

Older adults' neural tracking of interrupted speech is a function of task difficulty

Ira Kurthen^{a,*}, Allison Christen^a, Martin Meyer^{b,c,d}, Nathalie Giroud^{e,f,g}

^a Department of Psychology, University of Zurich, Binzmuehlestrasse 14/21, Zurich 8050, Switzerland

^b Department of Comparative Language Science, University of Zurich, Switzerland

^c Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Switzerland

^d Cognitive Psychology Unit, University of Klagenfurt, Austria

^e Department of Computational Linguistics, Phonetics and Speech Sciences, University of Zurich, Switzerland

^f Competence Center for Language & Medicine, University of Zurich, Switzerland

^g Center for Neuroscience Zurich, University of Zurich, Switzerland

ARTICLE INFO

Keywords:

Age-related hearing loss
Neural speech tracking
Speech processing
Interrupted speech

ABSTRACT

Age-related hearing loss is a highly prevalent condition, which manifests at both the auditory periphery and the brain. It leads to degraded auditory input, which needs to be repaired in order to achieve understanding of spoken language. It is still unclear how older adults with this condition draw on their neural resources to optimally process speech. By presenting interrupted speech to 26 healthy older adults with normal-for-age audiograms, this study investigated neural tracking of degraded auditory input. The electroencephalograms of the participants were recorded while they first listened to and then verbally repeated sentences interrupted by silence in varying interruption rates. Speech tracking was measured by inter-trial phase coherence in response to the stimuli. In interruption rates that corresponded to the theta frequency band, speech tracking was highly specific to the interruption rate and positively related to the understanding of interrupted speech. These results suggest that older adults' brain activity optimizes through the tracking of stimulus characteristics, and that this tracking aids in processing an incomplete auditory stimulus. Further investigation of speech tracking as a candidate training mechanism to alleviate age-related hearing loss is thus encouraged.

1. Introduction

We live in a society which is getting older and older, a fact that makes the careful study of the competencies and vulnerabilities of older adults both relevant and necessary. The prevalence of age-related hearing loss (ARHL) is estimated at approximately 20% at age 60, 50% at age 70 and 70% to 80% at age 80 and older (Bisgaard and Ruf, 2017; Goman and Lin, 2016), which makes ARHL one of the most prevalent age-related conditions. One of the most disruptive consequences of ARHL is a detrimental effect on the understanding of spoken conversation, which hinders effective communication and can lead to social isolation (Mick et al., 2014; Weinstein and Ventry, 1982). The negative outcomes are many, with loneliness and social isolation mediating a negative relationship of hearing loss and cognitive decline (Maharani et al., 2019). Therefore, understanding the processes that lead to successful speech comprehension in older adults is key to helping them maintain their social relationships and cognitive stability.

Currently, the only evidence-based intervention for ARHL is the fitting of hearing aids. Although hearing aid use can ameliorate the lis-

tening situation and the quality of life of older adults tremendously (Johnson et al., 2018; Stark and Hickson, 2004), compliance is often poor because hearing aids are of limited use in difficult listening situations (Bertoli et al., 2009). These difficult listening situations are characterized by spectral or temporal distortions of the auditory signal, both of which can manifest either at the source or during the transmission of an auditory signal (Mattys et al., 2012). Examples include syllable deletion, segment elision, disfluency, an unfamiliar accent, time-compressed speech, speaking over the phone, concurrent speakers, background noise and reverberation. It has also been shown that, for temporal waveform distortions, increasing the sound level of the auditory signal (which is the main function of a hearing aid) is not an effective measure to maximize performance (Gordon-Salant and Fitzgibbons, 1993). Another common temporal distortion is the interruption of speech (Başkent et al., 2010; Bologna et al., 2018; Gordon-Salant and Fitzgibbons, 1993; Miller and Licklider, 1950; Profant et al., 2019; Saija et al., 2014; Wingfield et al., 1991), which can be the result of signal transmission errors from electronic devices, or which can manifest at the perceptual level as the result of poor signal transmission from the ear to the brain.

* Corresponding author.

E-mail address: ira.kurthen@uzh.ch (I. Kurthen).

<https://doi.org/10.1016/j.neuroimage.2022.119580>.

Received 16 June 2022; Received in revised form 14 August 2022; Accepted 18 August 2022

Available online 19 August 2022.

1053-8119/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Although interrupted or “gated” speech (i.e., speech which is interrupted by intervals of silence) results in incomplete speech segments, it can still be surprisingly easy to understand as long as a listener has normal hearing acuity. In certain circumstances, it is possible to understand nearly all of a speech signal of which 50% has been replaced by silence (Gilbert et al., 2007; Miller and Licklider, 1950). Wang and Humes (2010) investigated the influence of three parameters on interrupted speech understanding. These were on-duration (i.e., the absolute duration of the sound intervals), duty cycle (i.e., the relative duration of speech within each interruption cycle) and interruption rate (IR; the number of interruptions per second). They found that duty cycle primarily determined performance, but both IR and on-duration modified performance as well. Keeping the on-duration fixed, a higher IR results in a higher number of glimpses at the original signal. Understanding typically declines sharply when the IR becomes lower than the average syllabic rate of speech (~5 Hz; Bologna et al., 2018; Ding et al., 2017). Other factors that influence the understanding of interrupted speech are the presence of contextual (e.g., lexical or syntactic) information and the fundamental frequency of the speaker (Wang and Humes, 2010; Wingfield et al., 1991).

In their seminal study, Gordon-Salant and Fitzgibbons (1993) compared younger and older adults in interrupted speech understanding while including participants with clinically relevant peripheral hearing loss in each of the two age groups. They showed that both age and hearing loss influenced interrupted speech understanding. Similarly, in the study of Başkent et al. (2010), their data suggested that older, more hearing-impaired participants showed worse understanding of interrupted speech than younger participants with none or only a mild hearing loss, although this question was not statistically tested in the paper and the relation of age and hearing loss was not controlled for regarding this particular outcome variable. Shafiro et al. (2016) also found a difference in interrupted speech understanding between younger and older adults at IRs of 2 and 4 Hz. Taken together, it seems that younger and older adults differ on average in their ability to understand interrupted speech, and that hearing loss has an additional detrimental effect.

1.1. Cognitive repair mechanisms: domain-general?

If there is a difference in the percentage of remaining speech signal and understood speech signal, one must assume that some kind of repair of the missing input has taken place. The concept of repair mechanisms has been proposed by Başkent et al. (2010) and it is assumed that these repair mechanisms are somehow related to cognitive ability. Indeed, the positive influence of repair mechanisms associated with cognitive ability on speech understanding in adverse listening situations is a common research finding and is implied in several empirically grounded models of effortful listening (Arlinger et al., 2009; Pichora-Fuller et al., 2016; Rönnberg et al., 2013). In the Ease of Language Understanding (ELU) model (Rönnberg et al., 2013), the storage and processing aspects of working memory enable phonological, lexical, and semantic retrieval and pattern matching, which is important when listening to partially masked speech. The Framework for Understanding Effortful Listening (Pichora-Fuller et al., 2016) underscores the importance of attention and how it governs the allocation of cognitive capacity to cope with listening demand.

The most common test for the level of listening demand, or effortful listening, is the speech-in-noise paradigm. The paradigm of interrupted speech is less frequently employed and therefore the findings on the influence of cognition on understanding interrupted speech are few. In a study by Benard et al. (2014), receptive vocabulary predicted understanding of speech interrupted at a rate of 2.5 Hz in an age-diverse sample. Bologna et al. (2018) tested younger and older participants on understanding interrupted speech and additionally administered a cognitive test battery encompassing processing speed, working memory, inhibitory control, and visual linguistic closure, which is a measure of the ability to infer written language that is presented behind visual ob-

structions (see Zekveld et al., 2007). Only processing speed and visual linguistic closure predicted understanding of interrupted speech, both across age groups.

The paradigm of visual linguistic closure can be considered a visual analogue of interrupted speech. As such, the relationship between visual linguistic closure and understanding of interrupted speech is particularly interesting. The finding of such a relationship by Bologna et al. (2018) raises the question of whether cognitive repair mechanisms are rooted within the auditory domain or whether they draw upon a domain-general cognitive resource. Indeed, the ELU posits both a modality-specific and a modality-independent capacity (Rönnberg et al., 2013). In our study, we continued this line of thought and investigated whether there was a correlation between visual linguistic closure and interrupted speech understanding as well as between visual linguistic closure and a measurement of the neural tracking of interrupted speech. A correlation between tests of interrupted language understanding in the auditory and visual domains would point to a domain-general capacity, and finding that correlation again between visual linguistic closure and neural speech tracking would provide evidence for neural speech tracking being a candidate mechanism involved in the manifestation of such a domain-general capacity in the auditory domain.

1.2. Cortical tracking mechanisms during speech processing

Aging is associated with the loss of fine inner and outer hair cells in the cochlea. This loss impairs the ability to perceive and discriminate sounds. It is thus productive to focus research on central neural processes that support the processing of degraded auditory signals. A better understanding of how older adults' brain activity processes incomplete speech signals will inform better interventions for ARHL. In the current study, we aimed to investigate the brain processes that underlie older adults' understanding of interrupted sentences. Specifically, we investigated whether neural tracking of interrupted speech would take place and if yes, whether there was a statistical relationship between the difficulty of understanding the interrupted speech signal and the strength of neural speech tracking.

The notion of an alignment between the speech signal and the neural activity related to its processing is not new. The TEMPO model (Ghitza, 2011) as well as the later model by Giraud and Poeppel (2012) posit a mechanism by which intrinsic cortical oscillations become synchronized to the inherent, quasi-rhythmic properties of speech (Peelle et al., 2013). This mechanism has been called “entrainment”, and many following studies have been dedicated to elucidate its importance for speech processing. However, it is very difficult to provide actual evidence for an intrinsic oscillator, and many scientific results that have been framed in an entrainment context can be equally well explained by superposition of evoked responses (Alexandrou et al., 2020; Obleser and Kayser, 2019). For this reason, we refer to the alignment of the neural response to the speech signal in our study as “neural speech tracking”, which is a more general term and which does not assume an intrinsic oscillator as the underlying source of the observed neural activity.

At the heart of this alignment between speech and neural oscillations is the notion of different time scales. It is presumed that linguistic features operating on different time scales (e.g., prosody, phrase structure, syllables, phonemes) are tracked by functionally distinct components of brain activity (“frequency bands”), which share the temporal resolution of the corresponding linguistic units (Ding et al., 2016; 2017; Giraud and Poeppel, 2012; Keitel et al., 2018; Poeppel, 2003).

The reason why this mechanism should take place is that an alignment between neural frequency-related activity and these specific time scales in the speech signal optimizes the firing rate of neuronal assemblies in such a way that they are maximally excitable when there is an acoustic signal to process, and they recover when there is less acoustic signal to process (Giraud and Poeppel, 2012). The interrupted speech paradigm fits particularly well within this framework, as one can make

sure that there are periods in the speech signal which are ideally suited for neuronal recovery (i.e., the silent period).

Neural speech tracking is a robust phenomenon and it exhibits considerable inter-individual variability (Lam et al., 2018). Studies employing transcranial alternating current stimulation (tACS) have provided evidence that neural speech tracking is not simply an epiphenomenon of speech processing, but that it serves a causal role (Riecke et al., 2018; Wilsch et al., 2018; Zoefel et al., 2018). In these studies, an electric alternating current signal was applied to participants' brains while they performed speech-related tasks. The frequency of the alternating current carried information about the temporal envelope of the speech stimuli presented. In all three studies, participants' task performance significantly differed with regard to the phase angle of the alternating current that was used for stimulation. While these results are intriguing, the spatial resolution of tACS is limited (Yang et al., 2021) and it is unclear where exactly in the brain the stimulation takes effect (Obleser and Kayser, 2019).

A current question is whether neural speech tracking changes in adverse listening situations and, if yes, whether it changes in a qualitative or in a quantitative way. Haegens and Zion Golumbic (2018) conclude that neural speech tracking is primarily initiated by automatic, bottom-up processes, but that it can be modulated by top-down control (see also Lakatos et al., 2016; 2019; Park et al., 2015; Petersen et al., 2017). Because ARHL alters the encoding of the speech signal in the brainstem (Anderson et al., 2013; Bidelman et al., 2019; 2014), which then serves as the basis for acoustic analyses in the auditory cortex, it is important to investigate mechanisms that optimize processing of the auditory input. If neural speech tracking is indeed causally related to speech decoding, the most obvious mechanism to enhance decoding would be to increase speech tracking. Research shows that the relationship is not that simple. Petersen et al. (2017) found that better hearing thresholds were not related to increased speech tracking, but to *decreased* tracking of a competing speech signal, which had to be ignored. A recent study by Presacco et al. (2019) found that hearing loss modulated neither speech tracking at the midbrain nor at the cortical level, but instead modulated the reciprocal connections between lower and higher levels of the auditory pathway.

These results, however, were all obtained using a speech-in-noise paradigm. As an alternative, interrupted speech is an interesting test case for speech tracking for several reasons. An adverse listening condition may not always arise from noise continuously masking speech, but also from noise interrupting speech at intervals, like during an unstable telephone connection. Interrupted speech rather than speech-in-noise is a better model of such a situation. Additionally, whenever the speech signal is resumed after a period of silence, sharp edges in the signal are created. These sharp edges are hypothesized to trigger phase resets of ongoing neural oscillations (Doelling et al., 2014; Giraud and Poeppel, 2012; Lakatos et al., 2019), and thereby induce an "entrainment" of said ongoing neural oscillations to the speech signal. It is possible that the more sharp edges an interrupted speech signal contains, a stronger correspondence between neural activity and the speech signal is observed, which might explain the high rates of understanding in conditions with a high IR (Gilbert et al., 2007).

1.3. Study design and hypotheses

Building upon previous research, our study aimed to investigate neural tracking of interrupted speech in healthy older adults. Specifically, we investigated whether neural speech tracking would occur in the frequencies corresponding to the IRs of the interrupted speech signal.¹ Further, we investigated neural speech tracking as a function of hearing and cognitive ability. During our experiment, older adults listened to sen-

tences interrupted with silent periods in different IRs. We aimed to minimize the distortions of the auditory signal caused by peripheral hearing impairment by sampling older adults who would be considered normal hearing or having only very mild hearing loss in clinical practice. This ensured that auditory deprivation—as much as possible—did not confound the neural processes under study. Because past research showed that of the parameters determining interrupted speech, the duty cycle was the most important parameter predicting interrupted speech understanding (Wang and Humes, 2010), we primarily manipulated the duty cycle to obtain different stimulus conditions. When faced with the decision to either keep on-duration or the length of the silent interval fixed, we decided to fix the length of the silent interval in order to have better control over how much signal would be missing. In three of the four conditions, the silent interval that masked the sentences was 100 ms long, which is approximately the median length of a syllable (Greenberg et al., 2003). We implemented a fourth experimental condition with a silent interval length of 50 ms. This condition has the same duty cycle (0.5) as one of the other three conditions, but it is easier due to the shorter silent interval length. This fourth condition therefore allows to investigate the relative influence of silent interval length on neural speech tracking and on interrupted speech understanding.

We expected that an increase in neural speech tracking would be related to better understanding of the interrupted sentences. Each of our experimental conditions exhibited a different IR, and therefore a singular frequency in which speech tracking would ideally take place in order to maximize excitability of the underlying neuronal assemblies (Giraud and Poeppel, 2012). Also, the interrupted speech paradigm is akin to an auditory frequency-tagging paradigm (e.g., Bharadwaj et al., 2014; Buiatti et al., 2009; Weisz and Lithari, 2017), with which an auditory steady-state response can be elicited (Picton et al., 2003). For this reason, one would expect that in a certain condition, we observe a neural response of the auditory cortex which is particularly strong in the IR of that condition.

For each condition, Fig. 1 illustrates the relationships between the IR frequency, the IR period, and the lengths of the silent and sound intervals by plotting these parameters onto an example waveform from each condition.

Since, in all sentences in each condition, the silent intervals always start and end at the same points in time relative to sentence onset, one can measure speech tracking in the specific IR by means of inter-trial phase coherence (ITPC; Delorme and Makeig, 2004), which is also known as phase-locking factor (Jervis et al., 1983; Tallon-Baudry et al., 1996). This method extracts the phase in each trial in specified frequencies for specified time points and quantifies their similarity across trials. ITPC can range between 0 and 1. For a concise, non-technical description of ITPC, see Roach and Mathalon (2008). Another possible measure to quantify the neural response to a stimulus in a certain frequency is spectral power (e.g., Ding et al., 2016). In our case, however, power would not constitute the best measure for neural tracking. Power reflects both the number of neurons involved in stimulus processing as well as their temporal synchronization (Werkle-Bergner et al., 2009). Therefore, it is reasonable to choose a measure which reflects only temporal synchronization, i.e., phase (Lachaux et al., 1999). Also, an integral part of our study is the investigation of inter-individual differences in older adults. It has been shown that cortical surface area is related to the strength of ERP amplitudes in healthy older adults (Giroud et al., 2019). Thus, it is reasonable to assume that inter-individual differences in brain structure due to age-related cortical atrophy would also confound a measurement of spectral power on the scalp. Therefore, ITPC as a measure of temporal synchronization irrespective of signal strength is a more appropriate measure of neural tracking than spectral power in the current study.

Additionally, we aimed to investigate whether there would be modulation of speech tracking by a domain-general cognitive ability related to the repair of missing sensory input. We hypothesized that performance in interrupted sentence understanding would correlate between

¹ In the remainder of the text, we will refer to this phenomenon as "neural tracking in the IRs".

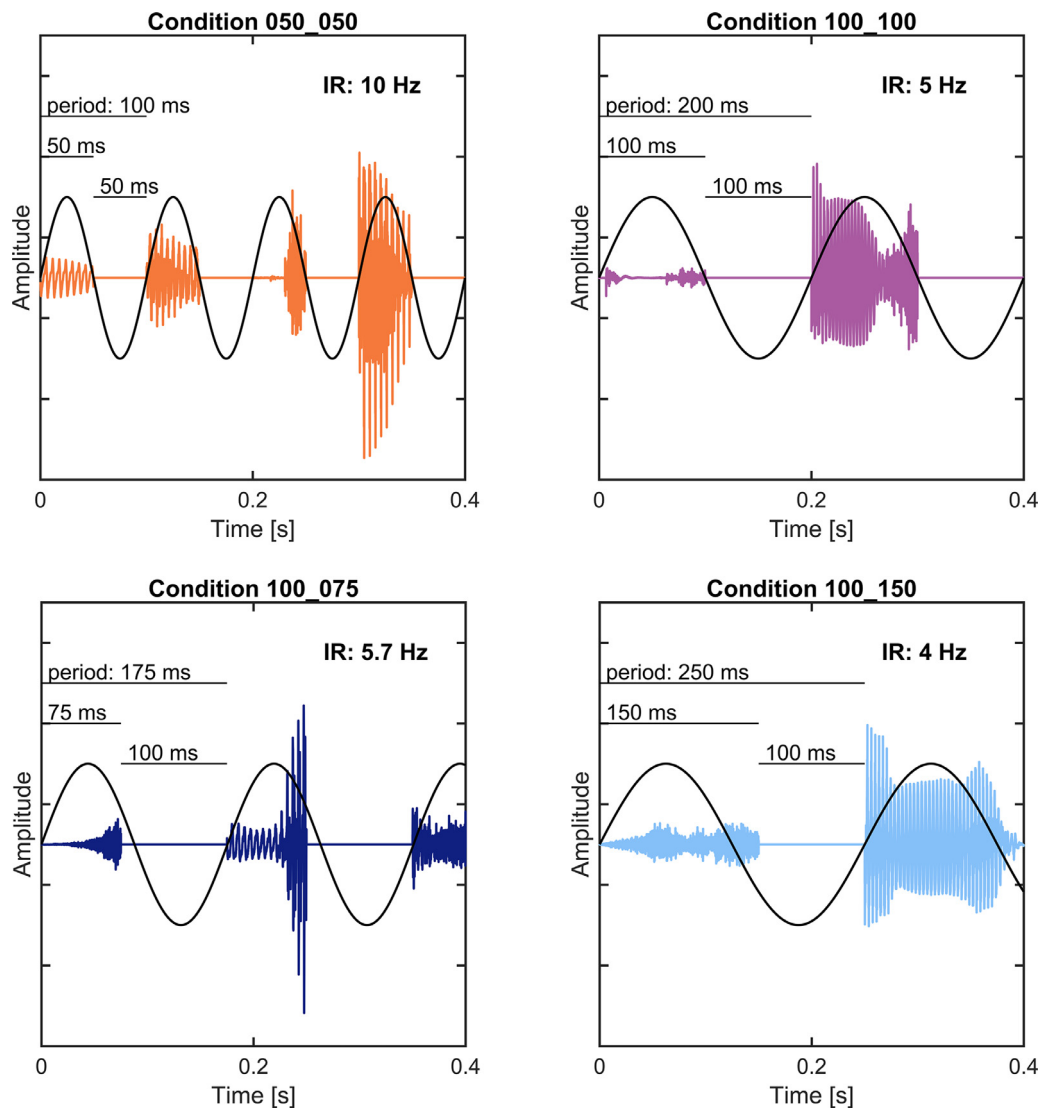


Fig. 1. Stimulus examples. *Note.* This figure illustrates the four experimental conditions. Each condition is defined by the lengths of their sound segments and silent intervals. These two lengths add up to the period of the IR. The IR frequency is further illustrated by plotting a corresponding oscillation over the example waveform of each condition. *Note* that the y axes do not contain tick labels because they represent arbitrary numbers between 0 and 1.

the visual and auditory domains. Based on the assumption that this correlation would indicate a shared cognitive resource, we hypothesized that higher ITPC would also be associated with better interrupted sentence understanding in the visual domain. In this case, ITPC would reflect recruitment of a domain-general cognitive resource rather than an exclusively auditory mechanism. Previous studies of interrupted speech understanding have shown considerable inter-individual differences in the percentage of understood words, although the interruption parameters were the same (Bologna et al., 2018; Wang and Humes, 2010). One can therefore consider the percentage of understood words as an indicator of the individual difficulty of the task for each participant. We hypothesized that task performance would be positively related to ITPC because it reflects how well a participant can engage repair mechanisms.

2. Materials and methods

2.1. Participants

The sample consisted of 26 older adults (mean age = 69.5 yrs, $SD = 3.8$ yrs, range 64–75 yrs, 13 females). One additional participant was tested but excluded from the final analysis because of floor behav-

ioral scores. All participants were right-handed as assessed by the Annett Hand Preference Questionnaire (Annett, 1970) and reported no past or present psychiatric or neurological disorders. Their native language was Swiss German and none had learnt another language before the age of seven. None wore a hearing aid nor reported having tinnitus. Their average hearing thresholds in the frequencies 0.5, 1, 2, and 4 kHz did not exceed 30 dB HL. They also stated that their hearing ability did not differ between the two ears. To only include participants not affected by mild cognitive impairment or dementia, participants were asked to completed the Montreal Cognitive Assessment (Nasreddine et al., 2005) and were only invited to participate in the study if they had scored 26 points or more. The ethics committee of the Canton of Zurich approved the study (application no. 2017-00284). Written informed consent was obtained from all participants. Participants were compensated for their participation.

2.2. Hearing tests

The computer-based hearing tests were administered via a custom MATLAB software built on the MAP auditory toolbox (Meddis et al., 2013). We measured absolute pure-tone hearing thresholds in the fre-

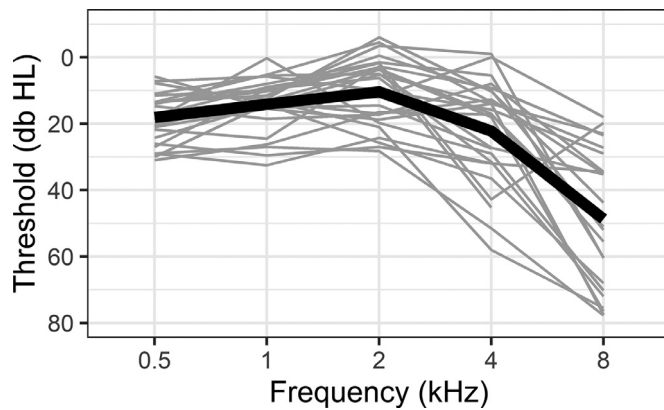


Fig. 2. Hearing thresholds. *Note.* Hearing thresholds for each participant (gray lines) and the group mean (black line) at all tested frequencies. In three participants, the 8 kHz tone was not audible at the maximum level of presentation (80 dB HL).

quencies 0.5, 1, 2, 4, and 8 kHz. The measurement procedure and the stimuli have already been described in detail elsewhere (Giroud et al., 2018; Lecluyse and Meddis, 2009; Lecluyse et al., 2013). Participants' pure-tone hearing thresholds are visualized in Fig. 2. The average absolute hearing threshold (i.e., pure-tone average; PTA) for each participant was calculated by averaging the thresholds for 0.5, 1, 2, and 4 kHz. Stimuli were controlled via sound card (RME Babyface Pro, RME, Haimhausen, Germany) and presented via loudspeaker with linear frequency response (8030B Studio Monitor, Genelec, Iisalmi, Finland). We used a head supporting device (SR Research Head Support, SR Research, Ottawa, Ontario, Canada) to ensure that each participant sat at the same distance from the loudspeaker. The distance between the center of the loudspeaker and participants' ears was 65 cm. The stimulus level was calibrated with a sound level meter (XL2, NTi Audio, Schaan, Liechtenstein).

2.3. Visual linguistic closure

To assess visual linguistic closure, we used the Text Reception Threshold Test (TRT; Zekveld et al., 2007). This test was developed and has been used as a visual equivalent to the Speech Reception Threshold test (SRT). In our view, the TRT shows at least as many similarities with the interrupted speech paradigm as with the SRT and, indeed, the TRT has already been used as a visual analogue of interrupted speech (Krull et al., 2013). Both the TRT and the interrupted speech paradigm mask only a certain proportion of the whole language stimulus (the TRT with black bars, the interrupted speech paradigm with silence), while the SRT typically masks the whole stimulus. We administered the TRT in the TRT_{WORD} version (Besser et al., 2012) because the original TRT presents all the words of the visual stimuli sentence at the same time. Due to the inherently temporal nature of acoustics, only one chunk can be presented at a specific point in time. We therefore judged the TRT_{WORD} version, in which each word is presented and then disappears before the next word is presented, as better corresponding to the interrupted speech paradigm.

During testing, participants sat in the EEG cabin with a microphone in front of them. In total, 30 sentences (10 practice trials, 20 test trials) taken from test list number 4 of the German matrix sentence test *Oldenburger Satztest* (OLSA; Wagener et al., 1999) were presented on the computer screen, masked by black bars. The practice trials served two purposes: We ensured that the participants understood the test procedure, and we also ensured that they had adequate vision for the test. For a more detailed description of the TRT_{WORD} , see Besser et al. (2012).

Table 1
Experimental conditions .

Silent interval length	Duty cycle		
	$\frac{3}{7}$	0.5	0.6
100 ms	100 ms : 75 ms (IR 5.7 Hz)	100 ms : 100 ms (IR 5 Hz)	100 ms : 150 ms (IR 4 Hz)
50 ms		50 ms : 50 ms (IR 10 Hz)	

Note. This table shows the four experimental conditions. In each cell, the lengths of the silent intervals and the sound segments are contrasted with the colon. The sum of these two values represents the period of the respective IR, which is noted in the parentheses below in Hertz.

2.4. Stimuli for EEG experiment

The stimuli for the EEG experiment consisted of spoken sentence material (Swiss German) with silent intervals inserted. In total, the main EEG experiment required 123 sentence stimuli. The sentences were recorded by a female native speaker of Swiss German and were normalized to 70 dB SPL (mean F_0 = 226.81 Hz, SD = 11.69 Hz; mean duration = 3.9 s, SD = 0.35 s). The sentences contained topics related to everyday life in Switzerland. The 123 sentences were assigned to four experimental conditions (30 sentences per condition) and a practice condition (3 sentences).

To create the stimuli for the conditions, the audio signal of the sentences was partially set to zero. This manipulation was performed after normalization to 70 dB SPL so that the audible segments in each sentence would be presented with the same intensity on average. Our conditions followed an incomplete 2 (Length of silent interval) $\times$ 3 (Duty cycle) within-subjects design. Please see Table 1 for an overview of the experimental conditions and Fig. 1 for a visualization of the stimuli and the corresponding IRs. In three of the four conditions, the silent interval that masked the sentences was 100 ms long. The duty cycle took values of $\frac{3}{7}$, 0.5, and 0.6, which resulted in non-masked speech segments of 75 ms, 100 ms, and 150 ms, respectively. Therefore, the sentences in these three conditions were interrupted at rates of 4 Hz, 5 Hz, and 5.7 Hz. The fourth condition had a silent interval of 50 ms and a duty cycle of 0.5, which resulted in non-masked speech segments of 50 ms and an IR of 10 Hz.

In the following, the four conditions will be referred to as “050_050”, “100_100”, “100_075”, and “100_150”. These numbers refer to the lengths of the silent intervals and non-masked speech segments in each condition. Stimuli were presented block-wise.

2.5. EEG Experiment: sentence repetition task

For the EEG experiment, participants were seated in an EEG cabin in front of a computer screen and a microphone. After a short instruction in which they could see the effects of their movement on the EEG signal, they were asked to keep as still as possible in the experiment. Then, the procedure for the main EEG task was explained to them. Their task was to repeat as much as possible of the whole sentence heard into the microphone, without gaps. They completed a short practice block and then four experimental blocks in total, each of which required approximately 12 min. After the practice session, participants were actively encouraged to adjust the presentation volume of the stimuli (i.e., louder or softer) to ensure optimal audibility for that individual. If participants refused initially, we asked them a second time to make sure they were not just polite. If they refused a second time, we kept the volume at 70 dB SPL. If they asked for an increase, we increased the volume in steps of 1 dB SPL. When participants were comfortable with the volume, it was kept at the resulting sound level for the rest of the experiment. We recorded the dB increases for each participant. Except for the first trial of each block, the experimenter controlled the start of the trials. Participants' recorded responses were scored according to the number of words they

correctly repeated. In the case of compound nouns, each compound was scored individually. No feedback was provided during the experiment. Participants' answers were aggregated over each block by calculating the ratio between the number of correctly repeated words and the total number of words. Consequently, each participant's score in each block of the sentence repetition task could range between 0 and 1. This score was used as a measure for individual task difficulty. After each block, participants had a 1-min break. Block order was counterbalanced between participants, but stimulus order within the blocks was fixed to facilitate the evaluation of participants' responses.

2.6. EEG recording and analysis

Participants' EEG was recorded continuously from 128 Ag/AgCl electrodes (BioSemi ActiveTwo, Amsterdam, The Netherlands) with an ActiveTwo AD-box amplifier system (BioSemi ActiveTwo, Amsterdam, The Netherlands) and was digitized at a sampling rate of 512 Hz. Data were online band-pass filtered between 0.1 and 100 Hz and impedances were reduced to below 25 k Ω . Data were analyzed in MATLAB Release 2016b (The MathWorks, Inc., Natick, Massachusetts, United States) using the FieldTrip Toolbox (Oostenveld et al., 2011). For pre-processing, data were re-referenced to Cz and then band-pass filtered between 0.1 and 40 Hz with a non-causal zero-phase two-pass 4th (8th) order Butterworth IIR filter with -6 (-12) dB half-amplitude cutoff. A non-causal zero-phase two-pass 4th (8th) order Butterworth IIR band-stop filter with -6 (-12) dB half-amplitude cutoff was applied between 48 and 52 Hz in order to eliminate artifacts resulting from electric interference. Then, data were visually screened for noisy channels, which were then removed. To follow, the continuous EEG was segmented into trials starting 2 s before sentence onset and lasting until the end of the sentence (mean trial duration = 5.9 s, $SD = 0.35$ s). Trials containing gross artifacts (i.e., containing large, non-systematic spikes in the EEG related to muscular activity as a result of movement or coughing) were removed manually. The data were then re-referenced to an average reference and an independent component analysis (ICA) was applied to remove eye movements and eye blinks (Jung et al., 2000). After this, the noisy channels were interpolated using spline interpolation (Perrin et al., 1987).

2.7. ITPC

To extract ITPC, a time-frequency analysis was conducted using Morlet wavelets with 7 cycles in frequencies from 3 to 18 Hz and in a time window between -2 and 3 s, with a step size of 10/512 ms. In order to achieve a frequency resolution of 0.1 Hz (which was necessary to obtain phase values for one of our IR frequencies, 5.7 Hz), data were zero-padded until 10 s. The outputs of the wavelet analysis were complex Fourier-spectra.

From these complex Fourier-spectra, ITPC was calculated using the following formula (adapted from Delorme and Makeig, 2004):

$$ITPC_{f,t,c} = \frac{1}{n} \sum_{k=1}^n \frac{F_{k,f,t,c}}{|F_{k,f,t,c}|}$$

in which n is the number of trials and $F_{k,f,t,c}$ is the complex Fourier-spectrum of trial k at frequency f and time t at channel c .

For each channel, ITPC values in each of the four interruption rates were averaged across the time window between 1 and 2 s after stimulus onset and were afterwards exported to a.csv file for further statistical analysis. Fig. 3 shows the topographies of ITPC values during this time window (upper row) and the time-frequency representation of ITPC in each condition (lower row). The lower bound of this time window of analysis was chosen in order to exclude phase-locking due to the N1-P2 complex after stimulus onset (the yellow blobs at stimulus onset visible in Fig. 3). Because FieldTrip clips the complex Fourier-spectra at half the length of the Gaussian taper of the wavelet, our maximum window of analysis would have lasted until 2.127 s after stimulus onset (i.e., the last time point for which the complex Fourier-spectrum was computed

at 4 Hz). For reasons of simplicity and readability, we chose to restrict the analysis to 2 s after stimulus onset.

Furthermore, in order to reduce the levels of the channel dimension of the EEG data while still remaining free of assumptions regarding the topography of our effects to avoid 'double dipping' (Kriegeskorte et al., 2009), channels were grouped together so that they formed nine clusters (left anterior, left central, left posterior, medial anterior, medial central, medial posterior, right anterior, right central, and right posterior). Cluster sizes were chosen so that all clusters would contain approximately the same number of electrodes. ITPC for each cluster was calculated by averaging across all channels of each cluster. This cluster-based LMEM analysis was favored instead of cluster-based permutation tests because it requires far less compute resources while still retaining the notion of no preconceptions about regions with significant effects.

2.8. Statistical analysis

Statistical tests were conducted in R, Version 3.6.2 (R Core Team, 2018) and FieldTrip, Version 20190419 (Oostenveld et al., 2011). We derived the p values for estimates in linear mixed-effects models (LMEM) using the Satterthwaite method implemented in the R package lmerTest (Kuznetsova et al., 2017).

3. Results

3.1. Behavioral results

We first investigated the behavioral results of our interrupted speech experiment. A LMEM was run to analyze differences in sentence repetition scores between the four experimental conditions. The model contained a fixed effect for condition and a random intercept per participant. As condition was a dummy-coded categorical variable with the 050_050 condition as the reference category, this resulted in three categorical predictors. All three condition predictors were significant (all $p < 0.001$). Post-hoc Tukey t -tests were conducted with the glht function of the multcomp package (Hothorn et al., 2008). All pairwise comparisons were significant (all $p < 0.001$). Please see Fig. 4 for a visualization of sentence repetition scores. Participants scored highest in the 050_050 condition ($M = 0.811$, $SD = 0.088$), followed by the 100_150 condition ($M = 0.681$, $SD = 0.121$), the 100_100 condition ($M = 0.461$, $SD = 0.137$), and the 100_075 condition ($M = 0.34$, $SD = 0.133$). The significant difference between conditions 050_050 and 100_150 shows that with regard to these specific conditions, the length of the silent interval outweighs the duty cycle. However, taking into account only the duty cycle with a fixed silent interval length, a monotonic pattern of increasing understanding with increasing amount of signal is observed.

3.2. Predicting sentence repetition scores from participant variables

As a next step, we investigated whether our participant-level variables (age, PTA, and TRT) would predict sentence repetition scores. To achieve this, we updated the previous LMEM, which predicted sentence repetition scores from condition, to include age, PTA, and TRT. Of the three participant-level variables, only PTA was predictive of sentence repetition scores ($b = -0.008$, $t(22) = -3.15$, $p = 0.005$). Because participants were encouraged to amplify or attenuate stimulus presentation, and because they had been able to work on the practice trials, we ensured that all participants were able to perceive the stimuli. Participants' final sound level after amplification and PTA were correlated (Pearson's $r = 0.493$, $p = 0.01$). We therefore assume that the participant-led volume increases were performed in accordance with their hearing ability. However, it is possible that some participants did not amplify the stimuli sufficiently. In this case, PTA would be predictive of sentence repetition scores not because of a genuine relationship between peripheral hearing ability and task performance, but simply because of audibility issues. As control variable for this possible confounding factor, we

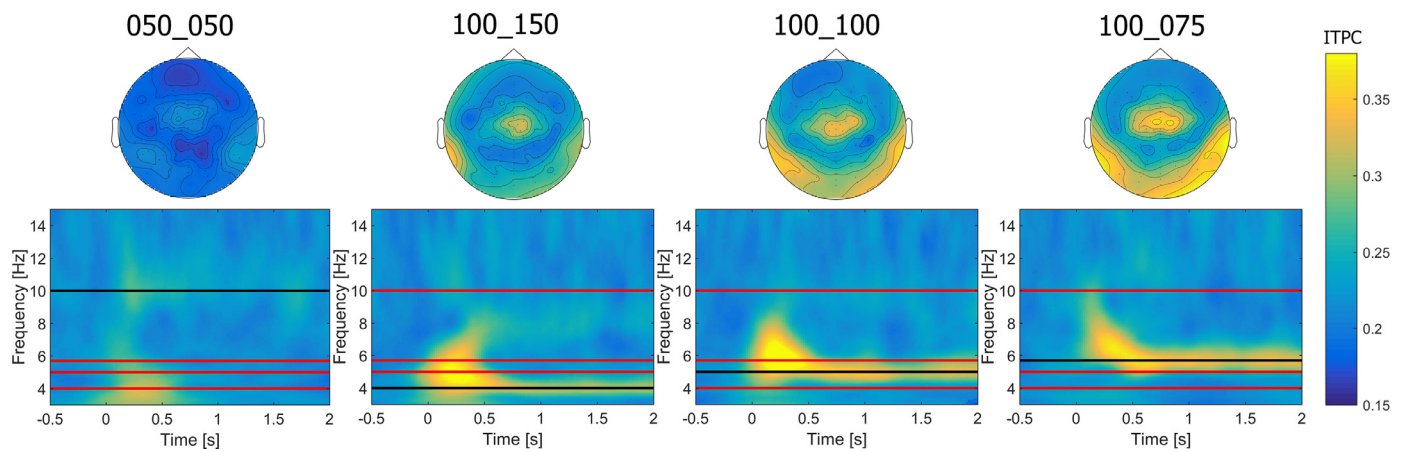


Fig. 3. ITPC topographies and time-frequency representations. *Note.* This figure shows topographies and time-frequency representations of ITPC for all four conditions. Upper row: Topographies of ITPC for the four conditions averaged across a time window of 1 and 2 s after sentence onset. Lower row: Time-frequency representations of ITPC in the medial central electrode cluster for the four conditions. The four IR frequencies are marked with horizontal lines. For each condition, the corresponding IR frequency is colored black, while the IR frequencies of the other conditions are colored red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

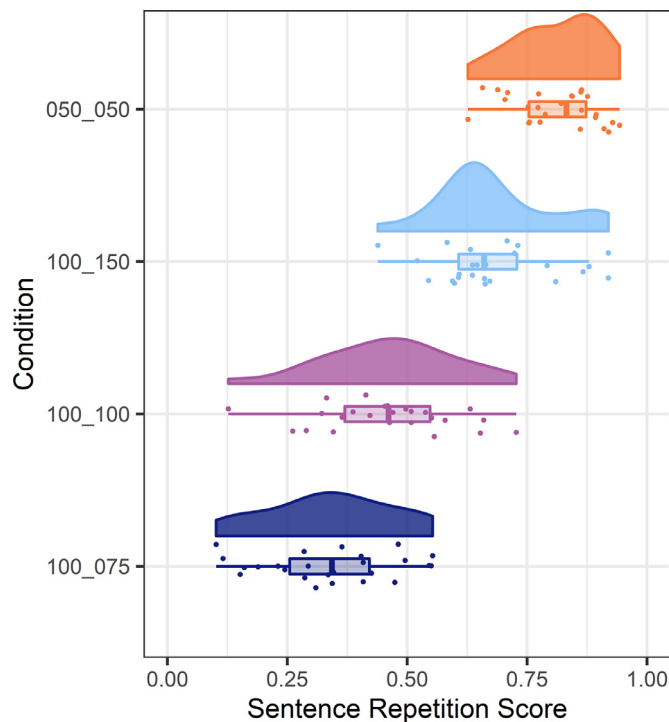


Fig. 4. Behavioral results. *Note.* This figure shows the distributions of sentence repetition scores for each condition. Single dots show individual data points, the box-and-whiskers plots show the median (thick line within the box), the inter-quartile range (IQR; the length of the box), the first (Q1) and the third (Q3) quartiles (left and right bounds of the box), $Q1 - 1.5 \cdot IQR$ (end of left whisker), and $Q3 + 1.5 \cdot IQR$ (end of right whisker). The distribution visualizations above show the kernel density estimation. Participants scored highest in the 050_050 condition, followed by the 100_150 condition, the 100_100 condition, and the 100_075 condition.

calculated residual attenuation (rAtt) by extracting and standardizing the residuals of the regression of attenuation on PTA. We then fitted a LMEM with condition, PTA and rAtt as predictors for sentence repetition scores. Even with this additional control variable, PTA remained a significant predictor ($b = -0.008$, $t(23) = -3.434$, $p = 0.002$) for sentence repetition scores. This finding suggests that raised peripheral hearing

thresholds have an effect on speech understanding which cannot simply be reversed via amplification. The relationship between condition, sentence repetition scores, and PTA is visualized in Figure S1.

3.3. ITPC by condition

We then investigated whether the stimuli elicited speech tracking in their specific IR frequency. Four data sets were created, one for each IR frequency (4, 5, 5.7, and 10 Hz), containing ITPC values in all conditions at that frequency. For each data set, ITPC values were included in a LMEM with condition, cluster, and trial count as fixed effects and with a random intercept for participant. For each model, the reference category of that predictor condition was chosen to match the respective IR frequency, so that ITPC in that frequency would be compared to the remaining conditions (e.g., for the 5 Hz data set, the reference condition was 100_100). We thus expected negative values for the condition estimates as these would indicate less ITPC in the conditions for which this frequency was not the IR frequency. Trial count per condition and participant was added to the analyses as a control variable because ITPC is directly related to the number of trials that are used to calculate it (Vinck et al., 2010). All clusters were included in the analysis in order to obtain an assumption-free estimate of how the effect of condition was distributed across the scalp, and to identify loci of especially strong ITPC on the topography without relying on visual inspection. The factor cluster was encoded via sum coding. For concision, we only present main effects here. A significant negative main effect of any condition points to significantly lower ITPC in that condition compared to the reference condition. Table S1 shows full model estimates including cluster interactions.

For the IR frequency of 4 Hz, ITPC was lower in all three tested conditions relative to the reference condition, 100_150 (050_050: $b = -0.06$, $t(875.96) = -17.48$, $p < 0.001$; 100_075: $b = -0.06$, $t(879.25) = -17.59$, $p < 0.001$; 100_100: $b = -0.06$, $t(874.41) = -17.860$, $p < 0.001$). For the IR frequency of 5 Hz, ITPC was lower in all three tested conditions relative to the reference condition, 100_100 (050_050: $b = -0.08$, $t(876.23) = -22.21$, $p < 0.001$; 100_075: $b = -0.04$, $t(879.36) = -10.88$, $p < 0.001$; 100_150: $b = -0.07$, $t(874.28) = 20.01$, $p < 0.001$). For the IR frequency of 5.7 Hz, ITPC was lower in all three tested conditions relative to the reference condition, 100_075 (050_050: $b = -0.09$, $t(875.03) = -25.92$, $p < 0.001$; 100_100: $b = -0.05$, $t(879.29) = -13.36$, $p < 0.001$; 100_150: $b = -0.09$, $t(878.29) = -25.59$, $p < 0.001$). For the IR frequency of 10 Hz, ITPC was lower in the 100_075 and in the 100_150 conditions relative to the reference condition, 050_050 (100_075: $b = -0.01$, $t(874.49) =$

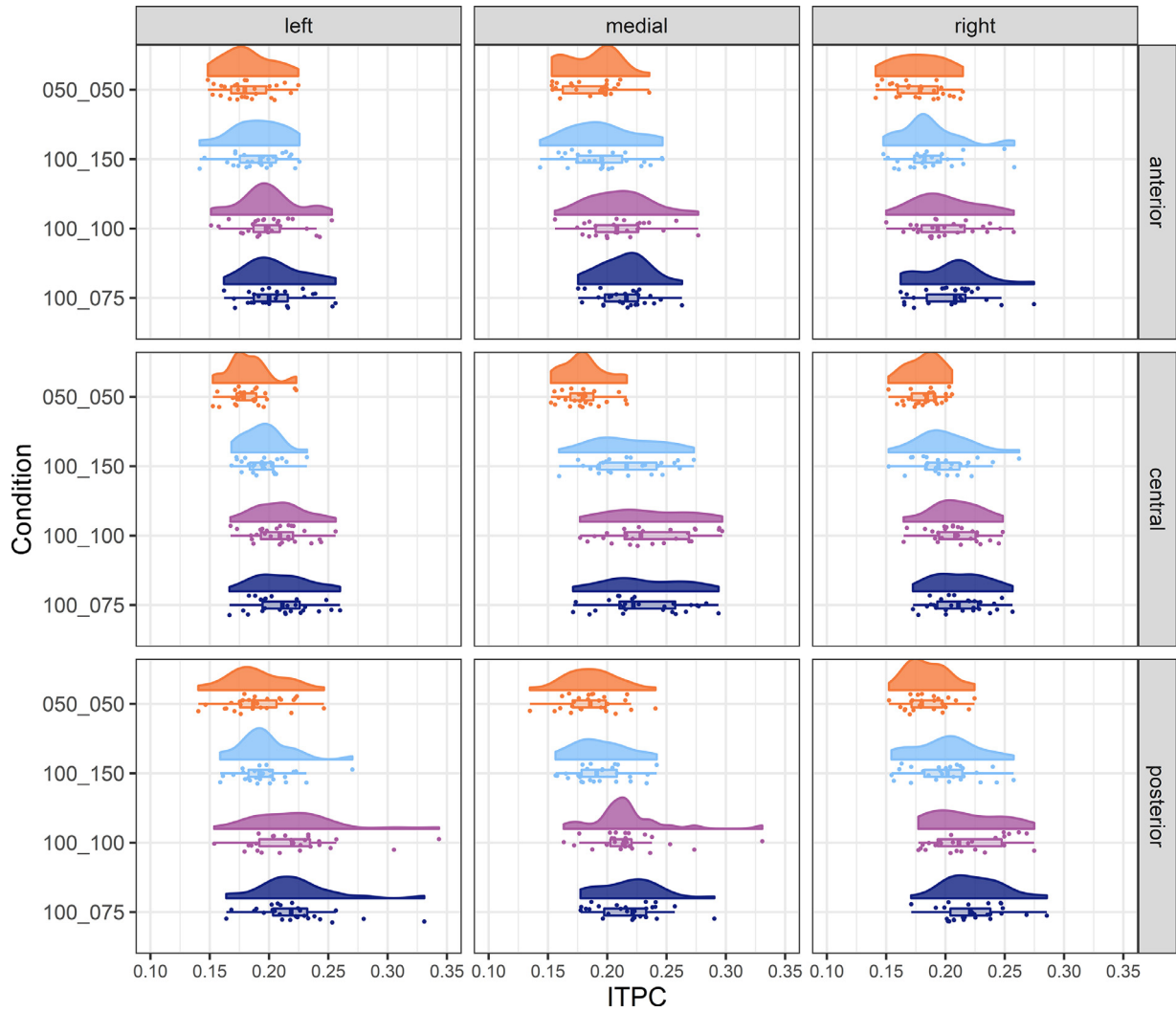


Fig. 5. Distribution of ITPC per cluster and condition. *Note.* This figure shows the distributions of ITPC for each condition in each cluster averaged across IR frequencies.

-4.40 , $p < 0.001$; 100_100 : $b = 0.00$, $t(876.06) = 1.23$, $p = 0.22$; 100_150 : $b = -0.01$, $t(875.20) = -3.33$, $p < 0.001$). This shows that in the three conditions with a silent interval length of 100 ms, ITPC was significantly higher in the IR corresponding to that specific condition than in all other tested frequencies. In the 050_050 condition, this mostly holds true, albeit with estimates that are much smaller than for the other comparisons. Therefore, we show that speech tracking is highly specific to the IR of the stimulus. Fig. 5 illustrates the distributions of ITPC in the four IR frequencies per condition at each cluster. Fig. 6 shows the distributions of each measured ITPC frequency averaged across clusters, dependent on whether the ITPC frequency corresponded to the condition's interruption rate.

3.4. Inter-individual variation in ITPC

We further aimed to explain inter-individual variation in ITPC. We were interested in whether age, peripheral hearing, and visual linguistic closure would predict the amount of ITPC in the IR frequencies. Furthermore, we aimed to include a measure of how challenged participants were by a specific condition. To this end, we used participants' sentence repetition score in each condition as a proxy for individual task difficulty. To explain inter-individual variation in ITPC, we fitted a LMEM with ITPC as criterion and all four participant variables (age, PTA, TRT, individual task difficulty) as predictors in ad-

dition to the control variables (trial count and residual attenuation), and a random intercept per participant. Because individual task difficulty strongly depended on the experimental condition, we also included main effects of condition and interaction effects of condition with individual task difficulty to the model. After initial model fitting, the assumption of normally distributed errors was not met. We therefore log-transformed ITPC, which resulted in normally distributed errors. Of the four participant-related predictors, only individual task difficulty ($b = -0.62$, $t(595.07) = -2.96$, $p = 0.003$) significantly predicted ITPC. Because the factor "condition" was dummy-coded, with the reference condition being the 050_050 condition, this negative coefficient indicates that in the condition with an IR of 10 Hz, lower individual task difficulty (= better performance) resulted in less ITPC. Additionally, all interaction effects between individual task difficulty and condition were significant (condition 100_075: $b = 0.379$, $t(922.37) = 1.98$, $p = 0.048$; condition 100_100: $b = 0.593$, $t(922.99) = 3.17$, $p = 0.002$; condition 100_150: $b = 0.525$, $t(921.49) = 2.74$, $p = 0.006$). These positive interaction coefficients reveal a positive relationship between sentence repetition scores and ITPC in these conditions. Higher sentence repetition scores, which represent lower individual task difficulty, were related to more ITPC. To better illustrate this finding, the relationship between individual task difficulty and ITPC in each condition at each electrode cluster is visualized in Fig. 7. Also, the two control variables trial count ($b = -0.02$, $t(799.98) = -4.93$, $p < 0.001$) and residual atten-

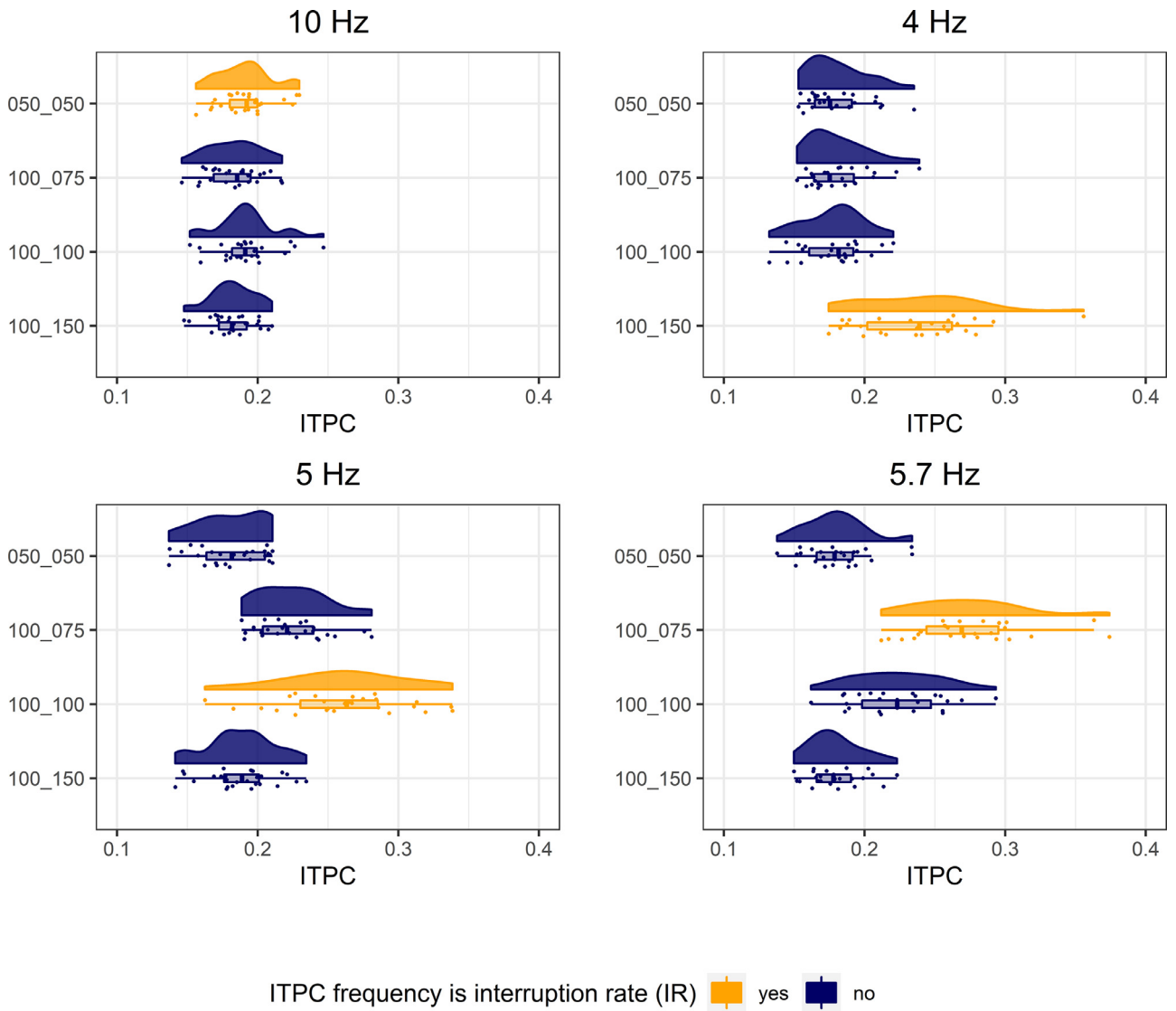


Fig. 6. ITPC in the IRs of all conditions. *Note.* This figure shows the distributions of each measured ITPC frequency averaged across clusters, dependent on whether the ITPC frequency corresponded to the condition's interruption rate.

uation ($b = 0.054$, $t(21.32) = 2.35$, $p = 0.03$) significantly predicted ITPC. This result confirms the importance of controlling for the number of trials that were used to calculate ITPC, and it also suggests that even small changes in audibility of the acoustic stimuli influences speech tracking.

There is, however, an alternative explanation for the significant effect of individual task difficulty. For the three conditions with a silent period of 100 ms, a higher difficulty goes hand in hand with a higher number of sound onsets in the acoustic signal. Since speech tracking is thought to be elicited by the presence of sharp edges in the acoustic signal, it is possible that not difficulty, but rather the number of sound onsets in the stimuli is reflected in higher ITPC. Therefore, an alternative explanation of the association between difficulty and ITPC could be that ITPC was solely elicited by the summation of cortical evoked potentials (P1-N1-P2), which were phase-locked to the onset of the sentence as well as to the onsets of every single sound snippet after each silent period. This issue cannot be resolved by simply controlling for the number of sound onsets because there is no variation in the number of sound onsets between the conditions. Thus, entering both variables into the model would result in a rank-deficient model matrix. We therefore chose to compare models with either difficulty or number of

sound onsets as fixed effects of interest but, because the two models were not nested, it was not possible to conduct a likelihood ratio test to test for a significantly better fit of one model compared to the other. However, it was still possible to compare the two models by means of the Akaike Information Criterion (AIC). We decided on AIC as comparison criterion because it is the most appropriate information criterion for comparing models when the “true model” is not part of the model ensemble (Vrieze, 2012). The models fitted for comparison were a LMEM predicting log-transformed ITPC with individual task difficulty and a random intercept per participant and one predicting log-transformed ITPC with the number of sound onsets and a random intercept per participant. Both models were fitted using Maximum Likelihood variance estimation. The model with individual task difficulty as predictor for ITPC exhibited a lower AIC (-257.16) than the model with number of sound onsets ($\text{AIC} = -180.38$). As a rule of thumb, an absolute difference of 10 between two models strongly favors the model with the smaller AIC (Posada and Buckley, 2004). In our case, the absolute difference between the two models' AICs was 76.78, thereby strongly favoring the difficulty model. We therefore concluded that ITPC was not higher in the more difficult conditions simply because of the co-occurrence of a higher number of sound onsets.

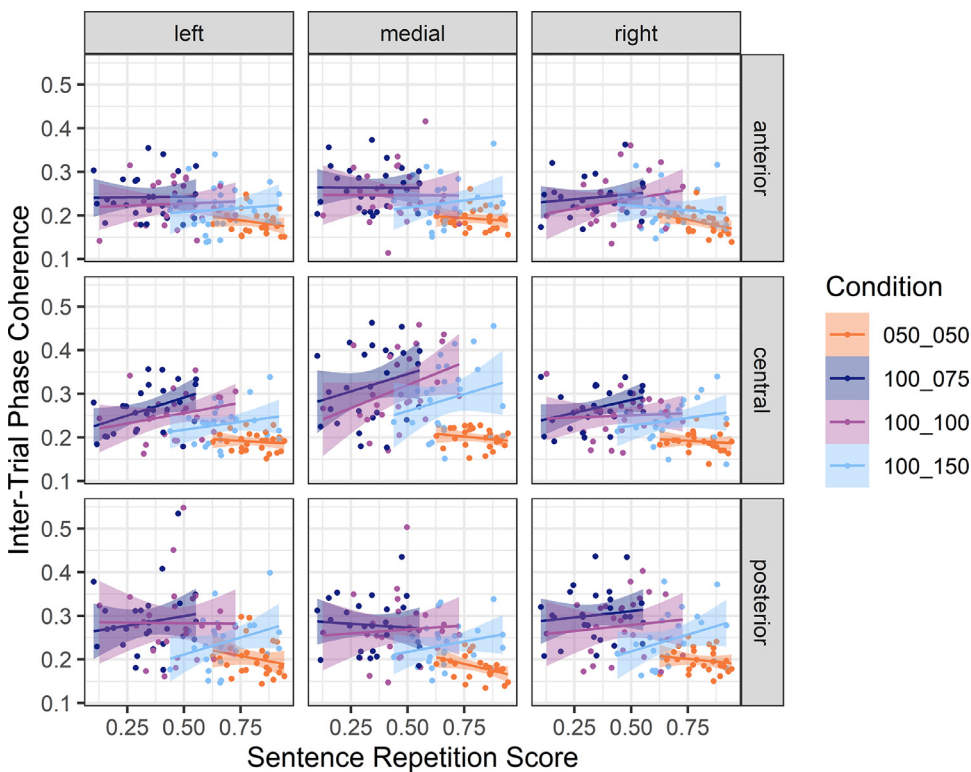


Fig. 7. Relationship of ITPC and Sentence Repetition Score per cluster and condition. *Note.* This figure shows regression plots for the relationship between sentence repetition score and ITPC in each condition at each electrode cluster. Regression lines are plotted with the formula ITPC sentence repetition score.

We also conducted an analysis on whether the pattern of increasing ITPC in increasingly difficult conditions was related to PTA. This analysis was suggested by a helpful reviewer. First, we visually analyzed the relationship between condition, PTA, and ITPC at each cluster by plotting each participant's ITPC per cluster and condition, with the conditions ranked for difficulty, and with connecting each participant's ITPC values with a line whose color represents their PTA (Figure S2). In this visualization, we observed the rising ITPC in more difficult conditions, which we already displayed as an average across participants in the topographical maps in Fig. 3. Visually, it was not obvious whether this pattern differed as a function of PTA. To statistically test this hypothesis, we first fitted linear models to the data points of each participant and each cluster. In other words, we approximated the plotted lines in Figure S2 by means of linear modeling. The purpose of this fitting was to approximate the degree to which ITPC increased for each participant in each cluster as a function of condition difficulty. To obtain these values, continuous values for condition difficulty were necessary, for which we took the mean difficulty over all participants as a proxy. The estimates for the coefficients of condition difficulty (i.e., the slopes of the linear models) were extracted and subjected to a LMEM with the slopes as criterion, PTA as a fixed effect, and random intercepts for participant and cluster. PTA was not predictive of the fitted slopes ($b = 0$, $t(24) = -0.27$, $p = 0.808$). We therefore did not find any evidence that the pattern of ITPC increasing as a function of condition difficulty was related to peripheral hearing. The relationship between PTA and slopes is visualized in Figure S3.

4. Discussion

This study investigated neural tracking of interrupted speech in healthy older adults with age-typical hearing. Specifically, we investigated whether neural speech tracking in the IRs of the interrupted speech signal would take place and if so, whether neural speech tracking could be expressed as a function of hearing and cognitive ability.

4.1. Repetition of interrupted speech

Our behavioral results complimented previous findings (Miller and Licklider, 1950; Wang and Humes, 2010) showing that the length of the silent interval is a key variable for successful understanding, in addition to duty cycle and IR. Participants scored highest in the 050_050 condition, although the duty cycle was higher in the 100_150 condition. However, taking into account only the duty cycle by keeping the length of the silent interval fixed at 100 ms, we observed a monotonic pattern of increasing understanding with increasing amount of signal as participants scored highest in the 100_150 condition, followed by the 100_100 condition and the 100_075 condition.

In a sample of older adults with normal hearing, we found that hearing thresholds were predictive of sentence repetition scores even when controlling for residual attenuation differences. We did not expect this result because our sample exhibited normal-for-age audiograms. However, because including residual attenuation differences did not remove PTA as a significant predictor, this finding indicates that raised pure-tone hearing thresholds can result in speech understanding difficulties than cannot be attributed to loudness (see also Lesica, 2018).

It was also one of our goals to ascertain whether visual linguistic closure would predict interrupted speech understanding (as in Bologna et al., 2018) and neural speech tracking. If it did, this would provide evidence for the involvement of a domain-general cognitive resource in the repair of the missing speech input. However, visual linguistic closure did not significantly predict the understanding of interrupted speech. This replicates the same null finding from Krull et al. (2013), who did not find such an association in their older participant group (but did find it in their younger participant group).

A possible explanation as to why we did not find an involvement of a domain-general cognitive resource in our study is that our paradigm may have been insufficient to trigger restoration of the missing speech signal. Top-down restoration of speech is often measured via the phonemic restoration paradigm (Saija et al., 2014). The term *phonemic restoration* in these studies refers to the difference in understanding speech interrupted by silence and speech interrupted by noise, the latter case be-

ing associated with better understanding, usually. The presence of noise is presumed to trigger restorative processes while the silence does not, or at least not as strongly.

Even though phonemic restoration is considered a cognitive process, studies that use cognitive tests to predict the extent of the phonemic restoration benefit are surprisingly rare. In a paradigm similar to ours, Jaekel et al. (2018) presented older adults with normal hearing stimuli interrupted by silence (their silent interval duration was 200 ms gaps, with a duty cycle of 0.5). They did not find effects of working memory or linguistic skills on sentence repetition. Bologna et al. (2018) found that visual linguistic closure actually predicted sentence repetition in speech interrupted by silence, but not speech interrupted by noise. In our study, the results for visual linguistic closure did not support the finding of Bologna et al. (2018) as we did not find a relationship between the current TRT_{WORD} version and the understanding of speech interrupted by silence. To further illustrate the relationship between the understanding of interrupted speech, visual linguistic closure and the type of interruption (silence vs. noise), it would be interesting to investigate speech tracking and cognition in relation to speech interrupted by noise as a potential trigger of cognitive restoration mechanisms.

4.2. Neural speech tracking of interrupted speech

We expected that the presentation of a speech signal interrupted by silence would elicit speech tracking in the specific IR frequency. We found significant main effects of condition when comparing speech tracking in the IR with speech tracking in the IRs of the other conditions, except when comparing speech tracking at 10 Hz between the 050_050 and 100_100 conditions. It is possible that this exception occurred because 10 Hz is a harmonic of 5 Hz, which was the IR of the 100_100 condition. Taking this into account, we argue that speech tracking in the IR occurred in all conditions. Speech tracking spanned across the scalp, as evidenced by significant main effects of condition, but it occurred most prominently in a medial, and, to a lesser degree, posterior central, a right posterior and a left posterior electrode cluster (see Table S1 and Fig. 5).

First, this result shows that healthy older adults' brains are susceptible to speech tracking, extending the results of Herrmann et al. (2019). To our best knowledge, we are the first to show that this speech tracking of natural interrupted speech occurs in healthy older adults and that it is highly sensitive even to small differences in frequency. Our 100_100 and 100_075 conditions had IR frequencies of 5 and 5.7 Hz, respectively, and even this small difference of 0.7 Hz could be reliably discriminated.

Nevertheless, comparing the 050_050 condition to the other conditions on speech tracking at the 10 Hz frequency yielded considerably smaller effects than the other comparisons. In other words, we found speech tracking a lot stronger in IR frequencies in the theta band (3–8 Hz), and a lot weaker in the alpha band (8–12 Hz). Two studies (Teng and Poeppel, 2019; Teng et al., 2017) recorded the magnetoencephalogram of participants listening to sounds modulated at theta, alpha, and gamma frequencies, and found speech tracking only for the theta- and gamma- modulated sounds, but not (Teng et al., 2017) or only to a lesser extent (Teng and Poeppel, 2019) for the alpha- modulated sounds. It is possible that our 050_050 condition with the IR frequency of 10 Hz did not elicit speech tracking to the same extent because the alpha band is not as sensitive to speech tracking as the theta band. Another explanation is that the 050_050 condition was simply too easy, so that strong speech tracking was not necessary to understand the stimuli.

In an auditory frequency-tagging study with young adults, Weisz and Lithari (2017) compared ITPC in response to amplitude-modulated sounds in different modulation rates. They found that ITPC in response to 4 Hz amplitude modulated sounds was stronger than ITPC in response to 10 Hz amplitude modulated sounds. Given the similarity between the interrupted speech paradigm and the frequency-tagging paradigm, their results concur with ours: We found stronger ITPC in the conditions with an IR of 4, 5, or 5.7 Hz in comparison with an IR of 10 Hz. Additionally,

we found evidence for a functional property of this neural response: Correctly repeating the interrupted sentences was related positively to ITPC in the conditions with IRs of 4, 5 and 5.7 Hz, and negatively related to ITPC in the condition with an IR of 10 Hz. We discuss this finding further below.

There is another possible explanation for the lowest ITPC values in the 050_050 condition. Assuming that evoked response amplitudes underlie ITPC, the faster IR rate and, correspondingly, the shorter duration of silence, might have resulted in smaller evoked response amplitudes simply due to neuronal refractoriness. However, in our view, there are two points that speak against this explanation. First, the frequency-tagging study by Weisz and Lithari (2017) demonstrated a nonlinear relationship between IR frequency and ITPC. ITPC in response to 4 Hz amplitude modulated sounds was stronger than ITPC in response to 10 Hz amplitude modulated sounds, but the ITPC response was by far the strongest to 40 Hz modulated sounds. If ITPC strength was directly related to neural refractoriness, the results of Weisz and Lithari (2017) should have demonstrated a monotonously decreasing relationship between modulation rate and ITPC. Second, assuming a monotonous increase in evoked response amplitudes as a function of IR, 4 Hz should have the strongest overall response, followed by 5 Hz and then by 5.7 Hz. However, the data show a reverse pattern for the three conditions with these IRs: In the 100_150 condition (IR: 4 Hz), ITPC is lower than in the 100_100 condition (IR: 5 Hz), which in turn is lower than ITPC in the 100_075 condition (IR: 5.7 Hz). We therefore conclude that ITPC values in the 050_050 condition were not lowest because of neural refractoriness.

4.3. Inter-individual differences in neural speech tracking of interrupted speech

Speech tracking in the specific IR of each condition showed considerable inter-individual variability (ITPC ranging between 0.11 and 0.55). This noticeable inter-individual variability in speech tracking is a common finding (Lam et al., 2018). In our study, one of our aims was to investigate possible sources of this variability.

We did not find evidence that higher hearing thresholds would be associated with less speech tracking. This is in line with our expectations, because our sample consisted of older adults with normal-for-age audiograms. Also, previous studies have presented evidence that hearing thresholds are related to speech tracking in a more subtle way than a direct correlation (Petersen et al., 2017; Presacco et al., 2019).

In our study, we measured visual linguistic closure in order to approach a domain-general cognitive mechanism which would contribute to repairing the missing language input. However, in addition to not predicting sentence repetition scores, visual linguistic closure did also not predict speech tracking. It is possible that we did not target the relevant variable. For example, as a measure of cognitive capacity, working memory span assessed via the reading span task (Daneman and Carpenter, 1980) has been more promising in cognitive hearing science (Arlinger et al., 2009).

We were also interested in whether ITPC would be modulated by task difficulty. To approximate a measure of (individual) task difficulty, we took the sentence repetition scores as a proxy, because it gives us a measure of how challenged a participant was in the task. In the 100_150, 100_100, and 100_075 conditions, speech tracking increased as individual task difficulty decreased (i.e., as sentence repetition performance increased). In the 050_050 condition however, speech tracking decreased as individual task difficulty decreased (i.e., as sentence repetition performance increased). This result can be explained in a framework outlining the functional roles of different frequency bands for cortical speech processing (Giraud and Poeppel, 2012). The IRs of 4, 5, and 5.7 Hz correspond to the theta frequency band of EEG frequencies (e.g., Ghitzia, 2011). In these conditions, we had fixed the silent interval length to 100 ms, which is approximately the median length of a syllable (Greenberg et al., 2003). Possibly, ITPC in theta-range frequen-

cies was positively related to the ability to reproduce the interrupted sentence because in the relevant conditions, higher ITPC allowed a better capturing of the syllable chunks that were actually present in the interrupted speech signal. The alpha frequency band, which includes 10 Hz, is prominently involved in cognitive abilities like attention and working memory (e.g., Klimesch, 2012; Klimesch et al., 1993); abilities, which are also relevant during speech perception in adverse listening conditions (Obleser et al., 2012). However, the alpha band appears less relevant in the context of basic speech processing (e.g., Teng and Poeppel, 2019), which might also be reflected in the weaker neural response to amplitude modulated sounds in 10 Hz compared to 4 Hz (Weisz and Lithari, 2017). Thus, stronger ITPC in the condition with the IR of 10 Hz might not be beneficial, but rather detrimental during a speech processing task, as our results suggest.

Interestingly, auditory frequency-tagging studies usually find the strongest neural response to amplitude modulated sounds with a modulation rate of 40 Hz (Galambos et al., 1981; Ross et al., 2005; Weisz and Lithari, 2017). In future studies, it would be interesting to also add a condition with a 40 Hz IR. However, this condition would be extremely easy to understand and it we would only expect behavioral variability in a specially selected stimulus set, where missing speech chunks of 12.5 ms (in a condition with a 40 Hz IR and a duty cycle of 0.5) actually introduce ambiguity.

It is worth noting that neural responses to sounds that are amplitude modulated in all of the mentioned frequencies (from 4 Hz to 40 Hz) seem to be preserved in older adults (Boettcher et al., 2001; Grose et al., 2009), in comparison to higher frequencies (Grose et al., 2009; Leigh-Paffenroth and Fowler, 2006).

We further explored the possibility of ITPC in the IR frequencies being observed only because of phase reset concurring with early cortical evoked potentials at every new sound segment. However, if this were true, (1) ITPC in the 10 Hz frequency in the 050_050 condition should be of similar size as in the other conditions, and (2) ITPC would have been better accounted for as a function of the number of sound onsets rather than a function of difficulty.

Regarding the first point, we have already discussed the difference in effect magnitudes between the 050_050 and the other conditions. Regarding the second point, a model predicting ITPC as a function of difficulty yielded a lower AIC than a model predicting ITPC as a function of the number of sound onsets. Clearly, this does not exclude the possibility that the ITPC signal did not partly stem from a summation of evoked potentials. Even so, a study by Doelling et al. (2019) found that an oscillator model better predicted neural tracking of rhythmic musical stimuli than an evoked responses model, even when the music notes contained sharp attacks, which should have fostered evoked responses.

To summarize, we showed that speech tracking in the IR occurs in healthy older adults, and that stronger speech tracking in the theta frequency band is related to better understanding of interrupted speech. Because hearing aids currently constitute the only evidence-based treatment for age-related hearing loss, there is a lack of treatments for hearing problems that do not arise out of audibility issues, but rather out of intelligibility issues. The sensitivity of speech tracking to the difficulty of a listening situation could qualify it as a mechanism which could possibly be trained or excited via brain stimulation methods such as transcranial alternating current stimulation (Riecke et al., 2015; Rufener et al., 2016; Zoefel et al., 2018) in order to improve speech understanding in older adults with hearing loss. Other studies have used acoustic stimulation to drive entrainment of known intrinsic cortical rhythms in other cognitive functions such as sleep (Lafon et al., 2017). This form of therapy would, however, be grounded on the conception that the mechanism underlying speech tracking is the entrainment of endogenous cortical rhythms, which is under debate (Obleser and Kayser, 2019).

Another application of this research is the integration of neuroscientific research findings into the engineering of hearing aids. This approach exploits the quantifiable similarity between the attended speech signal and neural signatures of speech processing. With these methods,

the focus of the listener's attention can be decoded and the attended speech signal can be selectively amplified. Technological advances like the "ear-EEG" (Kidmose et al., 2013) and the "cEEGgrid" (Debener et al., 2015; Mirkovic et al., 2016) allow for efficient and discreet recording of EEG, and advances in attention-decoding algorithms (Fiedler et al., 2016; Han et al., 2019) further improve the applicability of these devices. Our study provides additional evidence that speech tracking occurs in difficult listening situations in healthy older adults, which underscores the feasibility of this new technological feature of hearing aids.

5. Conclusion

In this study, we showed that interrupted speech elicits neural speech tracking in older adults with normal hearing and that neural speech tracking is highly specific to the IR. Although our sample consisted of normal hearing individuals who could amplify the stimuli if necessary, we found a negative relationship between hearing thresholds and the understanding of interrupted speech. This relationship suggests a critical role for even slightly elevated hearing thresholds in speech processing. Additionally, neural speech tracking in the theta frequency band was positively related to the understanding of interrupted speech. Therefore, neural speech tracking might be a candidate training mechanism for older adults with hearing loss.

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

Credit authorship contribution statement

Ira Kurthen: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Project administration, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Allison Christen:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft. **Martin Meyer:** Conceptualization, Resources, Writing – review & editing, Supervision, Funding acquisition. **Nathalie Giroud:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Funding acquisition.

Data Availability

Data will be made available on request.

Acknowledgments

This work was supported by grants from the Swiss National Science Foundation (Grant no. 172268 to IK and grant no. 169964 to MM) and a Postdoc "Forschungskredit" of the University of Zurich to NG (No. FK-19-072), and by the Sonova AG. During the work on her dissertation, IK was a pre-doctoral fellow of the International Max Planck Research School on the Life Course (LIFE). We thank Dr. Oliver Profant and Dr. Zbynek Bures for fruitful discussions in the early stages of this research.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.neuroimage.2022.119580](https://doi.org/10.1016/j.neuroimage.2022.119580)

References

- Alexandrou, A.M., Saarinen, T., Kujala, J., Salmelin, R., 2020. Cortical entrainment: what we can learn from studying naturalistic speech perception. *Lang., Cogn. Neurosci.* 35 (6), 681–693. doi:[10.1080/23273798.2018.1518534](https://doi.org/10.1080/23273798.2018.1518534).
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., Dreihobl, S., Kraus, N., 2013. Effects of hearing loss on the subcortical representation of speech cues. *J. Acoust. Soc. Am.* 133 (5), 3030–3038. doi:[10.1121/1.4799804](https://doi.org/10.1121/1.4799804).
- Annett, M., 1970. A classification of hand preference by association analysis. *Br. J. Psychol.* 61 (3), 303–321. doi:[10.1111/j.2044-8295.1970.tb01248.x](https://doi.org/10.1111/j.2044-8295.1970.tb01248.x).

- Arlinger, S., Lunner, T., Lyxell, B., Kathleen Pichora-Fuller, M.K., 2009. The emergence of cognitive hearing science. *Scand. J. Psychol.* 50 (5), 371–384. doi:10.1111/j.1467-9450.2009.00753.x.
- Başkent, D., Eiler, C.L., Edwards, B., 2010. Phonemic restoration by hearing-impaired listeners with mild to moderate sensorineural hearing loss. *Hear. Res.* 260 (1–2), 54–62. doi:10.1016/j.heares.2009.11.007.
- Benard, M.R., Mensink, J.S., Başkent, D., 2014. Individual differences in top-down restoration of interrupted speech: links to linguistic and cognitive abilities. *J. Acoust. Soc. Am.* 135 (2), EL88–EL94. doi:10.1121/1.4862879.
- Bertoli, S., Staehelin, K., Zemp, E., Schindler, C., Bodmer, D., Probst, R., 2009. Survey on hearing aid use and satisfaction in Switzerland and their determinants. *Int. J. Audiol.* 48 (4), 183–195. doi:10.1080/14992020802572627.
- Besser, J., Zekveld, A.A., Kramer, S.E., Rönnberg, J., Festen, J.M., 2012. New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J. Speech, Lang., Hear. Res.* 55, 194–209.
- Bharadwaj, H.M., Lee, A.K.C., Shinn-Cunningham, B.G., 2014. Measuring auditory selective attention using frequency tagging. *Front. Integr. Neurosci.* 8. doi:10.3389/fnint.2014.00006.
- Bidelman, G.M., Price, C.N., Shen, D., Arnott, S.R., Alain, C., 2019. Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. *Hear. Res.* 382, 107795. doi:10.1016/j.heares.2019.107795.
- Bidelman, G.M., Villafuerte, J.W., Moreno, S., Alain, C., 2014. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol. Aging* 35 (11), 2526–2540. doi:10.1016/j.neurobiolaging.2014.05.006.
- Bisgaard, N., Ruf, S., 2017. Findings from EuroTrak surveys from 2009 to 2015: hearing loss prevalence, hearing aid adoption, and benefits of hearing aid use. *Am. J. Audiol.* 26 (3S), 451–461. doi:10.1044/2017_AJA-16-0135.
- Boettcher, F.A., Poth, E.A., Mills, J.H., Dubno, J.R., 2001. The amplitude-modulation following response in young and aged human subjects. *Hear. Res.* 153 (1–2), 32–42. doi:10.1016/S0378-5955(00)00255-0.
- Bologna, W.J., Vaden, K.I., Ahlstrom, J.B., Dubno, J.R., 2018. Age effects on perceptual organization of speech: contributions of glimpsing, phonemic restoration, and speech segregation. *J. Acoust. Soc. Am.* 144 (1), 267–281. doi:10.1121/1.5044397.
- Buiatti, M., Pena, M., Dehaenelambertz, G., 2009. Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage* 44 (2), 509–519. doi:10.1016/j.neuroimage.2008.09.015.
- Daneman, M., Carpenter, P.A., 1980. Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19 (4), 450–466. doi:10.1016/S0022-5371(80)90312-6.
- Debener, S., Emkes, R., De Vos, M., Bleichner, M., 2015. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Sci. Rep.* 5 (1), 16743. doi:10.1038/srep16743.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21. doi:10.1016/j.jneumeth.2003.10.009.
- Ding, N., Melloni, L., Zhang, H., Tian, X., Poeppel, D., 2016. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* 19 (1), 158–164. doi:10.1038/nn.4186.
- Ding, N., Patel, A.D., Chen, L., Butler, H., Luo, C., Poeppel, D., 2017. Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* 81, 181–187. doi:10.1016/j.neubiorev.2017.02.011.
- Doelling, K.B., Arnal, L.H., Ghitz, O., Poeppel, D., 2014. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage* 85, 761–768. doi:10.1016/j.neuroimage.2013.06.035.
- Doelling, K.B., Assaneo, M.F., Bevilacqua, D., Pesaran, B., Poeppel, D., 2019. An oscillator model better predicts cortical entrainment to music. *Proc. Natl. Acad. Sci.* 116 (20), 10113–10121. doi:10.1073/pnas.1816414116.
- Fiedler, L., Obleser, J., Lunner, T., Graversen, C., 2016. Ear-EEG allows extraction of neural responses in challenging listening scenarios – a future technology for hearing aids? In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, Orlando, FL, USA, pp. 5697–5700. doi:10.1109/EMBC.2016.7592020.
- Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. *Proc. Natl. Acad. Sci.* 78 (4), 2643–2647. doi:10.1073/pnas.78.4.2643.
- Ghitza, O., 2011. Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Front. Psychol.* 2. doi:10.3389/fpsyg.2011.00130.
- Gilbert, G., Bergeras, I., Voillery, D., Lorenzi, C., 2007. Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues. *J. Acoust. Soc. Am.* 122 (3), 1336–1339. doi:10.1121/1.2756161.
- Giraud, A.-L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15 (4), 511–517. doi:10.1038/nn.3063.
- Giroud, N., Hirsiger, S., Muri, R., Kegel, A., Dillier, N., Meyer, M., 2018. Neuroanatomical and resting state EEG power correlates of central hearing loss in older adults. *Brain Struct. Funct.* 223 (1), 145–163. doi:10.1007/s00429-017-1477-0.
- Giroud, N., Keller, M., Hirsiger, S., Dellwo, V., Meyer, M., 2019. Bridging the brain structure-brain function gap in prosodic speech processing in older adults. *Neurobiol. Aging* 80, 116–126. doi:10.1016/j.neurobiolaging.2019.04.017.
- Goman, A.M., Lin, F.R., 2016. Prevalence of hearing loss by severity in the united states. *Am. J. Public Health* 106 (10), 1820–1822. doi:10.2105/AJPH.2016.303299.
- Gordon-Salant, S., Fitzgibbons, P.J., 1993. Temporal factors and speech recognition performance in young and elderly listeners. *J. Speech, Lang., Hear. Res.* 36 (6), 1276–1285. doi:10.1044/jshr.3606.1276.
- Greenberg, S., Carvey, H., Hitchcock, L., Chang, S., 2003. Temporal properties of spontaneous speech—a syllable-centric perspective. *J. Phon.* 31 (3–4), 465–485. doi:10.1016/j.wocn.2003.09.005.
- Grose, J.H., Mamo, S.K., Hall, J.W., 2009. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 30 (5), 568–575. doi:10.1097/AUD.0b013e3181ac128f.
- Haegens, S., Zion Golumbic, E., 2018. Rhythmic facilitation of sensory processing: a critical review. *Neurosci. Biobehav. Rev.* 86, 150–165. doi:10.1016/j.neubiorev.2017.12.002.
- Han, C., O'Sullivan, J., Luo, Y., Herrero, J., Mehta, A.D., Mesgarani, N., 2019. Speaker-independent auditory attention decoding without access to clean speech sources. *Sci. Adv.* 5 (5), eaav6134. doi:10.1126/sciadv.aav6134.
- Herrmann, B., Buckland, C., Johnsrude, I.S., 2019. Neural signatures of temporal regularity processing in sounds differ between younger and older adults. *Neurobiol. Aging* 83, 73–85. doi:10.1016/j.neurobiolaging.2019.08.028.
- Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. *Biom. J.* 50 (3), 346–363. doi:10.1002/bimj.200810425.
- Jaekel, B.N., Newman, R.S., Goupell, M.J., 2018. Age effects on perceptual restoration of degraded interrupted sentences. *J. Acoust. Soc. Am.* 143 (1), 84–97. doi:10.1121/1.5016968.
- Jervis, B.W., Nichols, M.J., Johnson, T.E., Allen, E., Hudson, N.R., 1983. A fundamental investigation of the composition of auditory evoked potentials. *IEEE Trans. Biomed. Eng.* BME-30 (1), 43–50. doi:10.1109/TBME.1983.325165.
- Johnson, C., Jilla, A., Danhauer, J., Sullivan, J., Sanchez, K., 2018. Benefits from, satisfaction with, and self-efficacy for advanced digital hearing aids in users with mild sensorineural hearing loss. *Semin. Hear.* 39 (02), 158–171. doi:10.1055/s-0038-1641742.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37 (2), 163–178. doi:10.1111/1469-8986.3720163.
- Keitel, A., Gross, J., Kayser, C., 2018. Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biol.* 16 (3), e2004473. doi:10.1371/journal.pbio.2004473.
- Kidmose, P., Looney, D., Jochumsen, L., Mandic, D.P., 2013. Ear-EEG from generic earpieces: a feasibility study. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, Osaka, pp. 543–546. doi:10.1109/EMBC.2013.6609557.
- Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16 (12), 606–617. doi:10.1016/j.tics.2012.10.007.
- Klimesch, W., Schimke, H., Pfurtscheller, G., 1993. Alpha frequency, cognitive load and memory performance. *Brain Topogr.* 5 (3), 241–251. doi:10.1007/BF01128991.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12 (5), 535–540. doi:10.1038/nn.2303.
- Krull, V., Humes, L.E., Kidd, G.R., 2013. Reconstructing wholes from parts: effects of modality, age, and hearing loss on word recognition. *Ear Hear.* 34 (2), e14–e23. doi:10.1097/AUD.0b013e31826d0c27.
- Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82 (13), 1–26. doi:10.18637/jss.v082.i13.
- Lachaux, J.-P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8 (4), 194–208.
- Lafon, B., Henin, S., Huang, Y., Friedman, D., Melloni, L., Thesen, T., Doyle, W., Buzsáki, G., Devinsky, O., Parra, L.C., A. Liu, A., 2017. Low frequency transcranial electrical stimulation does not entrain sleep rhythms measured by human intracranial recordings. *Nat. Commun.* 8 (1), 1199. doi:10.1038/s41467-017-01045-x.
- Lakatos, P., Barczak, A., Neymotin, S.A., McGinnis, T., Ross, D., Javitt, D.C., O'Connell, M.N., 2016. Global dynamics of selective attention and its lapses in primary auditory cortex. *Nat. Neurosci.* 19 (12), 1707–1717. doi:10.1038/nn.4386.
- Lakatos, P., Gross, J., Thut, G., 2019. A new unifying account of the roles of neuronal entrainment. *Curr. Biol.* 29 (18), R890–R905. doi:10.1016/j.cub.2019.07.075.
- Lam, N.H.L., Hultén, A., Hagoort, P., Schoffelen, J.-M., 2018. Robust neuronal oscillatory entrainment to speech displays individual variation in lateralisation. *Lang., Cogn. Neurosci.* 33 (8), 943–954. doi:10.1080/23273798.2018.1437456.
- Lecluyse, W., Meddis, R., 2009. A simple single-interval adaptive procedure for estimating thresholds in normal and impaired listeners. *J. Acoust. Soc. Am.* 126 (5), 2570–2579. doi:10.1121/1.3238248.
- Lecluyse, W., Tan, C.M., McFerran, D., Meddis, R., 2013. Acquisition of auditory profiles for good and impaired hearing. *Int. J. Audiol.* 52 (9), 596–605. doi:10.3109/14992027.2013.796530.
- Leigh-Paffenroth, E.D., Fowler, C.G., 2006. Amplitude-modulated auditory steady-state responses in younger and older listeners. *J. Am. Acad. Audiol.* 17 (08), 582–597. doi:10.3766/jaaa.17.8.5.
- Lesica, N.A., 2018. Why do hearing aids fail to restore normal auditory perception? *Trends Neurosci.* 41 (4), 174–185. doi:10.1016/j.tins.2018.01.008.
- Maharani, A., Pendleton, N., Leroi, I., 2019. Hearing impairment, loneliness, social isolation, and cognitive function: longitudinal analysis using english longitudinal study on ageing. *Am. J. Geriatr. Psychiatry* 27 (12), 1348–1356. doi:10.1016/j.jagp.2019.07.010.
- Mattys, S.L., Davis, M.H., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27 (7–8), 953–978. doi:10.1080/01690965.2012.705006.
- Meddis, R., Lecluyse, W., Clark, N.R., Jürgens, T., Tan, C.M., Panda, M.R., Brown, G.J., 2013. A computer model of the auditory periphery and its application to the study of hearing. In: Moore, B.C.J., Patterson, R.D., Winter, I.M., Carlyon, R.P., Gockel, H.E. (Eds.), *Basic Aspects of Hearing*. Springer New York, pp. 11–20.
- Mick, P., Kawachi, I., Lin, F.R., 2014. The association between hearing loss and so-

- cial isolation in older adults. *Otolaryngology-Head Neck Surg.* 150 (3), 378–384. doi:[10.1177/0194599813518021](https://doi.org/10.1177/0194599813518021).
- Miller, G.A., Licklider, J.C.R., 1950. The intelligibility of interrupted speech. *J. Acoust. Soc. Am.* 22 (2), 167–173. doi:[10.1121/1.1906584](https://doi.org/10.1121/1.1906584).
- Mirkovic, B., Bleichner, M.G., De Vos, M., Debener, S., 2016. Target speaker detection with concealed EEG around the ear. *Front. Neurosci.* 10. doi:[10.3389/fnins.2016.00349](https://doi.org/10.3389/fnins.2016.00349).
- Nasreddine, Z.S., Phillips, N.A., Bédrian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H., 2005. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53 (4), 695–699. doi:[10.1111/j.1532-5415.2005.53221.x](https://doi.org/10.1111/j.1532-5415.2005.53221.x).
- Obleser, J., Kayser, C., 2019. Neural entrainment and attentional selection in the listening brain. *Trends Cogn. Sci.* 23 (11), 913–926. doi:[10.1016/j.tics.2019.08.004](https://doi.org/10.1016/j.tics.2019.08.004).
- Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., Maess, B., 2012. Adverse listening conditions and memory load drive a common alpha oscillatory network. *J. Neurosci.* 32 (36), 12376–12383. doi:[10.1523/JNEUROSCI.4908-11.2012](https://doi.org/10.1523/JNEUROSCI.4908-11.2012).
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1–9. doi:[10.1155/2011/156869](https://doi.org/10.1155/2011/156869).
- Park, H., Ince, R., Schyns, P., Thut, G., Gross, J., 2015. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr. Biol.* 25 (12), 1649–1653. doi:[10.1016/j.cub.2015.04.049](https://doi.org/10.1016/j.cub.2015.04.049).
- Peelle, J.E., Gross, J., Davis, M.H., 2013. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex* 23 (6), 1378–1387. doi:[10.1093/cercor/bhs118](https://doi.org/10.1093/cercor/bhs118).
- Perrin, F., Pernier, J., Bertrand, O., Giard, M., Echallier, J., 1987. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr. Clin. Neurophysiol.* 66 (1), 75–81. doi:[10.1016/0013-4694\(87\)90141-6](https://doi.org/10.1016/0013-4694(87)90141-6).
- Petersen, E.B., Wöstmann, M., Obleser, J., Lunner, T., 2017. Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J. Neurophysiol.* 117 (1), 18–27. doi:[10.1152/jn.00527.2016](https://doi.org/10.1152/jn.00527.2016).
- Pichora-Fuller, M.K., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W., Humes, L.E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C.L., Naylor, G., Phillips, N.A., Richter, M., Rudner, M., Sommers, M.S., Tremblay, K.L., Wingfield, A., 2016. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear.* 37, 5S–27S. doi:[10.1097/AUD.0000000000000312](https://doi.org/10.1097/AUD.0000000000000312).
- Picton, T.W., John, M.S., Dimitrijevic, A., Purcell, D., 2003. Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int. J. Audiol.* 42 (4), 177–219. doi:[10.3109/14992020309101316](https://doi.org/10.3109/14992020309101316).
- Poeppl, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun.* 41 (1), 245–255. doi:[10.1016/S0167-6393\(02\)00107-3](https://doi.org/10.1016/S0167-6393(02)00107-3).
- Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53 (5), 793–808. doi:[10.1080/10635150490522304](https://doi.org/10.1080/10635150490522304).
- Presacco, A., Simon, J.Z., Anderson, S., 2019. Speech-in-noise representation in the aging midbrain and cortex: effects of hearing loss. *PLoS One* 14 (3), e0213899. doi:[10.1371/journal.pone.0213899](https://doi.org/10.1371/journal.pone.0213899).
- Profant, O., Jilek, M., Bures, Z., Vencovsky, V., Kucharova, D., Svobodova, V., Korynta, J., Syka, J., 2019. Functional age-related changes within the human auditory system studied by audiometric examination. *Front. Aging Neurosci.* 11. doi:[10.3389/fnagi.2019.00026](https://doi.org/10.3389/fnagi.2019.00026).
- Riecke, L., Formisano, E., Herrmann, C.S., Sack, A.T., 2015. 4-Hz transcranial alternating current stimulation phase modulates hearing. *Brain Stimul.* 8 (4), 777–783. doi:[10.1016/j.brs.2015.04.004](https://doi.org/10.1016/j.brs.2015.04.004).
- Riecke, L., Formisano, E., Sorger, B., Başkent, D., Gaudrain, E., 2018. Neural entrainment to speech modulates speech intelligibility. *Curr. Biol.* 28 (2), 161–169.e5. doi:[10.1016/j.cub.2017.11.033](https://doi.org/10.1016/j.cub.2017.11.033).
- Roach, B.J., Mathalon, D.H., 2008. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophr. Bull.* 34 (5), 907–926. doi:[10.1093/schbul/sbn093](https://doi.org/10.1093/schbul/sbn093).
- Ross, B., Herdman, A., Pantev, C., 2005. Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cereb. Cortex* 15 (12), 2029–2039. doi:[10.1093/cercor/bhi078](https://doi.org/10.1093/cercor/bhi078).
- Rufener, K.S., Oechslin, M.S., Zaehle, T., Meyer, M., 2016. Transcranial alternating current stimulation (tACS) differentially modulates speech perception in young and older adults. *Brain Stimul.* 9 (4), 560–565. doi:[10.1016/j.brs.2016.04.002](https://doi.org/10.1016/j.brs.2016.04.002).
- Rönnberg, J., Lunner, T., Zekveld, A.A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M.K., Rudner, M., 2013. The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7 (31), 1–17. doi:[10.3389/fnsys.2013.00031](https://doi.org/10.3389/fnsys.2013.00031).
- Sajja, J.D., Akyürek, E.G., Andringa, T.C., Başkent, D., 2014. Perceptual restoration of degraded speech is preserved with advancing age. *J. Assoc. Res. Otolaryngol.* 15 (1), 139–148. doi:[10.1007/s10162-013-0422-z](https://doi.org/10.1007/s10162-013-0422-z).
- Shafiro, V., Sheft, S., Risley, R., 2016. The intelligibility of interrupted and temporally altered speech: effects of context, age, and hearing loss. *J. Acoust. Soc. Am.* 139 (1), 455–465. doi:[10.1121/1.4939891](https://doi.org/10.1121/1.4939891).
- Stark, P., Hickson, L., 2004. Outcomes of hearing aid fitting for older people with hearing impairment and their significant others. *Int. J. Audiol.* 43 (7), 390–398. doi:[10.1080/14992020400050050](https://doi.org/10.1080/14992020400050050).
- Tallon-Baudry, C., Bertrand, O., Delpeuch, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J. Neurosci.* 16 (13), 4240–4249. doi:[10.1523/JNEUROSCI.16-13-04240.1996](https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996).
- R Core Team, 2018. R: A Language and Environment for Statistical Computing.
- Teng, X., Poeppel, D., 2019. Theta and gamma bands encode acoustic dynamics over wide-ranging timescales. *Cereb. Cortex* 30 (4), 2600–2614. doi:[10.1093/cercor/bhz263](https://doi.org/10.1093/cercor/bhz263).
- Teng, X., Tian, X., Rowland, J., Poeppel, D., 2017. Concurrent temporal channels for auditory processing: oscillatory neural entrainment reveals segregation of function at different scales. *PLoS Biol.* 15 (11), e2000812. doi:[10.1371/journal.pbio.2000812](https://doi.org/10.1371/journal.pbio.2000812).
- Vinck, M., Lima, B., Womelsdorf, T., Oostenveld, R., Singer, W., Neuenschwander, S., Fries, P., 2010. Gamma-phase shifting in awake monkey visual cortex. *J. Neurosci.* 30 (4), 1250–1257. doi:[10.1523/JNEUROSCI.1623-09.2010](https://doi.org/10.1523/JNEUROSCI.1623-09.2010).
- Vrieze, S.I., 2012. Model selection and psychological theory: adiscussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* 17 (2), 228–243. doi:[10.1037/a0027127](https://doi.org/10.1037/a0027127).
- Wagener, K., Kühnel, V., Kollmeier, B., 1999. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Z. Audiol.* 38 (1), 4–15.
- Wang, X., Humes, L.E., 2010. Factors influencing recognition of interrupted speech. *J. Acoust. Soc. Am.* 128 (4), 2100–2111. doi:[10.1121/1.3483733](https://doi.org/10.1121/1.3483733).
- Weinstein, B.E., Ventry, I.M., 1982. Hearing impairment and social isolation in the elderly. *J. Speech, Lang., Hear. Res.* 25 (4), 593–599. doi:[10.1044/jshr.2504.593](https://doi.org/10.1044/jshr.2504.593).
- Weisz, N., Lithari, C., 2017. Amplitude modulation rate dependent topographic organization of the auditory steady-state response in human auditory cortex. *Hear. Res.* 354, 102–108. doi:[10.1016/j.heares.2017.09.003](https://doi.org/10.1016/j.heares.2017.09.003).
- Werkle-Bergner, M., Shing, Y.L., Müller, V., Li, S.-C., Lindenberger, U., 2009. EEG gamma-band synchronization in visual coding from childhood to old age: evidence from evoked power and inter-trial phase locking. *Clin. Neurophysiol.* 120 (7), 1291–1302. doi:[10.1016/j.clinph.2009.04.012](https://doi.org/10.1016/j.clinph.2009.04.012).
- Wilsch, A., Neuling, T., Obleser, J., Herrmann, C.S., 2018. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage* 172, 766–774. doi:[10.1016/j.neuroimage.2018.01.038](https://doi.org/10.1016/j.neuroimage.2018.01.038).
- Wingfield, A., Aberdeen, J.S., Stine, E.A.L., 1991. Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *J. Gerontol.* 46 (3), P127–P129. doi:[10.1093/geronj/46.3.P127](https://doi.org/10.1093/geronj/46.3.P127).
- Yang, D., Shin, Y.-I., Hong, K.-S., 2021. Systemic review on transcranial electrical stimulation parameters and EEG/fNIRS features for brain diseases. *Front. Neurosci.* 15, 629323. doi:[10.3389/fnins.2021.629323](https://doi.org/10.3389/fnins.2021.629323).
- Zekveld, A.A., George, E.L., Kramer, S.E., Goverts, S.T., Houtgast, T., 2007. The development of the text reception threshold test: a visual analogue of the speech reception threshold test. *J. Speech, Lang., Hear. Res.* 50 (3), 576–584. doi:[10.1044/1092-4388\(2007/040\)](https://doi.org/10.1044/1092-4388(2007/040)).
- Zoefel, B., Archer-Boyd, A., Davis, M.H., 2018. Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. *Curr. Biol.* 28 (3), 401–408.e5. doi:[10.1016/j.cub.2017.11.071](https://doi.org/10.1016/j.cub.2017.11.071).