The Noise-Resilient Brain: Resting-State Oscillatory Activity Predicts Words-In-Noise Recognition

Thomas Houweling[a,*], Robert Becker[a], Alexis Hervais-Adelman[a]

[a]*Neurolinguistics, Department of Psychology, University of Zürich*

Binzmühlestrasse 14, 8050, Zürich, Switzerland

Abstract

The role of neuronal oscillations in the processing of speech has recently come to prominence. Since resting-state (RS) brain activity has been shown to predict both task-related brain activation and behavioural performance, we set out to establish whether inter-individual differences in spectrally-resolved RS-MEG power are associated with variations in speech in noise recognition. In a frequency range between 21 and 29 Hz, significant positive correlations were observed between voxelwise power and resilience to noise in a large left-lateralised perisylvian cluster spanning from inferior frontal gyrus to temporo-parietal junction and peaking in left posterior superior temporal gyrus (pSTG). Smaller areas of association were also found around the right pSTG (21-29Hz) and bilateral pSTG (30-40Hz). These findings are spatially and spectrally consistent with the neural substrates of phonological processing. We propose that increased RS power in auditory cortices and the left perisylvian region can partly explain improved resilience to noise.

Keywords

Resting-state; MEG; speech perception; words-in-noise; oscillations; power; rapid auditory processing; phoneme; STG; HCP

Highlights

- Power of resting MEG activity predicts Words-In-Noise recognition performance
- Significant associations in higher beta and lower gamma frequency band
- Strongest in left-lateralised perisylvian cluster peaking in posterior STG
- Effects are spectrally and spatially consistent with phoneme-level processing

*Correspondence to: thomas.houweling@uzh.ch.

Neurolinguistics, Department of Psychology, University of Zürich,

Binzmühlestrasse 14, box 5, CH-8050, Zürich, Switzerland,

1. Introduction

Over the last decade, the scope of neuroscientific inquiry has considerably broadened, as the traditional approach to studying the brain through its task-elicited activation, is increasingly assisted and complemented by the interrogation of its task-independent properties (Biswal et al., 2009). Among the factors contributing to the rise in interest in spontaneous brain activity has been the realisation that resting-state (RS) brain activity has the potential to index a number of intrinsic features of the brain. Further, measures of RS brain activity have been shown to be predictive not only of task-related brain activation (Cole, Ito, Bassett, & Schultz, 2016; Mennes et al., 2010, 2011; Tavor et al., 2016), but also of behavioural outcomes (Finn et al., 2016; Fox, Snyder, Vincent, & Raichle, 2007; Fox, Snyder, Zacks, & Raichle, 2006; Mennes et al., 2011). Contrary to traditional hypothesis-driven task-based approaches that aim at characterising how – where and when – the brain processes the given stimulus property, as indexed by brain activation, hypothesis- and task-free approaches aim to characterise inherent features of the brain that shape how information from the outside world is processed. To this extent, the brain's RS activity may be conceived of as a structuring context in which evoked activity occurs.

In this study we used magnetoencephalography (MEG) to characterise the brain's RS activity in terms of spectrally-resolved power. MEG records magnetic field perturbations produced by neuronal electrical activity at high temporal resolution, enabling the decomposition of the broadband signal into spectrally-resolved components. It has been shown that different frequency ranges are functionally distinct (Buzsáki & Draguhn, 2004). Spectrally-specific modulations in amplitude/power of the electromagnetic field associated with neuronal activity are thought to reflect the degree of synchronisation among neuronal populations. Here, we tested whether individual differences in spectrally-resolved MEG power are able to predict inter-individual variability in a speech perception task. Specifically, we tested whether spectrally-resolved RS MEG power is predictive of the ability to perceive speech in noise, as measured by a Words-In-Noise (WIN) test. WIN assesses individuals' ability to recognise isolated monosyllabic words embedded

in background noise. The ability to comprehend speech in noise, and acoustically-degraded speech more generally, is known to be highly variable across individuals (Mattys, Davis, Bradlow, & Scott, 2012), and several hypotheses have been advanced regarding cognitive determinants of speech in noise and degraded speech comprehension (Rönnberg et al., 2013; Zekveld, Rudner, Johnsrude, & Rönnberg, 2013). Here, however, we aimed to examine the relationship between RS activity and WIN performance, while controlling for a battery of other, general cognitive and demographic factors. In this way, we hope to make progress in elucidating the neural bases of speech-in-noise comprehension that are determined by intrinsic brain properties, dissociated from cognitive factors.

It is thought that a first step in the analysis of sensory signals is their sampling and quantisation. Cortical oscillations, which are rhythmic fluctuations in neuronal synchrony at specific time scales, have been proposed as a key mechanism responsible for the processing of the incoming continuous speech stream. According to the influential asymmetric sampling in time hypothesis (AST; Giraud et al., 2007; Giraud & Poeppel, 2012; Poeppel, 2003), a "*principled relation*" (i.e., a quasi-isomorphism) exists between the time scales at which meaningful acoustic events in speech occur and those at which the speech perception system operates, by means of cortical oscillations. This hypothesis holds that speech processing is temporally-asymmetric in non-primary auditory areas, with left-hemisphere (LH) mechanisms (i.e., low-γ oscillations in the ~20-50Hz frequency range) preferentially extracting information over shorter (20-50ms) temporal windows and right-hemisphere (RH) mechanisms (i.e., θ oscillations in the ~3-7Hz frequency range) over longer (~150-300ms) windows (Poeppel, 2001, 2003). The slower time scale corresponds to the duration of slow spectrotemporal fluctuations associated with the alternation of syllables and the latter corresponds to a processing or sampling rate suitable for capturing the rapid acoustic features relevant for phonemic identification, such as formant transition and voice onset time (Delattre, Liberman, & Cooper, 1955; Kewley-Port, 1982; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Lisker & Abramson, 1964; Rosen, 1992).

The AST suggests that the evident functional asymmetry of non-primary auditory cortices is related to differences in the distribution of the centre frequency at which neuronal ensembles spontaneously synchronise. These are relatively more skewed towards synchronising at a θ rate in the RH and skewed towards synchronising at a low-γ pace in the LH. There is indeed some evidence for spectral peaks in θ and low-γ frequency ranges in resting electrophysiological activity within primary auditory areas, as observed, for instance, by Lakatos and collaborators (2005) by means of intracranial recordings in macaques. These spontaneous, ongoing oscillations were also shown to predict the activity evoked by noise bursts. Further, consistent with the slight asymmetry in spontaneous activity implied by the AST, it has been shown by means of simultaneous EEG and fMRI recordings in humans that spontaneous activity in the low-γ range at rest correlates best with LH auditory cortical synaptic activity (as indexed by the BOLD signal), whilst RS EEG power within the θ range correlates best with synaptic activity in the RH (Giraud et al., 2007). Finally, Lehongre and collaborators (2011) have shown asymmetries in peak oscillatory responses to rhythmic auditory stimulation in planum temporale and posterior superior temporal sulcus: whilst stronger responses where elicited in the RH by periodicities above 55Hz, the strongest responses in the LH were elicited by periodic auditory stimulation in the 25-35Hz range.

We propose that individual differences in WIN performance may be related to baseline excitability of the speech perception system. While there is evidence for a number of oscillatory regimes to be involved in the processing of language – with θ and low-γ being of paramount interest – evidence that the amplitude of spectrally-resolved spontaneous activity is able to predict a rather high-level speech perception task as WIN recognition has not been reported thus far. To date, the wide majority of studies on task-free brain activity have focused on examining its functional connectivity – i.e., the temporal correlation between spatially remote brain areas which is thought to represent communication – by means of functional magnetic resonance imaging (fMRI). By measuring coherent spontaneous low-frequency (i.e., <0.1Hz) fluctuations of the blood oxygenation level dependent signal, studies making use of fMRI have

demonstrated the existence of a number of large-scale neural networks, which highly resemble the activation and deactivation maps associated with performing perceptual, motor or cognitive tasks (see Van Den Heuvel & Pol, 2011 for a review). Here, in contrast, we use a relatively simple representation of the brain's RS characteristics in a bid to predict a specific behavioural ability. Importantly, rather than measuring information flow over large and distributed networks, as would be the case in connectivity analyses, we aimed at indexing local brain activity and how it relates to individual behavioural differences.

2. Material and methods

2.1. Participants and procedures. We analysed resting-state MEG data from 89 participants made available as part of the Human Connectome Project (HCP; Van Essen et al., 2013). Mean age was 28 years and 7 months (SD = 4 years, range = 22-35 years). Forty-one participants were female. All of the subjects had normal hearing. The full HCP protocol of participants undergoing MEG is typically completed in a three-day visit (see WU-Minn Consortium Human Connectome Project, 2017 for details). RS MEG data acquisitions always preceded WIN test by at least one full day.

2.2. Speech material. The Words-In-Noise test (WIN; Wilson, 2003; Wilson et al., 2007) was used to test participants' ability to recognise single words presented amid varying levels of background noise. The WIN establishes individual signal-to-noise ratio (SNR) thresholds for correct word recognition and, by so doing, measures how much difficulty a person might have listening to speech in a noisy environment. Specifically, participants are presented with isolated monosyllabic words embedded in varying levels of multitalker babble. They are asked to listen to and then repeat the words. The task becomes increasingly difficult as the background noise gets louder, thus reducing the SNR. Multitalker babble involves several (i.e., six; three male and three female) speakers talking simultaneously about various topics with all of the conservations being unintelligible (Sperry, Wiley, & Chial, 1997). Multitalker babble is the most common

environmental noise encountered by listeners in everyday life and is, therefore, highly ecologically valid (Plomp, 1978). Furthermore, it is more sensitive to mild hearing loss as compared to speech-spectrum noise (Findlay, 1976).

Thirty-five words are presented in blocks of five items whilst SNR decreases monotonically through seven levels (26, 22, 18, 14, 10, 6, 2 and -2 dB SNR). The test ends when all the seven lists of items have been completed or when all the five items of a given SNR are not correctly recognised. The examiner assigns one point for each item correctly recognised and no points to incorrect responses. The best raw score a participant can obtain is 35 (all correct) and the worst is zero (none correct). Raw scores are converted back to dB SNR threshold via the formula: *SNR threshold (dB) = 26 dB -.8 * raw score.* SNR thresholds can therefore take values from 26 dB (none correct) to -2 dB (all correct). In the sample we analysed SNR thresholds ranged from 7.6 to 2 dB (mean=4.7, SD=1.3). One subject had a missing WIN score and therefore scores from 88 participants were analysed. Of note, participants composing the current sample have been tested through either one of two versions of the NIH Toolbox WIN test (V1 and V2). Participants tested through V1 had stimuli delivered via speakers, whilst the majority of V2 participants had stimuli delivered monaurally through headphones, and the better score (lowest SNR) out of the two ears was used as participant score. V1 and V2 scores were normed to be directly comparable across the whole 1200-subjects release. We verified this is the case in our limited subset of 88 participants by testing for mean differences across WIN score in the two groups using the independent 2-group Mann-Whitney U Test (two-way). The test revealed that participants tested through WIN v1 (median = 5.2 dB SNR) did not perform differently from participants tested with v2 (median = 4.4 dB SNR), W=846, p=.939. Note, however, that some of the participants labelled as 'v2' could have actually been tested with v1 WIN.

2.2.1. Deconfounding procedure. We aimed to establish whether resting MEG power is specifically associated with the ability to recognise words in noise. However, several unspecific factors and variables including demographics (e.g., age) and general cognitive abilities (e.g., working memory, processing

speed), may contribute to (i.e., confound) each participants' SNR threshold (Frisina & Frisina, 1997; Humes et al., 1994). For this reason, we exploited the wide range of participant measures and information collected as part of the HCP to control for potential contributions of these factors. A detailed description of the measures collected as part of the assessment are described in detailed in Van Essen and collaborators (2013). In short, these measures include demographic information, neurological/psychiatric diagnoses, description of subjects' habits (e.g., sleep, drug consumption, etc.), and a wide range of perceptual, cognitive and emotional tests and questionnaires. Out of the 478 original subject measures (reported in the "open access" and "restricted" subject information spreadsheets which can be found here: http://humanconnectome.org/data), we selected those meeting the criteria outlined in Table 1 (adapted from Smith et al., 2015). This procedure resulted in the selection of 132 measures. Of these measures, nine were found to be significantly correlated with WIN SNR after FDR correction for multiple comparisons. Specifically, four measures of tobacco consumption over the seven days preceding assessment were negatively associated with resilience to noise (positively associated with WIN SNR): number of days using tobacco ($r=.297$, $p_{(FDR)}=.038$), average weekday cigarette consumption ($r=.312$, $p_{(FDR)}=.042$), average weekend tobacco consumption ($r=.306$, $p_{(FDR)}=.035$), and number of cigarettes smoked ($r=.306$, $p_{(FDR)}=.043$). This is likely due to the fact that tobacco consumption is associated with hearing loss (Cruickshanks et al., 1998). Further, five measures positively predicted resilience to noise (negatively correlated with WIN SNR), namely both age-adjusted and -unadjusted scores on NIH list sorting test ($r=-.314$, $p_{(FDR)}=.038$, and $r=-.312$, $p_{(FDR)}=.043$, respectively), both age-adjusted and -unadjusted scores on NIH flanker task ($r=-.299$, $p_{(FDR)}=.039$, and $r=-.279$, $p_{(FDR)}=.048$, respectively), and unadjusted score on NIH dexterity test ($r=-.292$, $p_{(FDR)}=.036$). It is interesting to note that working memory and executive functions have been highlighted as cognitive contributors to degraded speech comprehension (Rönnberg et al., 2013) and that response inhibition seems to be particularly required to perceive speech-in-noise when noise is verbal as compared to non-verbal (Rouleau & Belleville, 1996), as is the case in the

context of the WIN test (multitalker babble). We nonetheless regressed out every available variable, including those that were not significantly individually associated with the measure of interest because the subject measures taken together are sources of individual differences that we wished to remove, as we aimed at establishing relationships that are specific to WIN. In order to avoid overfitting in the regression, principal component analysis (PCA) was used to reduce the dimensionality of the set of confounds. Before entering the PCA, variables were mean-centred (by subtracting the sample mean from each individual score) and by using rank-based inverse normal transformation we ensured gaussian-like distribution of variables. The statistical rationale for transformation of precipitation variables towards a normal distribution prior to PCA is to facilitate a meaningful quantification of variance, and to stabilize the variance (reduce heteroskadacity). The ecological rationale for transformation of precipitation is that the meaning of the units changes across the range of values. Twenty-five principal components, accounting for 80% of the total variance of the set of 132 measures, were regressed out from the original WIN score following the steps described by Smith and collaborators (2015).

Table 1. **Criteria for the selection of subject measures entering the PCA (adapted from Smith et al., 2015).**

Criterion	*Formal requirement*
Subject measure data need to have enough variability	1. Standard deviation (*SD*) > 0 2. At least 49 valid subject scores 3. The number of values in the largest group of unique values does not need to exceed 80% of the total number of valid values
Subject measure data need to have excessively high leverage points	4. If x_s is a subject measure value for subject $_s$, and $y_s=(x_s-median(x_s))^2$, the subject measure does not need to contain any x_s for which $max(y_s) > 100 \times mean(y_s)$.

2.3. RS-MEG data. RS-MEG data was acquired in three runs of approximately six minutes per subject during which participants are asked to remain supine with eyes open. Details on acquisition protocols and hardware specifications can be found in WU-Minn Consortium Human Connectome Project, 2017. Briefly,

subjects were scanned on a whole head MAGNES 3600 (4D Neuroimaging, San Diego, CA) system housed in a quiet, darkened and magnetically shielded room. The MEG system includes 248 magnetometer channels together with 23 reference channels. Data were recorded at a sampling rate of 2.035 kHz and resampled at 200 Hz. Raw MEG data underwent the standard HCP MEG data pre-processing pipeline based on independent component analysis (ICA) aimed at identifying and removing artefacts, bad channels and bad segments (2s each). Additionally, single shell volume conduction models (one per participant) were defined in the MEG-system based head coordinate system from segmentation of anatomical MRI T1 images. We computed the inverse solution using Linearly Constrained Minimum Variance (LCMV) Beamformer filters. These allow source models to be defined on a regular 3D grid in normalized MNI-space with a resolution of 8mm, aligning the subjects in source space. Welch's method was used to obtain the power spectral density (PSD) of each of the source-reconstructed and standardised (over time to mean=0 and SD=1) time series session data (three per participant). The two-sided PSD was computed on eight equally-long time series windows with 50% overlap from 1 to 40 Hz, in 1-Hz bins at each source point (3559 voxels of size 8*8*8mm each). PSDs of the three sessions were then averaged within participants for each voxel.

2.3.1. Definition of frequency bands. In order to accommodate the inherent similarity of the power distribution in neighbouring frequencies, *k*-means clustering (1000 replicates, using correlation distances) was used to cluster the 40 frequency bins into 6 wider bins based on the similarity of the spatial distribution of voxelwise power. Evaluation of the optimal clustering solution was based on a two-step procedure. First, whole-brain voxelwise power cross-correlation matrix was inspected visually in order to determine a range of *k*s to test. Once the range of clusters to evaluate was established (4 to 6), the optimal solution was determined using two measures of internal clustering validation measures, namely Dunn's Index (DI; Dunn, 1973), and Silhouette Index (SI; Rousseeuw, 1987). Both indices indicated that *k*=6 represented the best clustering solution. The final frequency bins – mainly corresponding to conventional

frequency bands – were as follows: 1-4Hz ('δ'), 5-7Hz ('θ'), 8-15 ('α'), 16-20 ('low-β'), 21-29Hz ('high-β'), and 30-40Hz ('low-γ').

2.4. Statistical Analyses. Within-subject means of voxelwise power in each frequency band were correlated with the standardised deconfounded WIN score. We evaluated the statistical significance of the clusters by means of permutation testing. Because the sample analysed here is composed by subgroups of monozygotic and dizygotic siblings in addition to non-related participants, we took into account the possible effect of family structure on the null distribution of correlations generated by permutation by randomly shuffling WIN scores (10000 times) across participant labels in a manner constrained by family structure. Multi-level block permutation (Winkler, Webster, Vidaurre, Nichols, & Smith, 2015) was performed using FSL's PALM utility (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM), and the exchangeability block structure specified in the HCP documentation (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM/ExchangeabilityBlocks). For each frequency band (n=6) and voxel (n=3559), a *t*-value was calculated for both the actual and the data-driven null distribution; *t*-values were then compared to produce *p*-values. Below we report results that are significant at a voxelwise level of $p<.05$, corrected for multiple comparisons using threshold-free cluster enhancement (TFCE), a statistical approach that determines significance by weighting each voxel's effect magnitude by the 'support' received from neighbouring voxels in terms of both spatial extent and effect magnitude (Mensen & Khatami, 2013; Smith & Nichols, 2009). The general principle of TFCE is to increase the value (*t*-value in the current instance) of highly supported voxels and to decrease the value of weakly supported voxels.

3. Results

RS activity was found to be predictive of WIN performance in two adjacent frequency ranges (Table 2, Figure 1). In both the high-β (21-29Hz) and low-γ (30-40Hz) bands, we observed significant negative

correlations between power and WIN SNR, indicative of a positive association between amplitude of the oscillatory activity and the ability to be resilient to noise-induced breakdown of intelligibility. In the high-β frequency band, we observed a large cluster of significant association in the LH spanning the inferior frontal gyrus until the temporo-parietal junction. This large cluster comprises a number of cortical regions that have been consistently recognized as being involved in the processing of various features of speech. Most of these cortical sites are associated with processing along the ventral 'speech recognition' stream of dual stream models of speech perception (Hickok & Poeppel, 2007). The peak effect was observed in the posterior superior temporal gyrus (pSTG). We also observed a more constrained cluster in terms of both spatial extent and magnitude of the effect in the RH, centred on the pSTG. In the low-γ band, effects were spatially more constrained, but were again characterised by a slight leftward asymmetry. Whilst significant correlations were confined to the pSTG in the right hemisphere, they were slightly more extended anteriorly and inferiorly in the left hemisphere.

4. Discussion

This analysis has revealed that topographically- and spectrally-resolved RS MEG power – an index of spontaneous neuronal synchronisation at various time scales – can predict behavioural performance. This analytical approach offers an alternative and complementary perspective on mechanisms of brain functioning as compared to activation studies. Indeed, whilst in the latter type of studies, individual differences in brain responses to stimulation are generally treated as random noise and often ascribed to confounding 'volatile' factors, such as participants' momentary alertness level, here we specifically tested the functional relevance of such differences. Of note, the observed effects cannot be accounted for by a spurious association with volatile confounds, since the electrophysiological and behavioural assessments took place in different contexts, day(s) apart. Furthermore, we removed the potential confounding effects

Table 2. Loci of significant correlations between RS brain activity and WIN. Anatomical labels were obtained labels using in-house routines to assign identity according to the nearest labelled coordinate in the AAL template (Tzourio-Mazoyer et al., 2002), Brodmann area labels for the peak co-ordinates were extracted, where available, from the BA template provided with MRIcron (https://www.nitrc.org/projects/mricron).

Frequency band (Hz)	*Cluster ID*	*N voxels*	*AAL Label*	*Brodmann Area*	*x,y,z (MNI)*	*r(uncorr.)*	*p(fwe)*
high-β (21-29Hz)	1	402	Left superior temporal gyrus	42	-52,-40,16	0.362	0.026
			Left anterior insula		-28,16,8	0.255	0.033
			Left Rolandic operculum		-44,0,16	0.263	0.034
			Left superior occipital gyrus	18	-20,-72,32	0.272	0.037
			Left inferior parietal lobule	7	-28,-48,56	0.257	0.038
			Right middle cingulate gyrus	23	4,-16,40	0.187	0.041
			Right middle cingulate gyrus		4,-32,48	0.226	0.041
			Left precuneus	7	-4,-56,56	0.274	0.041
			Left middle occipital gyrus	19	-28,-80,0	0.263	0.043
	2	15	Right superior temporal gyrus	22	52,-32,8	0.299	0.044
low-γ (30-40Hz)	1	29	Right superior temporal gyrus	9	-44,-64,16	0.203	0.042
			Left middle temporal gyrus		-28,-48,64	0.204	0.044
			Left superior parietal lobule	7	-20,-72,48	0.185	0.047
	2	8	Left superior parietal lobule	2	60,-32,16	0.218	0.04

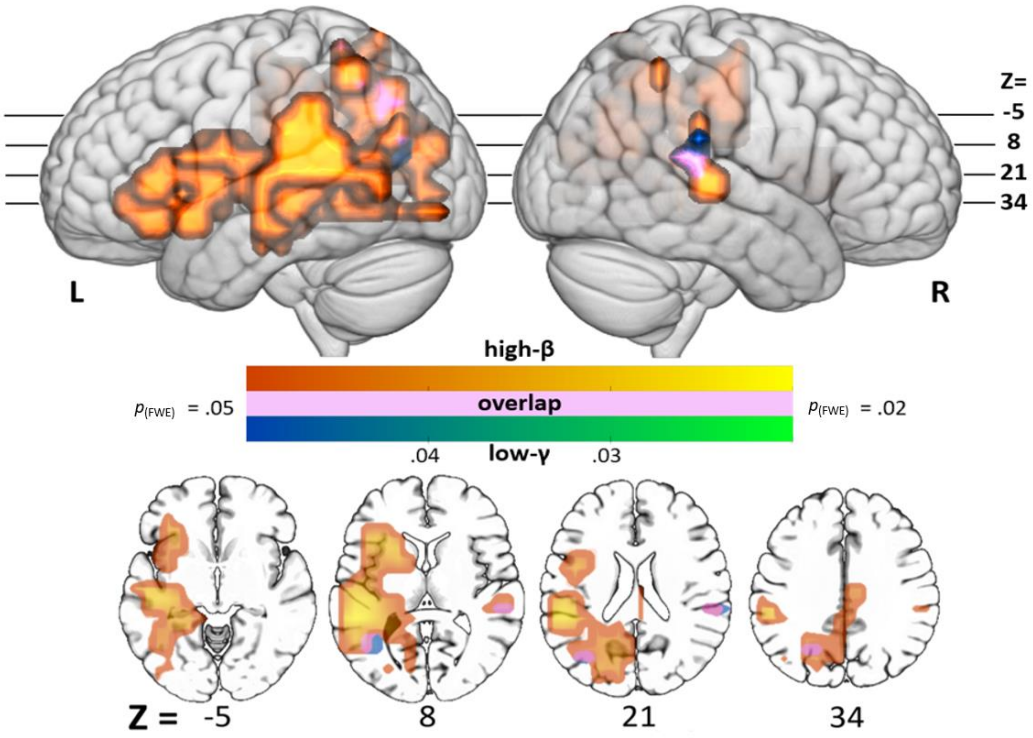

Figure 1. Axial (top panel) and surface (bottom panel) view of the correlation clusters.

of a substantial array of cognitive and demographic factors, ensuring that the results we report are specific to the relevant behavioural domain, namely speech processing. By excluding potential confounds from even related cognitive domains, such as verbal working memory and vocabulary we may legitimately conclude that these brain-behaviour relationships relate to the auditory and speech-processing specific aspects of the task. Thus, we propose that individual differences in WIN performance are related to baseline excitability levels of the speech perception system, and that the locus of this effect is early in the speech-processing hierarchy. Specifically, we suggest that increased power in spectrally- and topographically-specific MEG power at rest, by determining a suitable context in which relevant acoustic features (i.e., cues to phonemic identity) are processed, are associated with better WIN recognition abilities.

This proposal is consistent with a number of lines of evidence. The locus of the LH high-β peak is consistent with the region in which categorical phoneme perception has been demonstrated in fMRI (e.g., Specht, Baumgartner, Stadler, Hugdahl, & Pollmann, 2014; see Turkeltaub & Branch Coslett, 2010 for a review). Evidence from intracranial recordings also supports this. For example, Chang and colleagues (2010) carried out intracranial recordings of STG responses to synthesised phonemic continua and found neural populations in pSTG responding to phonemic cues important for phoneme recognition. In addition, they observed populations of neurons responding selectively and categorically to single phonemes, providing a strong physiological correspondence to the behavioural categorical perception phenomenon. A further study (Mesgarani, Cheung, Johnson, & Chang, 2014) has shown that neuronal populations discretely responding to phoneme categories are hierarchically clustered in a way that is strikingly similar to the traditional clustering of phonemes performed by phoneticians and linguists. Mesgarani and colleagues (2014) conclude that selectivity is determined primarily by acoustic cues for manner and secondarily by place of articulation (and, in general, by articulatory properties). Thus, not only is the leftward perisylvian lateralisation of the effects in line with speech-related processes, but the topography of the peak effects

suggests a specific involvement of the spontaneous activity highlighted here in phoneme-level processing: in both the high-β and low-γ bands, peak effect was found in the pSTG (in the LH and RH, respectively).

Brain-behaviour relationships observed in this study were confined to two neighbouring frequency bands (21-29 and 30-40Hz). Different lines of inquiry indicate that neural activity within this range of the spectrum is important for the parsing and processing of acoustical information on short time scales relevant for phoneme recognition, i.e., phonetic features. Among the work supporting this hypothesis, relatively strong evidence comes from a growing body of research on exogenous modulation of neural excitability. Recent studies using transcranial alternating current stimulation (tACS) have begun to provide causal evidence for the role of oscillations in this frequency range for phoneme perception. The oscillatory current produced by tACS is assumed to instantaneously synchronize (i.e., to entrain) the firing the targeted neuronal populations, enhancing oscillatory power at the stimulation frequency (Helfrich et al., 2014). A recent study using tACS reported the induction of phoneme recognition deficits in healthy participants as a result of applying a 40Hz current to auditory areas (Rufener, Zaehle, Oechslin, & Meyer, 2016). In a separate study, Rufener, Zaehle and colleagues (2016) found that applying 40Hz stimulation to the bilateral auditory cortex reduced the rate of perceptual learning during a phoneme categorisation task, while applying 6Hz tACS stimulation did not affect perceptual learning. Importantly, whilst this result was replicated in a comparable sample of young participants (20-35 years), it was shown that perceptual learning of older participants (60-75 years) was enhanced, rather than hindered, by 40Hz stimulation and the presumed tACS-induced enhancement of 40Hz activity, with 6Hz stimulation again having no significant effect (Rufener, Oechslin, Zaehle, & Meyer, 2016). Intriguingly, Rufener and colleagues suggest that the effect of perturbing low-γ activity may depend on (i.e., interact with) its 'pre-existing' (i.e., RS) level.

The idea that there may be such a thing as an optimal resting-state level for phoneme processing is further supported by experiments on developmental dyslexia (DD). DD is a common disability defined as sub-

normal acquisition of reading skills and phonological abilities within the context of otherwise normal hearing and cognitive abilities. Rufener and colleagues (Rufener, Krauel, Meyer, Heinze, & Zaehle, 2019) showed that 40Hz tACS stimulation over the bilateral auditory cortex improved phoneme-categorisation accuracy in a sample of patients suffering from DD. Thus, by altering the putatively optimal balance of synchronous endogenous activity in healthy individuals unaffected by phonological deficits, sampling of rapid acoustic events is altered and performance is worsened. In contrast, when the equilibrium of oscillatory activity is sub-optimal, as might be the case in the context of healthy aging (Voytek et al., 2015) and DD (Hancock, Pugh, & Hoeft, 2017), abnormal, or sub-optimal, temporal sampling is remediated by the external stimulation, and performance improves. Taken together, evidence from these tACS studies strongly supports a causal involvement of low-γ activity in the rapid auditory processing hypothesised to be essential for phoneme extraction and identification.

A large body of evidence suggests that the phonological deficits that characterise DD may arise from abnormal sampling of the speech stream into the short time windows relevant for the processing of phonetic cues, (Chobert, François, Habib, & Besson, 2012; Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007; Raschle, Stering, Meissner, & Gaab, 2014; Snowling, 1998). Consequently, DD is increasingly characterised as a disruption of rapid auditory processing (RAP). Lehongre and colleagues (Lehongre et al., 2011) tested the auditory steady state response (ASSR) of individuals affected by DD and a group of controls. ASSR reflects an intrinsic and stable property of the brain that is relevant for auditory processing (Baltus & Herrmann, 2015). Lehongre and colleagues (2011) found that the magnitude of left lateralised ASSR to 30Hz amplitude-modulated stimuli was reduced compared to controls, and that the magnitude of the ASSR in planum temporale at 30Hz in both DD and control participants predicted performance on phonological tasks. They take this to be a reflection of an intrinsic property of the auditory system and conclude that an abnormal oscillatory rhythm in left auditory areas in the range 25-35Hz is responsible for DD. This demonstrates the relevance of this frequency range at this cortical area, for speech

perception, consistent with our finding that RS activity in this frequency region at this anatomical locus predicts WIN performance. Further, deviance from normal ASSRs in the 25-35Hz range were observed by Lehongre and collaborators also in the left prefrontal cortex and in the right posterior superior temporal cortex, showing a high spatial similarity with regions that were correlated with WIN recognition abilities in our sample of healthy participants. Unfortunately, the boundaries of this frequency range exactly matches the centre of two different frequency bands defined in this study (21-29Hz and 30-40Hz, respectively), slightly complicating the comparison. Nevertheless, the ASSR is known to peak higher in the RH than LH (Lehongre et al., 2011; Ross, Herdman, & Pantev, 2005), and that the effect we see may be related to this asymmetry. Indeed, it is possible that our results indicate that higher power at the peak ASSR frequency of the auditory system at rest is beneficial for speech perception. Under the assumption that the human auditory system has co-evolved with speech and is thus highly tuned to the acoustic properties of the latter, it may not be too much of a stretch to suggest that higher resting-state activity at the peak of the local neural tuning may provide a superior neural context for speech stimulus processing, underpinning improved speech in noise perception.

We acknowledge that the present study is subject to a number of limitations. One issue is the task that was used to evaluate speech in noise perception, namely the words in noise test. This presents participants only with a series of unconnected monosyllabic words. Use of such a set precluded investigation of another oscillatory component that has recently been tied to speech comprehension, namely the θ band (Ding, Melloni, Zhang, Tian, & Poeppel, 2015; Peelle & Davis, 2012). According to a number of reports and the predictions of the AST, θ underpins the syllabic parsing based on envelope cues, which are neither relevant nor present in the WIN. Further investigations of individual differences in performance on noisy connected speech materials would be necessary to evaluate whether RS θ may also contribute to the comprehension of acoustically challenging speech under slightly more ecologically valid conditions.

5. Conclusion

We have observed that analysis of spectrally-resolved RS-electrophysiology can provide valuable insights into the spatial and spectral properties of cerebral systems that are specific to WIN recognition. In this regard, the present results are consistent with the growing body of research indicating that the interactions between resting and evoked activity need to be better characterised in order to obtain a more complete and accurate account of cerebral functioning. In particular, beyond advancing our understanding of properties intrinsic to the brain that determine speech comprehension, a number of implications arise from this investigation. First, it suggests, together with a number of studies presented above, that altering resting oscillatory activity associated with RAP may constitute a viable approach for overcoming phonological processing deficits that arise from sub-optimal resting-state oscillatory activity patterns. This has already been shown to be the case within the realm of non-invasive electrical stimulation, but other interventions aimed at altering the balance of endogenous oscillatory activity associated with phonological processing can be conceived, such as EEG-informed neurofeedback, or closed-loop neurostimulation, targeting the spectral and spatial regions highlighted here. Furthermore, since the affordances of such interventions are inevitably tightly linked to pre-intervention activity levels, characterisation of RS activity in the normal population and in populations of patients affected by phonological impairments is a necessary step for the development of effective interventions. Thus, the scope of the analysis of neural correlates of RS power extends well beyond fundamental research and has important implications for applied research into how to support and enhance the auditory brain's resilience in the face of noisy signals.

Statement of Significance

We show that resting-state power, measured by MEG, in auditory cortices and left perisylvian areas is predictive of individual differences in speech in noise recognition performance. Intrinsic brain properties indexed by resting-state activity are behaviourally relevant, and provide further insight into the neural substrates of speech in noise processing capacity.

Acknowledgements

We are very grateful to the WU-Minn HCP Consortium for generating and publicly sharing data and implementing the procedures needed to acquire, analyse, visualise and share these datasets.

Funding

Research funded by the Swiss National Science Foundation (grant PP_163726 awarded to A.H-A).

Declaration of Interests:

The authors declare no competing interests.

References:

Baltus, A., & Herrmann, C. S. (2015). Auditory temporal resolution is linked to resonance frequency of the auditory cortex. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *98*(1), 1–7. https://doi.org/10.1016/j.ijpsycho.2015.08.003

Biswal, B. B., Mennes, M., Zuo, X., Gohel, S., Kelly, C., Smith, S. M., … Milham, M. P. (2009). Toward discovery science of human brain function. *PNAS*, *107*(10), 4734–4739. https://doi.org/10.1073/pnas.0911855107

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*(5679), 1926–1929. https://doi.org/10.1126/science.1099745

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience*, *13*(11), 1428–1432. https://doi.org/10.1038/nn.2641

Chobert, J., François, C., Habib, M., & Besson, M. (2012). Deficit in the preattentive processing of syllabic duration and VOT in children with dyslexia. *Neuropsychologia*, *50*(8), 2044–2055. https://doi.org/10.1016/j.neuropsychologia.2012.05.004

Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nat Neurosci*, *19*(12), 1718–172. https://doi.org/doi:10.1038/nn.4406

Cruickshanks, K. J., Klein, R., Klein, B. E. K., Wiley, T. L., Nondahl, D. M., & Tweed, T. S. (1998). Cigarette Smoking and Hearing Loss: The Epidemiology of Hearing Loss. *Jama*, *279*(21), 1715. https://doi.org/10.1001/jama.279.21.1715

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic Loci and Transitional Cues for Consonants. *The Journal of the Acoustical Society of America*, *27*(4), 769–773. https://doi.org/10.1121/1.1908024

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, *3*(3), 32–57. https://doi.org/10.1080/01969727308546046

Findlay, R. C. (1976). Auditory Dysfunction Accompanying Noise-Induced Hearing Loss. *Journal of Speech and Hearing Disorders*, *41*(3), 374–380. https://doi.org/10.1044/jshd.4103.374

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., … Constable, T. R. (2016). Functional connectome fingerprinting: Identifying individuals based on patterns of brain connectivity Emily. *Nat Neurosci*, *18*(11), 1664–1671. https://doi.org/10.1038/nn.4135.Functional

Fox, M. D., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2007). Intrinsic Fluctuations within Cortical Systems Account for Intertrial Variability in Human Behavior. *Neuron*, *56*(1), 171–184. https://doi.org/10.1016/j.neuron.2007.08.023

Fox, M. D., Snyder, A. Z., Zacks, J. M., & Raichle, M. E. (2006). Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci*, *9*(1), 23–25. https://doi.org/10.1038/nn1616

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research*, *106*(1–2), 95–104. https://doi.org/10.1016/S0378-5955(97)00006-3

Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study. *Restorative Neurology and Neuroscience*, *25*(3–4), 295–310. https://doi.org/17943007

Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous Cortical Rhythms Determine Cerebral Specialization for Speech Perception and Production. *Neuron*, *56*(6), 1127–1134. https://doi.org/10.1016/j.neuron.2007.09.038

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Hancock, R., Pugh, K. R., & Hoeft, F. (2017). Neural Noise Hypothesis of Developmental Dyslexia. *Trends in Cognitive Sciences*, *21*(6), 434–448. https://doi.org/10.1016/j.tics.2017.03.008

Helfrich, R. F., Rach, S., Trautmann-Lengsfeld, S. A., Schneider, T. R., Herrmann, C. S., & Engel, A. K. (2014). Entrainment of Brain Oscillations by Transcranial Alternating Current Stimulation. *Current Biology*, *24*(3), 333–339. https://doi.org/10.1016/j.cub.2013.12.041

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, *8*(393), 393–402. https://doi.org/10.7554/eLife.14521

Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., & Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *Journal of Speech and Hearing Research*, *37*(April), 465–474. https://doi.org/10.1044/jshr.3702.465

Kewley-Port, D. (1982). Measurement of formant transitions in naturally produced stop consonant–vowel syllables. *The Journal of the Acoustical Society of America*, *72*(2), 379–389. https://doi.org/10.1121/1.388081

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., Schroeder, C. E., … Ulbert, I. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*, 1904–1911. https://doi.org/10.1152/jn.00263.2005

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, *72*(6), 1080–1090. https://doi.org/10.1016/j.neuron.2011.11.002

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431. https://doi.org/10.1037/h0020279

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Mennes, M., Kelly, C., Zuo, X., Martino, A. Di, Biswal, B., Castellanos, X., & Milham, M. P. (2010). Inter-Individual Differences in Resting State Functional Connectivity Predict Task-Induced BOLD Activity. *NeuroImage*, *50*(4), 1690–1701. https://doi.org/10.1016/j.neuroimage.2010.01.002

Mennes, M., Zuo, X.-N., Kelly, C., Di Martino, A., Zang, Y.-F., Biswal, B., … Milham, M. P. (2011). Linking inter-individual differences in neural activation and behavior to intrinsic brain dynamics. *NeuroImage*, *54*(4), 2950–2959. https://doi.org/10.1016/j.neuroimage.2010.10.046.Linking

Mensen, A., & Khatami, R. (2013). Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics. *NeuroImage*, *67*, 111–118. https://doi.org/10.1016/j.neuroimage.2012.10.027

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science*, *1006*(January), 1–6. https://doi.org/10.1126/science.1245994

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

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, *63*(2), 533–549. https://doi.org/10.1121/1.381753

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, *25*, 679–693. https://doi.org/https://doi.org/10.1207/s15516709cog2505_3

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as “asymmetric sampling in time.” *Speech Communication*, *41*(1), 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3

Raschle, N. M., Stering, P. L., Meissner, S. N., & Gaab, N. (2014). Altered neuronal response during rapid auditory processing and its relation to phonological processing in prereading children at familial risk for dyslexia. *Cerebral Cortex*, *24*(9), 2489–2501. https://doi.org/10.1093/cercor/bht104

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., … Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*(31), 1–17. https://doi.org/10.3389/fnsys.2013.00031

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *336*(1278), 367–373. https://doi.org/10.1098/rstb.1992.0070

Ross, B., Herdman, A. T., & Pantev, C. (2005). Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cerebral Cortex*, *15*(12), 2029–2039. https://doi.org/10.1093/cercor/bhi078

Rouleau, N., & Belleville, S. (1996). Irrelevant speech effect in aging: An assessment of inhibitory processes in working memory. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *51*(6), 356–363. https://doi.org/10.1093/geronb/51B.6.P356

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational And Applied Methematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Rufener, K. S., Krauel, K., Meyer, M., Heinze, H. J., & Zaehle, T. (2019). Transcranial electrical stimulation improves phoneme processing in developmental dyslexia. *Brain Stimulation*. https://doi.org/10.1016/j.brs.2019.02.007

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 Stimulation*, *9*(4), 560–565. https://doi.org/10.1016/j.brs.2016.04.002

Rufener, K. S., Zaehle, T., Oechslin, M. S., & Meyer, M. (2016). 40 Hz-Transcranial alternating current stimulation (tACS) selectively modulates speech perception. *International Journal of Psychophysiology*, *101*, 18–24. https://doi.org/10.1016/j.ijpsycho.2016.01.002

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement : Addressing problems of smoothing , threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E. J., Glasser, M., … Miller, K. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience*, *18*(11), 1565–1567. https://doi.org/10.1038/nn.4125.A

Snowling, M. (1998). Dyslexia as a Phonological Deficit: Evidence and Implications. *Child Psychology and Psychiatry Review*, *3*(1), 4–11. https://doi.org/10.1017/S1360641797001366

Specht, K., Baumgartner, F., Stadler, J., Hugdahl, K., & Pollmann, S. (2014). Functional asymmetry and effective connectivity of the auditory system during speech perception is modulated by the place of articulation of the consonant- A 7T fMRI study. *Frontiers in Psychology*, *5*(June), 1–10. https://doi.org/10.3389/fpsyg.2014.00549

Sperry, J. L., Wiley, T. L., & Chial, M. R. (1997). Word recognition performance in various background competitors. *Journal of the American Academy of Audiology*, *8*(2), 71–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9101453

Steinschneider, M., Volkov, I. O., Noh, M. D., Garell, P. C., & Howard, M. A. (1999). Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *Journal of Neurophysiology*, *82*(5), 2346–2357. https://doi.org/10.1152/jn.1999.82.5.2346

Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 1773–1776.

Turkeltaub, P. E., & Branch Coslett, H. (2010). Localization of sublexical speech perception components. *Brain and Language*, *114*(1), 1–15. https://doi.org/10.1016/j.bandl.2010.03.008

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Van Den Heuvel, M. P., & Pol, H. E. H. (2011). Exploring the brain network: A review on resting-state fMRI functional connectivity. *Psiquiatria Biologica*, *18*(1), 28–41. https://doi.org/10.1016/j.psiq.2011.05.001

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Voytek, X. B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1 / f Neural Electrophysiological Noise, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Wilson, R. H. (2003). Development of a Speech-in-Multitalker- Babble Paradigm to Assess Word-Recognition Performance. *Journal of the American Academy of Audiology*, *14*(9), 453–470.

Wilson, R. H., Carnell, C. S., & Cleghorn, A. L. (2007). The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *Journal of the American Academy of Audiology*, *18*(6), 522–529. https://doi.org/10.3766/jaaa.18.6.7

Winkler, A. M., Webster, M. A., Vidaurre, D., Nichols, T. E., & Smith, S. M. (2015). Multi-level block permutation. *NeuroImage*, *123*, 253–268. https://doi.org/10.1016/j.neuroimage.2015.05.092

WU - Minn Consortium Human Connectome Project. (2017). WU-Minn HCP 1200 Subjects Data Release: Reference Manual, *2017*(June), 1–169. https://doi.org/http://www.humanconnectome.org/documentation/S900/

Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, *134*(3), 2225–2234. https://doi.org/10.1121/1.4817926