].

Do these differences in temporal dynamics of the two neural filters challenge the existence of a systematic single-trial brain–brain relationship? Yes, but they also point to a potential benefit of independent neural filter solutions. If the two neural measures of auditory attention were indeed functionally unconnected as suggested by the current results, they would allow for a wider range of neural filter state configurations to flexibly adapt to the current task demands and behavioural goals. The co-existence of two independent but complementary filter mechanisms operating either via the selective amplification of relevant or via the controlled inhibition of irrelevant sounds, enables different modes of auditory attention to serve a listener's goal in the face of complex real-life listening situations[20,84,85].

Do age and hearing loss affect neural filter strategies?

The detrimental effects of increasing age and associated hearing loss on speech comprehension in noisy listening situations are well attested[86] and borne out by the current results. However, the extent to which the neural implementations of attentional filtering are affected by age and hearing loss, and in how far they may constitute neural markers of age-related speech comprehension problems, remains poorly understood[51].

As in a previous study on a subset of the current sample, we found the fidelity of alpha lateralization unchanged with age[14]. Other studies on auditory attention, however, have observed diminished and less sustained alpha lateralization for older compared to younger adults that were to some extend predictive of behaviour[66,87,88].

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 mirrored the attentional modulation observed for younger or older normal-hearing adults[45-47,89]. As revealed by follow-up analysis (see Supplementary Tables 18–19), 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[90,91].

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[92], 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.

In a representative, age-varying sample of listeners, we underscore the functional significance of lateralized alpha power and neural speech tracking to spatial attention. Our results point to the co-existence of two independent yet complementary neural filter mechanisms to be flexibly engaged depending on a listener's attentional goals. However, we see no direct, behaviourally relevant impact of alpha power modulation on early sensory gain processes.

Only for neural speech tracking, we establish a mechanistic link from trial-to-trial neural filtering during the concurrent sound input to the ensuing behavioural outcome. This link exists irrespective of age and hearing status, which points to the potency of neural speech tracking to serve as an individualized marker of comprehension problems in clinical settings and as a basis for translational neurotechnological advances.

This key advance notwithstanding, the notable absence of an association between alpha lateralization and listening behaviour also highlights the level of complexity associated with establishing statistically robust relationships of complex neural signatures and behaviour 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 mechanisms alone, and further jointly account for the variability in both neural states and behavioural outcomes[93].

Methods

Data collection

The analysed data are part of an ongoing large-scale study on the neural and cognitive mechanisms supporting adaptive listening behaviour in healthy middle-aged and older adults ("The listening challenge: How ageing brains adapt (AUDADAPT)"; https://cordis.europa.eu/project/rcn/197855_en.html). This project encompasses the collection of different demographic, behavioural, and neurophysiological measures across two time points. The analyses carried out on the data aim at relating adaptive listening behaviour to changes in different neural dynamics[35,36].

Participants and procedure

A total of $N = 155$ right-handed German native speakers (median age 61 years; age range 39–80 years; 62 males; see Supplementary Fig. 1 for age distribution) were included in the analysis. Handedness was assessed using a translated version of the Edinburgh Handedness Inventory[94]. All participants had normal or corrected-to-normal vision, did not report any neurological, psychiatric, or other disorders and were screened for mild cognitive impairment using the German version of the 6-Item Cognitive Impairment Test (6CIT[95]).

During the EEG measurement, participants performed six blocks of a demanding dichotic listening task (see Fig. 2 and Supplementary Methods for details on sentence materials).

As part of our the overarching longitudinal study on adaptive listening behaviour in healthy aging adults, prior to the EEG session, participants also underwent a session consisting of a general screening procedure, detailed audiometric measurements, and a battery of cognitive tests and personality profiling (see ref.[14] for details). Only participants with normal hearing or age-adequate mild-to-moderate hearing loss were included (see Supplementary Fig. 1 for individual audiograms). As part of this screening procedure an additional 17 participants were excluded prior to EEG recording due to non-age-related hearing loss or a medical history. Three participants dropped out of the study prior to EEG recording and an additional 9 participants were excluded from analyses after EEG recording: three due to incidental findings after structural MR acquisition, and six due to technical problems during EEG recording or overall poor EEG data quality. Participants gave written informed consent and received financial compensation (8€ per hour). Procedures were approved by the ethics committee of the University of Lübeck and were in accordance with the Declaration of Helsinki.

Dichotic listening task

In a recently established[35] linguistic variant of a classic Posner paradigm[34], participants listened to two competing, dichotically presented five-word sentences. They were probed on the sentence-final noun in one of the two sentences. All sentences had the

followed the same sentence structure and had an average length of 2512 ms (range: 2183–2963 ms).

Sentences were spoken by the same female talker. Root mean (mean) square intensity (– 26 dB Full Scale, FS) was equalized across all individual sentences and they were masked by continuous speech-shaped noise at a signal-to-noise-ratio of 0 dB. Noise onset was presented with a 50 ms linear onset ramp and preceded sentence onset by 200 ms. Each sentence pair was temporally aligned by the onset of the two task-related sentence-final nouns. This, however led to slight differences in the onset on the individual sentences. Crucially, the range and average sentence onset difference was similar for trials in which the probed (to-be-attended) sentence began earlier and those in which the unprobed (to-be-ignored) sentence began earlier (probed first: range: 0–580 ms, 162.1 ms ± 124.6; unprobed first: 0–560 ms, 180.6 ms ± 127.2). All participants listened to the same 240 sentence pairs but in subject-specific randomized order. In addition, across participants we balanced the assignment of sentences to the right and left ear, respectively. Details on stimulus construction and recording can be found in the Supplementary Methods.

Critically, two visual cues preceded auditory presentation. First, a spatial-attention cue either indicated the to-be-probed ear, thus invoking selective attention, or did not provide any information about the to-be-probed ear, thus invoking divided attention. Second, a semantic cue specified a general or a specific semantic category for the final word of both sentences, thus allowing to utilize a semantic prediction. Cue levels were fully crossed in a 2×2 design and presentation of cue combinations varied on a trial-by-trial level (Fig. 2a). The trial structure is exemplified in Figure 2b.

Each trial started with the presentation of a fixation cross in the middle of the screen (jittered duration: mean 1.5 s, range 0.5–3.5 s). Next, a blank screen was shown for 500 ms followed by the presentation of the spatial cue in the form of a circle segmented equally into two lateral halves. In selective-attention trials, one half was black, indicating the to-be-attended side, while the other half was white, indicating the to-be-ignored side. In divided-attention trials, both halves appeared in grey. After a blank screen of 500 ms duration, the semantic cue was presented in the form of a single word that specified the semantic category of both sentence-final words. The semantic category could either be given at a general (natural vs. man-made) or specific level (e.g. instruments, fruits, furniture) and thus provided different degrees of semantic predictability. Each cue was presented for 1000 ms.

After a 500 ms blank-screen period, the two sentences were presented dichotically along with a fixation cross displayed in the middle of the screen. Finally, after a jittered retention period, a visual response array appeared on the left or right side of the screen, presenting four word choices. The location of the response array indicated which ear (left or right) was probed. Participants were instructed to select the final word presented on the to-be-attended side using the touch screen. Among the four alternatives were the two actually presented nouns as well as two distractor nouns from the same cued semantic category. Note that because the semantic cue applied to all

four alternative verbs, it could not be used to post-hoc infer the to-be-attended sentence-final word.

Stimulus presentation was controlled by PsychoPy Standalone v2.0[96]. The visual scene was displayed using a 24" touch screen (ViewSonic TD2420) positioned within an arm's length. Auditory stimulation was delivered using in-ear headphones (EARTONE 3A) at sampling rate of 44.1 kHz. Following instructions, participants performed a few practice trials to familiarize themselves with the listening task. To account for differences in hearing acuity within our group of participants, individual hearing thresholds for a 500-ms fragment of the dichotic stimuli were measured using the method of limits. All stimuli were presented 50 dB above the individual sensation level. During the experiment, each participant completed 60 trials per cue-cue condition, resulting in 240 trials in total. The cue conditions were equally distributed across six blocks of 40 trials each (~ 10 min) and were presented in random order. Participants took short breaks between blocks.

Behavioural data analysis

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

EEG data analysis

The preprocessed continuous EEG data (see Supplementary Methods for details on data collection and preprocessing) were high-pass-filtered at 0.3 Hz (finite impulse response (FIR) filter, zero-phase lag, order 5574, Hann window) and low-pass-filtered at 180 Hz (FIR filter, zero-phase lag, order 100, Hamming window). The EEG was cut into epochs of –2 to 8 s relative to the onset of the spatial-attention cue to capture cue presentation as well as the entire auditory stimulation interval.

For the analysis of changes in alpha power, EEG data were down-sampled to f_s=250 Hz. Spectro-temporal estimates of single-trial data were then obtained for a time window of –0.5 to 6.5 s (relative to the onset of the spatial-attention cue) at frequencies ranging from 8 to 12 Hz (Morlet's wavelets; number of cycles = 6).

For the analysis of the neural encoding of speech by low-frequency activity, the continuous preprocessed EEG were down-sampled to f_s=125 Hz and filtered between f_c=1 and 8 Hz (FIR filters, zero-phase lag, order: $8f_s/f_c$ and $2f_s/f_c$, Hamming window). The

EEG was cut to yield individual epochs covering the presentation of auditory stimuli, beginning at noise onset until the end of auditory presentation.

Following EEG source and forward model construction (see Supplementary Methods for details), sensor-level single-trial data in each of our two analysis routines were projected to source space by matrix-multiplication of the spatial filter weights. To increase signal-to-noise ratio in source estimates and computationally facilitate source-level analyses, source-projected data were averaged across grid-points per cortical area defined according to the HCP functional parcellation template[97,98]. This parcellation provides a symmetrical delineation of each hemisphere into 180 parcels for a total of 360 parcels. We constrained the analysis of neural measures to an a priori defined, source-localized auditory region of interest (ROI) as well as one control ROI in the inferior parietal lobule (see Supplementary Methods for details). The described analyses were carried out using the Fieldtrip toolbox (v. 2017-04-28) in Matlab 2016b, and the Human Connectome Project Workbench software (v1.5) as well as FreeSurfer (v.6.0).

Attentional modulation of alpha power

Absolute source power was calculated as the square-amplitude of the spectro-temporal estimates. Since oscillatory power values typically follow a highly skewed, non-normal 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[99]. 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[12,14,15].

$$ALI = (\alpha\text{-}power_{ipsi} - \alpha\text{-}power_{contra})/(\alpha\text{-}power_{ipsi} + \alpha\text{-}power_{contra}) \qquad (1)$$

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–behaviour 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 Supplementary Methods for details) of the attended and ignored speech stream from EEG[100,101]. 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[102]:

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

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, I is the identity matrix , and m is a subject-specific scalar representing the mean of the trace of $R^T R$[103,104]. The same grid of ridge parameters ($\lambda = 10^{-5}, 10^{-4}, \ldots 10^{10}$) was used across subject, and *m* proved to be relatively stable across subjects (387.2 ± 0.18, mean ± sd). The optimal ridge value of λ = 1 was determined based on the average Pearson's correlation coefficient and mean squared error of the reconstructed and actually presented envelope across all trials and subjects.

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[105]. All analyses were performed using the multivariate temporal response function (mTRF) toolbox[100] (v1.5) for Matlab (v. 2016b).

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 analysed 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 248-ms sliding window (rectangular window, step size of 1 (8 ms) sample). 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[28] (see Fig. 4 and Supplementary Fig. 4).

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)[27]:

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

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 first assigned each sample to one of 100 bins covering the length of the original sentence in 1% increments. We then averaged across neighbouring bins using a centred rectangular 3% sliding window (1% overlap). 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 behaviour. 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[50].

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 sample size of N=155 participants, *p*-values for individual model terms are based on Wald *t-as-z*-values for linear models[106] 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[107].

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 (v.3.6.1)[108] using the packages lme4 (v. 1.1-23)[109], and sjPlot (v2.8.5)[110].

Model selection

To avoid known problems associated with a largely data-driven stepwise model selection that include the overestimation of coefficients[111] or the selection of irrelevant predictors[112], 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–behaviour model. The brain–behaviour 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 neural and behavioral data required to reproduce the analyses supporting this work, as well as the auditory stimuli used in this study are publicly available in the study's Open Science Framework repository (https://osf.io/nfv9e/). Source data are provided with this paper.

Code availability

Code for the analyses supporting this work is publicly available in the study's Open Science Framework repository (https://osf.io/nfv9e/).

References

1. Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* **25,** 975–979 (1953).
2. Obleser, J. & Erb, J. in *The Cognitive Neurosciences* (eds. Poeppel, D., Mangun, G. S. & Gazzaniga, M. S.) (MIT Press, 2020).
3. Broadbent, D. E. *Perception and communication*. (Pergamon Press, 1958). doi:10.1186/s12889-020-10047-2
4. Fernandez-Duque, D. & Johnson, M. L. Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science* **23,** 83–116 (1999).
5. Jensen, O. & Mazaheri, A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* **4,** 186 (2010).
6. Foxe, J. J. & Snyder, A. C. The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Front. Psychology* **2,** 154 (2011).
7. Händel, B. F., Haarmeier, T. & Jensen, O. Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J. Cogn. Neurosci.* **23,** 2494–2502 (2011).
8. Rihs, T. A., Michel, C. M. & Thut, G. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. *Eur. J. Neurosci.* **25,** 603–610 (2007).
9. Kerlin, J. R., Shahin, A. J. & Miller, L. M. Attentional gain control of ongoing cortical speech representations in a 'cocktail party'. *J. Neurosci.* **30,** 620–628 (2010).
10. Ahveninen, J., Huang, S., Belliveau, J. W., Chang, W.-T. & Hämäläinen, M. Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. *J. Cogn. Neurosci.* **25,** 1926–1943 (2013).
11. Müller, N. & Weisz, N. Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. *Cereb. Cortex* **22,** 1604–1613 (2011).
12. Wöstmann, M., Herrmann, B., Maess, B. & Obleser, J. Spatiotemporal dynamics of auditory attention synchronize with speech. *Proceedings of the National Academy of Sciences* **113,** 3873–3878 (2016).
13. Wöstmann, M., Vosskuhl, J., Obleser, J. & Herrmann, C. S. Opposite effects of lateralised transcranial alpha versus gamma stimulation on auditory spatial attention. *Brain Stimulation* **11,** 752–758 (2018).
14. Tune, S., Wöstmann, M. & Obleser, J. Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *Eur. J. Neurosci.* **48,** 2537–2550 (2018).
15. Haegens, S., Handel, B. F. & Jensen, O. Top-Down Controlled Alpha Band Activity in Somatosensory Areas Determines Behavioral Performance in a Discrimination Task. *J. Neurosci.* **31,** 5197–5204 (2011).
16. Bauer, M., Kennett, S. & Driver, J. Attentional selection of location and modality in vision and touch modulates low-frequency activity in associated sensory cortices. *J. Neurophysiol.* **107,** 2342–2351 (2012).
17. Worden, M. S., Foxe, J. J., Wang, N. & Simpson, G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* 1–6 (2000). doi:https://doi.org/10.1523/JNEUROSCI.20-06-j0002.2000
18. Kelly, S. P., Lalor, E. C., Reilly, R. B. & Foxe, J. J. Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J. Neurophysiol.* **95,** 3844–3851 (2006).

19. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* **32,** 9–18 (2009).
20. Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H. & Lakatos, P. Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology* **20,** 172–176 (2010).
21. Henry, M. J. & Obleser, J. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 20095–20100 (2012).
22. Obleser, J. & Kayser, C. Neural Entrainment and Attentional Selection in the Listening Brain. *Trends Cogn. Sci.* **23,** 913–926 (2019).
23. Brodbeck, C., Hong, L. E. & Simon, J. Z. Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Curr. Biol.* **28,** 3976–3982.e6 (2018).
24. Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J. & Lalor, E. C. Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. *Curr. Biol.* **28,** 1–11 (2018).
25. Zion Golumbic, E. M. Z. *et al.* Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a '"Cocktail Party"'. *Neuron* **77,** 980–991 (2013).
26. Ding, N. & Simon, J. Z. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 11854–11859 (2012).
27. Mesgarani, N. & Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485,** 233–236 (2012).
28. O'Sullivan, J. A. *et al.* Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cereb. Cortex* **25,** 1697–1706 (2014).
29. Horton, C., D'Zmura, M. & Srinivasan, R. Suppression of competing speech through entrainment of cortical oscillations. *J. Neurophysiol.* **109,** 3082–3093 (2013).
30. Keitel, C. *et al.* Stimulus-Driven Brain Rhythms within the Alpha Band: The Attentional-Modulation Conundrum. *J. Neurosci.* **39,** 3119–3129 (2019).
31. Gundlach, C., Moratti, S., Forschack, N. & Müller, M. M. Spatial Attentional Selection Modulates Early Visual Stimulus Processing Independently of Visual Alpha Modulations. *Cereb. Cortex* **16,** 637–18 (2020).
32. Henry, M. J., Herrmann, B., Kunke, D. & Obleser, J. Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nat Commun* **8,** 15801 (2017).
33. Lakatos, P. *et al.* Global dynamics of selective attention and its lapses in primary auditory cortex. *Nat. Neurosci.* **19,** 1707–1717 (2016).
34. Posner, M. I. Orienting of attention. *Quarterly Journal of Experimental Psychology* **32,** 3–25 (1980).
35. Alavash, M., Tune, S. & Obleser, J. Modular reconfiguration of an auditory control brain network supports adaptive listening behavior. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 660–669 (2019).
36. Alavash, M., Tune, S. & Obleser, J. Large-scale cortical connectivity dynamics of intrinsic neural oscillations support adaptive listening behavior. *bioRxiv* 2021.02.22.432315 (2021). doi:10.1101/2021.02.22.432315
37. Presacco, A., Simon, J. Z. & Anderson, S. Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J. Neurophysiol.* **116,** 2356–2367 (2016).
38. Sohoglu, E., Peelle, J. E., Carlyon, R. P. & Davis, M. H. Predictive Top-Down Integration of Prior Knowledge during Speech Perception. *J. Neurosci.* **32,** 8443–8453 (2012).

39. Peelle, J. E., Gross, J. & Davis, M. H. Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cereb. Cortex* **23,** 1378–1387 (2013).
40. Obleser, J. & Weisz, N. Suppressed Alpha Oscillations Predict Intelligibility of Speech and its Acoustic Details. *Cereb. Cortex* **22,** 2466–2477 (2012).
41. Wöstmann, M., Lim, S.-J. & Obleser, J. The human neural alpha response to speech is a proxy of attentional control. *Cereb. Cortex* **27,** 3307–3317 (2017).
42. Broderick, M. P., Anderson, A. J. & Lalor, E. C. Semantic Context Enhances the Early Auditory Encoding of Natural Speech. *J. Neurosci.* **39,** 7564–7575 (2019).
43. Broadbent, D. E. & Gregory, M. Accuracy of recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology* **16,** 359–360 (1964).
44. Kimura, D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology/Revue canadienne de psychologie* **15,** 166–171 (1961).
45. Decruy, L., Vanthornhout, J. & Francart, T. Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hearing Research* **393,** 107961 (2020).
46. Presacco, A., Simon, J. Z. & Anderson, S. Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS ONE* **14,** e0213899–26 (2019).
47. Petersen, E. B., Wöstmann, M., Obleser, J. & Lunner, T. Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J. Neurophysiol.* **117,** 18–27 (2017).
48. Bonacci, L. M., Bressler, S. & Shinn-Cunningham, B. G. Nonspatial Features Reduce the Reliance on Sustained Spatial Auditory Attention. *Ear and Hearing* **41,** 1635–1647 (2020).
49. Fiedler, L., Wöstmann, M., Herbst, S. K. & Obleser, J. Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *NeuroImage* **186,** 33–42 (2019).
50. Bell, A., Fairbrother, M. & Jones, K. Fixed and random effects models: making an informed choice. *Qual Quant* **53,** 1051–1074 (2018).
51. Passow, S. *et al.* Human aging compromises attentional control of auditory perception. *Psychol Aging* **27,** 99–105 (2012).
52. Anderson, S., White-Schwoch, T., Parbery-Clark, A. & Kraus, N. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing Research* **300,** 18–32 (2013).
53. Banerjee, S., Snyder, A. C., Molholm, S. & Foxe, J. J. 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,** 9923–9932 (2011).
54. Waschke, L., Tune, S. & Obleser, J. Local cortical desynchronization and pupil-linked arousal differentially shape brain states for optimal sensory performance. *eLife Sciences* **8,** 1868–27 (2019).
55. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A. & Poeppel, D. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93,** 480–490 (2017).
56. van Ede, F., Köster, M. & Maris, E. Beyond establishing involvement: quantifying the contribution of anticipatory α- and β-band suppression to perceptual improvement with attention. *J. Neurophysiol.* **108,** 2352–2362 (2012).
57. Ding, N. & Simon, J. Z. Cortical entrainment to continuous speech: functional roles and interpretations. *Front. Hum. Neurosci.* **8,** 13367 (2014).

58. Hamilton, L. S. & Huth, A. G. The revolution will not be controlled: natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience* **27,** 1–10 (2018).
59. Sassenhagen, J. How to analyse electrophysiological responses to naturalistic language with time-resolved multiple regression. *Language, Cognition and Neuroscience* **0,** 1–17 (2018).
60. Ding, N. & Simon, J. Z. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J. Neurophysiol.* **107,** 78–89 (2012).
61. Pernet, C. R., Sajda, P. & Rousselet, G. A. Single-trial analyses: why bother? *Front. Psychology* **2,** 322 (2011).
62. Riecke, L., Formisano, E., Sorger, B., Başkent, D. & Gaudrain, E. Neural Entrainment to Speech Modulates Speech Intelligibility. *Curr. Biol.* **28,** 1–9 (2018).
63. Zoefel, B., Archer-Boyd, A. & Davis, M. H. Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr. Biol.* **28,** 401–408.e5 (2018).
64. Thut, G. **α**-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* **26,** 9494–9502 (2006).
65. Bengson, J. J., Mangun, G. R. & Mazaheri, A. The neural markers of an imminent failure of response inhibition. *NeuroImage* **59,** 1534–1539 (2012).
66. Dahl, M. J., Ilg, L., Li, S.-C., Passow, S. & Werkle-Bergner, M. Diminished pre-stimulus alpha-lateralization suggests compromised self-initiated attentional control of auditory processing in old age. *NeuroImage* **197,** 414–424 (2019).
67. Hong, X., Sun, J., Bengson, J. J., Mangun, G. R. & Tong, S. Normal aging selectively diminishes alpha lateralization in visual spatial attention. *NeuroImage* **106,** 353–363 (2015).
68. Leenders, M. P., Lozano-Soldevilla, D., Roberts, M. J., Jensen, O. & De Weerd, P. Diminished Alpha Lateralization During Working Memory but Not During Attentional Cueing in Older Adults. *Cereb. Cortex* **28,** 21–32 (2018).
69. Mok, R. M., Myers, N. E., Wallis, G. & Nobre, A. C. Behavioral and neural markers of flexible attention over working memory in aging. *Cereb. Cortex* **26,** 1831–1842 (2016).
70. Wöstmann, M., Alavash, M. & Obleser, J. Alpha Oscillations in the Human Brain Implement Distractor Suppression Independent of Target Selection. *J. Neurosci.* **39,** 9797–9805 (2019).
71. Iemi, L. *et al.* Spontaneous neural oscillations influence behavior and sensory representations by suppressing neuronal excitability. *bioRxiv* 2021.03.01.433450 (2021). doi:10.1101/2021.03.01.433450
72. Kloosterman, N. A. *et al.* Humans strategically shift decision bias by flexibly adjusting sensory evidence accumulation. *eLife Sciences* **8,** (2019).
73. Hauswald, A., Keitel, A., Chen, Y.-P., Rösch, S. & Weisz, N. Degradation levels of continuous speech affect neural speech tracking and alpha power differently. in (John Wiley & Sons, Ltd, 2020). doi:10.1111/ejn.14912
74. Clayton, M. S., Yeung, N. & Cohen Kadosh, R. The many characters of visual alpha oscillations. *Eur. J. Neurosci.* **48,** 2498–2508 (2018).
75. Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J. & Tiesinga, P. Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nat. Neurosci.* **17,** 1031–1039 (2014).
76. Sadaghiani, S. & Kleinschmidt, A. Brain Networks and α-Oscillations: Structural and Functional Foundations of Cognitive Control. *Trends Cogn. Sci.* **20,** 805–817 (2016).

77. Wöstmann, M. & Obleser, J. Acoustic detail but not predictability of task-irrelevant speech disrupts working memory. *Front. Hum. Neurosci.* **10,** 201–9 (2016).
78. van Kerkoerle, T. *et al.* Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences* **111,** 14332–14341 (2014).
79. Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A. & Jensen, O. Layer-Specific Entrainment of Gamma-Band Neural Activity by the Alpha Rhythm in Monkey Visual Cortex. *Curr. Biol.* **22,** 2313–2318 (2012).
80. Chapeton, J. I., Haque, R., Wittig, J. H., Jr, Inati, S. K. & Zaghloul, K. A. Large-Scale Communication in the Human Brain Is Rhythmically Modulated through Alpha Coherence. *Curr. Biol.* **29,** 2801–2811.e5 (2019).
81. Antonov, P. A., Chakravarthi, R. & Andersen, S. K. Too little, too late, and in the wrong place: Alpha band activity does not reflect an active mechanism of selective attention. *NeuroImage* **219,** 117006 (2020).
82. Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O. & Jensen, O. Probing cortical excitability using rapid frequency tagging. *NeuroImage* **195,** 59–66 (2019).
83. Teoh, E. S. & Lalor, E. C. EEG decoding of the target speaker in a cocktail party scenario: considerations regarding dynamic switching of talker location. *J Neural Eng* **16,** 036017–30 (2019).
84. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* **32,** 9–18 (2009).
85. Herrmann, B., Henry, M. J., Haegens, S. & Obleser, J. Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage* **124,** 487–497 (2016).
86. Pichora-Fuller, M. K. Using the Brain when the Ears are Challenged Helps Healthy Older Listeners Compensate and Preserve Communication Function. 1–14 (2017).
87. Rogers, C. S., Payne, L., Maharjan, S., Wingfield, A. & Sekuler, R. Older adults show impaired modulation of attentional alpha oscillations: Evidence from dichotic listening. *Psychol Aging* **33,** 246–258 (2018).
88. Getzmann, S., Klatt, L.-I., Schneider, D., Begau, A. & Wascher, E. 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 (2020).
89. Fuglsang, S. A., Märcher-Rørsted, J., Dau, T. & Hjortkjær, J. Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. *J. Neurosci.* **40,** 2562–2572 (2020).
90. Decruy, L., Vanthornhout, J. & Francart, T. Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *J. Neurophysiol.* **122,** 601–615 (2019).
91. Presacco, A., Simon, J. Z. & Anderson, S. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J. Neurophysiol.* **116,** 2346–2355 (2016).
92. Verschueren, E., Vanthornhout, J. & Francart, T. The Effect of Stimulus Choice on an EEG-Based Objective Measure of Speech Intelligibility. *Ear and Hearing* **41,** 1586–1597 (2020).
93. Waschke, L., Kloosterman, N. A., Obleser, J. & Garrett, D. D. Behavior needs neural variability. *Neuron* **109,** 751–766 (2021).
94. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. **9,** 97–113 (1971).

95. Jefferies, K. & Gale, T. M. in *Cognitive Screening Instruments* 209–218 (Springer, London, 2013). doi:10.1007/978-1-4471-2452-8_11
96. Peirce, J. W. PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods* **162,** 8–13 (2007).
97. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536,** 171–178 (2016).
98. Keitel, A. & Gross, J. Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLoS Biol* **14,** e1002498–22 (2016).
99. Smulders, F. T. Y., Oever, ten, S., Donkers, F. C. L., Quaedflieg, C. W. E. M. & van de Ven, V. Single-trial log transformation is optimal in frequency analysis of resting EEG alpha. *Eur. J. Neurosci.* **44,** 94–14 (2018).
100. Crosse, M. J., Di Liberto, G. M., Bednar, A. & Lalor, E. C. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front. Hum. Neurosci.* **10,** 604 (2016).
101. Lalor, E. C. & Foxe, J. J. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur. J. Neurosci.* **31,** 189–193 (2010).
102. Hoerl, A. E. & Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **12,** 55–67 (1970).
103. Biesmans, W., Das, N., Francart, T. & Bertrand, A. 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,** 402–412 (2017).
104. Fiedler, L. *et al.* Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. *J Neural Eng* **14,** 036020 (2017).
105. Haufe, S. *et al.* On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* **87,** 96–110 (2014).
106. Luke, S. G. Evaluating significance in linear mixed-effects models in R. *Behav Res* **49,** 1494–1502 (2017).
107. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* **57,** 289–300 (1995).
108. R Core Team. R: A language and environment for statistical computing. (2019). doi:https://www.R-project.org/
109. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Soft.* **67,** 1–48 (2015).
110. Lüdecke, D. & Lüdecke, D. Data Visualization for Statistics in Social Science [R package sjPlot version 2.6.1]. (2018). doi:10.5281/zenodo.1308157
111. Tibshirani, R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* **56,** 267–288 (1996).
112. Derksen, S. & Keselman, H. J. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology* **45,** 265–282 (1992).

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.

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. analysed the data under supervision by J.O.; S.T. wrote the first draft; all authors contributed to the final version of the manuscript.

Competing interests: The authors declare no competing interests.

Figure legends

Figure 1. Schematic illustration of addressed research questions. 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 behavioural performance (Q1). In line with (Q2), we also examined the degree to which a spatial (and semantic) cue 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 behavioural performance. Furthermore, we investigated the impact of age, hearing loss, and probed ear on listening success and its underlying neural strategies. Source data are provided as a Source Data file.

Figure 2. Experimental design and behavioural benefit from informative cues.
(a) Visualization of used 2×2 design[35]. Levels of spatial and semantic cue differed on a trial-by-trial basis. Note that the effects of the semantic cue were of secondary importance to the current analyses. Top row shows the informative [+] cue levels, bottom row the uninformative [–] cue levels.
(b) Schematic representation of the trial structure. 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) Left: Accuracy per cue-cue combination. Coloured dots are individual (N=155 participants) trial-averages, black dots and vertical lines show group means with bootstrapped 95% confidence intervals (CI). Right: Individual cue benefits displayed separately for the two cues (top: spatial cue, bottom: semantic cue). Black dots indicate individual (N=155) mean accuracy with bootstrapped 95 % CI. Histograms show the distribution of difference in accuracy for informative vs. uninformative levels. OR: Odds ratio parameter estimate from generalized linear mixed-effects models; two-sided Wald test (FDR-corrected); spatial cue: $p=1.36\times10^{-24}$; semantic cue: $p=.68$.
(d) Left: Response speed per cue-cue combination. Coloured dots show individual (N=155 participants) mean speed, black dots and vertical lines show group means with bootstrapped 95% CI. Right: Individual cue benefits displayed separately for the two cues (top: spatial cue, bottom: semantic cue). Black dots indicate individual (N=155) mean speed with bootstrapped 95% CI. Box plots in (c) and (d) show median centre line, 25th to 75th percentile hinges, whiskers indicate minimum and maximum within 1.5 × interquartile range. β: slope parameter estimate from general linear mixed-effects models; two-sided Wald test (FDR-corrected); spatial cue: $p=4.49\times10^{-48}$; semantic cue: $p=2.49\times10^{-9}$. Source data are provided as a Source Data file.

Figure 3. Informative spatial cue elicits increased alpha power lateralization before and during speech presentation.
(a) Grand-average (N=155 participants) whole-trial attentional modulation of 8–12 Hz auditory alpha power. Purple traces show the grand-average alpha lateralization index (ALI) for the informative (solid dark purple line) and uninformative spatial cue (dashed light purple line), each collapsed across semantic cue levels. Error bands indicate ± 1 SEM. Positive values indicate relatively higher alpha power in the hemisphere ipsilateral to the attended/-probed sentence compared to the contralateral hemisphere. Shaded grey area shows the time window of sentence presentation. Brain models show auditory region of interest (red).

(b) ALI during sentence presentation (3.5–6.5 s) shown separately per spatial cue condition and probed ear (left plot) for *N*=155 participants. Coloured dots show trial-averaged individual results, black dots and error bars indicate the grand-average and bootstrapped 95% confidence intervals. Box plots show median centre line, 25th to 75th percentile hinges, whiskers show minimum and maximum within 1.5 × interquartile range. For probed-right trials there was a significant difference in ALI between selective- and divided-attention trials (right plot). Black dots represent individual mean ALI values with bootstrapped 95% CI error bars. 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; two-sided Wald test (FDR-corrected, ***p=2.65×10^{-17}. Source data are provided as a Source Data file.

Figure 4. Neural speech tracking of attended and ignored sentences.
(a) Schematic representation of linear backward model approach. Linear backward models estimated on selective-attention trials. Onset envelopes are reconstructed via convolution of auditory EEG responses with estimated backward models and compared actual envelopes to assess neural tracking strength and decoding accuracy (see Supplementary Methods).
(b) Left: Grand average (*N*=155 participants, 95% confidence interval (CI) error bands) forward-transformed temporal response functions (TRFs) for attended (green) and ignored (yellow) speech in the left and right auditory ROI. Right: Single-subject (*N*=155 participants; 95% CI error bars) mean Pearson correlation of reconstructed and presented envelopes shown separately for attended (top, green) and ignored speech (bottom, yellow).
(c) Top: Grand average (*N*=155 participants) peri-stimulus time course of neural tracking index shown separately for selective (solid dark green curve) and divided attention (dashed light green curve) ± 1 SEM error band. Histograms show sentence and final word onsets. Shaded area indicates final word presentation interval used for statistical analysis. Bottom: Single-subject (*N*=155 participants) mean attended and ignored neural speech tracking during final word presentation for selective and divided attention, respectively.
(d) Left: Neural tracking index shown per spatial-attention condition and for trials in which cued/probed sentences started ahead of ('probed first') or after the distractor ('probed second'). Coloured dots represent single-subject average (*N*=155 participants), black dots and error bars indicate grand-average and bootstrapped 95 % CI. Box plots show median centre line, 25th to 75th percentile hinges, whiskers indicate minimum and maximum within 1.5 × interquartile range. Right: Significant difference in neural tracking between selective- and divided-attention trials in probed-second trials (top plot), and stronger neural tracking in probed-left trials. Black dots represent individual mean neural tracking index with bootstrapped 95% CI error bars for *N*=155 participants. 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=1.22×10^{-9}, *p=.0233 (two-sided Wald test, FDR-corrected). Source data are provided as a Source Data file.

Figure 5. Relationship of alpha lateralization and neural speech 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 speech tracking and alpha 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 with 95 % confidence interval, 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, two-sided Wald test (FDR-corrected).

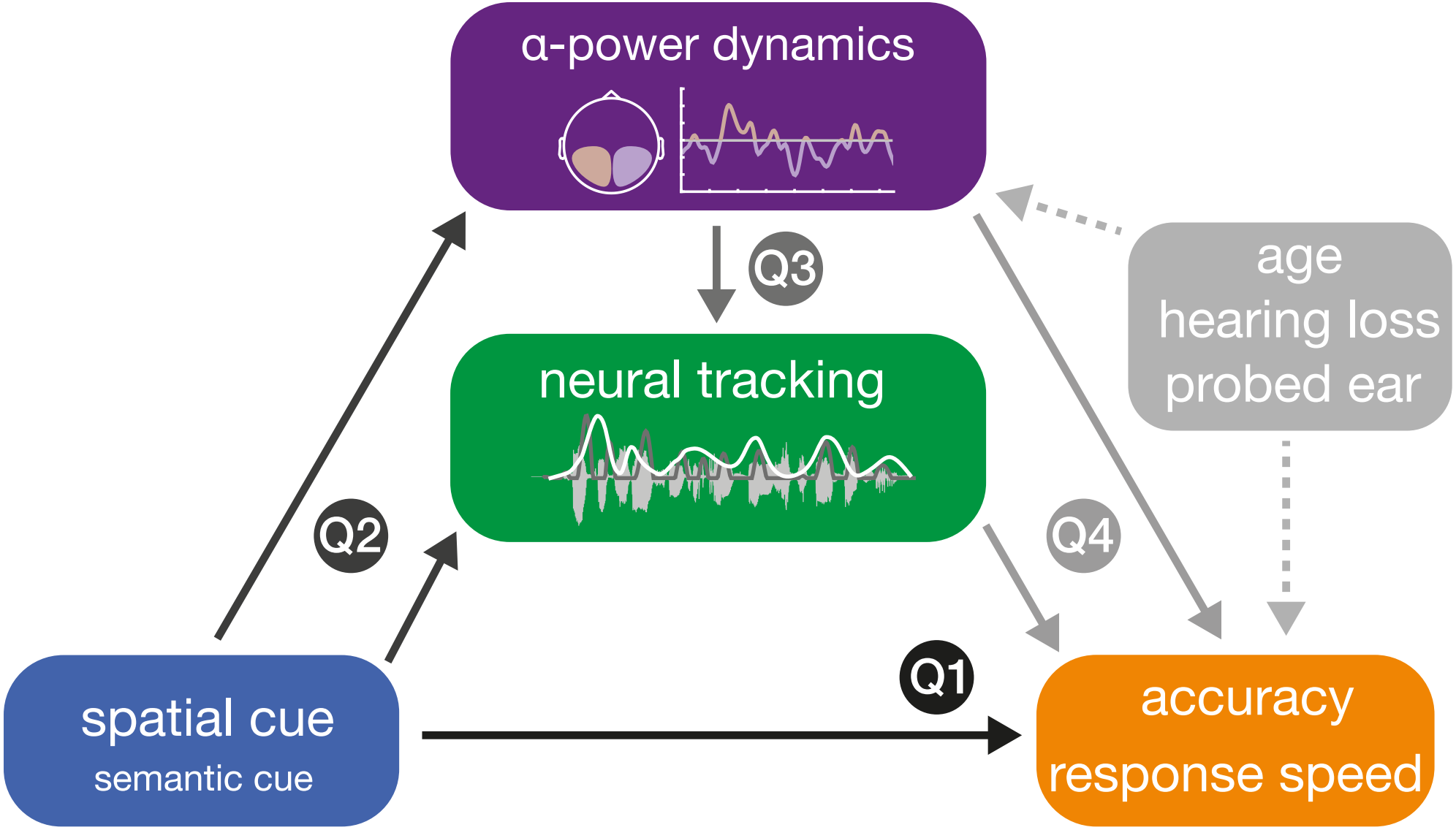
(c) Grand-average time courses of alpha lateralization and neural speech tracking during sentence presentation mapped to the same peri-stimulus time axis. Shown separately for selective-attention (darker, solid curves) and divided-attention trials (lighter, dashed curves). Error bands reflect ± 1 SEM. Note how the peak in neural speech tracking under selective attention precedes the peak in alpha lateralization.

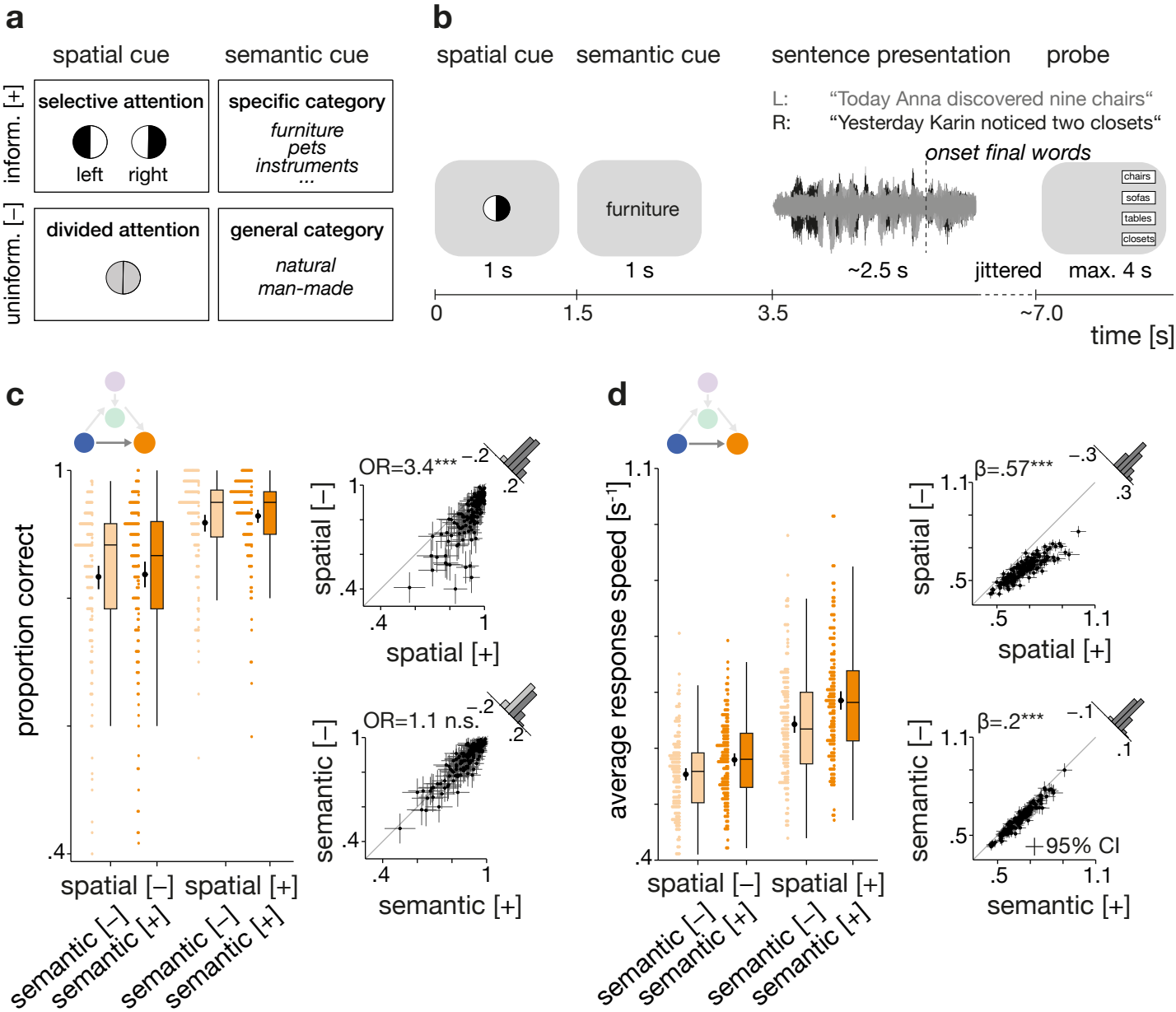
(d) Mean normalized cross-correlation of trial-averaged neural speech tracking and alpha lateralization time courses. Upper and lower bound of shaded area reflect the 97.5^{th} and 2.5^{th} percentile of surrogate data derived from 5000 independently permuted time courses of alpha power and neural speech tracking. Source data are provided as a Source Data file.

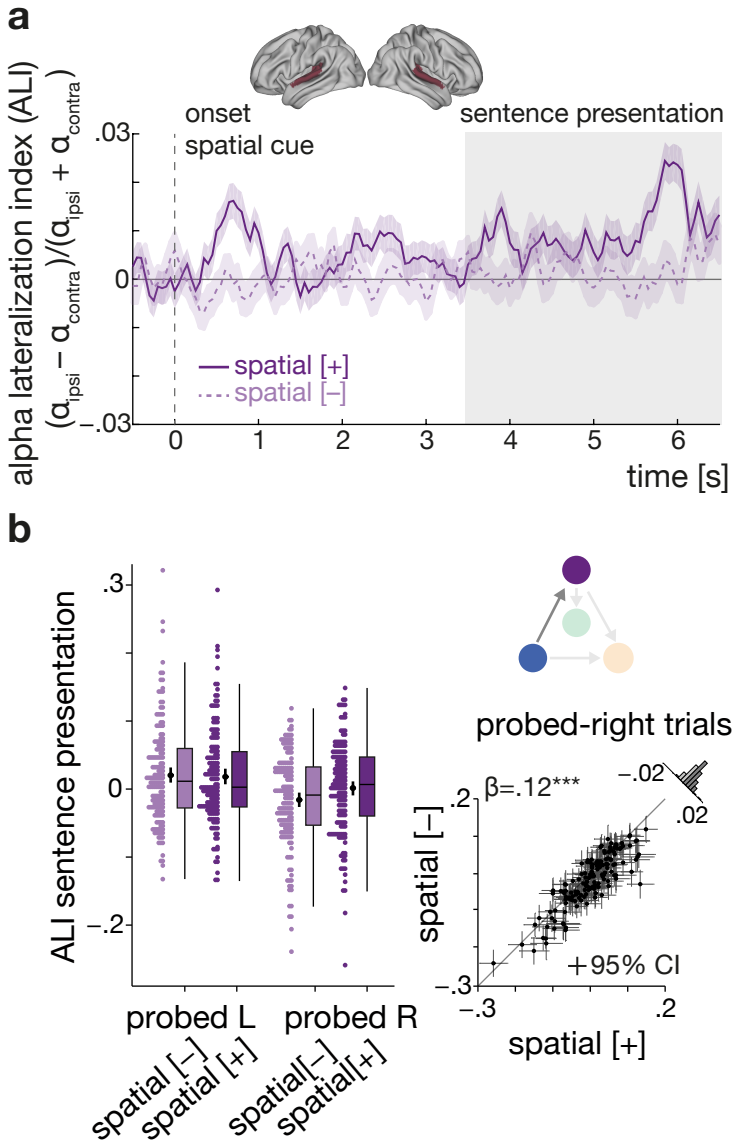
Figure 6. Neural speech tracking predicts listening behavior.

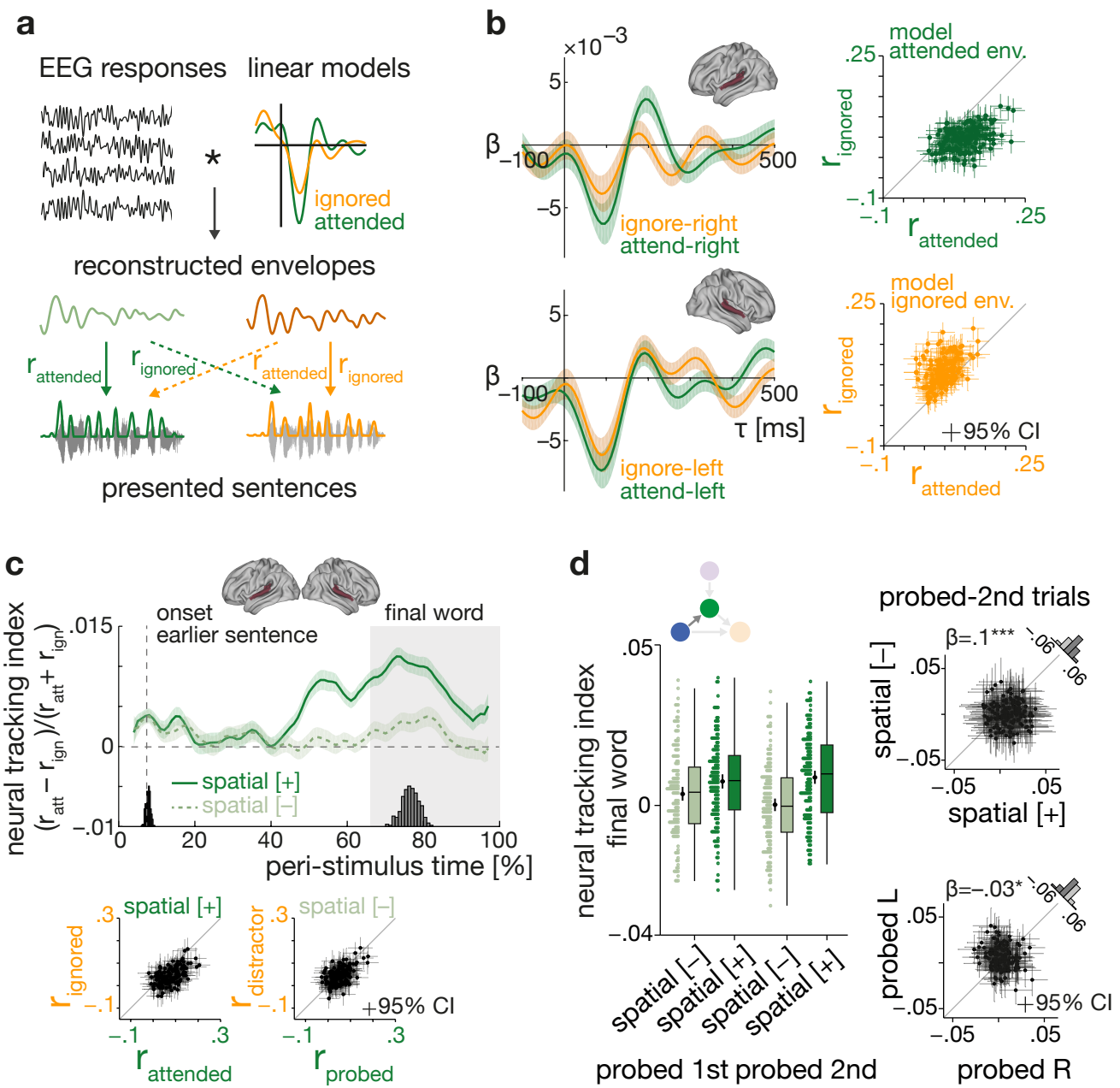
(a) Model predictions for the effect of neural tracking on behavior for *N*=155 participants. Left panel shows the predicted group-level fixed effect (green line ± 95% CI) of trial-to-trial variation in neural tracking on accuracy. Grey thin lines indicate estimated subject-specific slopes. Right panel 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 = slope parameter estimate from the corresponding general linear mixed-effects model, two-sided Wald test (FDR-corrected).

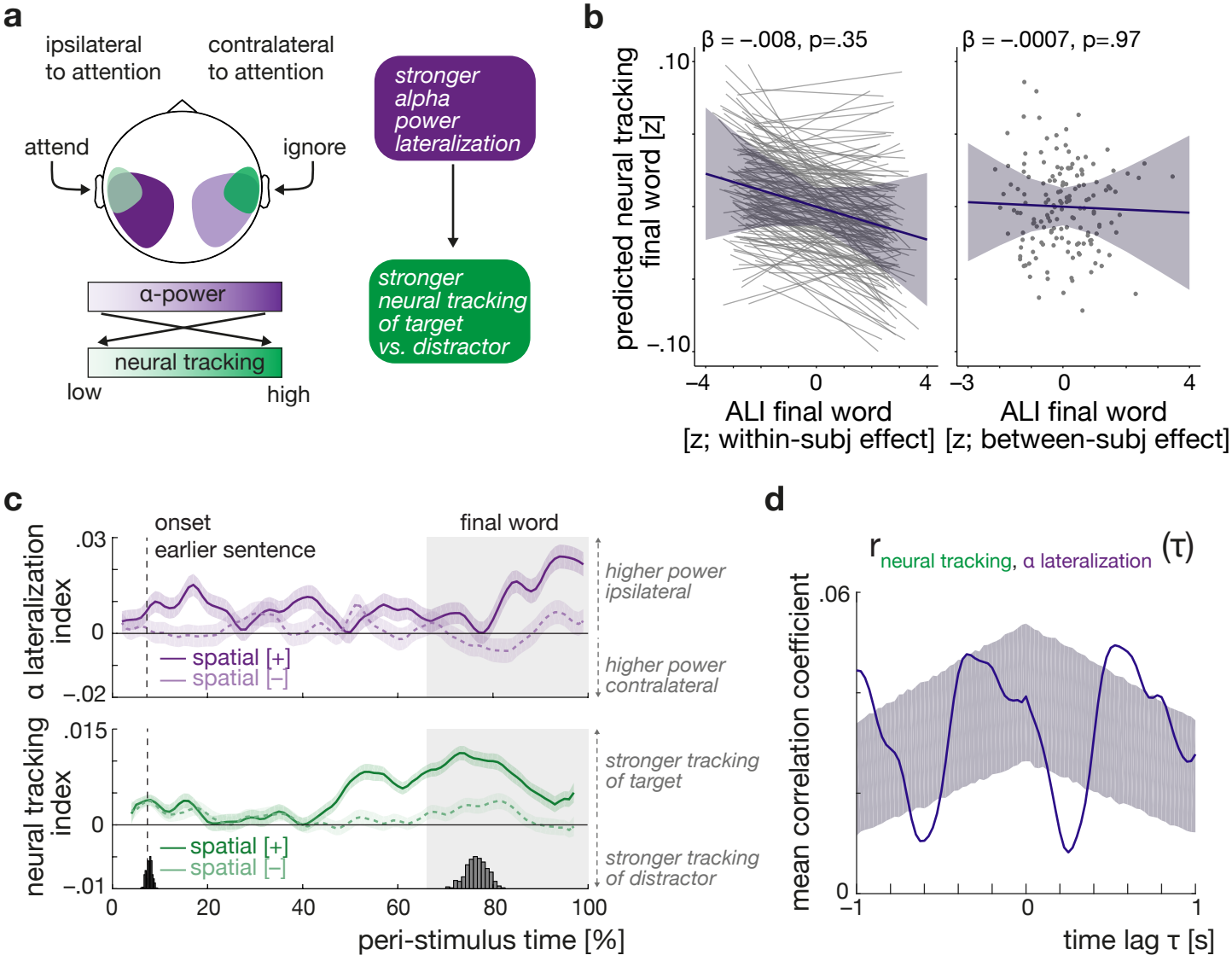
(b) Summary of results. Black arrows highlight statistically significant effects from (generalized) single-trial linear mixed-effects modelling. Grey arrow connecting alpha power dynamics with neural tracking reflects their systematic temporal delay observed in the exploratory cross-correlation analysis. Notably, changes in age and hearing loss did not modulate the fidelity of the two key neural measures. Source data are provided as a Source Data file.

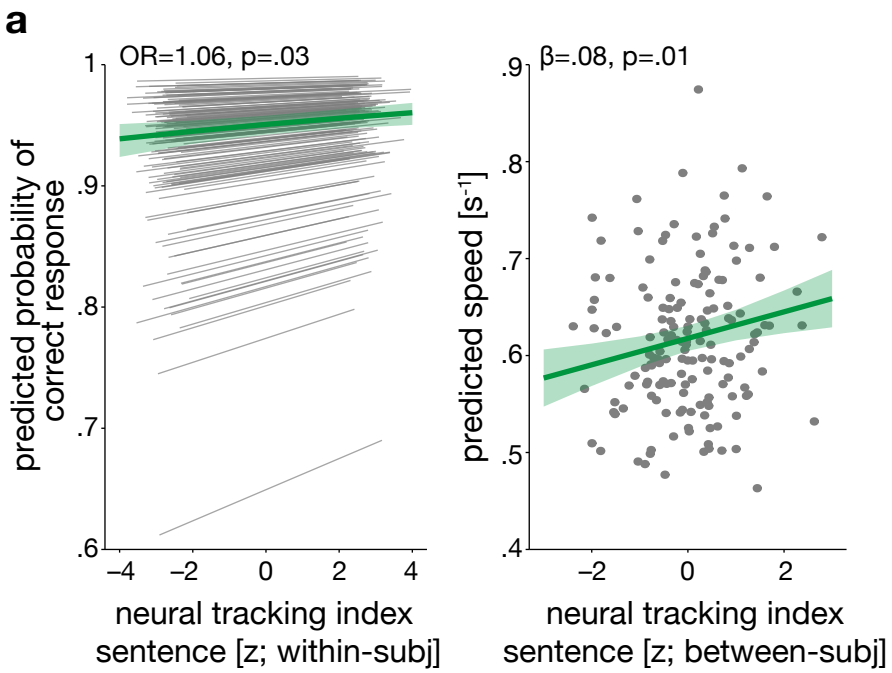

α-power dynamics
Q3
neural tracking
age
hearing loss
probed ear
Q2
Q4
Q1
spatial cue
semantic cue
accuracy
response speed

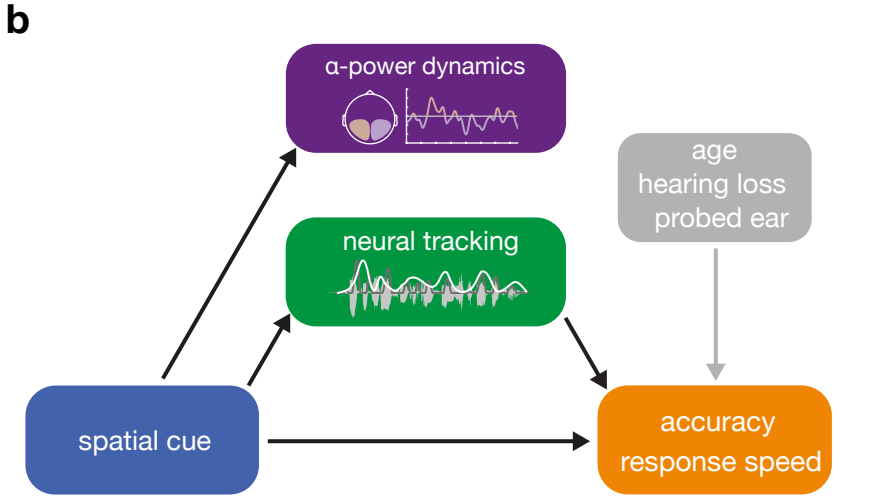

a
spatial cue
semantic cue
inform. [+]
uninform. [–]
selective attention
left
right
specific category
furniture
pets
instruments
...
divided attention
general category
natural
man-made
b
spatial cue
semantic cue
sentence presentation
probe
L: "Today Anna discovered nine chairs"
R: "Yesterday Karin noticed two closets"
onset final words
furniture
chairs
sofas
tables
closets
1 s
1 s
~2.5 s
jittered
max. 4 s
0
1.5
3.5
~7.0
time [s]
c
proportion correct
1
.4
spatial [–]
spatial [+]
semantic [–]
semantic [+]
OR=3.4***
–.2
.2
spatial [–]
spatial [+]
OR=1.1 n.s.
–.2
.2
semantic [–]
semantic [+]
d
average response speed [s⁻¹]
1.1
.4
β=.57***
–.3
.3
β=.2***
–.1
.1
.5
+95% CI


a
onset
spatial cue
sentence presentation
alpha lateralization index (ALI)
$(\alpha_{ipsi} - \alpha_{contra})/(\alpha_{ipsi} + \alpha_{contra})$
.03
0
−.03
0
1
2
3
4
5
6
time [s]
spatial [+]
spatial [−]
b
ALI sentence presentation
.3
0
−.2
probed L
probed R
spatial [−]
spatial [+]
spatial [−]
spatial [+]
probed-right trials
β=.12***
−.02
.02
.2
−.3
spatial [−]
+95% CI
−.3
.2
spatial [+]


a
EEG responses
linear models
ignored
attended
reconstructed envelopes
r attended
r ignored
r attended
r ignored
presented sentences
b
×10−3
5
β
−100
500
−5
ignore-right
attend-right
model
attended env.
.25
r ignored
−.1
−.1
r attended
.25
5
β
−100
500
−5
τ [ms]
ignore-left
attend-left
model
ignored env.
.25
r ignored
−.1
+95% CI
−.1
r attended
.25
c
neural tracking index
(r att − r ign)/(r att + r ign)
.015
onset
earlier sentence
final word
0
−.01
spatial [+]
spatial [−]
0
20
40
60
80
100
peri-stimulus time [%]
spatial [+]
.3
r ignored
−.1
−.1
.3
r attended
spatial [−]
.3
r distractor
−.1
+95% CI
−.1
.3
r probed
d
neural tracking index
final word
.05
0
−.04
spatial [−]
spatial [+]
spatial [−]
spatial [+]
probed 1st probed 2nd
probed-2nd trials
β=.1***
.06
.06
spatial [−]
.05
−.05
−.05
.05
spatial [+]
β=−.03*
.06
.06
probed L
.05
−.05
+95% CI
−.05
.05
probed R


a
ipsilateral to attention
contralateral to attention
attend
ignore
α-power
neural tracking
low
high
stronger alpha power lateralization
stronger neural tracking of target vs. distractor
b
β = −.008, p=.35
β = −.0007, p=.97
predicted neural tracking final word [z]
.10
−.10
−4
0
4
−3
0
4
ALI final word [z; within-subj effect]
ALI final word [z; between-subj effect]
c
onset earlier sentence
final word
α lateralization index
.03
0
−.02
spatial [+]
spatial [−]
higher power ipsilateral
higher power contralateral
neural tracking index
.015
0
−.01
spatial [+]
spatial [−]
stronger tracking of target
stronger tracking of distractor
0
20
40
60
80
100
peri-stimulus time [%]
d
r neural tracking, α lateralization (τ)
mean correlation coefficient
.06
0
−1
0
1
time lag τ [s]


a
OR=1.06, p=.03
predicted probability of correct response
1
.9
.8
.7
.6
−4
−2
0
2
4
neural tracking index sentence [z; within-subj]
β=.08, p=.01
predicted speed [s⁻¹]
.9
.8
.7
.6
.5
.4
−2
0
2
neural tracking index sentence [z; between-subj]


b
α-power dynamics
neural tracking
age
hearing loss
probed ear
spatial cue
accuracy
response speed