BRAIN COMMUNICATIONS

Inaccurate cortical tracking of speech in adults with impaired speech perception in noise

Marc Vander Ghinst,[1,2] Mathieu Bourguignon,[1,3,4] Vincent Wens,[1,5] Gilles Naeije,[1,6] Cecile Ducène,[1,2] Maxime Niesen,[1,2] Sergio Hassid,[2] Georges Choufani,[2] Serge Goldman[1,5] and Xavier De Tiège[1,5]

Impaired speech perception in noise despite normal peripheral auditory function is a common problem in young adults. Despite a growing body of research, the pathophysiology of this impairment remains unknown. This magnetoencephalography study characterizes the cortical tracking of speech in a multi-talker background in a group of highly selected adult subjects with impaired speech perception in noise without peripheral auditory dysfunction. Magnetoencephalographic signals were recorded from 13 subjects with impaired speech perception in noise (six females, mean age: 30 years) and matched healthy subjects while they were listening to 5 different recordings of stories merged with a multi-talker background at different signal to noise ratios (No Noise, +10, +5, 0 and −5 dB). The cortical tracking of speech was quantified with coherence between magnetoencephalographic signals and the temporal envelope of (i) the global auditory scene (i.e. the attended speech stream and the multi-talker background noise), (ii) the attended speech stream only and (iii) the multi-talker background noise. Functional connectivity was then estimated between brain areas showing altered cortical tracking of speech in noise in subjects with impaired speech perception in noise and the rest of the brain. All participants demonstrated a selective cortical representation of the attended speech stream in noisy conditions, but subjects with impaired speech perception in noise displayed reduced cortical tracking of speech at the syllable rate (i.e. 4–8 Hz) in all noisy conditions. Increased functional connectivity was observed in subjects with impaired speech perception in noise in Noiseless and speech in noise conditions between supratemporal auditory cortices and left-dominant brain areas involved in semantic and attention processes. The difficulty to understand speech in a multi-talker background in subjects with impaired speech perception in noise appears to be related to an inaccurate auditory cortex tracking of speech at the syllable rate. The increased functional connectivity between supratemporal auditory cortices and language/attention-related neocortical areas probably aims at supporting speech perception and subsequent recognition in adverse auditory scenes. Overall, this study argues for a central origin of impaired speech perception in noise in the absence of any peripheral auditory dysfunction.

1 Laboratoire de Cartographie fonctionnelle du Cerveau, UNI—ULB Neuroscience Institute, Université Libre de Bruxelles (ULB), Brussels 1070, Belgium
2 Service, d'ORL et de chirurgie cervico-faciale, CUB Hôpital Erasme, Université Libre de Bruxelles (ULB), Brussels 1070, Belgium
3 Laboratory of Neurophysiology and Movement Biomechanics, UNI—ULB Neuroscience Institute, Université Libre de Bruxelles (ULB), Brussels 1070, Belgium
4 Basque Center on Cognition, Brain and Language (BCBL), Donostia/San Sebastian 20009, Spain
5 Clinics of Functional Neuroimaging, Service of Nuclear Medicine, CUB Hôpital Erasme, Université Libre de Bruxelles (ULB), Brussels 1070, Belgium
6 Service de Neurologie, ULB-Hôpital Erasme, Université libre de Bruxelles (ULB), Brussels 1070, Belgium

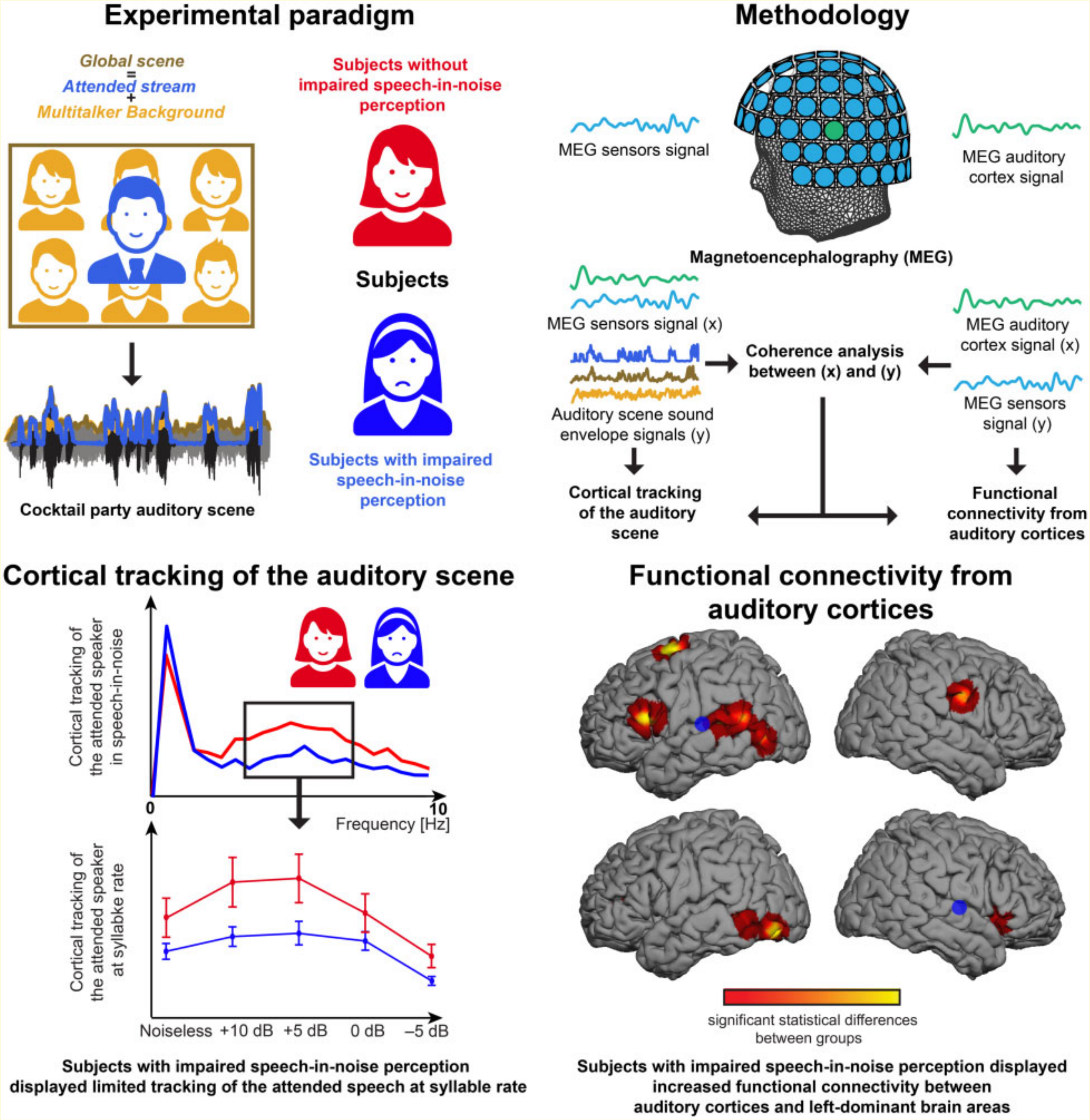
Correspondence to: Marc Vander Ghinst
Laboratoire de Cartographie fonctionnelle du Cerveau, UNI—ULB Neuroscience Institute
Université Libre de Bruxelles (ULB), 808 Lennik Street
1070 Brussels, Belgium.
Email: marc.vander.ghinst@erasme.ulb.ac.be

Received January 13, 2021. Revised June 5, 2021. Accepted June 8, 2021. Advance Access publication September 10, 2021

Keywords: magnetoencephalography; coherence analysis; functional connectivity; speech-in-noise

Abbreviations: dB=decibel; EHI =Edinburgh handedness inventory; FC =functional connectivity; FEF =frontal eye field; Hz =hertz; IFG =inferior frontal gyrus; IOG =inferior occipital gyrus; ISPiN =impaired speech perception in noise in the absence of any peripheral auditory dysfunction; ITG =inferior temporal gyrus; MEG =magnetoencephalography; MFG =middle frontal gyrus; MNE =minimum-norm estimates; MNI =Montreal Neurological Institute; MTG =middle temporal gyrus; Op =operculum; SiN =speech in noise; SNR =signal-to-noise ratio; STG =superior temporal gyrus; STS =superior temporal sulcus; TPJ =temporo-parietal junction

Graphical Abstract



Subjects with impaired speech-in-noise perception displayed limited tracking of the attended speech at syllable rate

Subjects with impaired speech-in-noise perception displayed increased functional connectivity between auditory cortices and left-dominant brain areas

Introduction

Experiencing difficulties to understand speech in noise (SiN) is a common symptom in adults with no peripheral hearing loss.[1–7] Hearing complaints in these subjects are comparable to those with mild to moderate hearing loss, leading to similar socio-professional difficulties.[8–10] Impaired speech perception in noise (ISPiN) in the absence of any peripheral auditory dysfunction is also known as the King–Kopetzky syndrome,[11] central auditory processing disorder,[12] hidden hearing loss[13] or obscure auditory dysfunction.[8]

Despite its high prevalence and socio-professional implications, no consensual pathophysiological mechanism explains ISPiN in subjects with normal peripheral auditory function. Several speculative mechanisms involving different elements of the auditory pathways have been proposed.[14] Lesion of outer hair cells[15,16] or loss of synapses between inner ear cells and auditory nerve fibres (i.e. cochlear synaptopathy)[17] might result in subtle electrophysiological changes not captured by typical evaluation of peripheral hearing but compatible with ISPiN clinical pattern. However, human studies failed to demonstrate such electrophysiological changes.[18,19] ISPiN has been associated with different neurological conditions such as attention deficit hyperactivity disorder,[20] dyslexia,[21,22] depression,[10,23] traumatic brain injury[24–26] or HIV-related neurocognitive disorders.[27] These findings suggest that ISPiN might be related to deficits in the central processing of auditory information.

In multi-talker auditory scenes (i.e. the '*cocktail party*' effect), auditory cortex activity selectively tracks the temporal envelope (i.e. amplitude modulations) of the attended speaker's voice rather than those of the global auditory scene.[28–31] This coupling occurs at frequencies below 10 Hz and decreases when the speech noise level increases.[30] Given that this frequency range matches with prosodic stress/phrasal/sentential (<1 Hz), word (1–4 Hz) and syllable (4–8 Hz) repetition rates, the corresponding cortical tracking of speech has been hypothesized to subserve the chunking of the continuous verbal flow into relevant hierarchical segments used for further speech recognition, up to a certain noise level.[30,32–34]

Behavioural studies have demonstrated that children aged <10 years have an inherent difficulty to understand speech in noisy conditions, such as multi-talker backgrounds.[35,36] Such behavioural difficulty has been related to an immaturity of the selective cortical tracking of the attended speech stream in noisy conditions, especially at the syllable rate.[31] Based on this observation, the present study tested the hypothesis that inaccurate low-frequency cortical tracking of SiN is a core mechanism of ISPiN. This assumption was tested using magnetoencephalography (MEG) in strictly selected adults with ISPiN and matched healthy subjects while they listened to connected speech recordings mixed with a multi-talker babble noise at different intensities. Coherence analysis quantified the frequency-specific cortical tracking of the slow fluctuations of the different elements of the auditory scene: (i) the *Attended stream* (i.e. attended speaker's voice), (ii) the *Multi-talker babble* noise and (iii) the *Global auditory scene* (i.e. the combination of *Attended stream and Multi-talker babble*). We then compared the spatial patterns of functional connectivity arising from the brain areas involved in SiN cortical tracking to unravel the possible impact of altered tracking on high-order SiN processing steps.

Methods

Methods were derived from Vander Ghinst et al.[30,31] and Wens et al.[37]

Participants

Thirteen subjects (mean age: 30 years, age range: 21–40 years, six females) who consulted at the Ear Nose and Throat Department of the CUB Hôpital Erasme for difficulties to understand speech in noisy backgrounds lasting for more than a year, were included based on the following stringent inclusion criteria: age <40 years; right-handedness [Edinburgh handedness inventory (EHI)],[38] native French speakers; absence of any history of neuropsychiatric, language or otologic disorder (including noise trauma and tinnitus); normal hearing according to pure tone audiometry [i.e. normal hearing thresholds (between 0–20 dB HL) for 250, 500, 1000, 2000, 3000, 4000, 6000 and 8000 Hz]; normal otomicroscopy and tympanometry; compatibility with MEG and MRI.

Thirteen right-handed (EHI) and native French-speaking healthy subjects, individually matched with subjects with ISPiN for age, sex and educational level (mean age: 29 years, age range 22–40 years, six females) were also recruited.

Participants' auditory perception was assessed with three separate subtests of a validated and standardized French language central auditory battery: (i) a dichotic test, (ii) a speech audiometry and (iii) a SiN audiometry[39] (see Vander Ghinst et al.[31] for details). Participants' attentional abilities were evaluated using two different subtests (i.e. visual scanning and divided attention tasks) of the computerized Attention Test Battery (TAP, Version 2.2).[40] The divided attention task relied on both visual and auditory modalities.

The study had prior approval by the CUB Hôpital Erasme Ethics Committee (REF: P2012/049). Participants gave written informed consent before participation.

Experimental paradigm

Participants underwent five listening and one rest (eyes opened, fixation cross) conditions each lasting 5 min in a randomized order. Listening conditions consisted of five different stories in French randomly selected from a set of six stories read by native French speakers (three females; http://www.litteratureaudio.com Accessed 27 August 2021) mixed with a continuous multi-talker babble of six native French speakers (three females) talking simultaneously in French.[41] Phrasal, word and syllable rates, assessed as the number of phrases, words or syllables divided by the corrected duration of the audio recording were 0.49, 3.39 and 5.56 Hz, respectively (mean phrasal, word and syllable rates across different stories). For phrases, the corrected duration was (trivially) the total duration of the audio recording. For words and syllables, the corrected duration was the total time during which the speaker was actually speaking, that is the total duration of the audio recording (here 5 min) minus the sum of all silent periods when the speech amplitude was below a tenth of the mean amplitude for at least 100 ms. A specific signal-to-noise ratio (SNR; signal: *Attended stream*, noise: *Multi-talker babble*) was randomly assigned to each story: a noiseless condition, +10, +5, 0 and −5 dB. Sound recordings were transmitted to a MEG-compatible flat-panel loudspeaker (Panphonics Oy, Espoo, Finland) and played at about 60 dB. Participants had to attend to the reader's voice and gaze at a fixation cross. At the end of each listening condition, participants were requested to verbally score the intelligibility of the attended stream (0 = totally unintelligible, 10 = perfectly intelligible) and to answer 16 yes/no forced-choice questions on the heard story.[42]

Data acquisition

Neuromagnetic signals were recorded (bandpass: 0.1–330 Hz, sampling rate: 1 kHz) using a whole-scalp-covering 306-channel MEG (Vectorview, Elekta Oy, Helsinki, Finland) installed in a light-weight magnetically shielded room (Maxshield, MEGIN, Helsinki, Finland; see De Tiège et al.[43] for details). The MEG device has 102 sensor chipsets, each comprising one magnetometer and two orthogonal planar gradiometers. Four head-tracking coils monitored participants' head position inside the MEG helmet. The location of the coils and at least 150 head-surface (on scalp, nose and face) points with respect to anatomical fiducials were digitized with an electromagnetic tracker (Fastrack, Polhemus, Colchester, VT). Audio signals (bandpass: 50–22 000 Hz, sampling rate: 44.1 kHz) were recorded (low-pass: 330 Hz) simultaneously to MEG signals for synchronization of the corresponding signals.

High-resolution 3D T1-weighted cerebral MRIs were acquired at 1.5 T (Intera, Philips, The Netherlands).

Data pre-processing

Continuous MEG data were preprocessed offline using the temporal extension of the signal space separation method (correlation limit, 0.9; segment length, 20 s) to suppress external inferences and correct for head movements.[44,45] Eyeblink, eye movement and heartbeat artefacts were removed from the band-pass filtered (0.1–45 Hz) data using independent component analysis (FastICA algorithm with dimension reduction to 30 and non-linearity *tanh*)[46] and visual inspection of the components (1–4 components/subject).[47]

Continuous MEG signals (and the synchronous audio signals in the listening conditions) were split into 2048-ms epochs with 1638-ms epoch overlap (frequency resolution: ∼0.5 Hz). MEG epochs exceeding 3 pT (magnetometers) or 0.7 pT/cm (gradiometers) were excluded to reject artefactual epochs. The number of artefact-free epochs was 742 ± 13 (mean ± SD across all participants and listening conditions). A two-way repeated-measures ANOVA did not reveal an effect of group ($F_{1,24} = 0.67$, $P = 0.42$) or condition ($F_{4,96} = 0.61$, $P = 0.67$), nor an interaction ($F_{4,96} = 2.04$, $P = 0.09$), on the number of artefact-free epochs.

Cortical tracking of speech streams in sensor space

The synchronization between the temporal envelope of wide-band (50–22 000 Hz) audio signals and artefact-free MEG signals was assessed with coherence analysis at frequencies in which speech temporal envelope is critical for speech comprehension (i.e. 0.1–20 Hz).[48] Coherence computation was based on all artefact-free epochs (2048-ms long), yielding a frequency resolution of ∼0.5 Hz.

For the four SiN conditions (*+10, +5, 0* and *−5 dB*), coherence was separately computed between MEG signals and three acoustic streams of the auditory scene: the *Attended stream* (Coh_{att}), the *Multitalker babble* (Coh_{noise}) and the *Global scene* (*Attended stream + Multitalker babble;* Coh_{global}). Sensor-level coherence maps were obtained using combined gradiometer signals as in Bourguignon et al.[49]

Sensor-level coherence maps were produced separately for frequencies matching with prosodic stress/phrasal/sentential (0.5 Hz), word (average across 1–4 Hz) and syllable (4–8 Hz) rhythms, which are henceforth referred to as *frequency ranges of interest*.

Cortical tracking of speech streams in source space

Individual MRIs were segmented using the Freesurfer software (Martinos Center for Biomedical Imaging, Boston, MA)[50] and manually coregistered to MEG coordinate systems. Then, a non-linear transformation from individual MRIs to the MNI brain was computed using the spatial normalization algorithm implemented in Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK).[51,52] This transformation was used to map a homogeneous 5-mm grid sampling the MNI brain volume onto individual brain volumes. For each subject and grid point, the MEG forward model corresponding to three orthogonal current dipoles was computed using the one-layer Boundary Element Method implemented in the MNE software suite (Martinos Centre for Biomedical Imaging, Boston, MA, USA).[53] The forward model was then reduced to its two first principal components. This procedure is justified by the insensitivity of MEG to currents radial to the skull, and hence, this dimension reduction leads to considering only the tangential sources. To simultaneously combine data from planar gradiometers and magnetometers for source estimation, sensor signals (and the corresponding forward-model coefficients) were normalized by their root-mean-square noise level, estimated from rest data (band-pass: 1–195 Hz). Coherence maps for each participant, listening condition (*Noiseless*, *+10*, *+5*, *0* and *−5 dB*), speech stream (*Attended stream*, *Multi-talker Babble* and *Global scene*) and frequency range of interest (<1, 1–4 and 4–8 Hz), were finally produced using Dynamic Imaging of Coherent Sources[54] with a Minimum-Norm Estimates (MNE) inverse solution.[55] Noise covariance was estimated from the rest data (band-pass: 1–195 Hz) and the regularization parameter was fixed in terms of the MEG sensor noise level.[37]

Seed-based functional connectivity mapping in source space

To search for any possible impact of altered cortical tracking of SiN on subsequent speech processing steps, we compared between groups the functional connectivity arising from the brain areas showing altered cortical tracking of SiN in ISPiN subjects. For that purpose, seed-based functional connectivity maps were computed using a source-space coherence analysis similar to that assessing the cortical tracking of speech. In that analysis, source space activity was estimated using MNE as above, and connectivity maps were obtained as the coherence between all sources and selected seed signals. The only major addition was to precede coherence estimation with signal orthogonalization[56] to correct for spatial leakage emanating from the seed, i.e. spurious inflation of connectivity due to zero-lag cross-correlations within the MEG forward model.[37,57] We focused here on seed(s) for which the cortical tracking of SiN was significantly different between groups (see Statistical Analyses section) and on the corresponding frequency range of interest. This procedure yielded one functional connectivity map for each participant, listening condition and seed.

Group-averaging

Maps of cortical tracking of speech and seed-based functional connectivity were averaged across participants. Note that averaging was straightforward given that individual maps were intrinsically coregistered to the MNI template. For illustration purposes, functional connectivity maps were computed (i) in the rest condition, (ii) in the Noiseless condition and (iii) as the mean of all noisy conditions (+10, +5, 0 and −5 dB).

Statistical analyses

Comparison of auditory perception and attention

Scores for auditory perception (dichotic test, speech audiometry and SiN audiometry) and attentional (visual scanning and divided attention) tests were compared between groups (i.e. subjects with ISPiN versus matched healthy subjects) using independent sample *t*-tests.

Effect of noise level on intelligibility ratings and comprehension scores

A two-way repeated-measures ANOVA was used to assess the effects of the listening condition (within-subject factor; *Noiseless*, *+10*, *+5*, *0* and *−5 dB*) and the group (between-subjects factor; subjects with ISPiN, healthy subjects) on intelligibility ratings and comprehension scores separately.

Significance of individual cortical tracking of speech

The statistical significance of individual-level coherence values assessing the cortical tracking of speech (for each listening condition, speech stream and frequency range of interest) was derived in the sensor space with surrogate-data-based maximum statistics. This approach intrinsically deals with the issue of multiple comparisons across sensors while preserving signals' temporal auto-correlation. For each participant, 1000 surrogate sensor-level coherence maps were computed as done for the genuine coherence maps but with speech stream signals replaced by random Fourier-phase surrogates.[58] The maximum coherence value across all sensors was extracted for each surrogate simulation, and the 95th percentile of this distribution of maximum coherence values yielded the significance threshold at $P < 0.05$.

Group-level cortical tracking of speech and comparison between speech streams and groups

The statistical significance of coherence values in group-level coherence maps assessing the cortical tracking of speech (for each listening condition, speech stream and frequency range of interest) was derived in source space with a non-parametric permutation test.[59] In practice, individual and group-level surrogate coherence maps were first computed as done for the genuine coherence maps, but with MEG signals in listening conditions replaced by the resting-state MEG signals (and speech stream signals unchanged). Group-level difference maps were obtained by subtracting the group mean of the genuine coherence maps with the corresponding surrogate maps. Under the null hypothesis that coherence maps are the same whatever the experimental condition, the genuine maps (i.e. during listening conditions) and the surrogate maps (i.e. at rest) are exchangeable at the individual level before taking the group-level difference.[59] To reject this hypothesis and to compute a threshold of statistical significance for the correctly labelled difference map, the permutation distribution of the maximum of the difference map's absolute value was computed for 10 000 permutations. The test assigned a P-value to each source in group-level coherence maps, equal to the proportion of surrogate values exceeding the corresponding source's difference value. Statistical significance was set to $P < 0.05$ since the P-value was intrinsically corrected for the multiple comparisons across all sources tested within each hemisphere.[59]

Permutation tests can be too conservative for sources other than the one with the maximum observed statistic.[59] For example, dominant coherence values in the right auditory cortex could bias the permutation distribution and over-shadow weaker coherence values in the left auditory cortex, even if these were highly consistent across subjects. Therefore, the permutation test described above was conducted separately for left- and right-hemisphere sources.

Coordinates of significant local maxima in group-level coherence maps were finally identified.[31]

Significance of selective cortical tracking of the attended speech

For each listening condition and frequency range of interest, we identified the cortical areas wherein activity reflected more the *Attended stream* than the *Global scene*. For that, we compared source-level Coh_{att} to Coh_{global} maps using the same permutation test described above, but now permuting the *Global scene and Attended stream* labels rather than the genuine and surrogate labels.

Such analysis was not done for contrasts involving the cortical tracking of the *Background babble* (i.e. Coh_{noise}) because such tracking was not consistently observed in subjects with ISPiN nor healthy subjects (see Results). This finding was in line with our previous studies relying on similar experimental paradigms.[30,31] Furthermore, as in our previous studies, the cortical tracking of speech streams was higher for the *Attended stream* than for the *Global scene.*[30,31,60] The next analyses about cortical tracking of speech, therefore, focused on the *Attended stream*.

Group comparison of the cortical tracking of the attended speech

To identify cortical areas exhibiting a modulation of the cortical tracking of the *Attended stream* due to ISPiN, we compared Coh_{att} maps between subjects with ISPiN and healthy subjects using the same permutation test described above, permuting on the labels *subjects with ISPiN* and *healthy subjects*.

Effect of the listening condition, group and hemispheric lateralization on cortical tracking of the attended speech

In this between-subject design, we used a 3-way repeated-measures ANOVA to compare the cortical tracking of speech between subjects with ISPiN and healthy subjects with additional factors of hemisphere (left versus right) and five different listening conditions (*Noiseless*, *+10*, *+5*, *0* and *−5 dB*). The dependent variable was the maximal Coh_{att} value within a sphere of 10 mm radius around the maximum of the group-level difference map in each hemisphere.

Comparison of seed-based functional connectivity between groups of participants

We assessed the effect of group (subjects with ISPiN versus healthy subjects) or listening conditions (*Noiseless*, *+10*, *+5*, *0* and *−5 dB*) on each connection included in the seed-based functional connectivity maps, using a mass-univariate two-way ANOVA of their Fisher-transformed coherence values. A statistical mask was then built at a 5% significance level, with Bonferroni correction for the effective number of independent connections in those maps, to reveal the significant regions reflecting ANOVA main effect (group and SNR conditions) and interaction (group × SNR conditions). Of note, the correction for multiple comparisons relied here on a parametric estimate rather than a non-parametric method used for the cortical tracking of speech because of our use of mass-univariate ANOVA, analogously to standard protocols in Statistical Parametric Mapping of activation maps.[61] The Bonferroni factor was set here as the effective number of degrees of freedom in MNE maps, i.e. 60, estimated as the rank of the forward model.[37] Subsequently, *post hoc* mass-univariate t-test maps were produced (with the same Bonferroni factor) to examine specific differences between each group/condition.

Data availability

Data are available upon reasonable request to the authors and after approval of Institutional authorities (i.e. CUB Hôpital Erasme and Université libre de Bruxelles).

Results

Comparison of auditory perception and attention between groups of participants

Performance for dichotic ($t_{24} = 1.21$; $P = 0.24$; subjects with ISPiN, 70.8 ± 15; mean ± SD; healthy subjects, 77.5 ± 13.6) and speech perception in silence ($t_{24} = 0.23$; $P = 0.82$; 28.8 ± 0.9; 28.8 ± 0.8) tests did not differ between groups, while those for SiN audiometry was significantly ($t_{24} = 5.86$; $P < 0.0001$) poorer in subjects with ISPiN (23.8 ± 1.8) than in healthy subjects (27.2 ± 1.1).

In the *visual scanning* task, reaction times did not differ between groups, whether the target was present ($t_{24} = 0.11$; $P = 0.91$) or not ($t_{24} = 0.1$; $P = 0.93$). The error rate did not differ between both groups, whether the target was present ($t_{24} = 0.8$; $P = 0.46$) or not ($t_{24} = 0.36$; $P = 0.72$).

In the *divided attention* task, reaction times did not differ between groups (auditory task, $t_{24} = 1.2$; $P = 0.2$; visual task, $t_{24} = 1.4$; $P = 0.17$). The omission rate did not differ between groups either (auditory task, $t_{24} = 0$; $P = 1$; visual task, $t_{24} = 1.1$; $P = 0.27$).

Effect of listening conditions on intelligibility ratings and comprehension scores

Figure 1 displays the intelligibility ratings and comprehension scores in all listening conditions in both groups.

The ANOVA performed on intelligibility ratings and comprehension scores revealed a statistically significant effect of listening condition (ratings, $F_{4,96} = 178$, $P < 0.0001$; scores, $F_{4,96} = 46.6$, $P < 0.0001$) and group (ratings, $F_{1,24} = 12.0$, $P = 0.002$; scores, $F_{1,24} = 9.63$, $P = 0.0048$), and a significant interaction (ratings, $F_{4,96} = 2.67$, $P = 0.037$; scores, $F_{4,96} = 3.58$, $P = 0.0092$). *Post hoc* comparisons demonstrated that comprehension scores were higher in healthy subjects than in subjects with ISPiN only in conditions with an SNR below +5 dB (see Fig. 1 for detailed *P*-values). Intelligibility rating was higher in healthy subjects than in subjects with ISPiN in intermediate SNR conditions (*+10*, *+5* and *0 dB*) but not in extreme conditions (*Noiseless* and *−5 dB*).

Significance of individual cortical tracking of speech

In the *noiseless* condition, all participants displayed statistically significant Coh_{att} at all frequencies (<1, 1–4 and 4–8 Hz), except for one healthy subject and three subjects with ISPiN at 1–4 Hz (Table 1). Fig. 2 presents the group-averaged coherence spectra. In line with previous literature, the coherence peaked at 0.5 Hz and was sustained in the range of 2–8 Hz. In further analyses, we focused on the a priori defined frequency ranges.

Table 1 Participants with significant cortical tracking of speech

	Healthy subjects			Subjects with ISPiN		
Condition	**Attended**	**Global**	**Noise**	**Attended**	**Global**	**Noise**
<1 Hz						
Noiseless	13		–	13		–
10 dB	13	13	1	13	13	2
5 dB	13	13	0	13	13	0
0 dB	13	12	1	13	13	2
–5 dB	13	6	0	12	7	1
1–4 Hz						
Noiseless	12		–	10		–
10 dB	12	11	2	12	12	0
5 dB	11	11	1	13	12	5
0 dB	11	8	6	12	9	5
–5 dB	9	6	8	11	5	5
4–8 Hz						
Noiseless	13		–	13		–
10 dB	13	13	1	13	13	2
5 dB	13	13	4	13	13	2
0 dB	13	12	7	12	12	4
–5 dB	8	6	5	7	6	4

Number of healthy subjects and subjects with ISPiN showing statistically significant coherence assessing the cortical tracking of speech (surrogate-data-based statistics) in at least one sensor for each audio signal, condition and frequency range of interest.

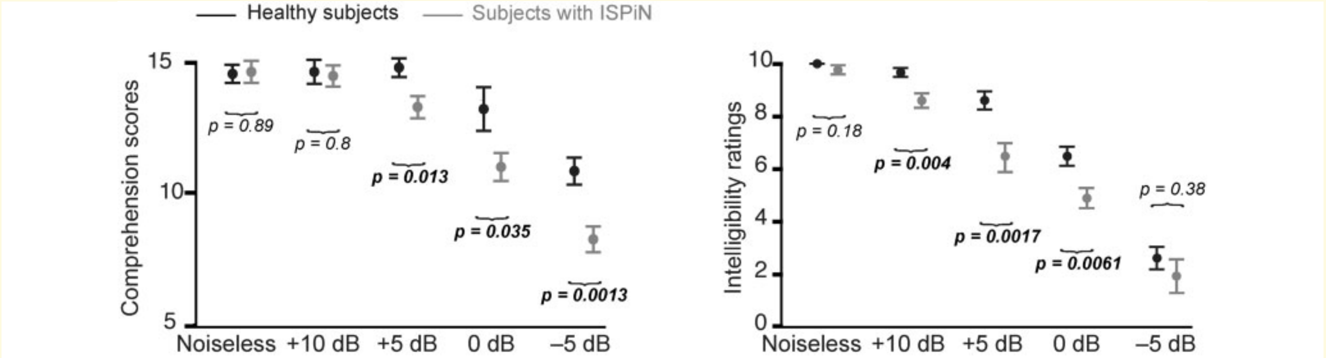


Figure 1 Effect of listening conditions on intelligibility ratings and comprehension scores. Comprehension scores (left) and intelligibility ratings (right) in healthy subjects (black) and subjects with IPSiN (grey). Dots indicate the mean and bar the standard deviation. Comprehension scores are reported in the number of questions (16) answered correctly, and intelligibility ratings ranged from 0 (totally unintelligible) to 10 (perfectly intelligible). Horizontal brackets indicate the outcome of *post hoc* paired *t*-tests between groups in each condition. Significant *P*-values are emphasized in bold.

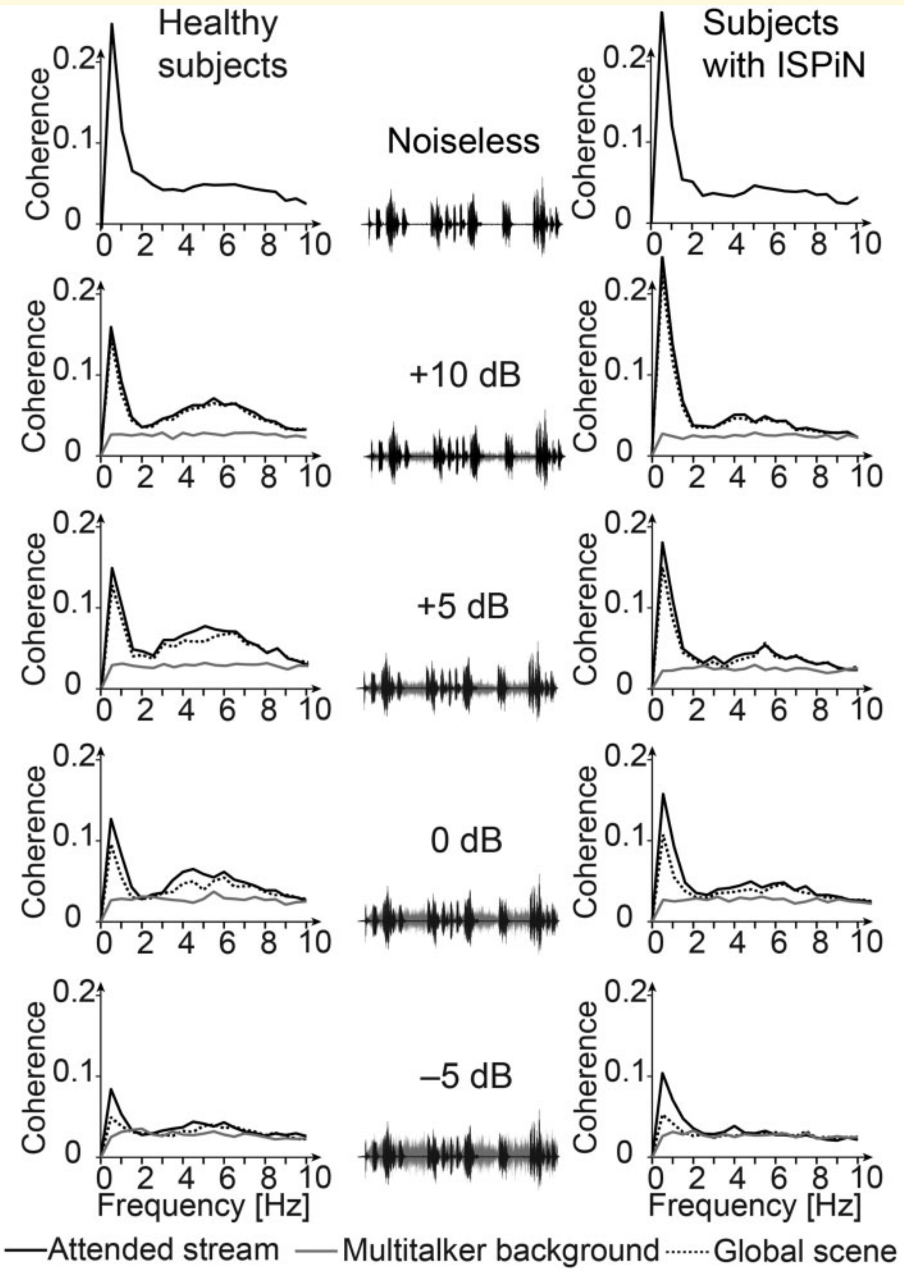


Figure 2 Spectra of cortical tracking of speech in the five listening conditions and corresponding sound excerpts. Group-averaged coherence spectra are shown separately for healthy subjects (left column) and subjects with ISPiN (right column), and when estimated with the temporal envelope of the different components of the auditory scene: the *Attended stream* (black connected trace), the *Multi-talker babble* (grey connected trace) and the *Global scene* (grey dotted trace). Each spectrum represents the group mean of the maximum coherence across all sensors. The sound excerpts showcase the *Attended stream* (black traces) and the *Multitalker babble* (grey traces) and their relative amplitude depending on the SNR.

Maximum Coh_{att} peaked at MEG sensors covering bilateral temporal areas (Supplementary Fig. 1).

In the SiN conditions, almost all participants displayed significant sensor-level Coh_{att} and Coh_{global}, except at $-5\,dB$ where fewer did (Table 1). Coherence local maxima were located for all frequency ranges of interest and in both groups at bilateral superior temporal sulcus (STS) and supratemporal auditory cortices (AC) (Fig. 3, Supplementary Table 1).

Significant Coh_{noise} was observed in a limited number of participants, justifying why further analyses concentrated on Coh_{att} and Coh_{global}.

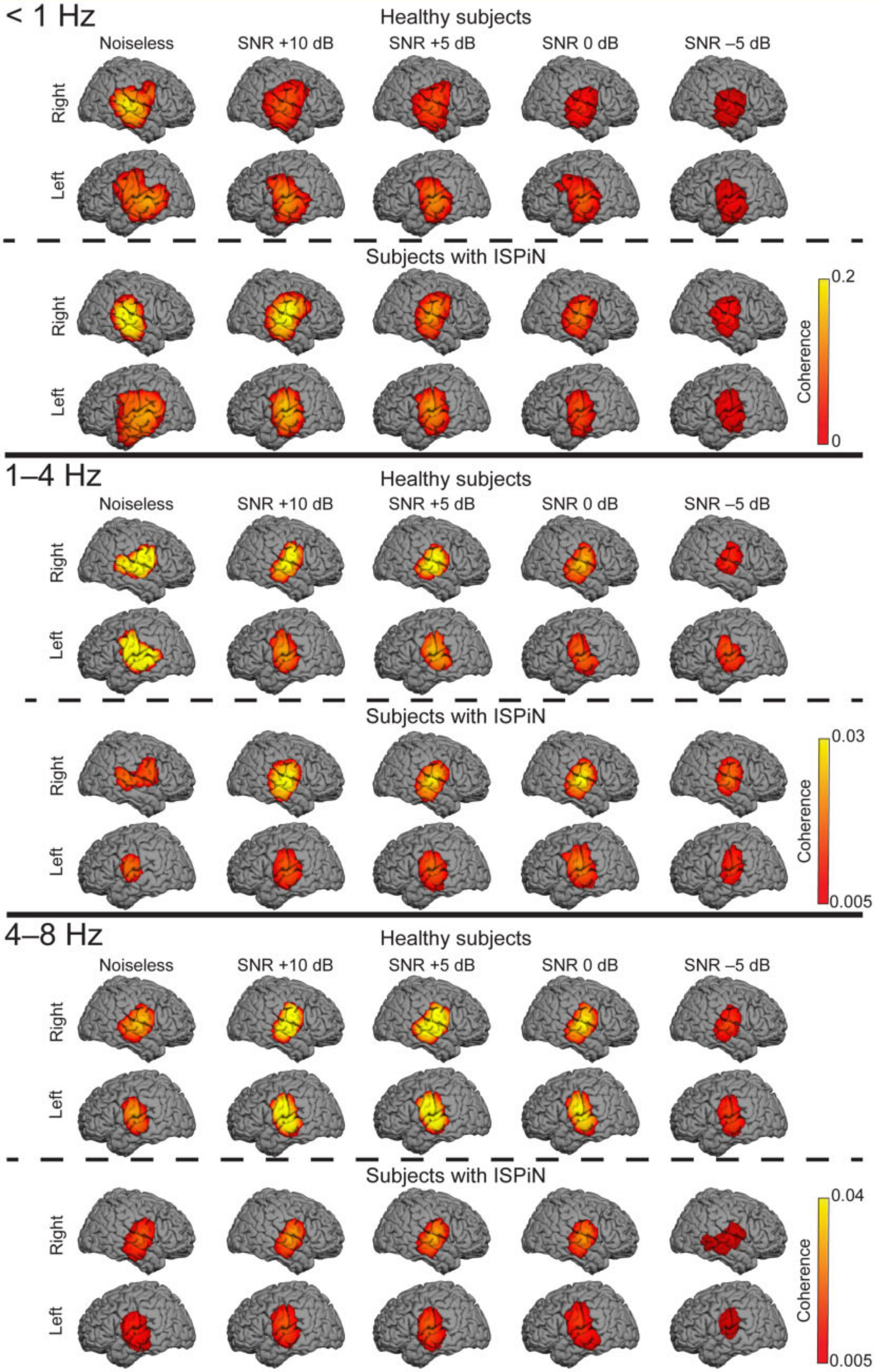


Figure 3 Cortical tracking of the *Attended stream* at <1, 1–4 and 4–8 Hz. The group-level coherence maps were masked statistically above the significance level (maximum-based permutation statistics). One source distribution is displayed for each possible combination of the group (healthy subjects, *top* panel; subjects with ISPiN, *bottom* panel) and listening condition (from left to right, *Noiseless*, +10, +5, 0 and −5 dB).

Group-level cortical tracking of speech

In source space, group-level Coh_{att} *and* Coh_{global} at <1 Hz peaked at bilateral neocortical areas around STS and the right operculum (Op), while it peaked at bilateral AC at 1–4 and 4–8 Hz (Fig. 3 and Supplementary Table 1).

Significance of selective cortical tracking of the attended speech

In both healthy subjects and subjects with ISPiN, Coh_{att} was higher than Coh_{global}, i.e. MEG signals tracked more the *Attended stream* than the *Global scene*. Cortical areas showing this effect were the bilateral AC/STS for every listening condition and frequency range of interest ($Ps < 0.05$), except, for some instances, for the *+5 dB* condition (in the left hemisphere at 1–4 Hz in healthy subjects, in the right hemisphere at 4–8 Hz in healthy subjects). Conversely, Coh_{global} did not exceed significantly Coh_{att} in any listening condition, frequency range of interest and group ($Ps > 0.5$).

Group comparison of cortical tracking of the attended speech

The comparison of Coh_{att} between healthy subjects and subjects with ISPiN (Fig. 4) revealed strikingly different patterns at the different frequencies investigated. At frequencies <1 Hz, Coh_{att} was higher in subjects with ISPiN than in healthy subjects at the right AC at *+10 dB* ($P = 0.005$) and *0 dB* ($P = 0.018$), and the same trend was seen at *+5 dB* ($P = 0.068$). At 1–4 Hz, Coh_{att} was significantly higher in healthy subjects than in subjects

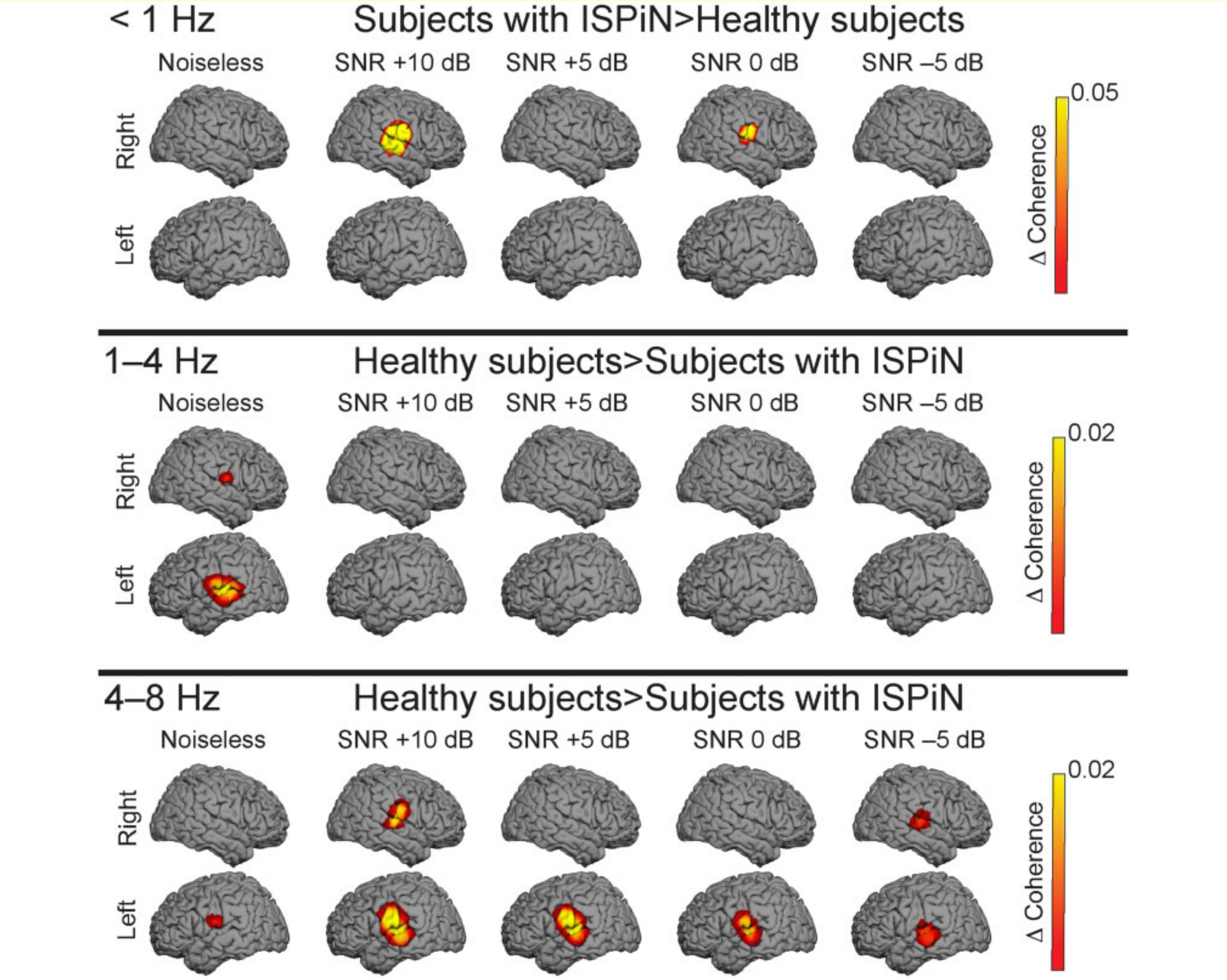


Figure 4 Group comparison of the cortical tracking of the attended speech. Contrast maps indicating where the cortical tracking of the attended speech (Coh_{att}) is higher in subjects with ISPiN with IPSiN than in healthy subjects at <1 Hz (*top* panel), and in healthy subjects than in subjects with ISPiN at 1–4 and 4–8 Hz (*middle* and *bottom* panels), for each listening condition (*Noiseless*, *+10*, *+5*, *0* and *−5 dB*). The group-level difference coherence maps were masked statistically above the significance level (maximum-based permutation statistics).

with ISPiN at bilateral AC in *Noiseless* (left, $P = 0.01$; right, $P = 0.048$). At 4–8 Hz, Coh_{att} was significantly higher in healthy subjects than in subjects with ISPiN at bilateral AC in all listening conditions ($Ps < 0.05$), except in right AC in *Noiseless* ($P = 0.19$), at *+5 dB* ($P = 0.07$) and at *0 dB* ($P = 0.22$).

Overall, data demonstrate that both groups do track the attended stream rather than the global scene and that such tracking localizes mainly at bilateral AC/STS. Also, the way noise modulates the tracking in both groups seems to depend on the frequency range considered, and possibly on the hemisphere. Accordingly, we next used an ANOVA to determine how group, listening condition and hemispheric lateralization impact the cortical tracking of speech in the three frequency ranges of interest separately.

Effect of group, listening condition and hemispheric lateralization on cortical tracking of speech

At <1 Hz, the ANOVA revealed a significant main effect of noise level on Coh_{att} ($F_{4,96} = 27.3$, $P < 0.0001$), a significant interaction between listening condition and hemispheric lateralization ($F_{4,96} = 7.00$, $P = 0.0001$), a marginally significant interaction between hemispheric lateralization and group ($F_{1,24} = 3.61$, $P = 0.069$) and no other effects ($Ps > 0.1$). The main effect of the listening condition was explained by a decrease in Coh_{att} as the noise level increased (Fig. 5). The interaction between listening condition and hemispheric lateralization was explained by a faster decrease in Coh_{att} in the right than the left STS with increasing noise level (Figs 2 and 5). *Post hoc* comparisons (see Fig. 5 for details) revealed that Coh_{att} at right STS decreased as soon as the *Multitalker babble* was added and it further diminished significantly as noise increased. In contrast, Coh_{att} in the left STS decreased significantly only in the two noisiest conditions. The marginal interaction between group and lateralization was explained by higher Coh_{att} in subjects with ISPiN than in healthy subjects in the right ($t_{24} = 2.19$, $P = 0.038$) but not left ($t_{24} = 0.25$, $P = 0.8$) STS.

At 1–4 Hz, the ANOVA revealed a significant main effect of the listening condition ($F_{4,96} = 4.02$, $P = 0.0046$), a significant effect of hemispheric lateralization ($F_{1,24} = 7.49$, $P = 0.012$), a significant interaction between listening condition and hemispheric lateralization ($F_{4,96} = 3.15$, $P = 0.018$), a significant interaction between listening condition and group ($F_{4,96} = 3.44$, $P = 0.011$) and no other significant effects or interactions ($Ps > 0.05$). The interaction between listening condition and hemispheric lateralization was explained by a stronger drop in Coh_{att} in left than right AC as soon as the noise was added [$Coh_{att}(Noiseless) - Coh_{att}(+10\,dB)$ in left versus right hemisphere, $t_{25} = 2.36$, $P = 0.027$] and by the reverse effect in the noisiest condition [$Coh_{att}(0\,dB) - Coh_{att}(-5\,dB)$ in right versus left hemisphere, $t_{25} = 2.22$, $P = 0.035$]. The interaction between listening condition and group was explained by higher values in healthy subjects than subjects with ISPiN in *Noiseless* only ($t_{24} = 2.28$, $P = 0.032$). The main effect of hemispheric lateralization was explained by higher Coh_{att} at right (mean ± SD coherence, 0.0321 ± 0.0106) than left AC (0.0272 ± 0.0101).

At 4–8 Hz, the ANOVA revealed a significant main effect of listening condition ($F_{4,96} = 30.28$, $P < 0.0001$), a significant effect of group ($F_{1,24} = 4.13$, $P = 0.05$), a significant interaction between listening condition and group ($F_{4,96} = 2.6$, $P = 0.04$) and no other significant main effects or interactions ($Ps > 0.1$). The main effect of the group was explained by significantly higher Coh_{att} in healthy subjects (mean ± SD coherence, 0.0427 ± 0.024) than in subjects with ISPiN (0.028 ± 0.0097). The interaction between listening condition and group was explained by a stronger modulation of Coh_{att} by noise in healthy subjects than in subjects with ISPiN (Figs 2 and 3). In healthy subjects, Coh_{att} (mean across hemispheres) was significantly higher in intermediate (*+10* and *+5 dB*) than in extreme (*Noiseless*, *0* and *−5 dB*) listening conditions, and in *Noiseless* and at *0 dB* than at *−5 dB* (Fig. 5). Some of these effects were also found in subjects with ISPiN: Coh_{att} was higher at *+5 dB* than in *Noiseless*, and lower at *−5 dB* than in all other conditions (Fig. 5).

Comparison of seed-based functional connectivity between groups of participants

Since a robust between-group difference in SiN cortical tracking was observed at 4–8 Hz, we limited functional connectivity analyses to this frequency range. We focused on mapping functional connectivity with seeds placed at bilateral AC with MNI coordinates [−65 −14 14] mm for left AC and [66 −11 8] mm for right AC. These coordinates were the mean across groups and conditions of the coordinates of Coh_{att} local maxima, which all were anyway in close proximity (see Supplementary Table 1).

Functional connectivity maps showed a similar topographical pattern in both groups. However, functional connectivity values for both seeds were higher in Subjects with ISPiN in every listening condition, but also in the rest condition (see Supplementary Fig. 2). These differences in functional connectivity value therefore irrespective of the task partly explain the results described below.

Functional connectivity from the left AC was significantly different between groups with the right Op, the left inferior frontal gyrus (IFG), the left frontal eye field (FEF), the left superior temporal gyrus (STG), the left middle temporal gyrus (MTG) and the left temporo-parietal junction (TPJ). The connectivity from the right AC

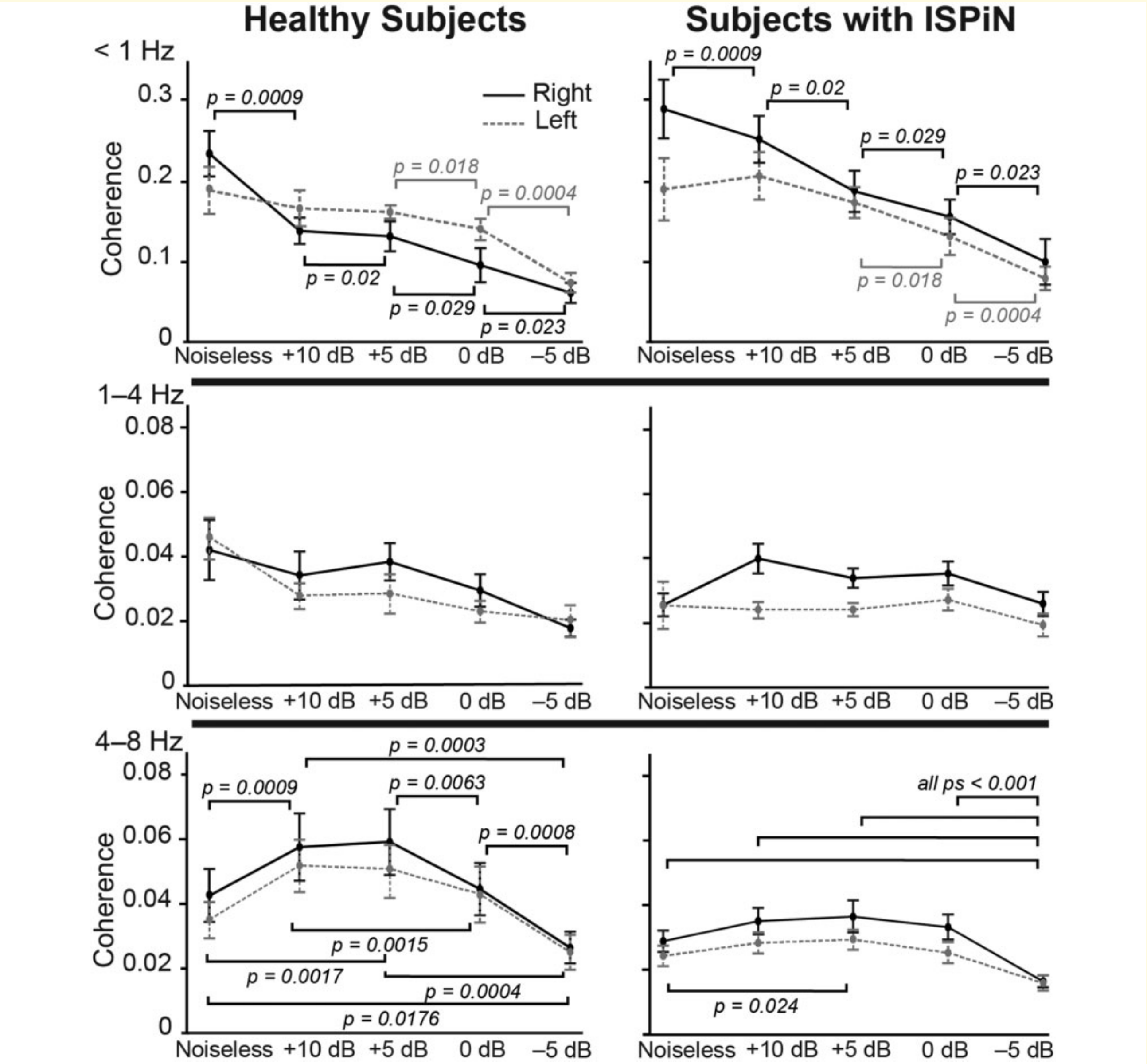


Figure 5 Effect of noise, group and hemisphere on the cortical tracking of the *Attended stream* at <1, 1--4 and 4–8 Hz. Dots indicate mean and bars the standard error on the mean of coherence values at the right (solid lines) and left (dashed lines) cortical area of peak group-level coherence (STS at <1 Hz, and AC at 1–4 and 4–8 Hz) in healthy subjects (left) and subjects with ISPiN (right). Horizontal brackets illustrate the different *post hoc t*-tests.

was significantly different between groups with the right IFG, the left middle frontal gyrus (MFG), the left inferior temporal gyrus (ITG) and the left inferior occipital gyrus (IOG) (Fig. 6). For both seed locations, these differences were explained by higher functional connectivity values in subjects with ISPiN than in healthy subjects (see Fig. 7 for functional connectivity values and MNI coordinates). Subsequent *post hoc t* maps showed significantly higher functional connectivity in subjects with ISPiN compared to healthy subjects, mainly between left AC and left FEF in every SiN condition, and left AC and left IFG in the two noisiest conditions (i.e. *0* and *−5 dB*). Furthermore, in the noisiest conditions (i.e. *0* and *−5 dB* with the seed at left AC, and *+5*, *0* and *−5 dB* with the seed at right AC), cortical areas showing significantly higher functional connectivity in subjects with ISPiN were only located in the left hemisphere.

A significant effect of noise emerged in the right middle occipital gyrus [MNI coordinates: (36 −87 4) mm] for the seed placed in the left AC, and in the left lingual gyrus [(−9 −84 −5) mm], the right IFG [(55 13 24) mm], the left precentral gyrus [(51 −6 44) mm] and the

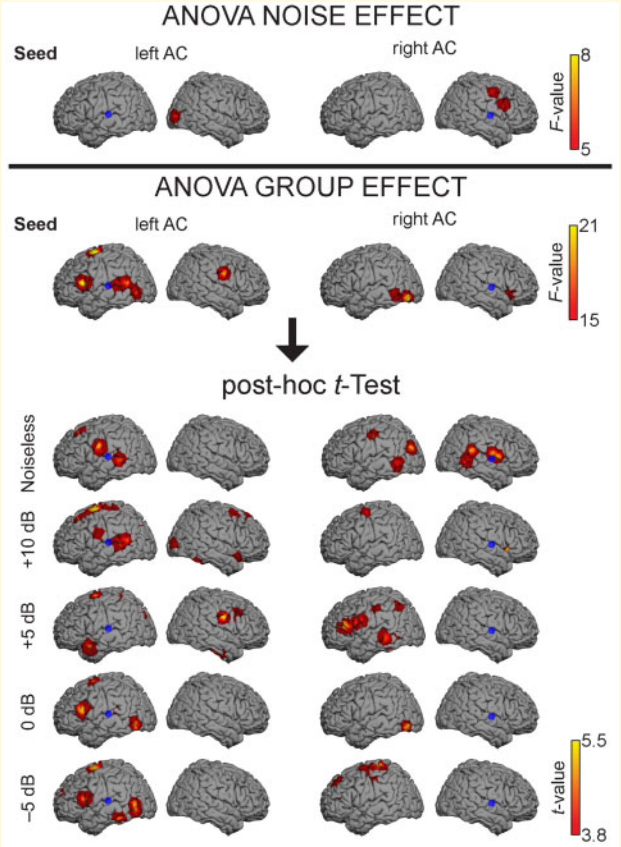


Figure 6 Effect of noise and group on functional connectivity. Brain regions showing a significant effect of noise and group (healthy subjects versus subjects with ISPiN) on functional connectivity between the left or right AC seed (blue dot) and the rest of the cortex. Seed-based maps are masked statistically using a mass-univariate two-way ANOVA at $P < 0.05$ Bonferroni corrected for the number of degrees of freedom. Subsequent *post hoc t* maps illustrate the between-group difference for every condition. Positive *t* values show that subjects with ISPiN exhibit higher functional connectivity values than healthy subjects (see Fig. 7 for details). Note that the precise anatomical location is distorted by the 3D surface rendering, we refer to Fig. 7 for the MNI coordinates of the significant local maxima.

right thalamus [(20 −12 15) mm] for the seed placed in the right AC.

No significant interaction between group and noise was observed with the seed placed in left or right AC.

Discussion

This MEG study performed in a group of highly selected subjects with ISPiN provides novel insights into the pathophysiology of ISPiN. Compared with healthy subjects, those with ISPiN mainly displayed (i) no behavioural deficit in attentional abilities as assessed by visual scanning and divided (visual and auditory modalities) attention tasks, (ii) reduced cortical tracking of speech at 4–8 Hz in SiN but also in *noiseless* conditions and (iii) increased functional connectivity at 4–8 Hz between AC and language/attention-related brain areas.

Behavioural tests favour a specific auditory processing disability hypothesis

In noise, the behavioural scores (SiN audiometry, intelligibility rating and comprehension) were significantly poorer in subjects with ISPiN than in healthy subjects, while they were similar in silence. This confirms the specific SiN impairment in the ISPiN group thereby ruling out a subjective underestimation of their auditory abilities.[62,63] Both groups also had comparable dichotic performances further suggesting that ISPiN is a specific auditory processing disability rather than a global central auditory processing disorder.[64]

The substantial variability in SiN performances in subjects with normal peripheral hearing is related to individual selective attentional abilities[65] in both visual and auditory modalities. A reduction in attentional control has, e.g. been suggested to explain SiN comprehension deficits in older people with normal hearing thresholds.[66] Our results do not support this hypothesis in subjects with ISPiN, since attentional testing did not disclose any significant difference in both visual scanning and divided attention tasks between both groups.

Left-lateralized selective tracking of speech in subjects with ISPiN

In ‘cocktail party’ settings, neural activity in auditory cortices selectively tracks the attended speaker’s voice rather than the global auditory scene.[28–30] We found such preferential tracking also in subjects with ISPiN, and furthermore, that they had a higher level of tracking at <1 Hz than healthy subjects. This selective tracking of speech is partly subsequent to the selective suppression of noise-related acoustic features in the cortical responses.[67] An alteration of this suppression mechanism is therefore not the source of ISPiN pathophysiology, since significant Coh_{noise} was seldom observed in both groups.

The cortical tracking of speech at <1 Hz in the left hemisphere is typically more robust to noise than that in the right hemisphere, probably reflecting a filtering process at the semantic processing level.[30,60,68] Interestingly, this functional asymmetry that is considered to promote speech recognition in adverse auditory scenes does not account for ISPiN pathophysiology since the way noise affected <1 Hz Coh_{att} in the right more than left STG was similar between groups.

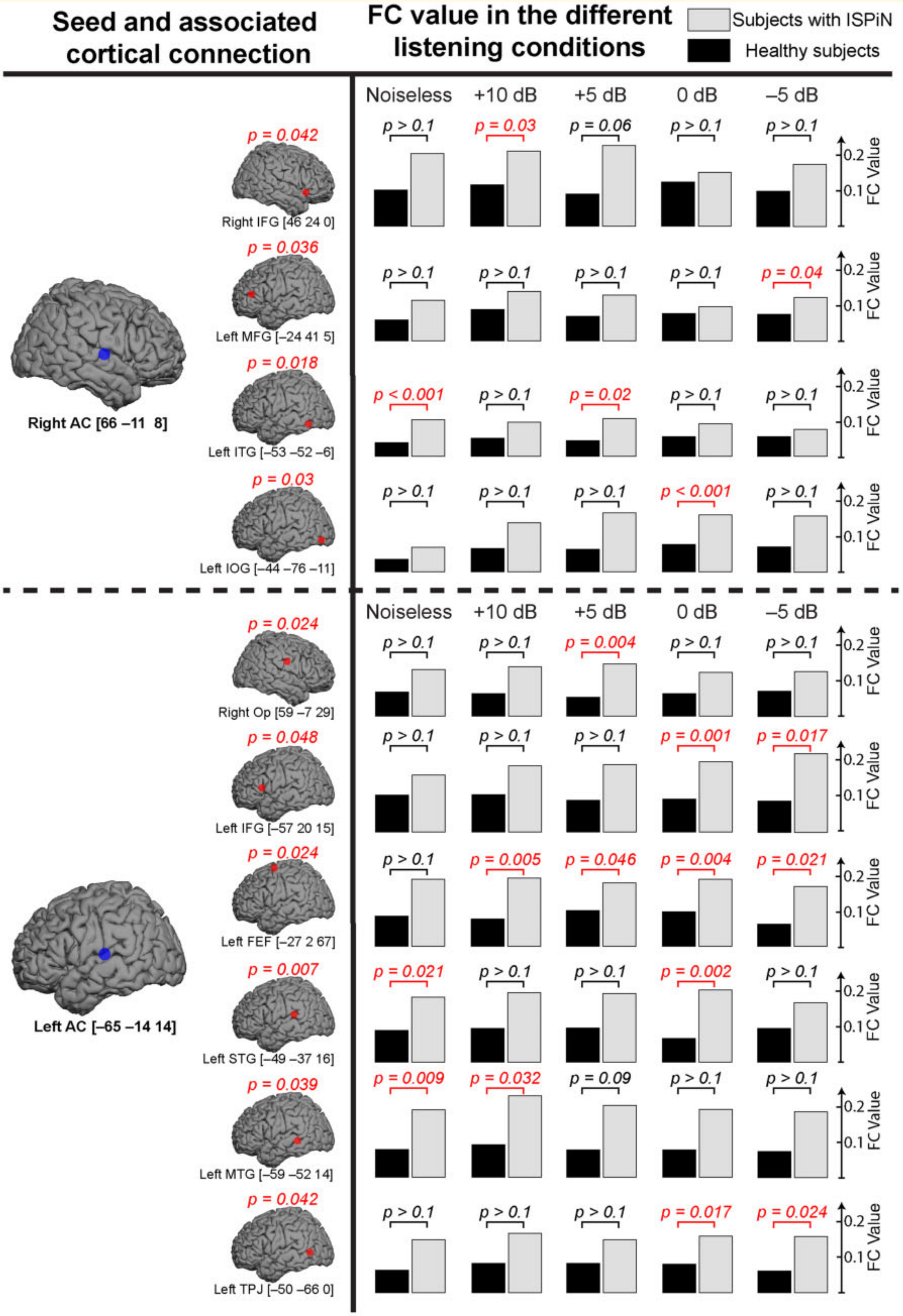


Figure 7 Comparison of seed-based functional connectivity (FC) between groups of participants. Left row: cortical regions (red dot) reflecting significant ANOVA main group effect (healthy subjects versus subjects with ISPiN), with their associated MNI coordinates in mm and corrected *P*-values, for seeds placed in the right and left AC (blue dot). Right row: Corresponding functional connectivity values and associated *post hoc t*-tests for each condition in both groups (healthy subjects/subjects with ISPiN).

Inaccurate cortical tracking of speech at 4–8 Hz in subjects with ISPiN

When listening to connected speech, ongoing auditory cortex oscillations align with speech rhythms at frequencies matching the syllable rate, i.e. at 4–8 Hz.[69–72] This cortical tracking is considered to be involved in speech perception by parsing incoming continuous speech into discrete syllabic units,[32,73,74] and by chunking different acoustic features at the syllable timescale to support the build-up of an auditory stream.[75,76] Speech intelligibility has been related to the magnitude of this cortical tracking,[77–79] which can moreover be exogenously enhanced with transcranial currents conveying speech-envelope information resulting in increased speech comprehension performance.[80,81] SiN comprehension is also improved with transcranial alternating current carrying information on the 4–8 Hz—but not the 1–4 Hz—content of the targeted speech envelope.[82] Taken together, these data highly suggest that cortical tracking of speech in this frequency range plays a key functional role in SiN comprehension. They also fit with our previous finding of inaccurate cortical tracking of speech at 4–8 Hz in children aged <10 years who typically display lower SiN processing abilities than adults.[31]

The results obtained in subjects with ISPiN bring additional empirical evidence supporting the critical role of an accurate auditory cortex tracking of the attended speech at 4–8 Hz for the successful understanding of SiN.

Cochlear synaptopathy has been raised as one possible etiopathogenic mechanism of ISPiN.[13,83] However, it develops mainly after a sound trauma—history of such trauma was missing in our subjects—and its electrophysiological signature in animal models (i.e. reduced cochlear responses to suprathreshold sound levels) is seldom replicated in human studies.[7,18,19,63,84–86] Furthermore, previous studies have shown that the cortical tracking of speech is increased in case of peripheral hearing loss,[87,88] suggesting that a reduced tracking rather reflects a central deficit. In addition, central pathologies with speech perception deficits, such as dyslexia and schizophrenia, are characterized by a decreased cortical tracking—especially at 1–4 Hz—of rhythmic auditory stimuli and connected speech.[89–94] Based on these considerations, our data do not support the involvement of cochlear synaptopathy in the mechanisms of ISPiN occurring in the absence of any history of sound trauma.

Enhanced functional connectivity between AC and language-related cortical areas in ISPiN

Increased functional connectivity during speech processing within an extended left-lateralized language-related cortical network was found in subjects with ISPiN compared with healthy subjects. The identified regions (i.e. IFG, MTG, STG, ITG, TPJ) corresponded to major nodes of the language processing network[95–97] that are critical for speech recognition.[95,98–100] Left-hemisphere dominance of SiN cortical processing has already been highlighted in healthy adults[30,68,101,102]; these left cortical regions being more resilient to acoustic degradations of speech signals.[103] However, a specific noise-related increase in functional connectivity was only highlighted with left IFG and left IOG.

The increased recruitment of left-dominant language-related brain areas observed in subjects with ISPiN during speech processing in the 4–8 Hz frequency range might underscore a mechanism of compensation for the inaccurate cortical tracking of speech at the syllable rate to support speech recognition. Indeed, increased functional connectivity was observed within brain areas contributing to syntaxic/semantic processing.[100,104,105] Alternatively, the increased recruitment may reflect differences in the intrinsic functional brain architecture of subjects with ISPiN as suggested by our data obtained at rest. An excessive intrinsic functional integration between low- and high-level linguistic regions could potentially alter the hierarchical processing of linguistic information, which would only become clinically apparent in adverse auditory scenes. This interpretation is more in line with the finding that the increased recruitment was equally present at rest and in all listening conditions.

In addition, in every SiN condition, subjects with ISPiN demonstrated stronger functional connections between left AC and left FEF. FEFs, which enable the planning and control of eye movements, are also involved in spatially directed attention, even in the absence of gaze changing (i.e. 'covert attention'[106]). Specifically, left FEF is implicated in auditory selective attention by controlling auditory spatial attention in a purely top-down manner.[76,107,108] Our results remain, however, puzzling since spatial cues, which lead to FEF activation, were lacking in our tasks. Interestingly, left FEF is anticipatively activated in the preparation of an auditory location task.[109] We hypothesize, therefore, that the increased functional connectivity between left AC and left FEF indicates a preparatory activity reflecting the additional need in 'orientated' auditory attention in subjects with ISPiN in noisy auditory scenes.

Conclusions

ISPiN is characterized by an inaccurate auditory cortical tracking of speech at the syllable rate in adverse auditory scenes and increased functional connectivity between auditory cortices and language/attention-related neocortical areas to support SiN recognition. These results argue for a central origin of ISPiN.

Supplementary material

Supplementary material is available at *Brain Communications* online.

Acknowledgements

We thank Nicolas Grimault and Fabien Perrin (Centre de Recherche en Neuroscience, Lyon, France) for providing us their entire audio material.

Funding

Marc Vander Ghinst, Gilles Naeije and Maxime Niesen were supported by a research grant from the Fonds Erasme (Brussels, Belgium). Mathieu Bourguignon was supported by the Program Attract of Innoviris (grant 2015-BB2B-10), Spanish Ministry of Economy and Competitiveness (grant PSI2016-77175-P) and Marie Skłodowska-Curie Action of the European Commission (grant 743562). Gilles Naeije and Xavier De Tiège are Post-doctorate Clinical Master Specialist at the Fonds de la Recherche Scientifique (FRS-FNRS, Brussels, Belgium). This study and the MEG project at the CUB Hôpital Erasme were financially supported by the Fonds Erasme (Research Convention 'Les Voies du Savoir', Fonds Erasme, Brussels, Belgium).

Competing interests

The authors report no competing interests.

References

1. Gates GA, Cooper JC, Kannel WB, Miller NJ. Hearing in the elderly: The Framingham cohort, 1983–1985. Part I. Basic audiometric test results. Ear Hear. 1990;11(4):247–256.
2. Chia E-M, Wang JJ, Rochtchina E, Cumming RR, Newall P, Mitchell P. Hearing impairment and health-related quality of life: The Blue Mountains Hearing Study. Ear Hear. 2007;28(2):187–195.
3. Zhao F, Stephens D. A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? Audiol Med. 2007;5(2):119–124.
4. Tremblay KL, Pinto A, Fischer ME, et al. Self-reported hearing difficulties among adults with normal audiograms: The Beaver Dam Offspring Study. Ear Hear. 2015;36(6):e290–e299.
5. Roup CM, Post E, Lewis J. Mild-gain hearing aids as a treatment for adults with self-reported hearing difficulties. J Am Acad Audiol. 2018;29(6):477–494.
6. Spankovich C, Gonzalez VB, Su D, Bishop CE. Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination Survey 1999–2002. Hear Res. 2018;358:30–36.
7. Parthasarathy A, Hancock KE, Bennett K, DeGruttola V, Polley DB. Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. eLife. 2020;9:e51419.
8. Saunders GH, Haggard MP. The clinical assessment of obscure auditory dysfunction–1. Auditory and psychological factors. Ear Hear. 1989;10(3):200–208.
9. Zhao F, Stephens D. Hearing complaints of patients with King-Kopetzky syndrome (obscure auditory dysfunction). Br J Audiol. 1996;30(6):397–402.
10. Pryce H, Metcalfe C, Hall A, Claire LS. Illness perceptions and hearing difficulties in King-Kopetzky syndrome: What determines help seeking? Int J Audiol. 2010;49(7):473–481.
11. King PF. Psychogenic deafness. J Laryngol Otol. 1954;68(9):623–635.
12. American Speech-Language-Hearing Association. (Central) auditory Processing Disorders. https://www.asha.org/policy/tr2005-00043/. Accessed 2005.
13. Schaette R, McAlpine D. Tinnitus with a normal audiogram: Physiological evidence for hidden hearing loss and computational model. J Neurosci. 2011;31(38):13452–13457.
14. Pienkowski M. On the etiology of listening difficulties in noise despite clinically normal audiograms. Ear Hear. 2017;38(2):135–148.
15. Badri R, Siegel JH, Wright BA. Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. J Acoust Soc Am. 2011;129(2):852–863.
16. Hoben R, Easow G, Pevzner S, Parker MA. Outer hair cell and auditory nerve function in speech recognition in quiet and in background noise. Front Neurosci. 2017;11:157.
17. Kujawa SG, Liberman MC. Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci. 2009;29(45):14077–14085.
18. Prendergast G, Guest H, Munro KJ, et al. Effects of noise exposure on young adults with normal audiograms I: Electrophysiology. Hear Res. 2017;344:68–81.
19. Smith SB, Krizman J, Liu C, White-Schwoch T, Nicol T, Kraus N. Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. Hear Res. 2019;371:66–74.
20. Blomberg R, Danielsson H, Rudner M, Söderlund GBW, Rönnberg J. Speech processing difficulties in attention deficit hyperactivity disorder. Front Psychol. 2019;10:1536.
21. Ziegler JC, Pech-Georgel C, George F, Lorenzi C. Speech-perception-in-noise deficits in dyslexia. Dev Sci. 2009;12(5):732–745.
22. Dole M, Hoen M, Meunier F. Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration. Neuropsychologia. 2012;50(7):1543–1552.
23. Zhao F, Stephens D. Subcategories of patients with King-Kopetzky syndrome. Br J Audiol. 2000;34(4):241–256.
24. Dundon NM, Dockree SP, Buckley V, et al. Impaired auditory selective attention ameliorated by cognitive training with graded exposure to noise in patients with traumatic brain injury. Neuropsychologia. 2015;75:74–87.
25. Kraus N, Thompson EC, Krizman J, Cook K, White-Schwoch T, LaBella CR. Auditory biological marker of concussion in children. Sci Rep. 2016;6:39009.
26. Hoover EC, Souza PE, Gallun FJ. Auditory and cognitive factors associated with speech-in-noise complaints following mild traumatic brain injury. J Am Acad Audiol. 2017;28(4):325–339.
27. Zhan Y, Fellows AM, Qi T, et al. Speech in noise perception as a marker of cognitive impairment in HIV infection. Ear Hear. 2018;39(3):548–554.
28. Ding N, Simon JZ. Emergence of neural encoding of auditory objects while listening to competing speakers. Proc Natl Acad Sci USA. 2012;109(29):11854–11859.
29. Zion Golumbic EM, Ding N, Bickel S, et al. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". Neuron. 2013;77(5):980–991.
30. Vander Ghinst M, Bourguignon M, Op de Beeck M, et al. Left superior temporal gyrus is coupled to attended speech in a

Cocktail-Party Auditory Scene. J Neurosci. 2016;36(5): 1596–1606.
31. Vander Ghinst M, Bourguignon M, Niesen M, et al. Cortical tracking of speech-in-noise develops from childhood to adulthood. J Neurosci. 2019;39(15):2938–2950.
32. Giraud AL, Poeppel D. Cortical oscillations and speech processing: Emerging computational principles and operations. Nat Neurosci. 2012;15(4):511–517.
33. Ding N, Simon JZ. Robust cortical encoding of slow temporal modulations of speech. Adv Exp Med Biol. 2013;787:373–381.
34. Keitel A, Gross J, Kayser C. Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. PLoS Biol. 2018;16(3):e2004473.
35. Elliott LL. Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. J Acoust Soc Am. 1979;66(3): 651–653.
36. Moore DR, Ferguson MA, Edmondson-Jones AM, Ratib S, Riley A. Nature of auditory processing disorder in children. Pediatrics. 2010;126(2):e382–e390.
37. Wens V, Marty B, Mary A, et al. A geometric correction scheme for spatial leakage effects in MEG/EEG seed-based functional connectivity mapping. Hum Brain Mapp. 2015;36(11): 4604–4621.
38. Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971;9(1):97–113.
39. Demanez L, Dony-Closon B, Lhonneux-Ledoux E, Demanez JP. Central auditory processing assessment: A French-speaking battery. Acta Otorhinolaryngol Belg. 2003;57(4):275–290.
40. Zimmermann P, Fimm B. A test battery for attentional performance. In: Taylor & Francis Group (eds.) Applied neuropsychology of attention: Theory. New Fetter Lane, London: Psychology Press; 2002:110–151.
41. Grimault N, Perrin F. *Fonds Sonores V-1.0.*; 2005.
42. Ferstl EC, Walther K, Guthke T, von Cramon DY. Assessment of story comprehension deficits after brain damage. J Clin Exp Neuropsychol. 2005;27(3):367–384.
43. De Tiege X, Op de Beeck M, Funke M, et al. Recording epileptic activity with MEG in a light-weight magnetic shield. Epilepsy Res. 2008;82(2-3):227–231.
44. Taulu S, Simola J, Kajola M. Applications of the signal space separation method. IEEE Trans Signal Process. 2005;53(9): 3359–3372.
45. Taulu S, Simola J. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys Med Biol. 2006;51(7):1759–1768.
46. Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. Neural Netw. 2000;13(4-5): 411–430.
47. Vigário R, Särelä J, Jousmäki V, Hämäläinen M, Oja E. Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans Biomed Eng. 2000;47(5):589–593.
48. Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. J Acoust Soc Am. 1994;95(2): 1053–1064.
49. Bourguignon M, Piitulainen H, De Tiege X, Jousmäki V, Hari R. Corticokinematic coherence mainly reflects movement-induced proprioceptive feedback. Neuroimage. 2015;106:382–390.
50. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage. 2012;61(4):1402–1418.
51. Ashburner J, Neelin P, Collins DL, Evans A, Friston K. Incorporating prior knowledge into image registration. Neuroimage. 1997;6(4):344–352.
52. Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. Hum Brain Mapp. 1999;7(4):254–266.
53. Gramfort A, Luessi M, Larson E, et al. MNE software for processing MEG and EEG data. Neuroimage. 2014;86:446–460.
54. Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. Proc Natl Acad Sci USA. 2001;98(2):694–699.
55. Dale AM, Sereno MI. Improved localizadon of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. J Cogn Neurosci. 1993;5(2):162–176.
56. Brookes MJ, Woolrich MW, Barnes GR. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. Neuroimage. 2012;63(2):910–920.
57. Wens V. Investigating complex networks with inverse models: Analytical aspects of spatial leakage and connectivity estimation. Phys Rev E. 2015;91(1):012823.
58. Faes L, Pinna GD, Porta A, Maestri R, Nollo G. Surrogate data analysis for assessing the significance of the coherence function. IEEE Trans Biomed Eng. 2004;51(7):1156–1166.
59. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum Brain Mapp. 2002;15(1):1–25.
60. Destoky F, Philippe M, Bertels J, et al. Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. Neuroimage. 2019;184:201–213.
61. Friston KJ, Penny WD, Ashburner JT, Kiebel SJ, Nichols TE. Statistical parametric mapping: The analysis of functional brain images. Boston: Elsevier; 2011.
62. Saunders G, Haggard M. The clinical assessment of C "Obscure Auditory Dysfunction" (OAD) 2. Case control analysis of determining factors. Ear Hear. 1992;13(4):241–254.
63. Guest H, Munro KJ, Prendergast G, Millman RE, Plack CJ. Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. Hear Res. 2018;364:142–151.
64. Demanez L, Boniver V, Dony-Closon B, Lhonneux-Ledoux F, Demanez JP. Central auditory processing disorders: Some cohorts studies. Acta Otorhinolaryngol Belg. 2003;57(4):291–299.
65. Oberfeld D, Klöckner-Nowotny F. Individual differences in selective attention predict speech identification at a cocktail party. eLife. 2016;5:e16747.
66. Tun PA, O'Kane G, Wingfield A. Distraction by competing speech in young and older adult listeners. Psychol Aging. 2002; 17(3):453–467.
67. Khalighinejad B, Herrero JL, Mehta AD, Mesgarani N. Adaptation of the human auditory cortex to changing background noise. Nat Commun. 2019;10(1):2509.
68. Power AJ, Foxe JJ, Forde EJ, Reilly RB, Lalor EC. At what time is the cocktail party? A late locus of selective attention to natural speech. Eur J Neurosci. 2012;35(9):1497–1503.
69. Ding N, Simon JZ. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J Neurophysiol. 2012;107(1):78–89.
70. Gross J, Hoogenboom N, Thut G, et al. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS Biol. 2013;11(12):e1001752-
71. Koskinen M, Seppa M. Uncovering cortical MEG responses to listened audiobook stories. Neuroimage. 2014;100:263–270.
72. Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. Nat Neurosci. 2016;19(1):158–164.
73. Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud A-L. Speech encoding by coupled cortical theta and gamma oscillations. eLife. 2015;4:e06213.
74. Teng X, Tian X, Doelling K, Poeppel D. Theta band oscillations reflect more than entrainment: Behavioral and neural evidence demonstrates an active chunking process. Eur J Neurosci. 2018; 48(8):2770–2782.
75. Riecke L, Sack AT, Schroeder CE. Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. Curr Biol. 2015;25(24):3196–3201.

76. Shinn-Cunningham B, Best V, Lee AKC. Auditory object formation and selection. In: JC Middlebrooks, JZ Simon, AN Popper, RR Fay, eds. The auditory system at the cocktail party. Springer handbook of auditory research. Switzerland: Springer International Publishing; 2017:7–40.
77. Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron. 2007; 54(6):1001–1010.
78. Peelle JE, Gross J, Davis MH. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. Cereb Cortex. 2013;23(6):1378–1387.
79. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. Neuroimage. 2014;85: 761–768.
80. Riecke L, Formisano E, Sorger B, Başkent D, Gaudrain E. Neural entrainment to speech modulates speech intelligibility. Curr Biol. 2018;28(2):161–169.e5.
81. Wilsch A, Neuling T, Obleser J, Herrmann CS. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. Neuroimage. 2018;172: 766–774.
82. Keshavarzi M, Kegler M, Kadir S, Reichenbach T. Transcranial alternating current stimulation in the theta band but not in the delta band modulates the comprehension of naturalistic speech in noise. Neuroimage. 2020;210:116557.
83. Kujawa SG, Liberman MC. Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss. Hear Res. 2015;330:191–199.
84. Fulbright ANC, Prell CGL, Griffiths SK, Lobarinas E. Effects of recreational noise on threshold and suprathreshold measures of auditory function. Semin Hear. 2017;38(4):298–318.
85. Grinn SK, Wiseman KB, Baker JA, Le Prell CG. Hidden hearing loss? No effect of common recreational noise exposure on cochlear nerve response amplitude in humans. Front Neurosci. 2017; 11:465.
86. Le Prell CG, Siburt HW, Lobarinas E, Griffiths SK, Spankovich C. No reliable association between recreational noise exposure and threshold sensitivity, distortion product otoacoustic emission amplitude, or word-in-noise performance in a college student population. Ear Hear. 2018;39(6):1057–1074.
87. Millman RE, Mattys SL, Gouws AD, Prendergast G. Magnified neural envelope coding predicts deficits in speech perception in noise. J Neurosci. 2017;37(32):7727–7736.
88. Fuglsang SA, Märcher-Rørsted J, Dau T, Hjortkjær J. Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. J Neurosci. 2020;40(12): 2562–2572.
89. Lakatos P, Musacchia G, O'Connel MN, Falchier AY, Javitt DC, Schroeder CE. The spectrotemporal filter mechanism of auditory selective attention. Neuron. 2013;77(4):750–761.
90. Soltész F, Szűcs D, Leong V, White S, Goswami U. Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. PLoS ONE. 2013;8(10):e76608.
91. Molinaro N, Lizarazu M, Lallier M, Bourguignon M, Carreiras M. Out-of-synchrony speech entrainment in developmental dyslexia. Hum Brain Mapp. 2016;37(8):2767–2783.
92. Power AJ, Colling LJ, Mead N, Barnes L, Goswami U. Neural encoding of the speech envelope by children with developmental dyslexia. Brain Lang. 2016;160:1–10.
93. Di Liberto GM, Peter V, Kalashnikova M, Goswami U, Burnham D, Lalor EC. Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. Neuroimage. 2018;175:70–79.
94. Thiede A, Glerean E, Kujala T, Parkkonen L. Atypical MEG inter-subject correlation during listening to continuous natural speech in dyslexia. Neuroimage. 2020;216:116799.
95. Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci. 2007;8(5):393–402.
96. Friederici AD. The brain basis of language processing: From structure to function. Physiol Rev. 2011;91(4):1357–1392.
97. Hagoort P. MUC (memory, unification, control) and beyond. Front Psychol. 2013;4:416.
98. Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition. 2004;92(1-2):145–177.
99. Turken AU, Dronkers NF. The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. Front Syst Neurosci. 2011;5:1.
100. Price CJ. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage. 2012;62(2):816–847.
101. Alain C, Reinke K, McDonald KL, et al. Left thalamo-cortical network implicated in successful speech separation and identification. Neuroimage. 2005;26(2):592–599.
102. Bishop CW, Miller LM. A multisensory cortical network for understanding speech in noise. J Cogn Neurosci. 2009;21(9):1790–1805.
103. Davis MH, Johnsrude IS. Hierarchical processing in spoken language comprehension. J Neurosci. 2003;23(8):3423–3431.
104. Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. Proc Natl Acad Sci USA. 1997;94(26):14792–14797.
105. Fedorenko E, Duncan J, Kanwisher N. Language-selective and domain-general regions lie side by side within Broca's area. Curr Biol. 2012;22(21):2059–2062.
106. Wardak C, Ibos G, Duhamel J-R, Olivier E. Contribution of the monkey frontal eye field to covert visual attention. J Neurosci. 2006;26(16):4228–4235.
107. Bharadwaj HM, Lee AKC, Shinn-Cunningham BG. Measuring auditory selective attention using frequency tagging. Front Integr Neurosci. 2014;8:6.
108. Lewald J, Hanenberg C, Getzmann S. Brain correlates of the orientation of auditory spatial attention onto speaker location in a "cocktail-party" situation. Psychophysiology. 2016;53(10):1484–1495.
109. Lee AK, Rajaram S, Xia J, et al. Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. Front Neurosci. 2012;6:190.