Brain activity underlying the recovery of meaning from degraded speech: a functional near-infrared spectroscopy (fNIRS) study

Pramudi Wijayasiri[1,2], Douglas E H Hartley[1,2,3,4], and Ian M Wiggins[1,2,3]

[1] National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Ropewalk House, 113 The Ropewalk, Nottingham, NG1 5DU, United Kingdom

[2] Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, NG7 2UH, United Kingdom

[3] Medical Research Council (MRC) Institute of Hearing Research, School of Medicine, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

[4] Nottingham University Hospitals NHS Trust, Queens Medical Centre, Derby Road, Nottingham, NG7 2UH, United Kingdom

Corresponding author:

Ian M Wiggins

NIHR Nottingham Biomedical Research Centre

Ropewalk House

113 The Ropewalk

Nottingham

NG1 5DU

United Kingdom

E-mail: ian.wiggins@nottingham.ac.uk

Phone: +44 (0)115 823 2600

Abstract

The purpose of this study was to establish whether functional near-infrared spectroscopy (fNIRS), an emerging brain-imaging technique based on optical principles, is suitable for studying the brain activity that underlies effortful listening. In an event-related fNIRS experiment, normally-hearing adults listened to sentences that were either clear or degraded (noise vocoded). These sentences were presented simultaneously with a non-speech distractor, and on each trial participants were instructed to attend either to the speech or to the distractor. The primary region of interest for the fNIRS measurements was the left inferior frontal gyrus (LIFG), a cortical region involved in higher-order language processing. The fNIRS results confirmed findings previously reported in the functional magnetic resonance imaging (fMRI) literature. Firstly, the LIFG exhibited an elevated response to degraded versus clear speech, but only when attention was directed towards the speech. This attention-dependent increase in frontal brain activation may be a neural marker for effortful listening. Secondly, during attentive listening to degraded speech, the haemodynamic response peaked significantly later in the LIFG than in superior temporal cortex, possibly reflecting the engagement of working memory to help reconstruct the meaning of degraded sentences. The homologous region in the right hemisphere may play an equivalent role to the LIFG in some left-handed individuals. In conclusion, fNIRS holds promise as a flexible tool to examine the neural signature of effortful listening.

Keywords

functional near-infrared spectroscopy; fNIRS; listening effort; speech comprehension; noise vocoding; auditory cortex; inferior frontal gyrus; neuroimaging

1. Introduction

Listening to speech in the complex environments of daily life can be an effortful task, especially for individuals with hearing loss (Alhanbali et al., 2016; McGarrigle et al., 2014). The need to expend greater cognitive effort to understand an auditory message can have serious consequences: listeners may have more difficulty remembering what they have heard (McCoy et al., 2005; Rabbitt, 1990; Ward et al., 2016) and may perform more poorly in situations that require multitasking (Hicks and Tharpe, 2002; Picou et al., 2013; Sarampalis et al., 2009); in the longer term, feelings of stress and fatigue associated with the increased effort of listening may lead to greater need for recovery after work (Nachtegaal et al., 2009), increased incidence of sick-leave (Kramer et al., 2006), and, ultimately, social withdrawal and isolation, potentially accelerating cognitive decline in older adults (Lin et al., 2013). It is therefore critical to improve our understanding of the neurocognitive processes that underlie effortful listening, so that we may detect and remediate it more effectively in future. The present study aimed to establish the potential of the emerging brain-imaging technique functional near-infrared spectroscopy (fNIRS) to shed light on the brain activity that underlies effortful listening.

A non-invasive technique based on optical measurements, fNIRS is rapidly gaining popularity as a flexible tool for imaging the haemodynamic response to neuronal activity in the human brain (Boas et al., 2014). The technique works by illuminating the brain with infrared light through the intact scalp. By measuring the intensity of light that returns to the surface, changes in local concentrations of oxygenated (HbO) and de-oxygenated (HbR) haemoglobin are derived (Scholkmann et al., 2014). Since active brain regions demand the delivery of oxygen to support their metabolic needs (Iadecola, 2004), changes in HbO and HbR concentrations can be taken as indicators of cortical activation. Interest in using fNIRS to study central auditory processing has grown rapidly in recent years (Chen et al., 2015; Hassanpour et al., 2015; Hong and Santosa, 2016; Plichta et al., 2011; Pollonini et al., 2014; Sevy et al., 2010; van de Rijt et al., 2016; Wiggins et al., 2016). However, compared to more established brain-imaging techniques like functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), the application of fNIRS in hearing research is in its infancy and the fundamental capabilities of the technique are still being established. Despite this, the potential benefits of fNIRS are clear: it is quiet, affordable, mobile, relatively forgiving of subject motion, compatible with hearing aids and cochlear implants, and suitable for use with all age groups and in naturalistic testing environments.

A common theme in models and conceptual frameworks that seek to explain effortful listening is that as the quality of the auditory input is reduced, additional cognitive resources are recruited to help recover meaning (Pichora-Fuller, 2016; Ronnberg et al., 2013). Brain-imaging studies using fMRI have proven useful in elucidating some of the brain regions and networks responsible for performing this additional work (Adank et al., 2012; Davis and Johnsrude, 2003; Davis and Johnsrude, 2007; Eckert et al., 2016; Erb et al., 2013; Evans et al., 2016; Golestani et al., 2013; Hervais-Adelman et al., 2012; Obleser et al., 2007; Scott and McGettigan, 2013; Wild et al., 2012; Zekveld et al., 2006; Zekveld et al., 2014). The regions in question primarily lie beyond the auditory cortex in higher-order language-sensitive areas (Peelle et al., 2010), as well as other non-auditory areas that support performance monitoring and attention (Eckert et al., 2016). One region that has often, although not universally (Adank, 2012), been implicated in compensating for acoustic degradation to the speech signal is the left inferior frontal gyrus (LIFG). Across a range of studies, the LIFG has been shown to respond more strongly to degraded-yet-intelligible speech than to either clear speech or unintelligible noise (Adank et al., 2012; Davis and Johnsrude, 2003; Hervais-Adelman et al., 2012; Wild et al., 2012; Zekveld et al., 2006).

Importantly, Wild et al. (2012) showed that elevated LIFG activation depends critically on attention to speech, indicating that it reflects the active attempt to recover meaning from degraded speech, rather than an obligatory response resulting merely from exposure to a degraded signal. Subsequently, Wild et al. (2012) proposed that measuring activation in frontal, speech-sensitive regions using fMRI may provide a brain-based objective measure of listening effort. Such a measure could potentially be used to help evaluate and optimize the design and programming of hearing devices. The use of a brain-based objective measure may be advantageous compared to the intuitive approach of simply asking listeners to rate their perceived effort, since such subjective ratings are prone to bias and listeners may struggle to report mental effort independently of perceived performance or task difficulty (McGarrigle et al., 2014). Such a measure may also prove useful in cases where behavioural assessment is challenging, for example in fitting hearing devices to infants. However, practical applications may be compromised by methodological challenges inherent to fMRI, including high levels of scanner noise (Peelle, 2014) and incompatibility with implanted auditory prostheses. Could fNIRS, a neuroimaging modality that is silent, child-friendly, and compatible with cochlear implants (Bisconti et al., 2015; Chen et al., 2016; McKay et al., 2016; Olds et al., 2015; Sevy et al., 2010; van de Rijt et al., 2016), provide a more flexible alternative? Hassanpour et al. (2015) recently showed that high-density diffuse optical tomography (HD-DOT), a technique closely related to fNIRS, is capable of mapping distributed patterns of brain activation during speech perception, including in higher-order frontal regions. However, most commercially available fNIRS systems use a sparser arrangement of sources and detectors than in HD-DOT, which reduces the spatial resolution of the measurements and increases susceptibility to interference from physiological signals of extra-cerebral origin (Scholkmann et al., 2014). Based on the existing literature, it is unclear whether standard fNIRS systems are capable of measuring elevated frontal-lobe activation associated with the processing of degraded speech.

The aims of the present study were: (i) to establish whether fNIRS is a suitable imaging modality for measuring frontal-lobe activation during speech perception; (ii) to corroborate using fNIRS Wild et al.’s (2012) fMRI findings that suggest elevated activation in the LIFG during listening to degraded speech depends critically on attention. To achieve these aims, we designed an fNIRS experiment broadly based on the procedures described by Wild et al. (2012), in which we included a subset of the stimulus conditions that were used in their study. Briefly, we used fNIRS to measure frontotemporal brain activation while participants listened to clear or degraded (four-channel noise-vocoded) sentences in the presence of a non-speech auditory distractor. Listening to four-channel noise-vocoded speech is known to be an effortful task for normally hearing individuals, as revealed by behavioural measures (e.g. increased response time on a secondary task; Pals et al., 2013), self-reported effort (Pals et al., 2013) and pupil dilation (Winn et al., 2015). On each trial, participants were cued to attend either to the speech or the distractor. Based on the extant findings, we predicted elevated LIFG activation in response to degraded versus clear speech, but only when attention was directed towards the speech.

2. Materials and methods

2.1 Participants and ethical approval

Twenty undergraduate students from the University of Nottingham (mean age 20.3 years, range 18–22 years, 11 males) participated in the study after giving written informed consent. All participants were native speakers of English with no known hearing problems (self-reported), normal or corrected-to-normal vision, and no history of any motor or cognitive impairment. The majority of participants (17 out of 20) were right-handed as assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). The study was approved by the University of Nottingham Faculty of Medicine and Health Sciences Research Ethics Committee.

2.2 Equipment

Testing took place in a double-walled sound booth. Participants were seated at a distance of 75 cm from a visual display unit. Auditory stimuli were presented in the free-field using a Genelec (Iisalmi, Finland) 8030A loudspeaker. Presentation levels were measured at the listening position using a Brüel & Kjær (Nærum, Denmark) Type 2250 sound level meter with the participant absent. Brain activity was non-invasively measured using a Hitachi (Tokyo, Japan) ETG-4000 continuous-wave fNIRS system. The ETG-4000 measures simultaneously at wavelengths of 695 nm and 830 nm (sampling rate 10 Hz), and uses frequency modulation to minimize crosstalk between channels and wavelengths (Scholkmann et al., 2014). Participants entered their responses using an “RTbox” button box (Li et al., 2010). The experiment was implemented in MATLAB (MathWorks, Natick, MA) using the Psychtoolbox-3 extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

2.3 Experimental design

We simultaneously conducted an event-related fNIRS imaging experiment along with a behavioural task. A 2x2 factorial design was used with factors “speech clarity” (clear vs. degraded) and “attention” (attend speech vs. attend distractor). On every trial, participants were presented with a mixture of a sentence (either clear or degraded) and a non-speech distractor. A visual prompt presented 1 s before the onset of auditory stimulation cued participants to attend either to the speech or the distractor. Following the offset of stimulation, after a delay of 0.5 s, participants were questioned on the cued stimulus in a yes/no decision task. Each trial type was presented 20 times in random interleaved order. An additional 20 silent trials, randomly interspersed among the stimulation trials, were included to act as a baseline. On silent trials, instead of responding to a specific question, participants were instructed to press any button in response to an appropriately timed on-screen prompt. The stimulus-onset asynchrony (SOA; the time between the onset of auditory stimulation on one trial and the next) was randomly varied in the range 6–12 s. Randomizing the SOA helps to reduce the influence of preparatory and anticipatory factors and can improve the efficiency of event-related designs (Dale, 1999). The timing of the behavioural task set a lower bound for the SOA, while the range of SOAs was informed by a previous event-related fNIRS study (Plichta et al., 2007). The total duration of fNIRS imaging was approximately 15 minutes.

2.4 Speech stimuli

Speech material consisted of recordings of Bamford-Kowal-Bench sentences (Bench et al., 1979) spoken by a male talker. Twenty sentence lists were available, each comprising sixteen sentences and a total of fifty keywords. An example sentence with keywords underlined is: “The clown had a funny face.” For each participant, we selected a random subset of the available lists to use for testing.

We used noise vocoding to create sentences that were acoustically degraded, yet still mostly intelligible. Noise vocoding reduces spectral clarity while preserving temporal envelope cues (Shannon et al., 1995). Speech signals were filtered into four adjacent frequency bands spaced approximately equally along the basilar membrane (Greenwood, 1990). The four bands spanned an overall bandwidth of 180–8000 Hz. Zero-phase bandpass filtering was performed by filtering signals first in the forward and then in the reverse direction (MATLAB *filtfilt* function) using 6th-order digital elliptic filters. The amplitude envelope within each frequency band was extracted by half-wave rectification followed by low-pass filtering at 160 Hz (1st-order elliptic filter applied consecutively in the forward and reverse directions). Each envelope was then applied to a white-noise carrier and bandpass filtered using the same scheme as described above. The root-mean-square (RMS) output level after processing was matched to the input level on a within-band basis. Summation of the four frequency bands gave the final degraded speech stimulus. On clear-speech trials, the vocoder was bypassed but the speech signal was passed through the same filter-bank to ensure equivalent bandwidth on clear- and degraded-speech trials. Speech stimuli were presented at 65 dB SPL (A-weighted).

2.5 Non-speech distractor

The non-speech distractor comprised a sequence of three or more 200-ms narrow-band ramped noise bursts (described to participants as “chirps”). The distractor stimulus began 0.1 s after the sentence and ended at approximately the same time as the sentence. The temporal spacing of individual noise bursts was fixed within a trial, but varied randomly across trials (inter-burst silence: 100–180 ms).

To create each noise burst, broadband white noise was passed through a 1000-Hz fixed-bandwidth filter (6th-order Butterworth) with a centre frequency randomly selected in the range 4500–5500 Hz. Onset and offset linear ramp durations for the standard bursts were 190 ms and 10 ms, respectively, meaning that each burst had a gradual onset and an abrupt offset. The level of the standard noise bursts was approximately 60 dB SPL.

On 50% of trials, a randomly selected noise burst within the sequence (never the first or last) was designated as a target. The target burst was distinguished from the standard bursts by a reversal of the onset and offset ramp durations (resulting in an abrupt onset and gradual offset) and by a 4-dB level enhancement.

2.6 Speech intelligibility testing

Speech intelligibility tests were conducted at the start of the session, prior to fNIRS imaging. Sentences were presented one at a time, with the participant asked to repeat back as many words as they could. The number of keywords correctly reported was scored by the experimenter using a touchscreen interface. No feedback was provided to participants. To allow participants practice listening to noise-vocoded speech, the intelligibility of degraded speech was initially measured using two consecutively presented lists. Subsequently, the intelligibility of both clear and degraded speech was measured using one additional list per condition. During this later stage of testing, clear and degraded sentences were randomly interleaved to achieve listening conditions more representative of those encountered during fNIRS imaging.

2.7 Main task

In the main task, conducted simultaneously with fNIRS imaging, participants were asked a yes/no question after each trial that related to the cued stimulus. When cued to attend to the speech, participants were asked to identify whether a word shown on the display had featured in the sentence. On 50% of trials the word was a true keyword, while on the other 50% of trials the word was a replacement foil word. Foil words were chosen to rhyme with a keyword, and, where possible, to be semantically plausible (e.g. in the sentence “The green tomatoes are small”, the keyword “green” might have been replaced with the foil word “clean”). Foil words were equally likely to fall towards the start, middle or end of the sentence. When cued to attend to the non-speech distractor, “Target?” was presented on the display and participants were asked to identify whether a target had been present in the distractor stimulus. Participants gave their responses by pressing one of two buttons to signify a “Yes” or “No” response according to labels shown on the display. For even-numbered participants, the “Yes” and “No” buttons were placed on the right and left side, respectively; the sides were reversed for odd-numbered participants. Participants had up to 2 s to respond, otherwise a missed response was recorded.

Participants were given time to practise the main task before data collection began. In initial practice runs, the speech and distractor stimuli were presented separately to allow participants to practise the respective tasks in isolation. In later practice runs, the speech and distractor stimuli were presented simultaneously. To encourage selective attention to the cued stimulus, participants were explicitly instructed to ignore the un-cued stimulus.

Sensitivity (*d*’) for detecting true keywords (attend-speech trials) or the presence of a target (attend-distractor trials) was determined for each experimental condition by subtracting the z-transform of the false-alarm rate (i.e. the proportion of “Yes” answers in response to a foil word/absent target) from the z-transform of the hit rate (i.e. the proportion of “Yes” answers in response to a true keyword/present target).

2.8 Sentence recognition post-test

Following Wild et al. (2012), we conducted a surprise sentence recognition test immediately after fNIRS imaging to test participants’ memory for the sentences they had heard. Participants had to decide whether each of a series of presented sentences had featured during fNIRS imaging or was an unfamiliar sentence not heard previously. All sentences, regardless of how they had originally been presented, were played as clear speech. Twelve sentences out of 20 from each of the four experimental conditions were randomly selected and interspersed among 24 unfamiliar sentences (chosen randomly from all so-far unused lists). This gave an overall 2:1 ratio of familiar to unfamiliar sentences. Sensitivity (*d*') for detecting familiar sentences was determined based on the z-transform of the hit rate for each experimental condition minus the z-transform of the common false-alarm rate across all unfamiliar sentences.

2.9 fNIRS data acquisition

We used a 3 x 11 optode array comprising 17 emitters and 16 detectors, giving 52 measurement channels in total. The source-detector spacing was fixed at 30 mm. The international 10-20 system (Jasper, 1958) was used to guide optode placement over the frontal and (superior) temporal lobes. The central optode in the bottom row of the array was placed on the forehead over position Fpz with the outermost optodes in the bottom row aligned inferolaterally towards position T3/T4 (Fig. 1a).

To evaluate the consistency of optode placement across individuals, a 3D digitizer was used to record the positions of the optodes, as well as anatomical surface landmarks, on eight volunteers. The measured positions were registered to the "Colin 27" atlas brain (Collins et al., 1998) using the AtlasViewer tool (Aasted et al., 2015). Across individuals, the standard deviation in the position of each optode was 6.64 ± 0.53 mm (mean ± SD across optodes) (Fig. 1b), which we considered satisfactory relative to the 30-mm source-detector spacing. The mean optode positions across the eight volunteers were used as the basis for subsequent assessments of probe sensitivity and for data visualization.

< Please insert Figure 1 here>

We also used AtlasViewer to confirm that our optode array provided sensitivity to the cortical region of interest (ROI): the LIFG. First, a cortical sensitivity profile was calculated for each measurement channel by running the photon migration forward problem, i.e. by simulating the probabilistic path of photons as they traverse through the head from source to detector (Aasted et al., 2015). The forward problem was evaluated using the Monte-Carlo photon transport software tMCimg (Boas et al., 2002), with 1 x 10^7 simulated photons launched from each optode. We then computed the aggregate sensitivity profile for three channels (Channels 30, 40 & 51) overlying the LIFG (Fig. 1c). We pre-selected these three channels as our primary ROI for subsequent analyses, as studies of fNIRS test-retest reliability have consistently indicated that fNIRS response amplitude is more reliable when averaged across a small number of channels overlying a cortical ROI than when assessed on a single-channel basis (Plichta et al., 2006; Schecklmann et al., 2008; Wiggins et al., 2016).

2.10 fNIRS data analysis

The fNIRS recordings were analyzed in MATLAB using functions provided in the HOMER2 package (Huppert et al., 2009) together with custom scripts. Since poor optode–scalp contact can be a limiting factor affecting fNIRS data quality, we initially screened the measurements to exclude channels that suffered from unacceptable signal quality. To do this, we used the scalp coupling index (SCI) approach introduced by Pollonini et al. (2014). In this method, the raw fNIRS signals at each wavelength are bandpass filtered between 0.5 and 2.5 Hz to isolate the cardiac component and the strength of correlation between the two wavelengths is taken as a measure of the quality of optode–scalp contact. Since our probe array did not provide for spatially overlapping channels, we were motivated to exclude channels from the analysis as sparingly as possible. We therefore adopted a liberal threshold of SCI ≥ 0.32 to retain channels in the analysis, chosen to exclude only the worst 5% of channels across the entire dataset.

Processing of the data for the retained channels proceeded as follows. First, the raw intensity signals were converted to changes in optical density (Huppert et al., 2009). Next, we used wavelet filtering to suppress motion artefacts. Specifically, we used the *hmrMotionCorrectWavelet* function provided in the HOMER2 package, which implements a simplified form of the algorithm described by Molavi and Dumont (2012). This function applies a probability threshold to remove outlying wavelet coefficients, which are assumed to correspond to motion artefacts. We excluded coefficients lying further than 0.719 times the interquartile range below the first quartile or above the third quartile. Assuming a Gaussian distribution of wavelet coefficients, this is equivalent to the $\alpha = 0.1$ threshold adopted in previous evaluations of fNIRS motion-correction techniques (Brigadoi et al., 2014; Cooper et al., 2012).

Following motion-artefact correction, we band-pass filtered the optical density signals between 0.02 and 0.5 Hz to attenuate low-frequency drift and cardiac oscillations. At this point, the optical density signals were converted into estimated changes in the concentrations of HbO and HbR using the modified Beer-Lambert law (Huppert et al., 2009). We used a default value of 6 for the differential path-length factor at both wavelengths. Note that this does not account for the partial-volume effect associated with focal changes in absorption (Boas et al., 2001). However, this was not a major concern here as our interest was not in estimating absolute changes in haemoglobin concentrations, but rather in contrasting relative response magnitude across experimental conditions.

As a final stage of pre-processing, we applied the haemodynamic signal separation algorithm described by Yamada et al. (2012) to isolate the functional component of the haemodynamic signal. This algorithm aims to reduce systemic physiological interference by exploiting the fact that changes in the concentrations of HbO and HbR tend to be negatively correlated in the functional cerebral response, but positively correlated in systemic physiological fluctuations (Yamada et al., 2012). We have shown previously that application of this algorithm can markedly improve the reliability of group-level speech-evoked fNIRS responses (Wiggins et al., 2016).

We used the general linear model (GLM) approach to quantify fNIRS response amplitude on a channel-wise basis (Schroeter et al., 2004). The design matrix included a set of three regressors for each of the four experimental conditions plus an additional set for the silent trials. Each trial was modelled as a short epoch corresponding to the period of auditory stimulation (mean duration 1.64 s; audio muted on silent trials). The trial time-courses were convolved with the canonical dual-gamma haemodynamic response function (HRF) provided in SPM8 [http://www.fil.ion.ucl.ac.uk/spm] together with its temporal and dispersion derivatives. Inclusion of the derivative terms allows the model to capture responses that are shifted in time or have extended activation durations compared to the canonical HRF (Friston et al., 1998; Lindquist and Wager, 2007; Lindquist et al., 2009). Within each condition, the regressor corresponding to the temporal derivative was orthogonalized with respect to the regressor for the canonical HRF, and the regressor corresponding to the dispersion derivative was orthogonolized with respect to both the canonical and temporal-derivative regressors. Model estimation was performed using the two-stage ordinary least squares procedure described by Plichta et al. (2007), which incorporates a correction for serial correlation (Cochrane and Orcutt, 1949).

To quantify response amplitude free from bias by any systematic differences in latency/dispersion between conditions, we combined the beta weights for the non-derivative and derivative terms using the ‘derivative-boost’ technique (Calhoun et al., 2004). We post-normalized the relevant columns of the design matrix following Steffener et al. (2010). Note that some authors (e.g. Cignetti et al., 2016) imposed additional constraints when applying the derivative boost, such that the beta weight for the canonical HRF is boosted by the non-derivative term(s) only when the estimated response meets certain criteria, for example, having a time-to-peak within the range 4–6 s. No such constraint was imposed here, although additional analyses (not shown) confirmed that the key findings were unchanged if the derivative boost was selectively applied only to responses peaking within a certain time window.

At the single-subject level, contrast images were formed by subtracting the estimated response amplitude for the silent baseline condition from that for each of the four experimental conditions. These contrast images formed the basis for subsequent statistical testing. Note that while the convention in fNIRS studies is to report results for both HbO and HbR, here we report results for HbO only. This is because the haemodynamic signal separation algorithm assumes a linear relationship between HbO and HbR in the functional response (Yamada et al., 2012), meaning that the two chromophores become statistically redundant after application of this algorithm.

2.11 Statistical analyses

Analyses of variance (ANOVAs) were carried out using IBM SPSS Statistics for Windows Version 22.0 software (IBM Corp., Armonk, New York). The Greenhouse-Geisser correction for non-sphericity was applied where necessary. Follow-up pairwise comparisons were Bonferroni corrected to account for multiple comparisons. Where appropriate, error bars in figures were corrected to make them suitable for repeated-measures data (Field, 2009).

2.11.1 Behavioural data

Behavioural data from the main task and sentence recognition post-test were analysed using two-way repeated-measures ANOVAs (RM-ANOVAs) with within-subjects factors “speech clarity” (two levels: clear vs. degraded) and “attention” (two levels: attend speech vs. attend distractor). A different approach was used to analyse the speech intelligibility data, since the RM-ANOVA approach is ill-suited for use with proportional data. We analysed the proportion of keywords correctly identified using logistic regression implemented within the context of a generalized linear mixed-effects analysis (Dixon, 2008). This analysis was conducted in *R* software (R Core Team, 2016) using the ‘lme4’ package (Bates et al., 2015). A random effect (intercept only) of “participant” was included to account for individual variability in performance, and “list” was treated as a fixed effect with four levels (three sentence lists for degraded speech plus one list for clear speech). Follow-up pairwise comparisons were conducted on the estimated marginal means to test for differences in intelligibility across lists, with Tukey's method used to control the familywise error rate.

2.11.2 fNIRS data

To test for significant cortical activation in each experimental condition compared to the silent baseline, we conducted group-level random-effects analyses using one-tailed *t*-tests on the corresponding contrast images. To account for the multiple comparisons issue arising from separately testing for cortical activation in 52 measurement channels, we controlled the false discovery rate (FDR) (Benjamini and Hochberg, 1995). We used the original formulation of the FDR procedure, which assumes either independence or slight positive dependency among channels, as recommended for fNIRS data analysis by Singh and Dan (2006).

Analysis of response amplitude within the primary ROI was conducted using a two-way RM-ANOVA similar to those used to analyse the behavioural main task and sentence-recognition post-test data. The same approach was also used to analyse response amplitude in a number of secondary, post-hoc ROIs defined in a data-driven manner. The input data to these RM-ANOVAs were the contrast values for each experimental condition compared to the silent baseline condition, averaged across the constituent channels that made up each ROI.

To test for interregional haemodynamic timing differences between temporal and frontal areas, we reconstructed estimates of the haemodynamic response in each region by combining each individual's GLM beta weights for the canonical and temporal derivative terms with the corresponding basis functions (Steffener et al., 2010). Time-to-peak values were extracted and compared across regions using a paired-samples *t*-test.

To assess the effect of handedness on the lateralization of brain activation, we conducted mixed-design ANOVAs with "handedness" as a dichotomous between-subjects factor and "hemisphere" as a within-subjects factor. Separate ANOVAs were performed for superior temporal and inferior frontal cortex, in each case contrasting brain activation between homologous regions in the left and right hemispheres.

3. Results and discussion

3.1 Behavioural results

3.1.1 Speech intelligibility

Participants' ability to correctly understand noise-vocoded speech improved with practice, ultimately reaching ~85% correct (Fig. 2). Generalised linear mixed-effect analysis of the percentage of keywords correctly reported in each sentence list (three lists for degraded speech plus one for clear speech) showed a significant overall effect of list ($\chi^2(3) = 119.51$, $p < .001$). Follow-up pairwise comparisons revealed that intelligibility significantly improved between the first (mean 68.4%-correct) and second (mean 81.9%-correct) degraded-speech lists ($p < .001$), but not between the second and third (mean 85.2%-correct) lists ($p = .197$). Performance appears to have approached a plateau at this stage, indicating that participants had received sufficient practice. As expected, even after practice, intelligibility was significantly worse for degraded (mean 85.2%-correct) than for clear (99.9%-correct) speech ($p < .001$).

< Please insert Figure 2 here>

3.1.2 ***Main task***

Mean sensitivity was significantly above chance level in all conditions (Fig. 3), suggesting that participants performed the tasks as instructed. Overall, performance was better on attend-speech trials than on attend-distractor trials (RM-ANOVA main effect of attention: $F(1, 19) = 38.91$, $p < .001$). That is, participants were more accurate at distinguishing between true keywords and rhyming foil words than they were at detecting the presence of a target in the distractor.

Performance was better overall on clear-speech trials than on degraded-speech trials (RM-ANOVA main effect of speech clarity: $F(1, 19) = 18.28$, $p < .001$). Interestingly, the deleterious effect of speech degradation did not depend on where attention was directed (RM-ANOVA speech clarity x attention interaction: *F*(1, 19) = 0.03, *p* = .86, n.s.): participants were not only worse at the speech-based task when the speech was degraded, but also at detecting the presence of a target in the accompanying distractor. We propose that the distractor may have been more difficult to perceptually segregate from the degraded (noise-vocoded) speech because of the noise-like quality of both signals. Alternatively, it is possible that participants may have found the degraded speech difficult to ignore, and thus may have partly attended to both the degraded speech and the distractor, rather than solely to the latter as instructed.

< Please insert Figure 3 here>

3.1.3 Sentence recognition post-test

Sensitivity for recognizing sentences that had been attended during fNIRS imaging was significantly better than for those that had been ignored (RM-ANOVA main effect of attention: *F*(1,19) = 62.21, *p* < .001). Indeed, recognition of ignored sentences was close to chance level (*d'* close to zero; Fig. 3). The benefit for recognition performance of having attended to a sentence was similar for clear and degraded speech (RM-ANOVA speech clarity x attention interaction: *F*(1, 19) = 1.39, *p* = .25, n.s.).

Unexpectedly, participants' memory was significantly better for sentences that had originally been presented as degraded speech than for those presented as clear speech (RM-ANOVA main effect of speech clarity: *F*(1,19) = 12.79, *p* = .002). This runs contrary to the expectation that memory for degraded speech would be worse than for clear speech on the basis that cognitive resources that might otherwise have been available for memory encoding would instead be needed to help understand the degraded speech (McCoy et al., 2005; Rabbitt, 1990; Ward et al., 2016). However, we note that our sentence familiarity post-test may not be representative of memory for speech in general, especially as participants were unaware that they would later be asked about the sentences they had heard. It may also be pertinent that the BKB sentences used in the present study are relatively simplistic and repetitive, with the same semantic themes recurring regularly across lists. Under these specific circumstances, it is possible that the extra cognitive effort needed to understand degraded compared with clear sentences actually made them more, rather than less, memorable. Indeed, variation in the amount of attention that is paid while encoding items into memory has been shown to affect later recognition performance, and may in fact be a dominant confounding factor underlying previously reported list-length effects in recognition memory (Kinnell and Dennis, 2011).

3.2 fNIRS results

3.2.1 Cortical activation maps

Group-level activation maps for each experimental condition contrasted against the silent baseline are shown in Fig. 4. Note that in this initial analysis testing was performed on all individual fNIRS measurement channels. Statistically significant activation ($p < .05$, FDR corrected) was observed in all conditions in channels overlying the left (Ch#42) and right (Ch#32) superior temporal gyri (i.e. in the vicinity of left and right auditory cortex). Significant activation beyond the superior temporal cortex was observed only when listeners actively attended to degraded speech: the additionally activated channels were located over left inferior frontal cortex (Ch#30, 40) extending towards (bilateral) pre-motor cortex (Ch#2, 9). Note that two of the additionally activated channels (Ch#30, 40) lie within the pre-defined LIFG ROI.

< Please insert Figure 4 here>

3.2.2 Response amplitude

While the activation maps shown in Fig. 4 are useful for visualizing the extent of cortical activation in each condition, to address the experimental hypothesis requires a direct comparison of fNIRS response amplitude between conditions. Fig. 5a plots mean estimated response amplitude relative to silence for each experimental condition in the LIFG ROI. In support of the experimental hypothesis, response amplitude in the LIFG depended critically on the interaction between speech clarity and attention (RM-ANOVA speech clarity x attention interaction: $F(1,19) = 5.43$, $p < .05$). Compared to the silent baseline, the LIFG was activated only during attentive listening to degraded speech; the LIFG was not activated in response to clear speech (whether the participant attended to it or not), nor when participants attended to the auditory distractor in the presence of degraded speech.

To explore whether this response pattern was specific to the LIFG, Fig. 5b–d shows estimated response amplitude in several secondary (post-hoc) ROIs that were significantly activated in one or more of the experimental conditions, namely left and right STG and bilateral pre-motor cortex. None of the secondary ROIs showed a significant interaction between speech clarity and attention. Bilateral pre-motor cortex was preferentially activated on attend-speech trials compared to attend-distractor trials (RM-ANOVA main effect of attention: $F(1,18) = 9.77$, $p < .01$). However, there was little evidence that the strength of this activation depended on whether the speech was clear or degraded (RM ANOVA speech clarity x attention interaction: $F(1,18) = 0.25$, $p = .626$, n.s.). Compared to the silent baseline, both left and right STG were robustly activated to a similar degree in all experimental conditions. In the left STG, there was a trend towards stronger activation when listeners attended to speech compared to when they attended to the distractor, although this effect did not reach statistical significance (RM-ANOVA main effect of attention: $F(1,18) = 2.36$, $p = .142$).

These results suggest that, among the cortical regions covered by our array, the interaction between speech clarity and attention was specific to the LIFG. However, we note that a definitive demonstration of regional specificity requires a direct statistical comparison across regions (Nieuwenhuis et al., 2011), i.e. in the present case, a region x speech clarity x attention interaction.

The significance of this higher-order interaction could not be confirmed in our data (RM-ANOVA ROI x speech clarity x attention interaction: *F*(3,51) = 1.30, *p* = .286, n.s.). We attribute this to a lack of power to detect this higher-order interaction, given that the study was powered (with 0.8 power and a Type I error rate of 0.05) only to detect the two-way speech clarity x attention interaction in the primary LIFG ROI. It should also be noted that the differing number of channels included in the various ROIs (3 in the pre-defined primary LIFG ROI; 1–2 in the data-driven, post-hoc ROIs) may have influenced the power of the statistical tests, since fNIRS response amplitude is generally more reliable when averaged across a small number of channels overlying a cortical ROI than at single-channel level (Plichta et al., 2006; Schecklmann et al., 2008; Wiggins et al., 2016).

< Please insert Figure 5 here>

3.2.3 Relative timing of the haemodynamic response in temporal versus frontal cortex

In a related fMRI study of sentence comprehension under varying levels of masking noise, Zekveld et al. (2006) discussed the possibility that the latency of the haemodynamic response may vary between speech-sensitive regions in the temporal and frontal lobes. However, the use of a sparse-sampling paradigm (Hall et al., 1999) in that study precluded any direct test of this possibility. In a later study, Davis et al. (2011) showed using time-resolved sparse fMRI that such interregional timing differences during sentence comprehension do exist: the haemodynamic response peaked significantly earlier (by ~1 s on average) in superior temporal cortex than in left inferior frontal cortex. The fact that no equivalent difference was seen when assessing a low-level baseline contrast (signal-correlated noise versus rest) was taken to suggest that interregional variation in the timing of the haemodynamic response likely reflected changes in the timing of underlying neuronal activity, rather than being an artefact associated with, for example, interregional differences in vasculature (Handwerker et al., 2012).

Here, we aimed to take advantage of the higher temporal resolution and continuous silent imaging offered by fNIRS to confirm this finding of interregional haemodynamic timing differences between temporal and frontal areas. We focused on the condition in which listeners attended to degraded speech, since the LIFG was significantly activated (compared to the silent baseline) only in this condition. The group-mean reconstructed haemodynamic responses for the LIFG compared to bilateral superior temporal cortex (i.e. the average of left and right STG responses) are plotted in Fig. 6. Only participants for whom the estimated haemodynamic response for both regions showed a positive peak were included in the analysis ($N = 18$ out of 20). Responses peaked later in the LIFG (mean time-to-peak = 6.7 s, SD = 1.6 s) than in bilateral STG (mean time-to-peak = 5.2 s, SD = 1.0 s), as confirmed by a paired-samples *t*-test ($t(17) = 3.56$, $p < .01$)[1]. This result is in agreement with the findings of Davis et al. (2011), both in terms of the direction of the interregional timing difference and its order of magnitude.

Note that while the mean response for the LIFG in Fig. 6 appears to show signs of an early minor secondary peak around 2–3 s, this may be an artefact associated with the failure of the informed basis set to accurately capture the full shape of individual responses that deviate too far from the canonical HRF (Calhoun et al., 2004), and so the reader is cautioned against assigning physiological meaning to this component.

< Please insert Figure 6 here>

3.2.4 Handedness and the lateralization of cortical activation

All analyses presented thus far were conducted on the full sample of twenty participants, which included three left-handed individuals, all with a laterality index ≤-70 on the Edinburgh Handedness Inventory (Oldfield, 1971). Thus, the sample included left-handed participants approximately in

---

[1] The haemodynamic response similarly peaked significantly later in the LIFG than in the left STG (mean time-to-peak 6.9 vs. 5.4 s, $t(16) = 3.08$, $p < .01$) and right STG (mean time-to-peak 6.7 vs. 5.0 s, $t(16) = 3.50$, $p < .01$) when comparing against each hemisphere separately.

proportion to their frequency (roughly 10%) in the population as a whole (Willems et al., 2014). However, it is well known that the lateralization of cortical function can differ between left- and right-handed individuals. While roughly 95% of right-handers show left-hemispheric dominance for language, the proportion of left-handers who show the same pattern is lower (~75%), with the remaining 25% displaying either bilateral activation or right-hemisphere dominance (Pujol et al., 1999; Szaflarski et al., 2002). We therefore wished to explore whether there was any evidence of atypical lateralization of cortical activation in the subset of our sample who were left-handed.

We again focused on the contrast between attentive listening to degraded speech and silence, since the most extensive cortical activation was observed for this contrast (cf. Fig. 4). Fig. 7 plots mean contrast values for homologous regions in the left and right hemispheres, broken down by handedness. Data are shown for superior temporal and inferior frontal regions. In the superior temporal cortices, bilateral activation was observed, with no strong evidence that responses were lateralized to one hemisphere or the other in either left- or right-handers. Accordingly, a mixed-design ANOVA showed no significant main effect of hemisphere (*F*(1,18) = 0.002, *p* = .966), no significant main effect of handedness (*F*(1,18) = 0.52, *p* = .479), and no significant interaction between the two (*F*(1,18) = 0.15, *p* = .703). In contrast, inferior frontal activation appears to have been left-lateralized in right-handers, but right-lateralized in left-handers. This differential lateralization between left- and right-handers was confirmed by a significant handedness x hemisphere interaction in the mixed-design ANOVA (*F*(1,18) = 8.86, *p* = .008). Thus, it is possible that the role played by the LIFG in processing degraded speech is shifted to the homologous region in the right hemisphere in at least some left-handed individuals.

The strength of right-hemispheric lateralization of inferior frontal activation in our subset of left-handed participants is perhaps surprising, given that only around 25% of left-handed individuals are thought to display atypical language lateralization. However, given the small number of left-handed participants in the present study (*N* = 3), it would be inappropriate to generalize these findings to the wider population of left-handed individuals.

< Please insert Figure 7 here>

3.3 General discussion

Using fNIRS, we confirmed an important finding from the fMRI literature (Wild et al., 2012): the processing of degraded speech depends critically on attention. Specifically, the LIFG exhibited an elevated response to degraded versus clear speech, but only when attention was directed towards the speech. This attention-dependent increase in frontal brain activation may be a cortical marker for effortful listening. Our results additionally indicate that the homologous region in right inferior frontal cortex may play an equivalent role in at least some left-handed individuals.

The precise role that the LIFG plays in supporting the recovery of meaning from degraded speech is unclear. The region of elevated activation lies in the vicinity of Broca's area (Brodmann areas 44 and 45), classically considered a motor speech-production area (Nishitani et al., 2005). One possibility, therefore, is that elevated activation during listening to degraded speech corresponds to a process in which the degraded input is matched to internal articulatory templates (Hervais-Adelman et al., 2012; Watkins and Paus, 2004), reminiscent of the motor theory of speech perception (Liberman and Mattingly, 1985). However, beyond its role as a motor speech area, the LIFG is also known to be extensively involved in higher-order language processing at the phonological (Poldrack et al., 1999), syntactic (Kaan and Swaab, 2002) and semantic levels (Rodd et al., 2005; Wagner et al., 2001), as well as in verbal working memory (Badre and Wagner, 2007; Nixon et al., 2004; Rogalsky and Hickok, 2010). Thus, an alternative explanation for elevated LIFG activation is that it reflects the unification of linguistic and contextual knowledge to infer the meaning of a degraded speech signal, potentially drawing on working memory resources (Hagoort, 2005). Interestingly, recent studies exploiting the greater temporal resolution of EEG/magnetoencephalography have shown that contextual information being processed in the LIFG is able to modulate activity in lower-level sensory regions of the STG in a top-down manner (Gow and Olson, 2016; Sohoglu et al., 2012).

Our fNIRS data confirmed that, during attentive listening to noise-vocoded sentences, the haemodynamic response peaks significantly later in the LIFG than in superior temporal cortex. This corroborates a similar finding from an fMRI study of sentence comprehension in noise by Davis et al. (2011). While caution is needed in interpreting interregional haemodynamic timing differences (Handwerker et al., 2012; Lindquist et al., 2009), the combined results of both studies suggest that neuronal activation may peak later in the LIFG than in superior temporal cortex. It is noteworthy that in the present study the mean interregional latency difference of 1.5 s was comparable to the mean sentence duration (1.64 s). To the extent that the haemodynamic delay is a true reflection of differences in the timing of underlying neuronal activity, this suggests that neurocognitive computations being performed in the LIFG may persist beyond the period of auditory stimulation and into the response period. This hypothesis, if true, would seem compatible with the engagement of verbal working memory in the LIFG, and also with the subjective experience that the meaning of a degraded sentence often only becomes clear after a short period of active reflection.

An alternative possibility is that inferior frontal activation reflects not so much the process of recovering meaning from degraded speech, but rather a general decision-making process associated with response selection. Binder et al. (2004) conducted an fMRI study, in which listeners were tasked with discriminating between two synthesized speech syllables at varying signal-to-noise ratios, and found evidence for a role of the anterior insular and adjacent frontal operculum in decision making. The level of brain activation in these regions was positively correlated with response time. However, there are several reasons to suspect that the inferior frontal activation that we measured in the present study may not be directly comparable. Firstly, the areas highlighted by Binder et al. (2004) lie somewhat medially within the brain, and given the limited depth penetration of fNIRS (Strangman et al., 2013) it is unlikely that our measurements had much sensitivity to these regions. Secondly, Binder et al. (2004) observed anterior insular-opercular activation bilaterally, whereas inferior frontal activation in the present study was largely lateralized to the dominant hemisphere (cf. Fig. 7). Thirdly, the elevated activation that we observed in the LIFG seems to have been speech-specific. The LIFG was not significantly activated (compared to the silent baseline condition) during attentive listening to the non-speech distractor, despite the fact that this too presented a challenging listening task. Thus, it seems likely that our measurements were primarily sensitive to language-related processing taking place more laterally in inferior frontal cortex. Binder et al. (2004) also observed activation in these more lateral areas, in addition to the insular-opercular activation that was modulated by decision time.

The fact that the elevated LIFG activation that we observed in the present study was speech-specific suggests that it may be a neural marker for effortful listening to degraded speech, but not for effortful listening in general. Conceivably, this could reflect a distinction between situations in which listening is effortful because of the need to rely heavily on knowledge of language (presumably as in our speech-based task) and other situations in which listening is effortful because of high attentional demands (distractor-based task), with the two cases drawing on separate neural mechanisms (Peelle and Wingfield, 2016).

Aside from inferior frontal cortex, we observed significant bilateral activation of the superior temporal cortices in all conditions, and of pre-motor cortex more selectively. In left STG, there was a trend towards stronger activation when listeners attended to speech compared to when they attended to the distractor, which might suggest that our fNIRS measurements in this area were preferentially sensitive to cortical sub-regions involved in speech processing. However, this effect of attention did not reach statistical significance, and so this should be considered a speculative proposition requiring further investigation. We did not find any evidence that STG activation in either hemisphere was stronger for more intelligible speech, as might have been expected based on prior fMRI research (Binder et al., 2004; Davis and Johnsrude, 2003; Wild et al., 2012). However, it is possible that our inclusion of only two levels of speech clarity, differing relatively modestly in intelligibility (85% versus ~100% correct), was insufficient to reveal such an effect. As regards pre-motor cortex, we observed statistically significant activation of this region (compared to silence) only in the case that listeners attended to degraded speech. This would seem compatible with the suggestion that pre-motor cortex is part of the network selectively recruited to support effortful comprehension of degraded speech (Hervais-Adelman et al., 2012). However, a post-hoc analysis of response amplitude suggested that pre-motor cortex was in fact fairly similarly activated when listeners attended to either clear or degraded speech. As such, further work is needed to clarify the role of pre-motor cortex in the present task. We note that since pre-motor cortex was not an a priori ROI in the present study, the results in this region should be interpreted with caution and cannot be considered definitive.

The optode array used in the present study did not provide coverage of all cortical regions thought to be implicated in effortful listening. Like the LIFG, some other relevant regions are located superficially (i.e. close to the scalp) and are likely to be imageable using fNIRS. An example of such a region would be the angular gyrus of the left inferior parietal cortex, which has been shown to play a role in facilitating comprehension through the use of linguistic/semantic context (Golestani et al., 2013; Hartwigsen et al., 2015; Obleser and Kotz, 2010; Obleser et al., 2007) and which forms part of a functional fronto–temporal–parietal network supporting speech comprehension (Abrams et al., 2013). Other relevant regions, however, are probably located too deeply within the brain to be successfully imaged using fNIRS given the limited penetration depth of infrared light in biological tissue (Strangman et al., 2013). An example would be the previously mentioned medial aspects of the non-sensory-specific cingulo-opercular system, elevated activation in which is thought to reflect the need for and decision to allocate effort during a challenging task (Eckert et al., 2016).

A limitation of the paradigm used in the present study is that the measured haemodynamic response may have contained a contribution from neuronal activity associated with the planning or execution of the button press. We aimed to control for this by: (i) requiring a button press on every trial (including silent trials); and (ii) using an informed basis set in the analysis to allow us to assess response amplitude independently of variations in response latency/dispersion (as might have resulted from variations in the timing of the button press). However, we cannot say with certainty that there was no residual influence of activity related to the button press. Unfortunately, since we did not record response time, we are unable to assess whether there was any systematic difference in response time between conditions. The haemodynamic response estimates plotted in Fig. 6 may have been particularly susceptible to a possible influence of neuronal activity related to the button press, since those estimates do not reflect any explicit subtraction of the response on silent trials. Reassuringly the interregional timing difference that we observed between temporal and frontal cortex is similar to that previously observed by Davis et al. (2011): Davis et al. designed their fMRI paradigm such that the measured haemodynamic response would be minimally sensitive to their behavioural task, indicating that a later response in the LIFG than in superior temporal cortex during sentence recognition is unrelated to button-press behaviour.

Following Wild et al. (2012), we conducted a surprise post-test of participants' ability to correctly recognise sentences they had heard during the imaging session. Our results differed from those obtained by Wild et al., who reported that recognition of sentences originally presented as clear speech was good regardless of whether they were attended or not, whereas recognition of sentences originally presented as degraded speech depended on them having been attended. In contrast, here we found that attention to speech was critical for accurate recognition of familiar sentences, regardless of whether they were originally presented as clear or degraded speech. Thus, while participants in Wild et al.'s study appear to have processed clear speech even when it was not attended, our participants seem not to have done so. The reason for this discrepancy between studies is unclear. It may be related to the more simplistic sentence materials used here (as discussed in Sec. 3.1.3), although other procedural differences between studies may also have played a role (e.g. the present study included only a subset of the experimental conditions tested by Wild et al. and did not include a simultaneous visual distractor stimulus). Performance on the auditory distractor task in the present study (mean d' = 2.74, SD = 1.50) was comparable to the auditory (mean d' = 2.15, SD = 1.30) and visual (mean d' = 3.17, SD = 0.55) distractor tasks in Wild et al.'s study, suggesting that distractor task difficulty is unlikely to have been an influential factor. Regardless of the reason(s) for the discrepancy in sentence recognition performance, it is noteworthy that both studies found a highly concordant effect of attention on the neural processing of degraded speech in inferior frontal cortex.

To conclude, the present study (i) confirms that the processing of degraded speech in the LIFG depends critically on attention and (ii) demonstrates that fNIRS is a capable neuroimaging modality for investigating the amplitude and timing of the haemodynamic response to neuronal activation in this important region. In light of recent developments in fNIRS technology that enable wireless brain imaging of freely moving participants in natural environments (Pinti et al., 2015; Piper et al., 2014; von Luhmann et al., 2015), we propose that fNIRS holds promise as a flexible tool to study the neural signature of effortful listening.

Acknowledgements

The authors thank Dr Toru Yamada and Dr Shinji Umeyama for kindly providing the code for the haemodynamic signal separation algorithm.

This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Figure captions

Fig. 1. Optode positioning and region-of-interest definition. (a) Photograph of the optode array holder placed on a volunteer's head before insertion of the optodes. Red/blue colour coding on the holder indicates the locations of emitters and detectors, respectively. (b) Variability in digitized optode positions across eight volunteers after registration to an atlas brain. Optode positions for each volunteer are represented by different coloured dots. Variability was similar across the forehead and in the right hemisphere (not shown). (c) Aggregate sensitivity profile for the left inferior frontal gyrus (LIFG) region of interest (fNIRS channels 30, 40 and 51). The colour scale depicts relative sensitivity logarithmically from 0.001 to 1.

Fig. 2. Speech intelligibility scores. Error bars show ±1 SEM corrected for repeated measures. Statistical results are shown for post-hoc pairwise comparisons between consecutive pairs of lists (*** $p < .001$ corrected; n.s. not significant). Note that degraded-speech lists 1 and 2 were tested sequentially, while degraded-speech list 3 and the clear-speech list were tested in interleaved fashion.

Fig. 3. Mean sensitivity scores (*d*') for the main task (left) and sentence recognition post-test (right). Error bars show ±1 SEM corrected for repeated measures. An asterisk above a data point indicates that mean sensitivity was significantly greater than chance level (one-tailed *t*-tests against zero, $p < .05$ after Bonferroni correction for $N = 4$ comparisons).

Fig. 4. Group-level activation maps for each experimental condition contrasted against the silent baseline. Significantly activated channels are highlighted ($q < .05$, FDR corrected). L and R denote left and right hemisphere, respectively. Note that the maps are interpolated from single-channel results and the overlay on the cortical surface is for illustrative purposes only.

Fig. 5. Mean contrast values (i.e. estimated response amplitude relative to silence; arbitrary units) in (a) the primary LIFG ROI and (b–d) secondary (post-hoc) ROIs. Inset figures illustrate the fNIRS sensitivity profile for each region. Error bars show ±1 SEM corrected for repeated measures. Vertical lines with asterisks indicate significant simple effects of attention ($p < .05$, Bonferroni-corrected for comparisons at $N = 2$ levels of speech clarity).

Fig. 6. Haemodynamic response estimates for the condition in which listeners attended to degraded speech. The mean response is compared between the LIFG (Ch#30,40,51; dashed line) and bilateral STG (Ch#32,42; solid line). Individual responses were normalized to a peak value of 1 before averaging. Shading indicates ±1 SEM across participants. The inset figure plots mean time-to-peak for each region, with error bars showing ±1 SEM corrected for repeated measures (* $p < .01$, paired-samples *t*-test).

Fig. 7. Handedness and lateralization of cortical activation. Mean contrast values (arbitrary units) for the contrast between attentive listening to degraded speech and silence, broken down by handedness and hemisphere. Data are shown for superior temporal (upper panel; left-hemisphere Ch#42; right-hemisphere Ch#32) and inferior frontal (lower panel; left-hemisphere Ch#30,40,51; right-hemisphere Ch#23,34,44) regions. Error bars show ±1 SEM.

References

Aasted, C.M., Yücel, M.A., Cooper, R.J., Dubb, J., Tsuzuki, D., Becerra, L., Petkov, M.P., Borsook, D., Dan, I., Boas, D.A. 2015. Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. NEUROW 2, 020801-020801.

Abrams, D.A., Ryali, S., Chen, T., Balaban, E., Levitin, D.J., Menon, V. 2013. Multivariate activation and connectivity patterns discriminate speech intelligibility in Wernicke's, Broca's, and Geschwind's areas. Cereb Cortex 23, 1703-14.

Adank, P. 2012. The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. Brain and Language 122, 42-54.

Adank, P., Davis, M.H., Hagoort, P. 2012. Neural dissociation in processing noise and accent in spoken language comprehension. Neuropsychologia 50, 77-84.

Alhanbali, S., Dawes, P., Lloyd, S., Munro, K.J. 2016. Self-Reported Listening-Related Effort and Fatigue in Hearing-Impaired Adults. Ear Hear.

Badre, D., Wagner, A.D. 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia 45, 2883-2901.

Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1-48.

Bench, J., Kowal, A., Bamford, J. 1979. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. British journal of audiology 13, 108-12.

Benjamini, Y., Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289-300.

Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A., Ward, B.D. 2004. Neural correlates of sensory and decision processes in auditory object identification. Nature neuroscience 7, 295-301.

Bisconti, S., Shulkin, M., Hu, X., Basura, G.J., Kileny, P.R., Kovelman, I. 2015. fNIRS brain imaging investigation of phonological awareness and passage comprehension abilities in adult recipients of Cochlear Implants. Journal of Speech, Language, and Hearing Research.

Boas, D., Culver, J., Stott, J., Dunn, A. 2002. Three dimensional Monte Carlo code for photon migration through complex heterogeneous media including the adult human head. Optics express 10, 159-70.

Boas, D.A., Elwell, C.E., Ferrari, M., Taga, G. 2014. Twenty years of functional near-infrared spectroscopy: introduction for the special issue. NeuroImage 85, 1-5.

Boas, D.A., Gaudette, T., Strangman, G., Cheng, X., Marota, J.J., Mandeville, J.B. 2001. The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. NeuroImage 13, 76-90.

Brainard, D.H. 1997. The Psychophysics Toolbox. Spatial vision 10, 433-6.

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., Gagnon, L., Boas, D.A., Cooper, R.J. 2014. Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. NeuroImage 85, 181-191.

Calhoun, V.D., Stevens, M.C., Pearlson, G.D., Kiehl, K.A. 2004. fMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. NeuroImage 22, 252-257.

Chen, L.-C., Sandmann, P., Thorne, J., Herrmann, C., Debener, S. 2015. Association of Concurrent fNIRS and EEG Signatures in Response to Auditory and Visual Stimuli. Brain topography, 1-16.

Chen, L.-C., Sandmann, P., Thorne, J.D., Bleichner, M.G., Debener, S. 2016. Cross-Modal Functional Reorganization of Visual and Auditory Cortex in Adult Cochlear Implant Users Identified with fNIRS. Neural Plasticity 2016, 13.

Cignetti, F., Salvia, E., Anton, J.-L., Grosbras, M.-H., Assaiante, C. 2016. Pros and cons of using the informed basis set to account for hemodynamic response variability with developmental data. Frontiers in neuroscience 10.

Cochrane, D., Orcutt, G.H. 1949. Application of Least Squares Regression to Relationships Containing Auto- Correlated Error Terms. Journal of the American Statistical Association 44, 32-61.

Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C. 1998. Design and construction of a realistic digital brain phantom. IEEE transactions on medical imaging 17, 463-8.

Cooper, R.J., Selb, J., Gagnon, L., Phillip, D., Schytz, H.W., Iversen, H.K., Ashina, M., Boas, D.A. 2012. A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. Frontiers in neuroscience 6, 147.

Dale, A.M. 1999. Optimal experimental design for event-related fMRI. Human brain mapping 8, 109-14.

Davis, M.H., Johnsrude, I.S. 2003. Hierarchical processing in spoken language comprehension. The Journal of neuroscience : the official journal of the Society for Neuroscience 23, 3423-31.

Davis, M.H., Johnsrude, I.S. 2007. Hearing speech sounds: top-down influences on the interface between audition and speech perception. Hearing research 229, 132-47.

Davis, M.H., Ford, M.A., Kherif, F., Johnsrude, I.S. 2011. Does Semantic Context Benefit Speech Understanding through “Top–Down” Processes? Evidence from Time-resolved Sparse fMRI. Journal of cognitive neuroscience 23, 3914-3932.

Dixon, P. 2008. Models of accuracy in repeated-measures designs. Journal of Memory and Language 59, 447-456.

Eckert, M.A., Teubner-Rhodes, S., Vaden, K.I.J. 2016. Is Listening in Noise Worth It? The Neurobiology of Speech Recognition in Challenging Listening Conditions. Ear and Hearing 37, 101S-110S.

Erb, J., Henry, M.J., Eisner, F., Obleser, J. 2013. The Brain Dynamics of Rapid Perceptual Adaptation to Adverse Listening Conditions. The Journal of Neuroscience 33, 10688-10697.

Evans, S., McGettigan, C., Agnew, Z.K., Rosen, S., Scott, S.K. 2016. Getting the cocktail party started: masking effects in speech perception. Journal of cognitive neuroscience 28, 483-500.

Field, A. 2009. Discovering Statistics Using SPSS. 3rd ed. SAGE Publications, London.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R. 1998. Event-related fMRI: characterizing differential responses. NeuroImage 7, 30-40.

Golestani, N., Hervais-Adelman, A., Obleser, J., Scott, S.K. 2013. Semantic versus perceptual interactions in neural processing of speech-in-noise. NeuroImage 79, 52-61.

Gow, D.W., Olson, B.B. 2016. Sentential influences on acoustic-phonetic processing: a Granger causality analysis of multimodal imaging data. Language, Cognition and Neuroscience 31, 841-855.

Greenwood, D.D. 1990. A cochlear frequency-position function for several species--29 years later. The Journal of the Acoustical Society of America 87, 2592-605.

Hagoort, P. 2005. On Broca, brain, and binding: a new framework. Trends in cognitive sciences 9, 416-423.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W. 1999. "Sparse" temporal sampling in auditory fMRI. Human brain mapping 7, 213-23.

Handwerker, D.A., Gonzalez-Castillo, J., D'Esposito, M., Bandettini, P.A. 2012. The continuing challenge of understanding and modeling hemodynamic variation in fMRI. NeuroImage 62, 1017-1023.

Hartwigsen, G., Golombek, T., Obleser, J. 2015. Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension. Cortex; a journal devoted to the study of the nervous system and behavior 68, 100-110.

Hassanpour, M.S., Eggebrecht, A.T., Culver, J.P., Peelle, J.E. 2015. Mapping cortical responses to speech using high-density diffuse optical tomography. NeuroImage 117, 319-326.

Hervais-Adelman, A.G., Carlyon, R.P., Johnsrude, I.S., Davis, M.H. 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. Language and Cognitive Processes 27, 1145-1166.

Hicks, C.B., Tharpe, A.M. 2002. Listening effort and fatigue in school-age children with and without hearing loss. Journal of speech, language, and hearing research : JSLHR 45, 573-84.

Hong, K.S., Santosa, H. 2016. Decoding four different sound-categories in the auditory cortex using functional near-infrared spectroscopy. Hearing research 333, 157-66.

Huppert, T.J., Diamond, S.G., Franceschini, M.A., Boas, D.A. 2009. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. Applied optics 48, D280-98.

Iadecola, C. 2004. Neurovascular regulation in the normal brain and in Alzheimer's disease. Nature reviews. Neuroscience 5, 347-60.

Jasper, H.H. 1958. The ten twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology 10, 371-375.

Kaan, E., Swaab, T.Y. 2002. The brain circuitry of syntactic comprehension. Trends in cognitive sciences 6, 350-356.

Kinnell, A., Dennis, S. 2011. The list length effect in recognition memory: an analysis of potential confounds. Memory & Cognition 39, 348-363.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C. 2007. What’s new in Psychtoolbox-3. Perception 36, 1.

Kramer, S.E., Kapteyn, T.S., Houtgast, T. 2006. Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. Int J Audiol 45, 503-512.

Li, X., Liang, Z., Kleiner, M., Lu, Z.L. 2010. RTbox: a device for highly accurate response time measurements. Behavior research methods 42, 212-25.

Liberman, A.M., Mattingly, I.G. 1985. The motor theory of speech perception revised. Cognition 21, 1-36.

Lin, F.R., Yaffe, K., Xia, J., Xue, Q.L., Harris, T.B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H.N., Ferrucci, L., Simonsick, E.M. 2013. Hearing loss and cognitive decline in older adults. JAMA internal medicine 173, 293-9.

Lindquist, M.A., Wager, T.D. 2007. Validity and power in hemodynamic response modeling: A comparison study and a new approach. Human brain mapping 28, 764-784.

Lindquist, M.A., Meng Loh, J., Atlas, L.Y., Wager, T.D. 2009. Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. NeuroImage 45, S187-S198.

McCoy, S.L., Tun, P.A., Cox, L.C., Colangelo, M., Stewart, R.A., Wingfield, A. 2005. Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. The Quarterly journal of experimental psychology. A, Human experimental psychology 58, 22-33.

McGarrigle, R., Munro, K.J., Dawes, P., Stewart, A.J., Moore, D.R., Barry, J.G., Amitay, S. 2014. Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group ‘white paper’. Int J Audiol 53, 433-445.

McKay, C.M., Shah, A., Seghouane, A.K., Zhou, X., Cross, W., Litovsky, R. 2016. Connectivity in Language Areas of the Brain in Cochlear Implant Users as Revealed by fNIRS. Advances in experimental medicine and biology 894, 327-35.

Molavi, B., Dumont, G.A. 2012. Wavelet-based motion artifact removal for functional near-infrared spectroscopy. Physiological measurement 33, 259-70.

Nachtegaal, J., Kuik, D.J., Anema, J.R., Goverts, S.T., Festen, J.M., Kramer, S.E. 2009. Hearing status, need for recovery after work, and psychosocial work characteristics: results from an internet-based national survey on hearing. Int J Audiol 48, 684-91.

Nieuwenhuis, S., Forstmann, B.U., Wagenmakers, E.-J. 2011. Erroneous analyses of interactions in neuroscience: a problem of significance. Nature neuroscience 14, 1105-1107.

Nishitani, N., Schürmann, M., Amunts, K., Hari, R. 2005. Broca’s Region: From Action to Language. Physiology 20, 60-69.

Nixon, P., Lazarova, J., Hodinott-Hill, I., Gough, P., Passingham, R. 2004. The Inferior Frontal Gyrus and Phonological Processing: An Investigation using rTMS. Journal of cognitive neuroscience 16, 289-300.

Obleser, J., Kotz, S.A. 2010. Expectancy constraints in degraded speech modulate the language comprehension network. Cereb Cortex 20, 633-40.

Obleser, J., Wise, R.J., Dresner, M.A., Scott, S.K. 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. The Journal of neuroscience : the official journal of the Society for Neuroscience 27, 2283-9.

Oldfield, R.C. 1971. The Assessment and Analysis of Handedness: The Edinburgh Inventory. Neuropsychologia 9, 97-113.

Olds, C., Pollonini, L., Abaya, H., Larky, J., Loy, M., Bortfeld, H., Beauchamp, M.S., Oghalai, J.S. 2015. Cortical Activation Patterns Correlate With Speech Understanding After Cochlear Implantation. Ear Hear.

Pals, C., Sarampalis, A., Başkent, D. 2013. Listening Effort With Cochlear Implant Simulations. Journal of Speech, Language, and Hearing Research 56, 1075-1084.

Peelle, J.E. 2014. Methodological challenges and solutions in auditory functional magnetic resonance imaging. Frontiers in neuroscience 8, 253.

Peelle, J.E., Wingfield, A. 2016. Listening Effort in Age-Related Hearing Loss. Hearing Journal 69, 10-12.

Peelle, J.E., Johnsrude, I.S., Davis, M.H. 2010. Hierarchical processing for speech in human auditory cortex and beyond. Front Hum Neurosci 4.

Pelli, D.G. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial vision 10, 437-42.

Pichora-Fuller, M.K. 2016. How Social Psychological Factors May Modulate Auditory and Cognitive Functioning During Listening. Ear and Hearing 37, 92S-100S.

Picou, E.M., Ricketts, T.A., Hornsby, B.W. 2013. How hearing aids, background noise, and visual cues influence objective listening effort. Ear Hear 34, e52-64.

Pinti, P., Aichelburg, C., Lind, F., Power, S., Swingler, E., Merla, A., Hamilton, A., Gilbert, S., Burgess, P., Tachtsidis, I. 2015. Using Fiberless, Wearable fNIRS to Monitor Brain Activity in Real-world Cognitive Tasks. Journal of visualized experiments : JoVE.

Piper, S.K., Krueger, A., Koch, S.P., Mehnert, J., Habermehl, C., Steinbrink, J., Obrig, H., Schmitz, C.H. 2014. A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. NeuroImage 85, 64-71.

Plichta, M.M., Heinzel, S., Ehlis, A.C., Pauli, P., Fallgatter, A.J. 2007. Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: a parametric validation study. NeuroImage 35, 625-34.

Plichta, M.M., Herrmann, M.J., Baehne, C.G., Ehlis, A.C., Richter, M.M., Pauli, P., Fallgatter, A.J. 2006. Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable? NeuroImage 31, 116-24.

Plichta, M.M., Gerdes, A.B., Alpers, G.W., Harnisch, W., Brill, S., Wieser, M.J., Fallgatter, A.J. 2011. Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (fNIRS) study. NeuroImage 55, 1200-7.

Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E. 1999. Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex. NeuroImage 10, 15-35.

Pollonini, L., Olds, C., Abaya, H., Bortfeld, H., Beauchamp, M.S., Oghalai, J.S. 2014. Auditory cortex activation to natural speech and simulated cochlear implant speech measured with functional near-infrared spectroscopy. Hearing research 309, 84-93.

Pujol, J., Deus, J., Losilla, J.M., Capdevila, A. 1999. Cerebral lateralization of language in normal left-handed people studied by functional MRI. Neurology 52, 1038-43.

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabbitt, P. 1990. Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. Acta oto-laryngologica. Supplementum 476, 167-75; discussion 176.

Rodd, J.M., Davis, M.H., Johnsrude, I.S. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb Cortex 15, 1261-9.

Rogalsky, C., Hickok, G. 2010. The Role of Broca's Area in Sentence Comprehension. Journal of cognitive neuroscience 23, 1664-1680.

Ronnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., Dahlstrom, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M.K., Rudner, M. 2013. The Ease of Language

Understanding (ELU) model: theoretical, empirical, and clinical advances. Frontiers in systems neuroscience 7, 31.

Sarampalis, A., Kalluri, S., Edwards, B., Hafter, E. 2009. Objective measures of listening effort: effects of background noise and noise reduction. Journal of speech, language, and hearing research : JSLHR 52, 1230-40.

Schecklmann, M., Ehlis, A.-C., Plichta, M.M., Fallgatter, A.J. 2008. Functional near-infrared spectroscopy: A long-term reliable tool for measuring brain activity during verbal fluency. NeuroImage 43, 147-155.

Scholkmann, F., Kleiser, S., Metz, A.J., Zimmermann, R., Pavia, J.M., Wolf, U., Wolf, M. 2014. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. NeuroImage 85, 6-27.

Schroeter, M.L., Bucheler, M.M., Muller, K., Uludag, K., Obrig, H., Lohmann, G., Tittgemeyer, M., Villringer, A., von Cramon, D.Y. 2004. Towards a standard analysis for functional near-infrared imaging. NeuroImage 21, 283-90.

Scott, S.K., McGettigan, C. 2013. The neural processing of masked speech. Hearing research 303, 58-66.

Sevy, A.B., Bortfeld, H., Huppert, T.J., Beauchamp, M.S., Tonini, R.E., Oghalai, J.S. 2010. Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. Hearing research 270, 39-47.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M. 1995. Speech recognition with primarily temporal cues. Science 270, 303-4.

Singh, A.K., Dan, I. 2006. Exploring the false discovery rate in multichannel NIRS. NeuroImage 33, 542-549.

Sohoglu, E., Peelle, J.E., Carlyon, R.P., Davis, M.H. 2012. Predictive top-down integration of prior knowledge during speech perception. The Journal of neuroscience : the official journal of the Society for Neuroscience 32, 8443-53.

Steffener, J., Tabert, M., Reuben, A., Stern, Y. 2010. Investigating hemodynamic response variability at the group level using basis functions. NeuroImage 49, 2113-2122.

Strangman, G.E., Li, Z., Zhang, Q. 2013. Depth Sensitivity and Source-Detector Separations for Near Infrared Spectroscopy Based on the Colin27 Brain Template. PloS one 8.

Szaflarski, J.P., Binder, J.R., Possing, E.T., McKiernan, K.A., Ward, B.D., Hammeke, T.A. 2002. Language lateralization in left-handed and ambidextrous people: fMRI data. Neurology 59, 238-44.

van de Rijt, L.P.H., van Opstal, A.J., Mylanus, E.A.M., Straatman, L.V., Hu, H.Y., Snik, A.F.M., van Wanrooij, M.M. 2016. Temporal Cortex Activation to Audiovisual Speech in Normal-Hearing and Cochlear Implant Users Measured with Functional Near-Infrared Spectroscopy. Front Hum Neurosci 10, 48.

von Luhmann, A., Herff, C., Heger, D., Schultz, T. 2015. Toward a Wireless Open Source Instrument: Functional Near-infrared Spectroscopy in Mobile Neuroergonomics and BCI Applications. Front Hum Neurosci 9, 617.

Wagner, A.D., Pare-Blagoev, E.J., Clark, J., Poldrack, R.A. 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31, 329-38.

Ward, C.M., Rogers, C.S., Van Engen, K.J., Peelle, J.E. 2016. Effects of Age, Acoustic Challenge, and Verbal Working Memory on Recall of Narrative Speech. Experimental aging research 42, 97-111.

Watkins, K., Paus, T. 2004. Modulation of motor excitability during speech perception: the role of Broca's area. Journal of cognitive neuroscience 16, 978-87.

Wiggins, I.M., Anderson, C.A., Kitterick, P.T., Hartley, D.E. 2016. Speech-evoked activation in adult temporal cortex measured using functional near-infrared spectroscopy (fNIRS): Are the measurements reliable? Hearing research 339, 142-154.

Wild, C.J., Yusuf, A., Wilson, D.E., Peelle, J.E., Davis, M.H., Johnsrude, I.S. 2012. Effortful Listening: The Processing of Degraded Speech Depends Critically on Attention. Journal of Neuroscience 32, 14010-14021.

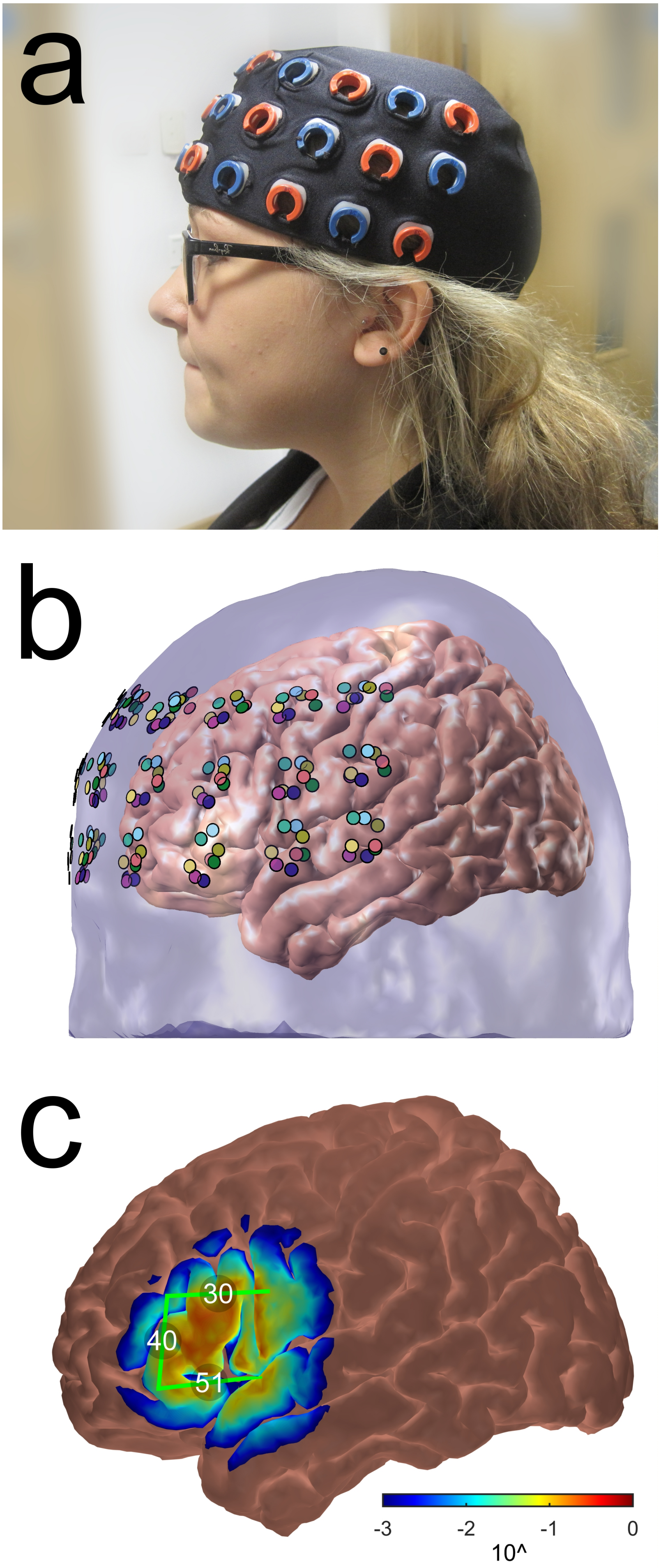
Willems, R.M., der Haegen, L.V., Fisher, S.E., Francks, C. 2014. On the other hand: including left-handers in cognitive neuroscience and neurogenetics. Nature reviews. Neuroscience 15, 193-201.

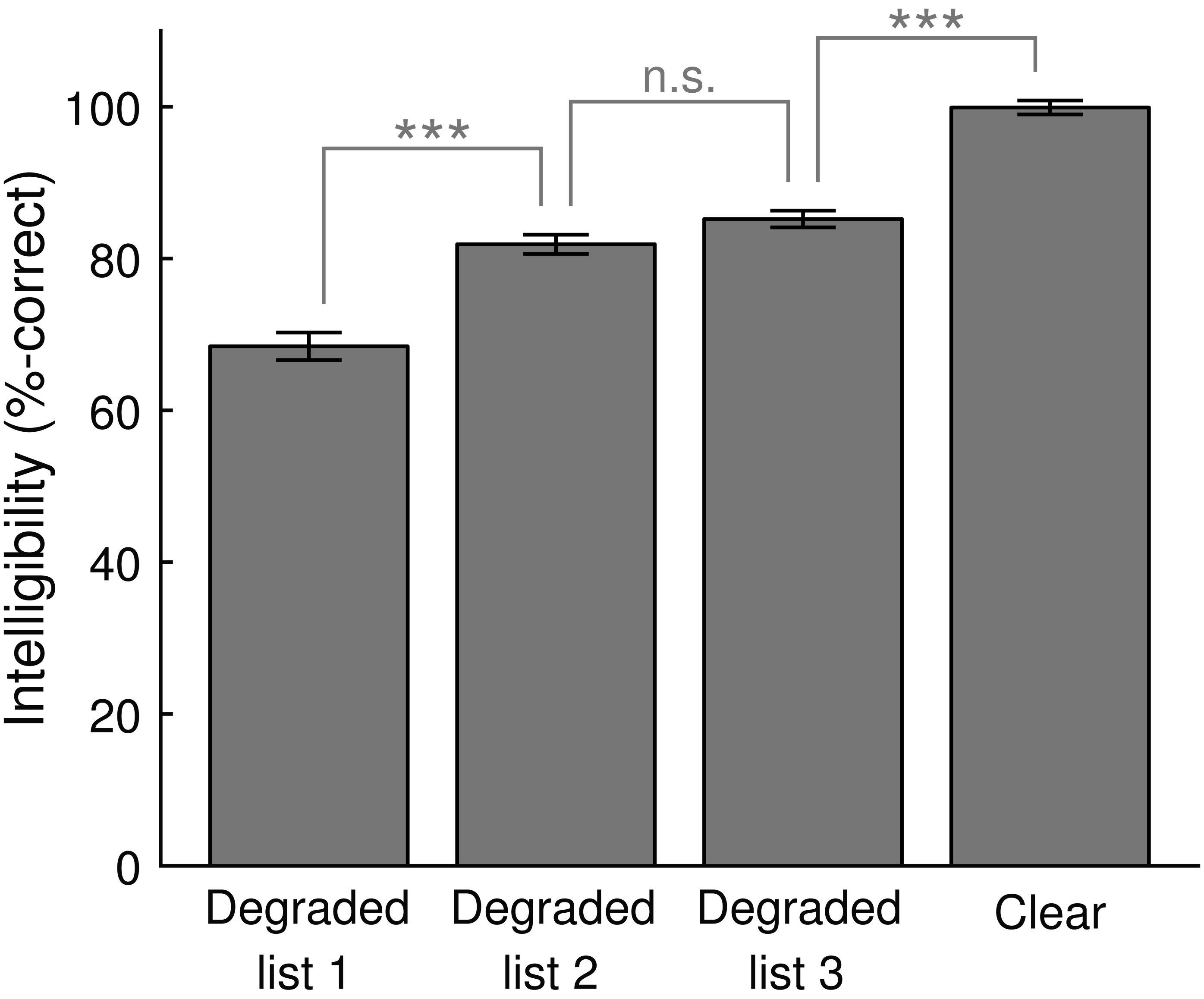
Winn, M.B., Edwards, J.R., Litovsky, R.Y. 2015. The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation. Ear Hear 36, e153-65.

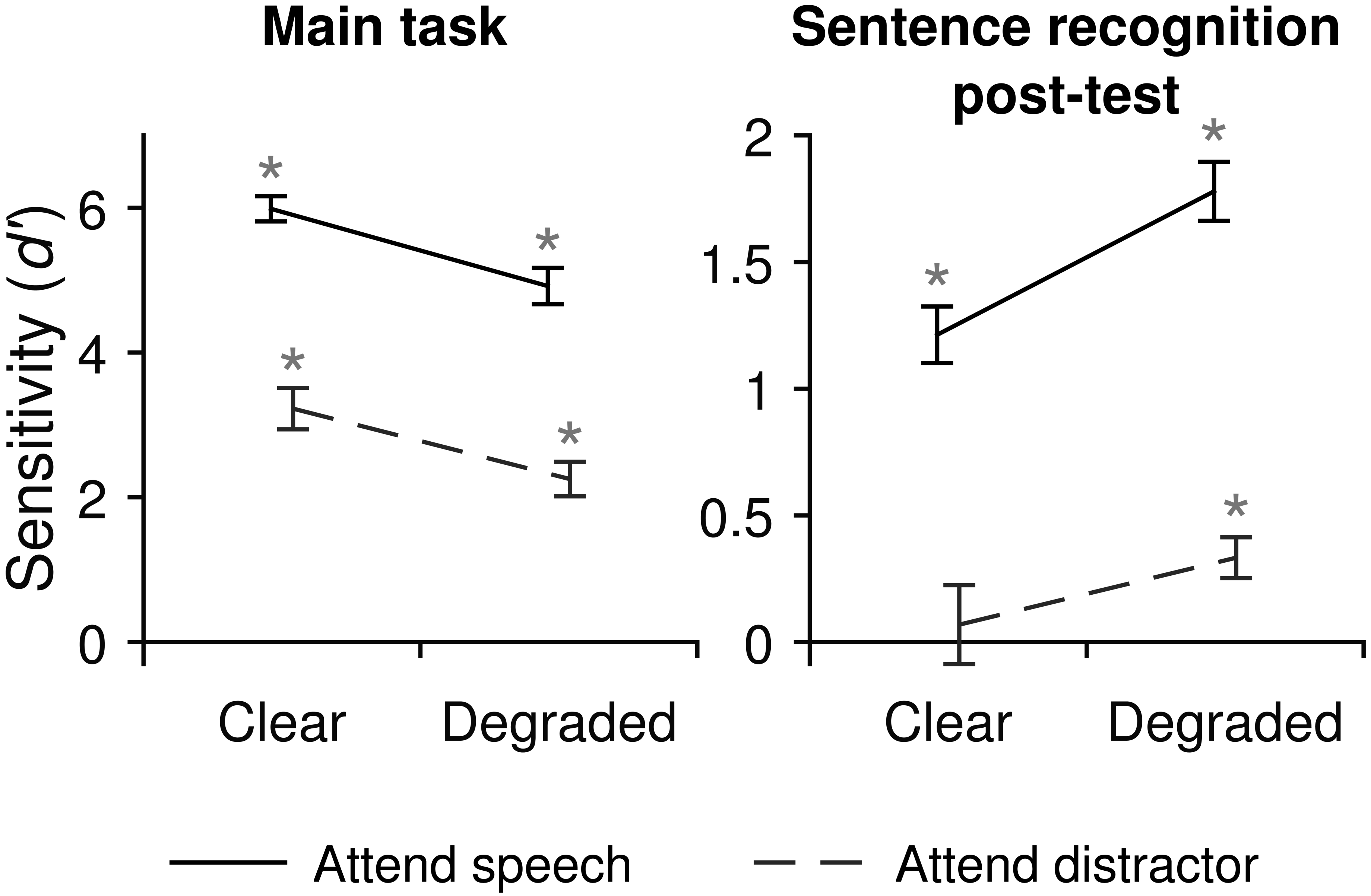
Yamada, T., Umeyama, S., Matsuda, K. 2012. Separation of fNIRS signals into functional and systemic components based on differences in hemodynamic modalities. PloS one 7, e50271.

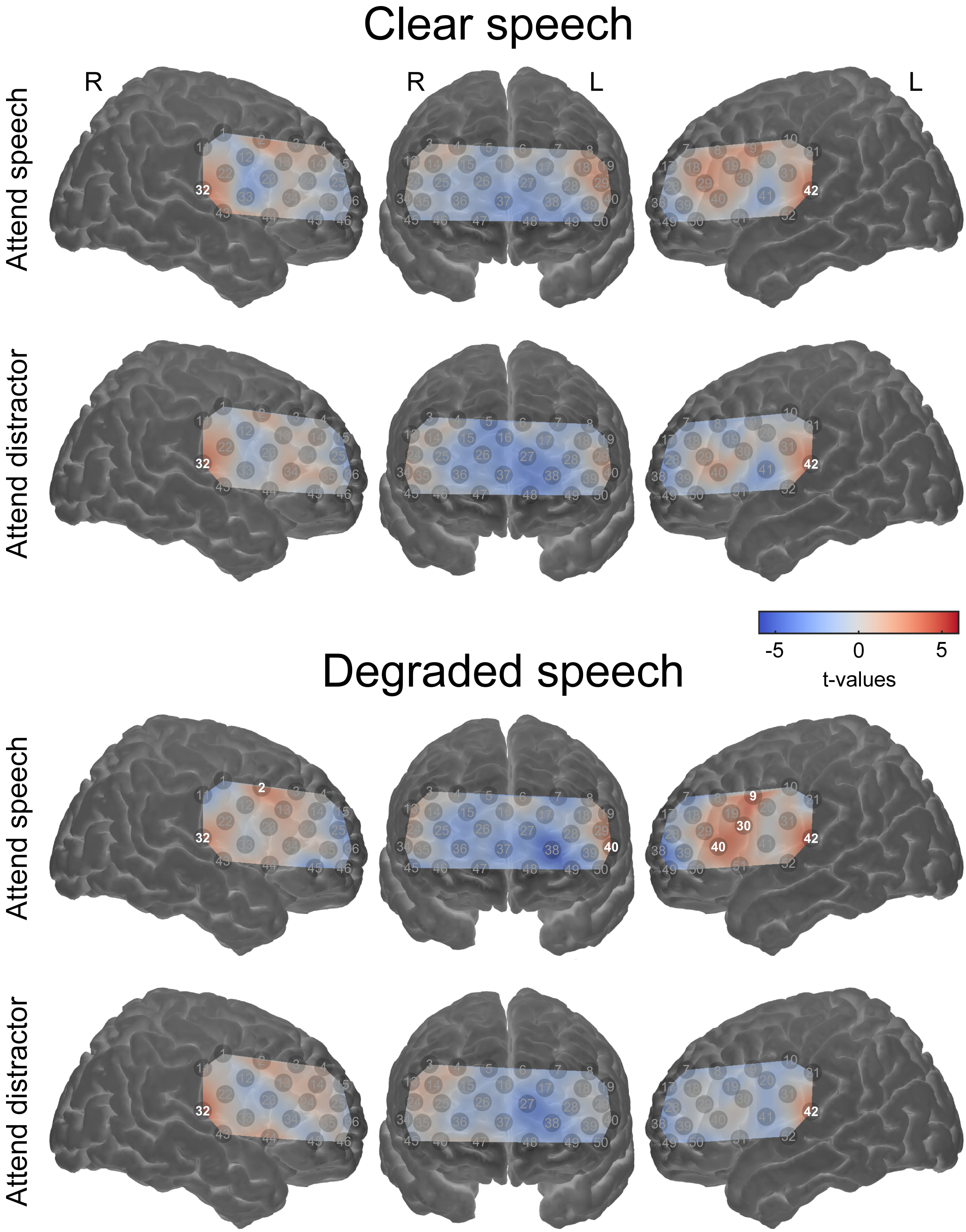
Zekveld, A.A., Heslenfeld, D.J., Festen, J.M., Schoonhoven, R. 2006. Top-down and bottom-up processes in speech comprehension. NeuroImage 32, 1826-36.

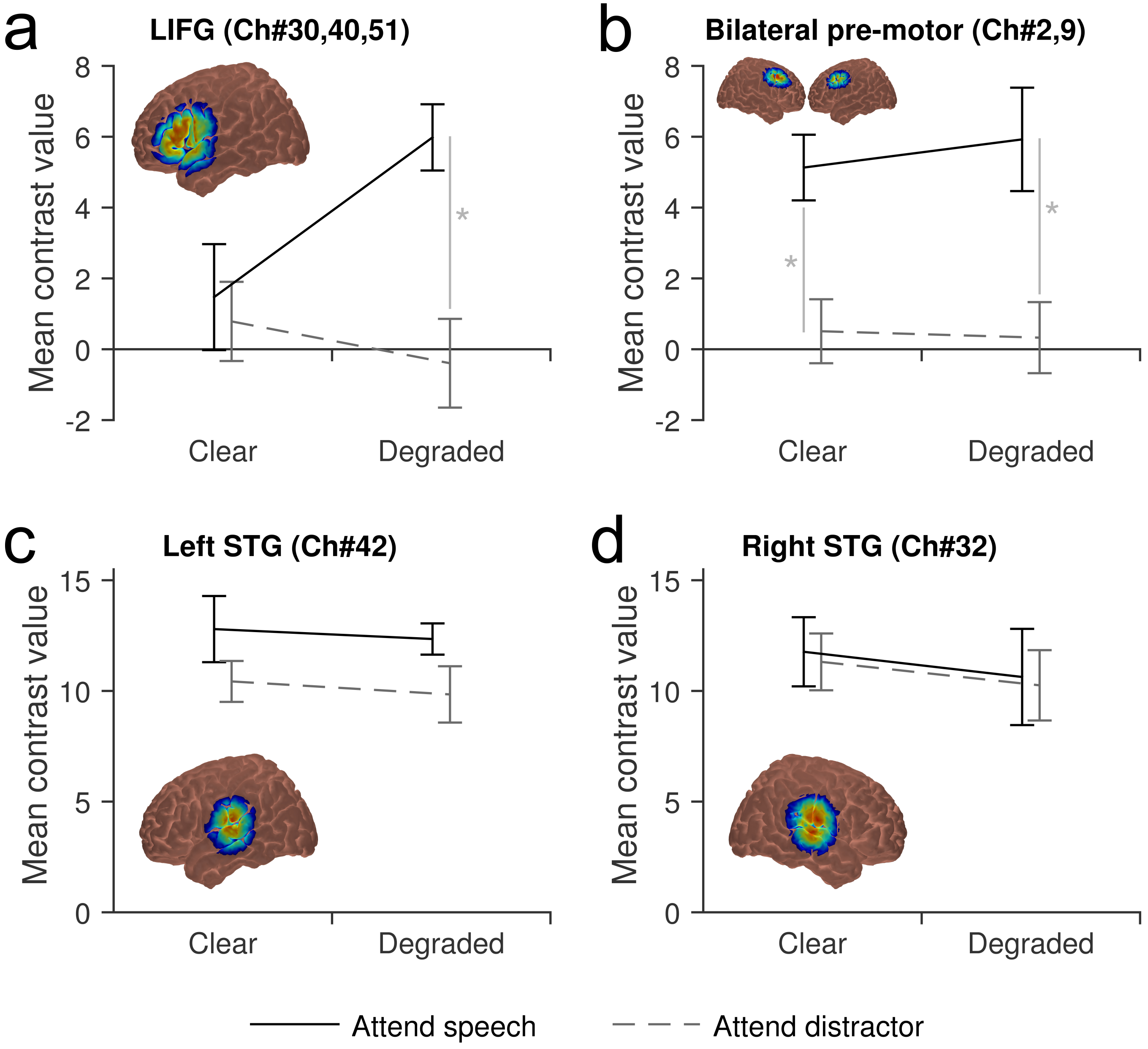
Zekveld, A.A., Heslenfeld, D.J., Johnsrude, I.S., Versfeld, N.J., Kramer, S.E. 2014. The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. NeuroImage 101, 76-86.

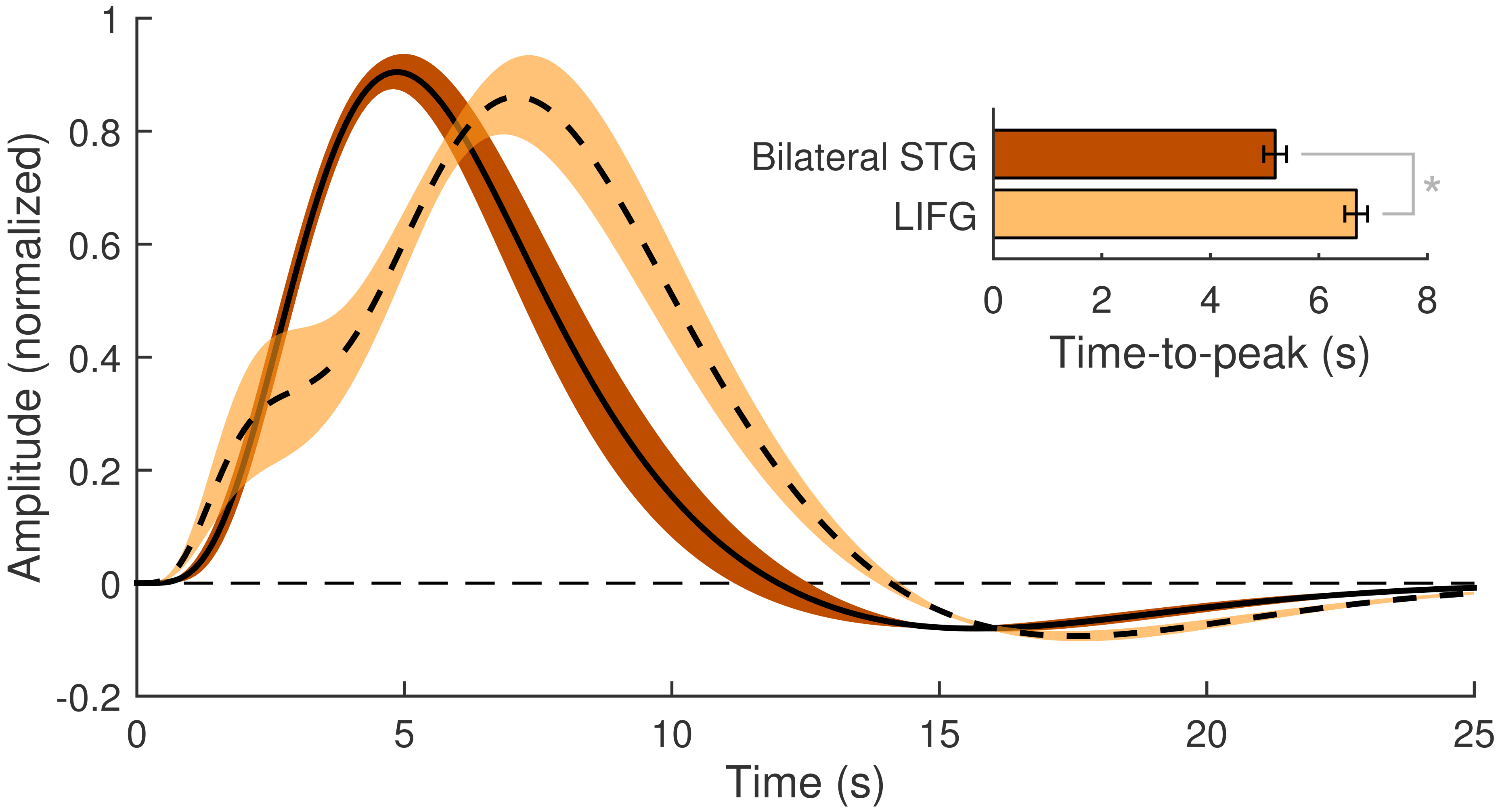

a
b
c
30
40
51
-3
-2
-1
0
10^

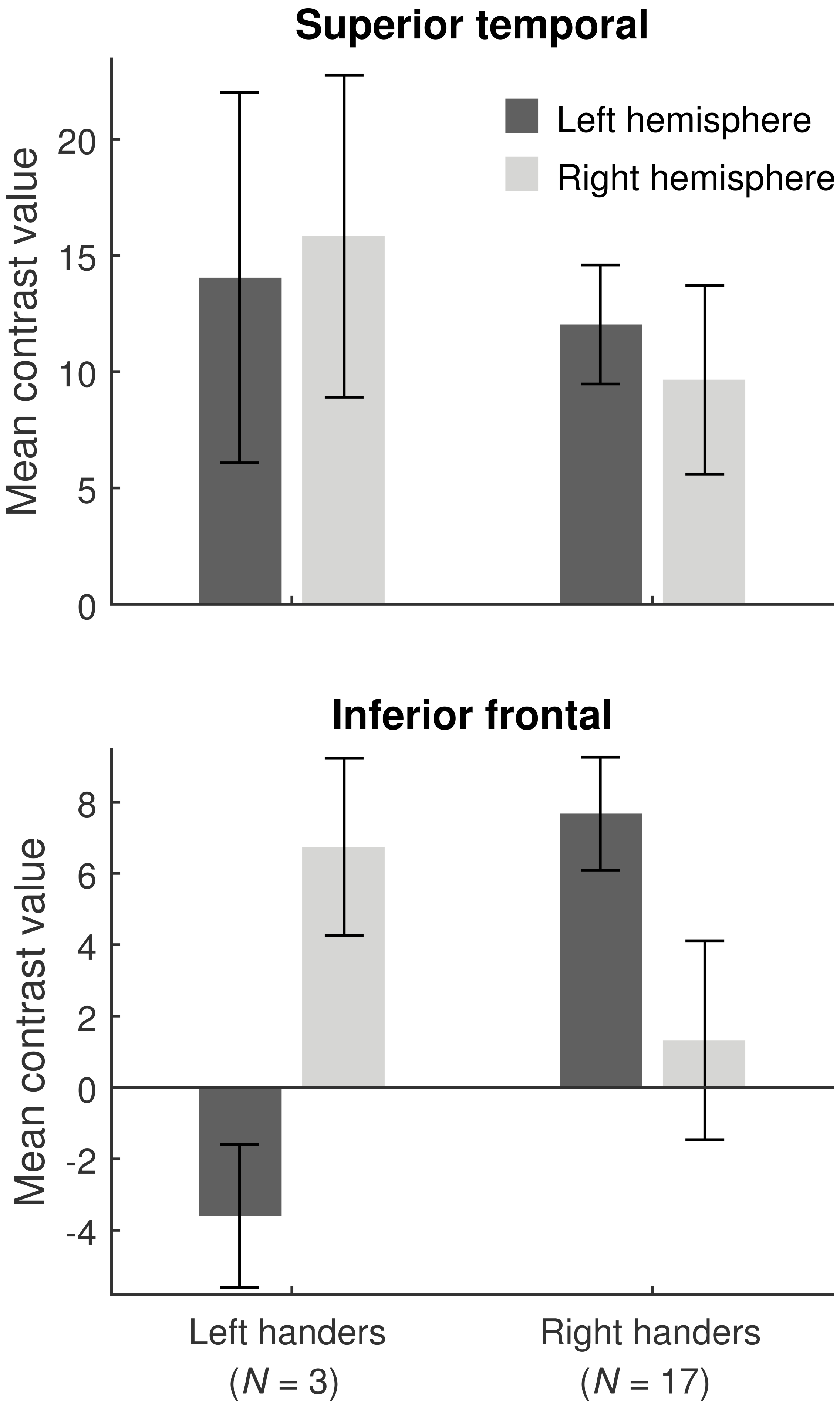

Intelligibility (%-correct)
100
80
60
40
20
0
***
n.s.
***
Degraded list 1
Degraded list 2
Degraded list 3
Clear


Main task
Sentence recognition post-test
Sensitivity (d′)
0
2
4
6
0
0.5
1
1.5
2
*
Clear
Degraded
Clear
Degraded
Attend speech
Attend distractor


Clear speech
R
R
L
L
Attend speech
Attend distractor
32
42
-5
0
5
t-values
Degraded speech
Attend speech
Attend distractor
2
32
40
9
30
40
42
32
42


a
LIFG (Ch#30,40,51)
b
Bilateral pre-motor (Ch#2,9)
c
Left STG (Ch#42)
d
Right STG (Ch#32)
Mean contrast value
8
6
4
2
0
-2
15
10
5
0
*
Clear
Degraded
Attend speech
Attend distractor


1
0.8
0.6
0.4
0.2
0
-0.2
Amplitude (normalized)
0
5
10
15
20
25
Time (s)
Bilateral STG
LIFG
*
0
2
4
6
8
Time-to-peak (s)


Superior temporal
Left hemisphere
Right hemisphere
Mean contrast value
0
5
10
15
20
Inferior frontal
Mean contrast value
-4
-2
0
2
4
6
8
Left handers
(*N* = 3)
Right handers
(*N* = 17)

Highlights

- The viability of event-related auditory fNIRS imaging is demonstrated
- Results corroborate important findings reported in the fMRI literature
- Processing of degraded speech in inferior frontal cortex depends on attention
- Haemodynamic responses peak later in frontal versus temporal speech-sensitive areas
- fNIRS holds promise for investigating the neural signature of effortful listening