Research Paper

Cortical correlates of speech intelligibility measured using functional near-infrared spectroscopy (fNIRS)

Rachael J. Lawrence [a, b, *], Ian M. Wiggins [a, c, d], Carly A. Anderson [a, c], Jodie Davies-Thompson [a, c, e], Douglas E.H. Hartley [a, b, c, d]

[a] National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, 113 The Ropewalk, Nottingham, NG1 5DU, United Kingdom
[b] Nottingham University Hospitals NHS Trust, Queens Medical Centre, Derby Road, Nottingham, NG7 2UH, United Kingdom
[c] Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, NG7 2UH, United Kingdom
[d] Medical Research Council (MRC) Institute of Hearing Research, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom
[e] Department of Psychology, Swansea University, Singleton Park, Swansea, SA2 8PP, United Kingdom

ARTICLE INFO

ABSTRACT

Functional neuroimaging has identified that the temporal, frontal and parietal cortex support core aspects of speech processing. An objective measure of speech intelligibility based on cortical activation in these brain regions would be extremely useful to speech communication and hearing device applications. In the current study, we used noise-vocoded speech to examine cortical correlates of speech intelligibility in normally-hearing listeners using functional near-infrared spectroscopy (fNIRS), a non-invasive, neuroimaging technique that is fully-compatible with hearing devices, including cochlear implants. In twenty-three normally-hearing adults we measured (1) activation in superior temporal, inferior frontal and inferior parietal cortex bilaterally and (2) behavioural speech intelligibility. Listeners heard noise-vocoded sentences targeting five equally spaced levels of intelligibility between 0 and 100% correct. Activation in superior temporal regions increased linearly with intelligibility. This relationship appears to have been driven in part by changing acoustic properties across stimulation conditions, rather than solely by intelligibility *per se*. Superior temporal activation was also predictive of individual differences in intelligibility in a challenging listening condition. Beyond superior temporal cortex, we identified regions in which activation varied non-linearly with intelligibility. For example, in left inferior frontal cortex, activation peaked in response to heavily degraded, yet still somewhat intelligible, speech. Activation in this region was linearly related to response time on a simultaneous behavioural task, suggesting it may contribute to decision making. Our results indicate that fNIRS has the potential to provide an objective measure of speech intelligibility in normally-hearing listeners. Should these results be found to apply similarly in the case of individuals listening through a cochlear implant, fNIRS would demonstrate potential for a clinically useful measure not only of speech intelligibility, but also of listening effort.

1. Introduction

Functional neuroimaging has identified critical components of the functional organization of speech and language processing in the human brain (Hickok and Poeppel, 2000; Scott and Johnsrude, 2003). Functional magnetic resonance imaging (fMRI) studies show

that temporal cortex responses to speech stimuli of varying intelligibility (Evans et al., 2014; Zekfeld et al., 2006; Belin et al., 2000). Furthermore, using positron emission tomography (PET) Scott and colleagues (2006) found a selective response to intelligible speech in the left anterior superior temporal sulcus by varying the intelligibility of speech parametrically using vocoded stimuli.

It has been suggested that neuroimaging of cortical language areas could provide an objective way to complement existing behavioural assessments of speech understanding in speech communication and hearing device applications (Anderson et al., 2017a). For example, such a measure could prove particularly

useful in very young children with a cochlear implant (CI), since accurate behavioural measures of speech perception to assess this group are currently lacking (Lawler et al., 2015). Unfortunately most neuroimaging techniques, including fMRI are not compatible with a CI. Furthermore, the risks associated with cumulative effects of radionuclide exposure limits the clinical usefulness of PET as a repeatable clinical tool to measure speech understanding in clinical populations.

Functional near-infrared spectroscopy (fNIRS) is an increasingly popular non-invasive optical technique that can be used safely and repeatedly for studying cortical function. This technique images the haemodynamic response to neuronal activity in the brain via the use of near-infrared light (Boas et al., 2014). Low-power near-infrared light is directed through the scalp and into the cortex; the intensity of the light returning to the surface of the scalp is then detected. Changes in the concentration of oxygenated haemoglobin (HbO) and deoxygenated haemoglobin (HbR) can be measured which are subsequently interpreted as an indirect reflection of neuronal activity. It has been shown that fNIRS can reliably measure cortical responses to speech in normally-hearing adults (Wiggins et al., 2016) and children (Blasi et al., 2014), at least at a group level. Furthermore, fNIRS has been shown to be suitable for use with both adult (Anderson et al., 2017b; Bisconti et al., 2016; Chen et al., 2016; McKay et al., 2016; Olds et al., 2016; Van de Rijt et al., 2016) and paediatric (Sevy et al., 2010) CI recipients.

Pollonini et al. (2014) used fNIRS in normally-hearing individuals to compare responses to speech with a baseline condition. They showed that the temporal cortex is more responsive to clear speech, than to scrambled speech or to environmental sounds. More recently, Defenderfer et al. (2017) used an event-related design in normally-hearing adults to investigate temporal lobe activation across different listening conditions using fNIRS. However, since intelligibility was not parametrically varied in these fNIRS studies, it was difficult to interpret whether there was an effect of speech intelligibility on cortical responsiveness, as opposed to another acoustic variable such as pitch or modulation envelope.

The aim of the current study was to investigate cortical correlates of speech intelligibility using fNIRS. We presented normally-hearing adults with noise-vocoded speech (Shannon et al., 1995) that were parametrically manipulated to vary speech intelligibility. A combination of normally-hearing participants and noise vocoding was used to assess the efficacy of fNIRS to detect cortical correlates of speech intelligibility in degraded listening conditions, while at the same time allowing us to readily target performance levels across the full range from 0 to 100% correct.

We principally targeted superior temporal brain regions, which previous fNIRS studies suggest may be sensitive to speech intelligibility (Olds et al., 2016; Pollonini et al., 2014). Additionally, we used a probe that provided coverage extending beyond superior temporal cortex, including the left inferior frontal cortex. Previous research suggests that activation in this frontal region might be expected to vary non-monotonically with speech intelligibility, potentially reflecting compensatory activity under more effortful listening conditions (Davis and Johnsrude, 2003; Wild et al., 2012; Wijayasiri et al., 2017). We aimed to assess correlations between fNIRS responses and speech intelligibility both within individuals (i.e. across different levels of acoustic degradation) and between individuals (i.e. assessing individual differences in intelligibility amongst listeners exposed to stimuli at a fixed level of acoustic degradation).

2. Materials and methods

2.1. Participants and ethical approval

Twenty-three healthy adult volunteers (median age 24.6 years,

range 18—38 years, 5 males, 18 females) were recruited to this study. The design was approved by the University of Nottingham Faculty of Medicine and Health Sciences Research Ethics Committee and all participants gave written informed consent. Participants had no known cognitive or psycho-motor impairments, and all self-reported normal hearing during a screening questionnaire. All participants were native English speakers with normal or corrected-to-normal vision. Most participants were right handed (20 out of 23) as assessed using the Edinburgh Handedness Inventory (Oldfield, 1971).

2.2. Equipment

Testing was conducted in a sound-attenuated room with the lighting dimmed. Participants were seated approximately 75 cm from a visual display unit. A Genelec 8030A loudspeaker mounted immediately above and behind the display was used to present the auditory stimuli, at a level of 65 dB SPL (A-weighted root-mean-square level averaged over the duration of each sentence, measured at the listening position using a Brüel & Kjær 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). A dense sound-absorbing screen was placed between the fNIRS equipment and the listening position, resulting in a steady ambient noise level of 38 dB SPL (A-weighted). During the main fNIRS task, 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. Speech stimuli

Stimuli were generated by applying eight-channel noise vocoding (Shannon et al., 1995) to recordings of Bamford-Kowal-Bench (BKB) sentences (Bench et al., 1979) spoken by a male talker. The eight channels were spaced approximately equally along the basilar membrane (Greenwood, 1990) and spanned an overall bandwidth of 180—8000 Hz. Channel filtering was performed using 6th-order digital elliptic filters applied consecutively in the forward and reverse directions to avoid phase distortion (MATLAB *filtfilt* function). Within-channel amplitude envelopes were extracted by half-wave rectification followed by zero-phase low-pass filtering at 160 Hz using a 1st-order elliptic filter (applied consecutively in the forward and reverse directions). Following subsequent manipulation (see below), each envelope was then applied to a white-noise carrier and bandpass filtered using the same filters as used for analysis. Input and output root-mean-square (RMS) levels were matched on a within-channel basis, before summation across channels to arrive at the final stimulus.

To parametrically vary intelligibility, we manipulated the depth of envelope modulation within each vocoder channel by raising the extracted envelopes to a fractional power (same exponent for all channels). Based on the results of pilot testing (see below for details), we used envelope exponents of 0.000, 0.149, 0.212, 0.297 and 1.000, chosen to target group-mean intelligibility levels of 0, 25, 50, 75 and 100% keywords correct, respectively. These five stimulation conditions are referred to as S_0, S_{25}, S_{50}, S_{75} and S_{100} throughout the remainder of the manuscript. Note that an envelope exponent of zero resulted in a steady speech-shaped noise containing no linguistic information, while an exponent of one left the original speech envelope unaltered (i.e. comparable to a standard eight-

channel vocoder).

The pilot study was carried out using seven adult volunteers (who did not take part in the main study). The volunteers were presented with BKB sentences processed as described above, though with one of seven different envelope exponents (1, 0.5, 0.25, 0.16, 0.125, 0.1 or 0). Following each sentence, the volunteer was instructed to repeat back what they had heard and the number of keywords correctly identified was scored. Each volunteer heard 16 sentences (one BKB list) per envelope exponent, with all trials interleaved in random order. We fitted a psychometric function to the group-average speech intelligibility data (see Supplementary Figure 1), and from this derived five envelope exponents for use in the main study that were expected to achieve group-average intelligibility levels of around 0, 25, 50, 75 and 100%. Volunteers in the pilot study received comparable practice listening to noise-vocoded sentences before testing began as did participants in the main study.

2.4. Speech intelligibility testing

All participants performed a behavioural test of speech intelligibility both before (pre-imaging session) and after the main fNIRS task (post-imaging session). During these tests participants were presented with a total of 80 sentences, 16 per stimulation condition, in random interleaved order. After each sentence, participants were asked to respond verbally by attempting to repeat what they had heard. The experimenter scored the number of keywords correctly reported using a touchscreen device. Before commencing data collection, participants completed a short familiarization session in which six sentences were presented for each stimulation condition. The purpose of this familiarization session was to ensure that participants understood the task and to provide them with some initial practice listening to the degraded speech stimuli.

2.5. Main fNIRS task

An event-related neuroimaging design was used in which each event corresponded to the presentation of a single sentence. Participants were presented with 20 sentences per stimulation condition, interspersed with 20 silent trials which acted as a baseline. All trial types were presented in randomized order. 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 (average SOA: 9 s; average offset-to-onset gap: 7.4 s), following a previous event-related fNIRS study conducted in our laboratory (Wijayasiri et al., 2017). Randomising the SOA helps to reduce the influence of preparatory and anticipatory factors and allows the response to multiple trial types to be deconvolved despite the possibility of temporal overlap in the haemodynamic activity elicited by consecutive trials (Dale, 1999). The imaging lasted approximately 18 min in total.

During imaging, participants were instructed to look at a fixation cross presented on a uniform background, while sitting as still as possible to minimize motion artefacts. Instead of participants repeating sentences verbally, 0.5 s after each stimulus ended a probe word appeared on the display. The probe word was, with equal probability, either a word that had featured in the presented sentence or a replacement foil word, chosen to rhyme with one of the true keywords. Foil words were equally likely to fall towards the start, middle, or end of the sentence. Participants were required to indicate by a button press whether the probe word had appeared in the sentence just heard. "Yes" and "No" labels were presented towards either side of the display accordingly. The orientation of these labels was fixed throughout an individual's session; however, the side that corresponded to a "Yes" response was alternated for even- and odd-numbered participants. Participants had up to 2 s to respond; otherwise a missed response was recorded. On silent trials, in lieu of the speech-based task, participants were simply randomly instructed to "Press yes" or "Press no". To ensure that participants were confident with the procedure, a short familiarization session was conducted before the optode array was placed on the participant's head.

2.6. fNIRS measurements

Measurements were made with a total of 30 optodes arranged in two 3 × 5 arrays (each containing 8 sources and 7 detectors). The arrays were placed on both sides of the head (Fig. 1a). We aimed to primarily measure cortical activation in the following regions bilaterally i) superior temporal cortex; ii) inferior frontal cortex and; iii) inferior parietal cortex. To ensure consistency of optode placement across individuals, we followed a fixed protocol. The International 10–20 System (Jasper, 1958) was used to guide and standardise optode placement: the central optical source in the bottom row of the array was positioned in vertical alignment with the preauricular point, with the uppermost source aligned towards position Cz. The inter-optode spacing was fixed at 30 mm.

To give an indication of the variability in probe placement across individuals, Fig. 1b shows optode locations measured on six volunteers who did not take part in the main study. A 3D digitizer was used to record the position of the optodes, as well as anatomical surface landmarks. The location of all optodes relative to the surface landmarks of the left and right tragus, nasion, inion and Cz were recorded for the purpose of estimating optode positioning relative to underlying cortical anatomy. The measured positions were registered to the "Colin 27" atlas brain (Collins et al., 1998) using the AtlasViewer tool (Aasted et al., 2015). The mean optode positions across the six volunteers were used as the basis for data visualization (Fig. 1b and c).

2.7. Analysis of fNIRS data

Analysis of the fNIRS data was performed in MATLAB (Math-Works, Natick, MA) using functions provided in the HOMER2 package (Huppert et al., 2009) together with custom scripts. The analysis proceeded in a similar manner to previous studies conducted in our laboratory (Anderson et al., 2017b; Wiggins et al., 2016; Wiggins and Hartley, 2015; Wijayasiri et al., 2017). In brief, the following steps were performed:

Exclusion of channels with poor signal quality: we used the scalp coupling index (SCI) to identify and exclude channels suffering from poor optode–scalp contact (Pollonini et al., 2014). We excluded channels with SCI <0.16, chosen to exclude only the worst 5% of channels in our dataset. In previous studies, we have found that excluding only a small percentage of the worst channels generally offers a good compromise between the wish to include only high-quality channels in the analysis vs. the need to maintain as large a sample size as possible for statistical power. The distribution of SCI values across our entire dataset is provided in Supplementary Figure 2.

Conversion to optical density: the measured light intensity levels were converted to optical density using the HOMER2 *hmrIntensity2OD* function, a standard step in fNIRS data analysis (Huppert et al., 2009).

Motion-artefact correction: motion artefacts were suppressed using the HOMER2 *hmrMotionCorrectWavelet* function, which implements a simplified form of the wavelet filtering algorithm described by Molavi and Dumont (2012). We excluded wavelet coefficients lying further than 0.719 times the interquartile range below/above the first/third quartile, respectively.
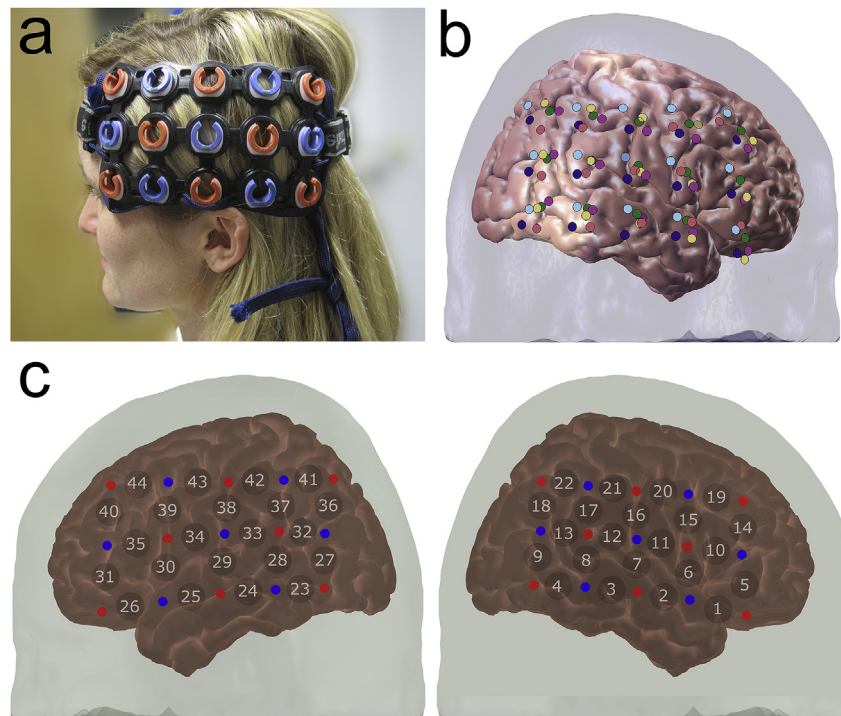
Fig. 1. Optode positioning. (a) Photograph of the optode array holder (without optodes inserted) placed on the head of one of the authors for demonstration purposes. Red/blue coding indicates optical sources/detectors, respectively. (b) Variability in digitized optode positions across the six volunteers. Optode positions for each volunteer are represented by different colour dots (right-hemisphere shown). (c) The mean optode positions across the six volunteers following registration to an atlas brain are represented by red/blue colour. The numbers indicate the position of each channel as referred to in the manuscript (channels 1–22 right-hemisphere, channels 23–44 left-hemisphere).

Bandpass filtering: the optical density signals were band-pass filtered between 0.02 and 0.5 Hz to attenuate low-frequency drift and cardiac oscillations.

Conversion to estimated changes in haemoglobin concentrations: optical density was converted to estimated changes in the concentration of HbO and HbR through application of the modified Beer-Lambert Law (Huppert et al., 2009). A default value of 6 was used for the differential path-length factor at both wavelengths. Note that the continuous-wave fNIRS system used in the present study allows for the estimation only of relative changes in haemoglobin concentrations and not absolute concentrations.

Isolation of the functional haemodynamic response: we applied the haemodynamic modality separation (HMS) algorithm described by Yamada et al. (2012) to isolate the functional component of the haemodynamic signal and suppress systemic physiological interference. This algorithm attempts to separate functional and systemic signals based on the assumption that the correlation between HbO and HbR will be different in each case. Although this approach does not accurately account for all statistical properties of the noise typically found in fNIRS data (Huppert, 2016), in previous studies we have found application of this algorithm to be beneficial to the detection of auditory cortical activation (Wiggins et al., 2016; Wijayasiri et al., 2017); in particular, application of the HMS algorithm was shown to substantially improve the test-retest reliability of auditory fNIRS measurements (Wiggins et al., 2016).

Quantification of response amplitude: to quantify fNIRS response amplitude, we used a general linear model (GLM) approach previously described in Wijayasiri et al. (2017). The GLM was applied to the continuous data collected over the duration of the imaging session. The design matrix included a set of three regressors (corresponding to the canonical haemodynamic response plus its first two temporal derivatives) for each experimental condition, plus a further set for the silent trials. Each trial was modelled as a short epoch corresponding to the actual period of auditory stimulation on that trial (mean duration 1.64 s; audio muted on silent trials). Within each condition, the canonical and temporal-derivative regressors were serially orthogonalized with respect to one another (Calhoun et al., 2004). 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). The 'derivative-boost' technique (Calhoun et al., 2004; Steffener et al., 2010) was used to estimate response amplitude: this technique calculates an amplitude value that is a function of both the canonical (nonderivative) and the derivative terms of the model; the resulting amplitude estimates are less affected by any systematic differences in latency or dispersion between conditions, compared to if the amplitude is estimated from the canonical term alone.

2.8. Statistical analyses

2.8.1. Behavioural data

Repeated-measures analyses of variance (RM-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.

The speech intelligibility data were analysed using a two-way RM-ANOVA with within-subjects factors "stimulation condition" (five levels: S_0, S_{25}, S_{50}, S_{75} and S_{100}) and "session" (two levels: pre- vs. post-imaging). The percent-correct scores were transformed to rationalized arcsine units (RAUs) prior to analysis to help equalize variance across different performance levels (Studebaker, 1985).

The main-task behavioural results were analysed using a one-way RM-ANOVA with within-subjects factor "stimulation condition" (five levels as listed above). Separate analyses were performed

for accuracy and response time. Polynomial contrasts were used to identify trends in the data across stimulation conditions. Note that, in keeping with Binder et al. (2004), response time data from all trials were included in the analysis, regardless of whether the participant's response was correct or incorrect.

2.8.2. fNIRS data

Our primary statistical analysis aimed to establish the relationship between speech intelligibility and fNIRS response amplitude on a channel-wise basis. The analysis began by contrasting the estimated response amplitude for each experimental condition against the silent baseline. The contrast values for each channel were then analysed using a linear mixed model (LMM) (West et al., 2006). The use of a LMM allowed us to include as a predictor variable subject-specific estimates of speech intelligibility. These subject-specific estimates were calculated as the mean percent-correct scores across the pre- and post-imaging sessions. To enable us to detect possible non-linear relationships between brain activation and speech intelligibility, we included the polynomial expansion of intelligibility up to 2nd order. Serial orthogonalization was applied to the polynomial terms such that the 0th order (constant), 1st order (linear) and 2nd order (quadratic) fixed effects were independent of one another and could be separately and straightforwardly interpreted (Büchel et al., 1998). The effect of the serial orthogonalization was such that maximal variance was at each stage assigned to the lower-order term (e.g. the constant term), with subsequent terms (e.g. the linear and quadratic terms) tested on the remaining variance. Each LMM additionally included a random intercept effect to model between-subject variability.

We performed secondary group-level analyses to obtain further insight into the nature of the fNIRS-measured cortical activation. Following the approach of Binder et al. (2004), we tested for a linear relationship between fNIRS response amplitude and response time to identify brain regions principally involved in decision-making, rather than sensory, processes during speech perception. This analysis was performed using a LMM similar to that employed in the primary analysis, but with the following differences: i) mean response time on the main task (conducted simultaneously with fNIRS imaging), rather than speech intelligibility, was used as the predictor variable; and ii) the polynomial expansion of mean response time was restricted to 1st order (linear effect).

Recognising that any systematic relationships observed between speech intelligibility and fNIRS response amplitude could, in theory, be driven by an effect of changing acoustic properties across stimulation conditions, rather than intelligibility *per se*, we ran a further analysis aimed at identifying regions of the brain that responded more strongly on trials that were perceived correctly versus those that were not. To achieve this, we re-ran the original GLM analysis, but now including additional regressors coding all "correct" and "incorrect" trials. We performed a group-level random-effects analysis on the resulting estimated response amplitudes by using two-tailed *t*-tests to compare correct versus incorrect trials on a channel-wise basis. Since we did not perform any orthogonalization of the new regressors with respect to the original condition-specific regressors, we were able to isolate variance that could be uniquely attributed to the veracity of trial-by-trial perception, after accounting for the effect of changing acoustic properties across conditions (Mumford et al., 2015).

Finally, we examined whether fNIRS response amplitude correlated with individual differences in behaviourally measured speech intelligibility. For this purpose, we focused on the S_{50} condition, in which speech intelligibility varied substantially between individuals and was unaffected by floor or ceiling effects. On a channel-wise basis, we calculated the Pearson correlation coefficient between (a) individually measured speech intelligibility and

(b) the fNIRS contrast between the S_{50} and S_0 conditions (i.e. the contrast between the response to the speech stimuli of interest and steady speech-shaped noise). This analysis is informative because it allows us to assess whether a one-shot fNIRS measurement (i.e. of a single stimulation condition) shows sensitivity to differences in speech intelligibility *between individuals*, as opposed to differences across stimulation conditions *within an individual*.

In all fNIRS analyses, we accounted for multiple comparisons by applying the false discovery rate (FDR) method (Benjamini and Hochberg, 1995) across channels. We used the original formulation of the FDR procedure, which offers greater statistical power, but which requires the assumption of either independence or slight positive dependency among channels. Although slight positive spatial correlation of channel-wise statistics is generally expected in fNIRS data (Singh and Dan, 2006), we cannot be certain that this assumption holds in all cases. Fortunately, the FDR procedure seems to be robust to the presence of deviations from the assumption of positive dependency (Groppe et al., 2011), especially as the number of tests increases (Clarke and Hall, 2009). We adopted an FDR-corrected threshold of $q < 0.05$, meaning that, of all the individual-channel effects we report as being statistically significant, on average around 1 in 20 may be expected to be erroneous (i.e. a false positive).

3. Results

3.1. Behavioural results

3.1.1. Speech intelligibility

At intermediate levels (S_{25}, S_{50}, and S_{75} conditions), mean intelligibility was between 5 and 14 percentage-points higher in the post-imaging session than in the pre-imaging session, suggesting a training effect over the course of the experiment (Fig. 2). This was confirmed by a RM-ANOVA performed on the (RAU-transformed) percentage of keywords correctly understood, which revealed a significant main effect of session ($F(1, 22) = 45.26$, $p < .001$), as well as a significant interaction between session and stimulation condition ($F(2.31, 50.72) = 8.95$, $p < .001$) (a training effect was observed only at intermediate intelligibility levels). Nonetheless, mean intelligibility scores across the pre- and post-imaging sessions were close to the targeted values: 0% for S_0, 22% for S_{25}, 50% for S_{50}, 74% for S_{75} and 99% for S_{100}. We therefore assumed that the average of each individual's pre- and post-imaging scores provided an appropriate measure of the intelligibility of the stimuli they were exposed to during the main fNIRS task.
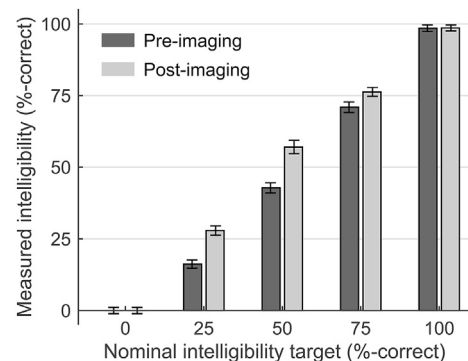
Fig. 2. **Mean speech intelligibility scores**. The mean score is shown for each of the five stimulation conditions in the pre- and post-imaging test sessions. Error bars show ±1 standard error of the mean (corrected for repeated measures).

3.1.2. Main fNIRS task

Accuracy and response time data for the main task (conducted simultaneously with fNIRS imaging) are plotted in Fig. 3. As expected, accuracy in distinguishing true keywords from rhyming foil words increased monotonically across the five stimulation conditions ($S_0 \cdots S_{100}$). A RM-ANOVA confirmed a significant effect of stimulation condition ($F(4, 88) = 145.17$, $p < .001$) with polynomial trend analysis revealing significant linear, quadratic, and cubic components ($p < .05$). Mean response time, in contrast, varied non-monotonically across stimulation conditions, peaking in the S_{25} condition. There was again a significant effect of stimulation condition ($F(2.06, 43.33) = 14.65$, $p < .001$) with significant linear, quadratic, cubic and 4th-order trends ($p < .05$). Missed responses were rare, occurring on less than 1% of all trials.

Accuracy on the main task was monotonically related to speech intelligibility as measured in the pre- and post-imaging sessions, although the relationship was curvilinear owing to the fact that chance performance in the main task was at 50%, even for a fully unintelligible stimulus (see Supplementary Figure 3).

3.2. fNIRS results

3.2.1. Channel-wise analysis of relationships with speech intelligibility

The results of the primary analysis testing for systematic relationships between fNIRS response amplitude and speech intelligibility are shown in Fig. 4. A corresponding table of statistical results is provided as supplementary information (Supplementary Table 1). Fig. 4a shows the 0th-order effect, i.e. overall activation or deactivation in response to sound vs. silence, irrespective of intelligibility. Statistically significant activation ($q < 0.05$, FDR corrected) was observed in channels overlying left (Ch#28,29,32,33) and right (Ch#2,7,12) superior temporal cortex, right pre/post-central gyrus (Ch#11,20), and inferior frontal cortex which was more prominent in the left (Ch#30,31,35,39) than in the right (Ch#14) hemisphere. A group of channels overlying right inferior parietal cortex (Ch#18,21,22) showed significant deactivation in response to sound vs. silence.

Fig. 4b identifies channels that exhibited a significant linear relationship with speech intelligibility (1st-order effect). Such channels were found to lie principally over bilateral superior temporal cortex and were more extensive in the left (Ch#25,28,29,32,33) than in the right (Ch#7,12) hemisphere. Activation in these channels increased as the speech became more intelligible. A single channel in right inferior parietal cortex (Ch#22) showed a significant linear relationship in the opposite direction, i.e. decreased activation in response to more intelligible speech.
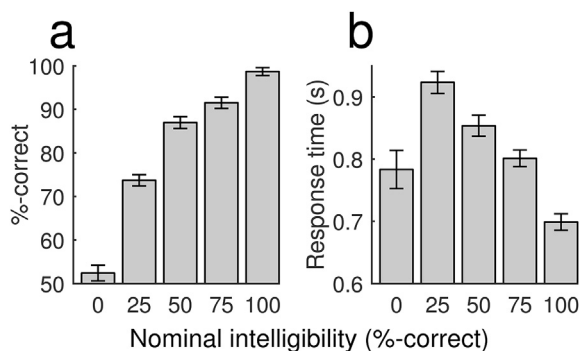
Fig. 4c identifies channels in which activation depended non-linearly on speech intelligibility (2nd-order/quadratic effect). An array of bilateral, posterior channels overlying middle temporal and inferior parietal regions (Ch#3,8,13,18,22,24,41,42) showed a significant positive quadratic relationship with intelligibility (reduced activation at intermediate intelligibility levels). A cluster of channels in left inferior frontal cortex (Ch#26,31,35) showed evidence of an "inverted-U" response profile (maximal activation at intermediate intelligibility levels), although this effect did not reach statistical significance in any individual channel after correcting for multiple comparisons.

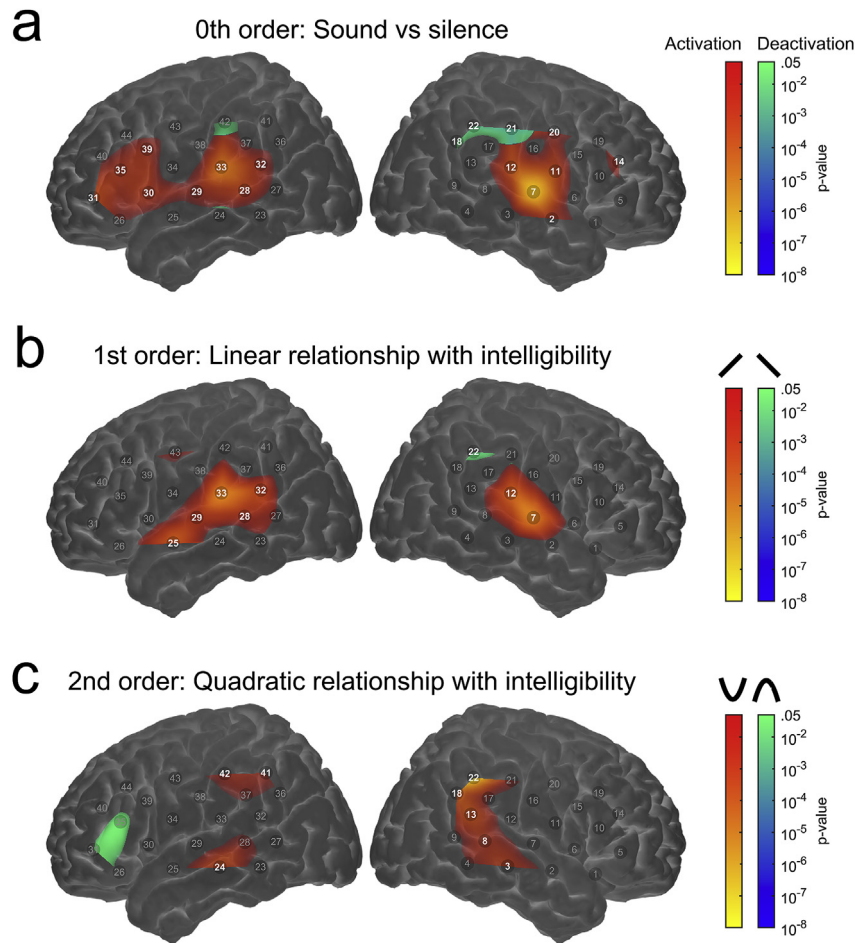#### 3.2.2. Response profiles in regions-of-interest

To provide a clearer picture of how fNIRS response amplitude varied with speech intelligibility in different parts of the brain, Fig. 5 plots mean fNIRS contrast values (relative to silence) against nominal intelligibility for select regions-of-interest (ROIs). These ROIs were defined in post-hoc manner by selecting clusters of adjacent channels that exhibited significant linear or quadratic relationships with speech intelligibility (Fig. 4b and c) and which all showed a qualitatively similar response profile within a given ROI. We gave the ROIs thus defined the following descriptive labels: left auditory (Ch#25,28,29,32,33); right auditory (Ch#7,12); left inferior frontal gyrus (LIFG, Ch#26,31,35); and bilateral posterior (Ch#3,8,13,18,22,24,41,42). Note that while the quadratic effect failed to reach statistical significance in any individual channel within the LIFG ROI (after multiple-comparisons correction), we nonetheless considered it appropriate to further investigate the response profile within this region on the basis that: i) three spatially contiguous channels all approached statistical significance; and ii) the quadratic effect was consistent with our *a priori* hypothesis for this region based on previous research (Wijayasiri et al., 2017; Wild et al., 2012).

In left and right auditory (i.e. superior temporal) regions, activation increased monotonically as intelligibility improved, with the most pronounced step occurring between the S_{75} and S_{100} conditions (Fig. 5a and b). In the LIFG, positive activation (relative to silence) was seen in all conditions (Fig. 5c), with the strength of that activation increasing monotonically as intelligibility was reduced from the S_{100} condition down to the S_{25} condition (i.e. as listening became more challenging). However, LIFG activation then fell off steeply in the S_0 condition, in which the stimuli were stripped of all task-relevant linguistic information. In the bilateral posterior ROI, covering an array of channels overlying middle temporal and inferior parietal regions, the response profile indicated a tendency towards deactivation relative to silence (Fig. 5d). The strength of deactivation was greatest at intermediate levels of intelligibility, peaking in the S_{50} condition.

Fig. 6 demonstrates grand-average event related haemodynamic time courses across the group of participants for reference. The mean fNIRS response to silent trials has been subtracted out to derive overlap-reduced event-related responses for each speech condition.

3.2.3. Relationship between fNIRS response amplitude and mean response time

We tested for channels that exhibited a linear relationship between fNIRS response amplitude and mean response time on the main task (conducted simultaneously with fNIRS imaging) (Fig. 7a). A corresponding table of statistical results is provided as supplementary information (Supplementary Table 2). Several channels in the left inferior frontal region showed a trend towards a positive linear relationship with response time ($p < .05$ uncorrected). This effect reached significance in Ch#35 after correction for multiple comparisons. Similarly, one channel (Ch#18) showed a significant

Fig. 3. Behavioural results from the main fNIRS task. Mean accuracy (a) and response time (b) is shown for each of the five stimulation conditions. Error bars show ±1 standard error of the mean (corrected for repeated measures).

Fig. 4. Channel-wise relationships between fNIRS response amplitude and speech intelligibility. Rows a–c show the results of statistical significance testing (uncorrected p-values, thresholded at $p < .05$) for 0th-order, 1st-order (linear) and 2nd-order (quadratic) effects, respectively. Individual channels exhibiting significant effects after correction for multiple comparisons ($q < 0.05$, FDR corrected) are highlighted. Note that the maps are interpolated from single-channel results and the overlay on the cortical surface is for illustrative purposes only.
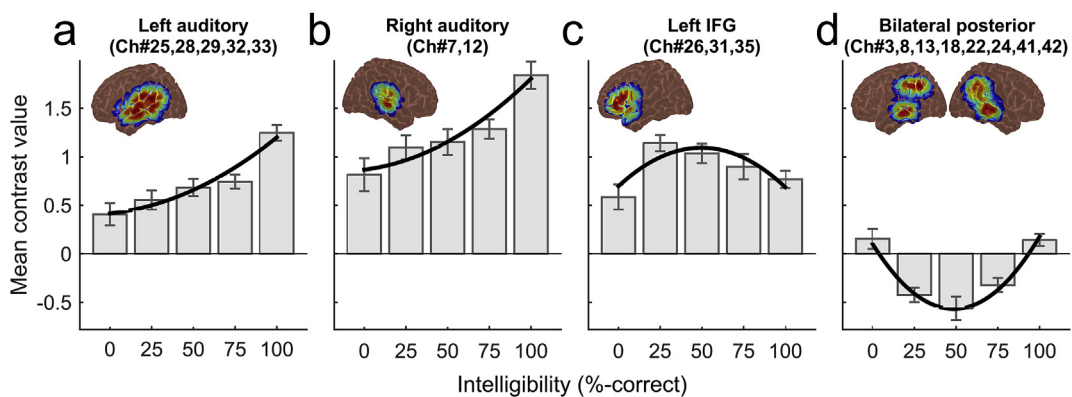

Fig. 5. Mean contrast values (i.e. estimated response amplitude relative to silence; arbitrary units) in select regions of interest. Inset figures illustrate the fNIRS sensitivity profile for each region. Error bars show ±1 standard error of the mean (corrected for repeated measures). Bold lines indicate the overall LMM fit for each region including polynomial expansion of individually measured speech intelligibility up to 2nd (quadratic) order.

negative linear relationship with mean response time; this channel, overlying right inferior parietal cortex, was amongst those that showed deactivation relative to silence at intermediate levels of intelligibility (Fig. 5d).

3.2.4. Relationship between fNIRS response amplitude and the veracity of trial-by-trial perception

Fig. 7b shows the results of channel-wise significance testing on the contrast between correctly versus incorrectly perceived trials, after accounting for the effects of changing acoustic properties
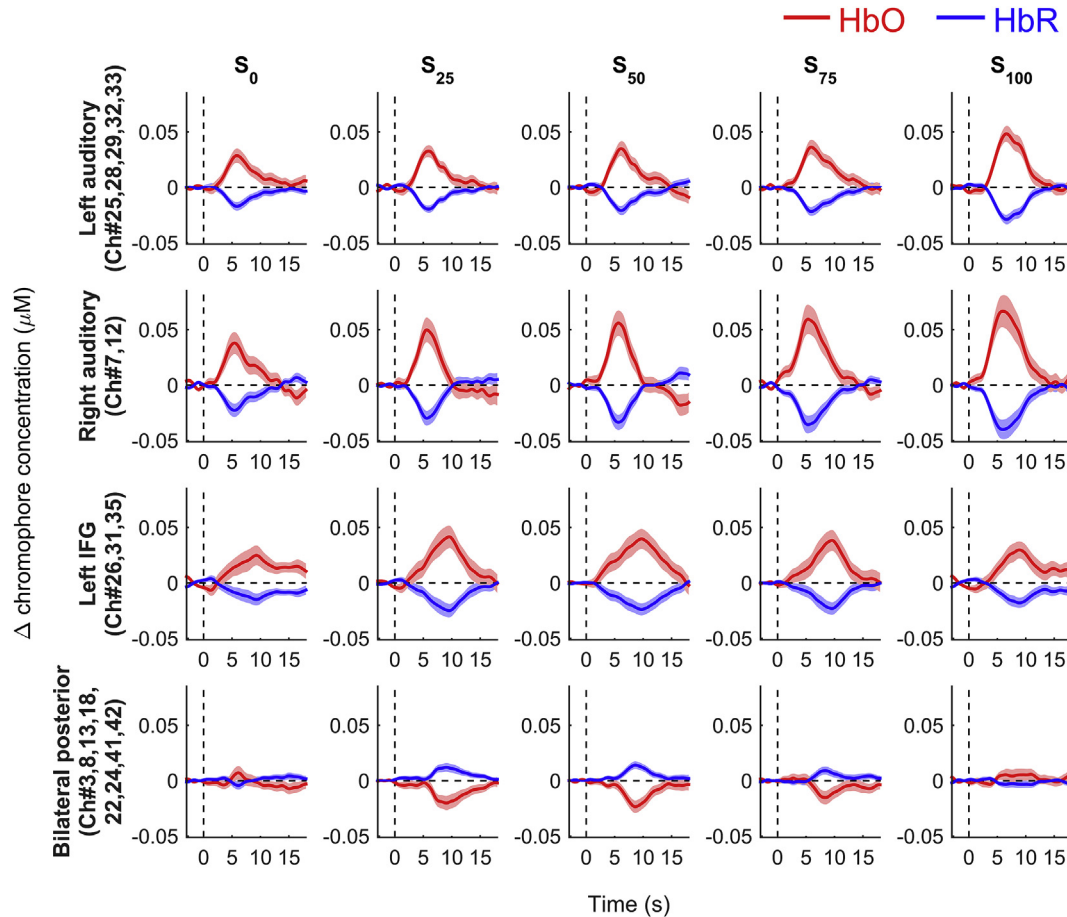
Fig. 6. Grand-average event-related haemodynamic time courses. The red and blue traces show estimated changes in the concentration of HbO and HbR, respectively. Shading indicates ±1 standard error of the mean across participants. Note that, prior to averaging across participants, the mean response to silent trials was subtracted out to derive overlap-reduced event-related responses.
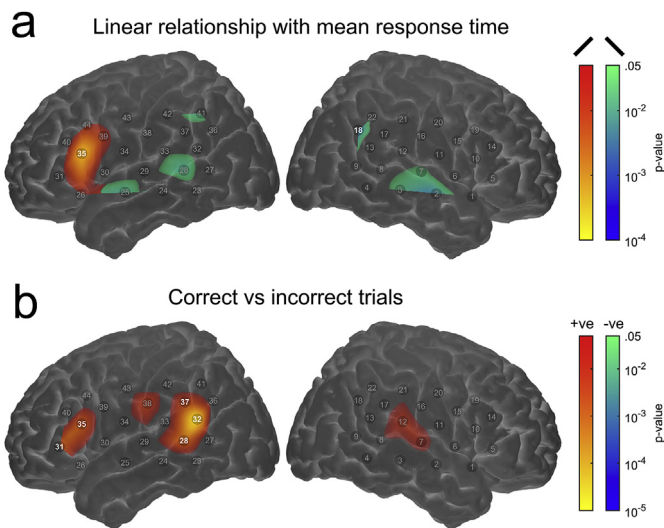

Fig. 7. Secondary group-level fNIRS analyses. (a) Testing for a linear relationship between fNIRS response amplitude and mean response time on the main task. (b) Testing for a difference between correctly and incorrectly perceived trials, after accounting for the effects of changing acoustic properties across conditions. In each case the colormap shows uncorrected *p*-values, thresholded at *p* < .05. Individual channels exhibiting a significant effect after correction for multiple comparisons (*q* < 0.05, FDR corrected) are highlighted. Note that the maps are interpolated from single-channel results and the overlay on the cortical surface is for illustrative purposes only.

across stimulation conditions. A corresponding table of statistical results is provided as supplementary information (Supplementary Table 2). Significantly greater activation on correct trials was observed in left inferior frontal cortex (Ch#31,35) and in left posterior superior temporal regions (Ch#28,32,37). Interestingly, channels overlying left auditory brain regions that had previously shown a significant linear relationship with speech intelligibility (Ch#25,29,33; Fig. 4b) did not show a difference between correct and incorrect trials in this analysis. This suggests that responses in these channels may have been driven more by changing acoustic properties across conditions, rather than by intelligibility *per se*. Channels overlying right auditory regions (Ch#7,12), which had also shown a significant linear relationship with speech intelligibility previously (Fig. 4b), showed a tendency towards greater activation on correct vs. incorrect trials (*p* < .05 uncorrected), however this effect did not reach significance in either individual channel after correction for multiple comparisons.

3.2.5. Correlation between fNIRS response amplitude and individual speech intelligibility

We found modest evidence that fNIRS response amplitude (specifically, the contrast between the S_{50} and S_0 conditions) correlated with individually measured speech intelligibility in the S_{50} condition (Fig. 8). Channels that showed a trend (*p* < .05 uncorrected) towards a positive correlation with individual speech intelligibility were clustered in left (Ch#25,29) and right (Ch#7,12) superior temporal cortex. However, the correlation did not reach
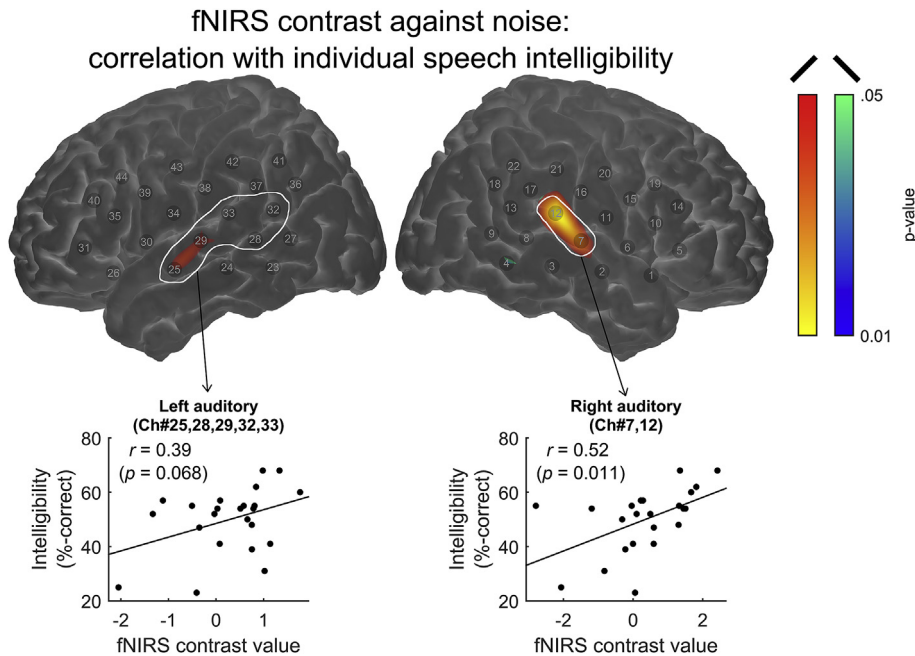
Fig. 8. Correlations between fNIRS amplitude (contrast between S_{50} and S_0 conditions) and individual speech intelligibility in the S_{50} condition. The main plot shows the results of significance testing on the channel-wise correlations (uncorrected p-values, thresholded at p < .05). The two subplots show the correlations in the left- and right-hemisphere auditory regions that showed a significant positive linear relationship with speech intelligibility in the primary group-level analysis (Fig. 4b).

statistical significance in any individual channel after correction for multiple comparisons. Recognising that fNIRS response amplitude is typically more reliable when averaged across a small number of channels overlying a cortical ROI than when assessed on a single-channel basis (Plitacha et al., 2007; Schecklmann et al., 2008; Wiggins et al., 2016), we went on to assess the correlation with individual speech intelligibility at a ROI level (Fig. 8 subplots). For this, we adopted the previously defined left and right auditory ROIs arising from the primary group-level analysis. Response amplitude in the right auditory ROI was significantly correlated with individual speech intelligibility ($r = .52$, $p = .011$). The corresponding correlation for the left auditory ROI was narrowly non-significant ($r = 0.39$, $p = .068$).

4. Discussion

We used noise vocoding to parametrically vary speech intelligibility whilst measuring cortical activation using fNIRS. The results confirm and extend the findings of previous fNIRS and fMRI studies. Specifically, we confirm that fNIRS-measured activation in superior temporal regions correlates with speech intelligibility (Pollonini et al., 2014; Olds et al., 2016), both within and, to a lesser extent, between individuals. A secondary analysis suggested that more posterior regions of left-hemisphere superior temporal cortex (Wernicke's area) may be sensitive to intelligibility *per se*, while regions lying closer to unilateral auditory cortex may be more responsive to changing acoustic properties across stimulation conditions. Beyond superior temporal cortex, we found that activation in left inferior frontal cortex peaked at intermediate levels of intelligibility, while an array of posterior channels (covering bilateral inferior parietal and middle temporal cortex) showed deactivation relative to silence that was again most pronounced at intermediate levels of intelligibility. These non-linear relationships with speech intelligibility may reflect effortful listening under challenging conditions (Wijayasiri et al., 2017; Wild et al., 2012).

4.1. Cortical correlates of speech intelligibility

Our results align with other recent fNIRS studies that reported a positive association between speech intelligibility and activation in superior temporal cortex (Pollonini et al., 2014; Olds et al., 2016). Previously, this effect has been probed by contrasting normal speech with discrete control stimuli designed to be either intelligible or unintelligible. Here, we parametrically manipulated the intelligibility of noise-vocoded sentences across the full range from 0 to 100% correct, allowing us to probe the relationship between speech intelligibility and fNIRS activation in greater depth. Consistent with previous fMRI studies (Binder et al., 2004; Davis and Jonsrude, 2003; Zekveld et al., 2006), we found that activation in superior temporal regions increased linearly as the speech became more intelligible. Interestingly, the intelligibility-sensitive region extended more anteriorly and posteriorly along the superior temporal gyrus in the left hemisphere compared with the right (Fig. 4b). While we did not directly test for inter-hemispheric differences, this pattern of results is consistent with left-hemispheric dominance for language processing in most normally-hearing individuals (Frost et al., 1999).

We found some evidence to suggest that superior temporal activation was related to speech intelligibility not just within individuals (i.e. across different levels of acoustic degradation), but also between individuals. That is, the amplitude of fNIRS activation within a right-hemisphere auditory ROI correlated with individual speech intelligibility in the challenging S_{50} listening condition (Fig. 8). Group-average intelligibility in this condition was 50% keywords correct, but individual performance varied from 23 to 68%. Thus, our results indicate that fNIRS-based measures of superior temporal activation have the potential to predict individual differences in speech intelligibility, even among a relatively homogenous population of normally-hearing listeners. However, such a relationship was not found in a recent fNIRS study of speech-in-noise perception (Defenderfer et al., 2017). In the present dataset, we also did not see a significant relationship between temporal-

lobe activation and individual speech intelligibility in the more challenging S_{25} condition. Overall, then, further research is needed to verify the ability of fNIRS to reliably predict individual differences in speech intelligibility.

Interestingly, although in our primary analysis we found that intelligibility-sensitive regions were more spatially extensive in the left hemisphere than in the right, right-hemisphere responses appear to have had greater ability to explain individual differences in speech intelligibility. Similarly, Pollonini et al. (2014) observed that right-hemisphere activation was more responsive to differences in intelligibility between stimulation conditions. Pollonini et al. (2014) speculated that this may be due to the stimuli evoking different responses in more superficial areas of the right cortex compared to the left, making the differences more easily detectable using fNIRS. In the present study, we did not formally test for inter-hemispheric differences, and so no firm conclusion can be drawn. However, the question of whether right-versus left-hemisphere fNIRS measurements have greater ability to predict individual speech intelligibility warrants consideration in future research.

Taking advantage of the event-related nature of the design, we conducted a secondary analysis comparing trials on which a correct response was given to those on which an incorrect response was given. By statistically factoring out variability in fNIRS activation that could be explained by changing acoustic properties across stimulation conditions, we aimed to identify cortical regions that were sensitive to intelligibility *per se*. We found that activation was significantly greater in left inferior frontal and posterior superior temporal cortex (approximately corresponding to Broca's and Wernicke's areas, respectively) on correct versus incorrect trials (Fig. 7). Channels lying closer to unisensory auditory cortex did not show significant sensitivity to the veracity of trial-by-trial perception, suggesting that the apparent linear relationship with intelligibility in these channels may have been driven more by changing acoustic properties of the stimuli across conditions. In the present study, the primary acoustic difference across stimulation conditions was an increase in amplitude-modulation depth. Another recent fNIRS study similarly found that activation was stronger on correctly perceived trials in a speech-in-noise task, although the channels in question were more anteriorly located compared to those showing a comparable effect in the present study (Defenderfer et al., 2017). It remains unclear to what extent increased activation in Broca's/Wernicke's area on correct trials reflects greater success in extracting linguistic information from a degraded speech signal versus the deployment of greater attention/ effort on those trials. Potential cortical correlates of effortful listening are discussed further below.

4.2. Potential cortical correlates of effortful listening

Convergent evidence from studies employing subjective, behavioural and physiological measures indicates that listening to noise-vocoded speech can be a cognitively demanding task (Pals et al., 2013; Winn et al., 2015) The LIFG has been identified as one brain region potentially involved in effortful listening on the basis that activation in this region is: i) greater for degraded-yet-somewhat-intelligible speech than for either clear speech or unintelligible noise (Davis and Johnsrude, 2003; Wild et al., 2012); and ii) critically dependent on attention to speech (Wijayasiri et al., 2017; Wild et al., 2012). Although the aforementioned has been postulated from studies involving normally-hearing subjects, increased activation in the left frontal cortex of CI users in response to non-speech auditory stimuli has been previously interpreted as a marker of increased attention to and effort of listening (Jiwani et al., 2016). Such a non-monotonic, "inverted U"-shaped relationship

with task difficulty is also observed in other objective markers of listening effort, including pupil dilation (Zekveld et al., 2014). Reduced effort under very difficult listening conditions is thought to reflect "giving up" when the task is deemed impossible or, equally, when the cost of completing the task is judged too high relative to the anticipated reward (Pichora-Fuller et al., 2016).

In the present study, we found that a cluster of channels overlying the LIFG showed a consistent trend towards a quadratic relationship with intelligibility (Fig. 4c). Analysis of response amplitude in these channels at ROI level (Fig. 5c) confirmed that activation was greatest at intermediate levels of intelligibility (peaking in the S_{25} condition), and smaller for trials in which the task was either impossible (S_0 condition) or relatively easy (S_{100} condition). This pattern of results is consistent with a role of the LIFG in supporting speech comprehension under effortful conditions. However, the precise nature of the computations being performed in the LIFG is unclear. The LIFG houses both language-selective and domain-general "multiple demand" regions in close proximity (Fedorenko et al., 2012). Thus, it is unclear whether the elevated activation that we observed in the LIFG is directly language-related, or whether it might reflect increased demand on a general cognitive resource such as working memory (Rogalsky and Hickok, 2010). Interestingly, in a secondary analysis we found that the strength of activation in the LIFG was linearly related to mean response time on the simultaneously conducted behavioural task (Fig. 6). This suggests that the response in LIFG is likely to reflect decision-making, rather than sensory, processes (Binder et al., 2004).

Unexpectedly, we also observed strongly significant quadratic relationships with intelligibility in an array of bilateral channels overlying middle temporal and inferior parietal regions (Fig. 4c). However, unlike the LIFG, these regions exhibited a tendency towards deactivation (relative to the silent baseline condition), with the strength of deactivation being greatest at intermediate levels of intelligibility (Fig. 5d). Interestingly, Wild et al. (2012) reported a similar pattern of deactivation in the right angular gyrus based on fMRI data (their Fig. 7d), which was furthermore found to be critically dependent on attention to speech. This provides further support that the deactivation we observed in the present study using fNIRS is likely to have been both task-relevant and neural in origin. We suggest that this selective deactivation under (presumably) more effortful listening conditions likely reflects sensitivity of our measurements to the default mode network (DMN). The DMN describes a distributed system of interconnected brain regions that are preferentially more active during "rest" (i.e. when an individual is engaged in self-generated thought) than during engagement in an external task (Buckner et al., 2008). The DMN includes, amongst other regions, inferior parietal cortex and lateral temporal cortex (Andrews-Hanna et al., 2014), aligning well with the channels in which we observed selective deactivation under challenging listening conditions. Furthermore, the strength of deactivation within the DMN has been shown to correlate with task difficulty (Mckiernan et al., 2003). Our results therefore indicate that fNIRS may have sensitivity to components of the DMN in inferior parietal/ lateral temporal regions, and that the strength of deactivation in corresponding channels may provide a marker of the attentional demands of a challenging listening task.

4.3. Potential clinical applications

The primary aim of cochlear implantation is to optimise an individual's ability to discriminate speech. CIs require programming to best suit the auditory requirements of the recipient. This programming process takes months, requires multiple sessions, and at present depends on behavioural speech perception testing.

However, this method of assessing speech understanding can be unreliable, especially in younger children and infants. Our findings contribute to an emerging body of evidence suggesting that fNIRS has the potential to objectively assess speech intelligibility at a cortical level through the measurement of activation in superior temporal brain regions (Olds et al., 2016; Pollonini et al., 2014).

Whereas previous studies have used the overall spatial extent of activation in superior temporal cortex as a predictive measure (Olds et al., 2016; Pollonini et al., 2014), our findings emphasize how clinically relevant information might be gleaned by assessing response amplitude in more specific cortical regions. For instance, our results tentatively suggest that activation in channels overlying unisensory auditory cortex might provide an indication of the fidelity with which complex modulated signals are transmitted from ear to cortex, while activation in higher-order auditory regions (e.g. Wernicke's area) might indicate how successful an individual is in extracting meaningful linguistic information from those signals. While further research is needed to confirm the validity and practical utility of these findings, such an approach could allow for differential diagnoses which might guide subsequent intervention/rehabilitation strategies.

Our results additionally suggest that, beyond superior temporal cortex, fNIRS could provide a brain-based marker for effortful listening. Possible forms this marker could take are: i) elevated activation in left inferior frontal cortex associated with recovering meaning from degraded speech (Wijayasiri et al., 2017; Wild et al., 2012); and/or ii) deactivation of the DMN (specifically, inferior parietal/lateral temporal cortex), presumably reflecting the suppression of self-generated thought under more effortful/attention-demanding listening conditions. Thus, fNIRS could in future provide an objective measure to help clinicians achieve the optimal goal of maximizing intelligibility while at the same time minimizing the cognitive load of listening (Pichora-Fuller et al., 2016). We appreciate that in deaf individuals, areas of the cortex may exhibit significant functional differences to the normally-hearing brain such as auditory regions of the brain becoming responsive to visual stimuli, so called cross-modal plasticity (Anderson et al., 2017b). Hence, the results of this study can obviously not be directly applied to deaf patients with or without a CI. However, our data does support an important goal for future research, which is to establish whether these non-linear patterns of activation in higher-level brain areas, observed here in normally hearing adults, apply equally in the case of adult and paediatric CI recipients.

Declarations of interest

None.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2018.09.005.

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.

Anderson, C.A., Lazard, D.S., Hartley, D.E.H., 2017a. Plasticity in bilateral superior temporal cortex: effects of deafness and cochlear implantation on auditory and visual speech processing. Hear. Res. 343, 138—149.

Anderson, C.A., Wiggins, I.M., Kitterick, P.T., Hartley, D.E.H., 2017b. Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. Proc. Natl. Acad. Sci. U.S.A. 114, 10256—10261.

Andrews-Hanna, J.R., Smallwood, J., Spreng, R.N., 2014. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann. N. Y. Acad. Sci. 1316, 29—52.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309—312.

Bench, J., Kowal, A., Bamford, J., 1979. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br. J. Audiol. 13, 108—112.

Benjamini, Y., Hochberg, Y., 1995. Controlling the flase discovery rate: a practical and powerful approach to multiple testing. J.R. Stat. Ser. B Methodol. 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. Nat. Neurosci. 7 (3), 295—301.

Bisconti, S., Shulkin, M., Hu, X., Basura, G.J., Kileny, P.R., Kovelman, I., 2016. fNIRS brain imaging investigation of phonological awareness and passage comprehension abilities in adult recipients of Cochlear Implants. J. Speech Lang. Hear. Res. 59 (2), 239e253.

Blasi, A., Lloyd-Fox, S., Johnson, M.H., Elwell, C., 2014. Test-retest reliability of functional near infrared spectroscopy in infants. Neurophotonics 1 (2) e025005.

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.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis. 10, 433—436.

Büchel, C., Holmes, A.P., Rees, G., Friston, K.J., 1998. Characterizing stimulus—response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8 (2), 140—148.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. 1124, 1—38.

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 (1), 252—257.

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 Plast. 13, 2016.

Clarke, S., Hall, P., 2009. Robustness of multiple testing procedures against dependence. Ann. Stat. 37, 332—358.

Cochrane, D., Orcutt, G.H., 1949. Application of least squares regression to relationships containing auto-corrected error terms. J. Am. Stat. Assoc. 44 (245), 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 Trans. Med. Imag. 17, 463—468.

Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8 (2e3), 109e114.

Davis, M.H., Johnsrude, I.S., 2003. Hierarchical processing in spoken language comprehension. J. Neurosci. 23 (8), 3423—3431.

Defenderfer, J., Kerr-German, A., Hedrick, M., Buss, A.T., 2017. Investigating the role of temporal lobe activation in speech perception accuracy with normal hearing adults: an event-related fNIRS study. Neuropsychologia 106, 31—41.

Evans, S., Kyong, J.S., Rosen, S., Golestani, N., Warren, J.E., McGettigan, C., Mourão-Miranda, J., Wise, R.J.S., Scott, S.K., 2014. The pathways for intelligible speech: multivariate and univariate perspectives. Cerebr. Cortex 24 (9), 2350—2361.

Fedorenko, E., Duncan, J., Kanwisher, N., 2012. Language-selective and domain-general regions lie side by side within Broca's area. Curr. Biol. 22, 2059—2062.

Frost, J.A., Binder, J.R., Springer, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Language processing is strongly left lateralized in both sexes; evidence from functional MRI. Brain 122, 199—208.

Greenwood, D.D., 1990. A cochlear frequency-position function for several species-29 years later. J. Acoust. Soc. Am. 87 (6), 2592e2605.

Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event-related brain potentials/fields II: simulation studies. Psychophysiology 48, 1726—1737.

Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. Trends Cognit. Sci. 4 (4), 131—138.

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. Appl. Optic. 48, D280—D298.

Huppert, T.J., 2016. Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. Neurophotonics 3, 010401-10.

Jasper, H.H., 1958. The ten twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 10, 371—375.

Jiwani, S., Papsin, B.C., Gordon, K.A., 2016. Early unilateral cochlear implantation promotes mature cortical asymmetries in adolescents who are deaf. Hum. Brain Mapp. 37, 135—152.

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

Lawler, C.A., Wiggins, I.M., Dewey, R.S., Hartley, D.E.H., 2015. The use of functional near-infrared spectroscopy for measuring cortical reorganisation in cochlear implant users: a possible predictor of variable speech outcomes? Cochlear Implants Int. 16 (1), S30–S32.

Li, X., Liang, Z., Kleiner, M., Lu, Z.L., 2010. RTbox: a device for highly accurate response time measurements. Behav. Res. Meth. 42, 212–225.

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. Adv. Exp. Med. Biol. 894, 327e335.

McKiernan, K.A., Kaufman, J.N., Thompson, J.N., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J. Cognit. Neurosci. 15, 394–408.

Molavi, B., Dumont, G.A., 2012. Wavelet-based motion artefact removal for functional near-infrared spectroscopy. Physiol. Meas. 33, 259–270.

Mumford, J.A., Poline, J.B., Poldrack, R.A., 2015. Orthogonalization of regressors in FMRI models. PLoS One 28 (10). https://doi.org/10.1371/journal.pone.0126255. eCollection 2015 (4):e0126255.

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., 2016. Cortical activation patterns correlate with speech understanding after cochlear implantation. Ear Hear. 37 (3) e160ee172.

Pals, C., Sarampalis, A., Baskent, D., 2013. Listening effort with cochlear implant simulations. J. Speech Lang. Hear. Res. 56 (4), 1075e1084.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vis. 10, 437–442.

Pichora-Fuller, M.K., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W.Y., Humes, L.E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C.L., Naylor, G., Phillips, N.A., Richter, M., Rudner, M., Sommers, M.S., Tremblay, K.L., Wingfield, A., 2016. Hearing impairment and cognitive mind: the framework for understanding effortful listening (FUEL). Ear Hear. 37, 5S–27S.

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 (2), 625–634.

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. Hear. Res. 309, 84–93.

Rogalsky, C., Hickok, G., 2010. The role of Broca's area in sentence comprehension. J. Cognit. Neurosci. 23, 1664e1680.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends Neuroci 26 (2), 100–107.

Scott, S.K., Rosen, S., Lang, H., Wise, R.J., 2006. Neural correlates of intelligibility in speech investigated with noise vocoded speech–a positron emission tomography study. J. Acoust. Soc. Am. 120 (2), 1075–1083.

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. Hear. Res. 270, 39–47.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primary temporal cues. Science 270, 303–304.

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

Steffener, Jason, Tabert, Matthias, Reuben, Aaron, Stern, Yaakov, 2010. Investigating hemodynamic response variability at the group level using basis functions. Neuroimage 49 (3), 2113e2122.

Studebaker, G.A., 1985. A "rationalized" arcsine transform. J. Speech Hear. Res. 28 (3), 455–462, 1985.

van de Rijt, L.P., van Opstal, A.J., Mylanus, E.A., Straatman, L.V., Hu, H.Y., Snik, A.F., 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.

West, B.T., Welch, K.B., Galecki, A.T., 2006. Linear Mixed Models: a Practical Guide Using Statistical Software. Chapman & Hall/CRC, Taylor & Francis Group, Florida.

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

Wiggins, I.M., Hartley, D.E.H., 2015. A synchrony-dependent influence of sounds on activity in visual cortex measured using functional near-infrared spectroscopy (fNIRS). PLoS One 31 (10) (3):e0122862.

Wijayasiri, P., Hartley, D.E.H., Wiggins, I.M., 2017. Brain activity underlying the recovery of the meaning from degraded speech: a functional near-infrared spectroscopy (fNIRS) study. Hear. Res. 351, 55–97.

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. J. Neurosci. 32, 14010–14021.

Winn, M.B., Edwards, J.R., Litovsky, R.Y., 2015 Jul-Aug. The impact of auditory spectral resolution on listening effort revealed by pupil dilation. Ear Hear. 36 (4), e153ee165.

Yamada, T., Umeyama, S., Matsuda, K., 2012. Separation of fNIRS signals into functional and systemic components based on differences in hemodynamic modalites. PLoS One 7 (11) e50271.

Zekveld, A.A., Heslenfeld, D.J., Festen, J.M., Schoonhoven, R., 2006. Top-down and bottom-up processes in speech comprehension. Neuroimage 32 (4), 1826e1836.

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, 76e86.