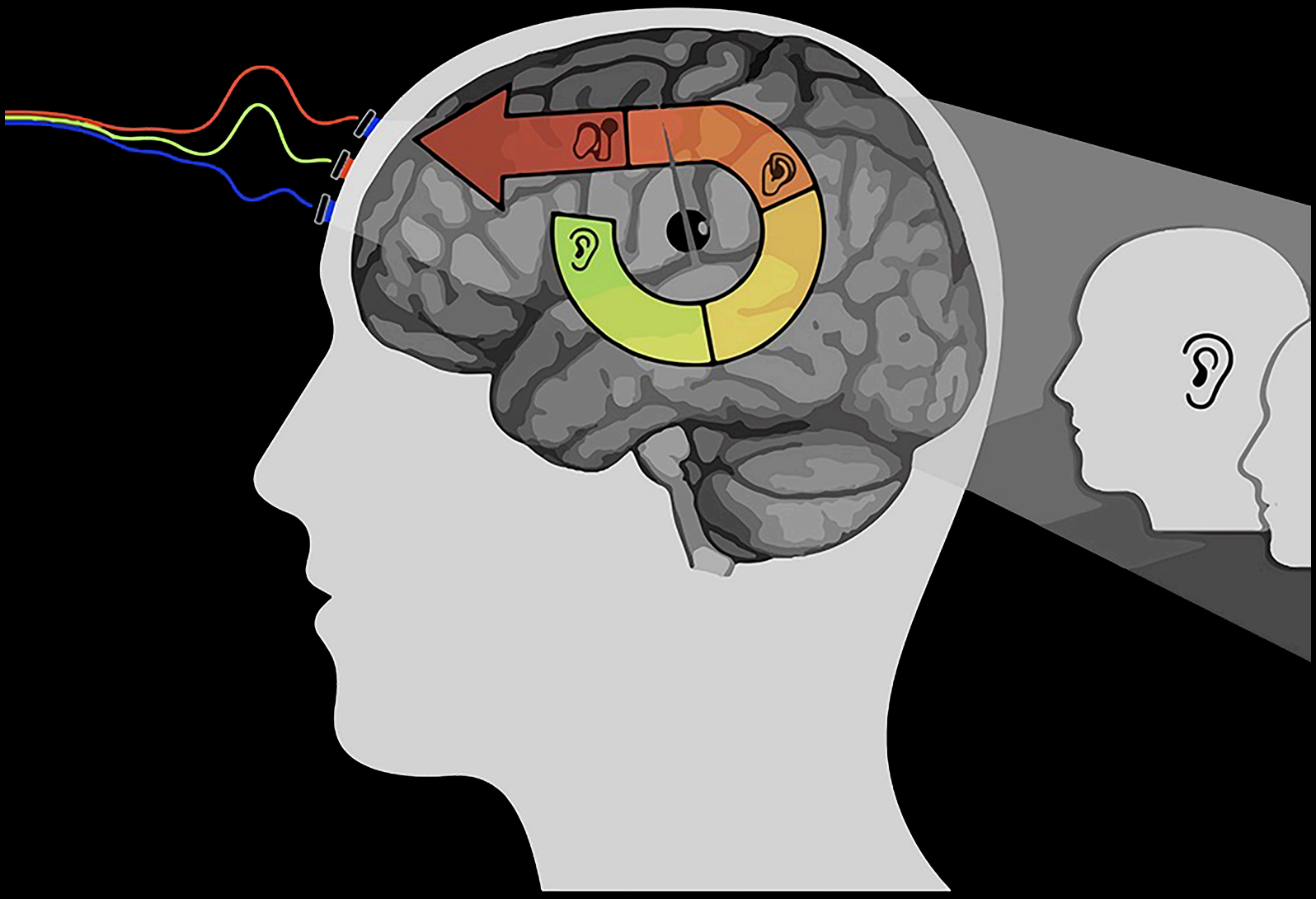
B. E. White
Front Cover, Vol. 238
Manuscript, Vol. 240, 118324

ELSEVIER

NeuroImage

Editor-in-Chief

Michael Breakspear



Available online at www.sciencedirect.com

ScienceDirect

NeuroImage 240 (2021) 118324

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

The cortical organization of listening effort: New insight from functional near-infrared spectroscopy

Bradley E. White [a,*], Clifton Langdon [b]

[a] Brain and Language Center for Neuroimaging, Gallaudet University, Washington, DC, USA
[b] Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA

ARTICLE INFO

Keywords:
Auditory cognition
Cortical organization
Listening effort
Attention
fNIRS
Hearing aids
Cochlear implants

ABSTRACT

Everyday challenges impact our ability to hear and comprehend spoken language with ease, such as accented speech (source factors), spectral degradation (transmission factors), complex or unfamiliar language use (message factors), and predictability (context factors). Auditory degradation and linguistic complexity in the brain and behavior have been well investigated, and several computational models have emerged. The work here provides a novel test of the hypotheses that listening effort is partially reliant on higher cognitive auditory attention and working memory mechanisms in the frontal lobe, and partially reliant on hierarchical linguistic computation in the brain's left hemisphere. We specifically hypothesize that these models are robust and can be applied in ecologically relevant and coarse-grain contexts that rigorously control for acoustic and linguistic listening challenges. Using functional near-infrared spectroscopy during an auditory plausibility judgment task, we show the hierarchical cortical organization for listening effort in the frontal and left temporal-parietal brain regions. In response to increasing levels of cognitive demand, we found (i) poorer comprehension, (ii) slower reaction times, (iii) increasing levels of perceived mental effort, (iv) increasing levels of brain activity in the prefrontal cortex, (v) hierarchical modulation of core language processing regions that reflect increasingly higher-order auditory-linguistic processing, and (vi) a correlation between participants' mental effort ratings and their performance on the task. Our results demonstrate that listening effort is partly reliant on higher cognitive auditory attention and working memory mechanisms in the frontal lobe and partly reliant on hierarchical linguistic computation in the brain's left hemisphere. Further, listening effort is driven by a voluntary, motivation-based attention system for which our results validate the use of a single-item post-task questionnaire for measuring perceived levels of mental effort and predicting listening performance. We anticipate our study to be a starting point for more sophisticated models of listening effort and even cognitive neuroplasticity in hearing aid and cochlear implant users.

1. Introduction

Everyday listening often occurs under a wide range of suboptimal and adverse listening conditions, such as competing background noise (e.g., the cocktail party effect), challenges from the spectral quality of the speech signal (e.g., listening through a telephone or through a hearing amplification or prosthetic device), accented speech, and complex or unfamiliar language use. These challenges impact our listening performance and the amount of cognitive energy we exert for comprehension (Blanco-Elorrieta et al., 2020; Mattys et al., 2012; Peelle, 2018; Van Engen and Peelle, 2014). The greater the challenge, the larger the demand on cognitive executive resources for information processing (within limits; Pichora-Fuller et al., 2016; Wingfield, 2016). Listening in the presence of auditory and linguistic challenges has been the focus of much research, and thus, several computational models have emerged. Here, we test novel hypotheses that listening effort is partially reliant on higher cognitive auditory attention and working memory mechanisms in the frontal lobe and partially reliant on hierarchical linguistic processing mechanisms in the left temporal-parietal cortices. We uniquely hypothesize that these models are robust and can be applied in ecologically relevant and coarse-grain contexts that rigorously control for acoustic and linguistic listening challenges. Therefore, as a unique design feature, the present study utilized simulated hearing aid (wide-bandwidth) and cochlear implant (narrow-bandwidth) speech signals that systematically varied in acoustic (spectral and temporal) and linguistic (syntactic) challenges which are experienced by actual hearing aid and cochlear implant users every day. Further, we investigate the use of a single-item post-task usability questionnaire for predicting listening performance, which, to

Abbreviations: fNIRS, Functional near-infrared spectroscopy; SO, Subject-relative clause; OS, object-relative clause; SMEQ, Subjective Mental Effort Question.

* Corresponding author.
E-mail address: bradley.white@gallaudet.edu (B.E. White).

https://doi.org/10.1016/j.neuroimage.2021.118324.
Received 10 March 2021; Received in revised form 17 June 2021; Accepted 28 June 2021
Available online 2 July 2021.

our knowledge, has never been done on a neurocognitive level. These findings serve as basis for further testing of these models with different populations that experience protracted effortful listening, such as hearing aid and cochlear implant users.

2. Background

2.1. Core speech and language processing

The neural networks involved in spoken language comprehension have been largely investigated by contrasting intelligible and unintelligible speech (Mattys et al., 2012). Functional neuroimaging studies consistently find evidence that intelligible sentences are processed by the bilateral temporal cortices, with occasional recruitment of the inferior frontal gyrus (IFG; Crinion et al., 2003; Davis et al., 2011; Davis and Johnsrude, 2003; Evans et al., 2016; Hassanpour et al., 2015; McGettigan et al., 2012; Obleser et al., 2007; Okada et al., 2010; Peelle et al., 2010; Rodd et al., 2005; Rodd et al., 2010). These regions come together to form pathways and function as hierarchical networks for information transfer and processing. Regions closer in proximity to the auditory cortex show increased sensitivity to the acoustic features of sound (Davis and Johnsrude, 2003; Friederici, 2011; Friederici and Gierhan, 2013; Peelle, 2019; Rauschecker, 2012; Rauschecker and Scott, 2009; Rauschecker and Tian, 2000; White, 2019), whereas higher level processing of language (e.g., syntax) has greater representation in posterior temporal cortices (e.g., perisylvian association cortex) and frontal lobe (particularly the bilateral IFG and superior temporal gyrus, STG; Caplan, 2007; Just et al., 1996).

Rapid processing of fine-grained acoustic features is constructed into larger units spanning phonemic segments (e.g., suprasegmental tone), and constructed into words and then sentences. Processing of the language signal does not solely proceed in this simplistic "bottom-up" sequence. Fine-grained processes of the speech signal also receive "top-down" predictive information. The network of language-related processing is posited to be mainly represented by a ventral and a dorsal flow. The ventral pathways in both human and non-human primates function in the decoding of complex sounds, such as speech sounds and their connection to meaning (Hickok and Poeppel, 2004; Rauschecker, 2012). The dorsal pathways, on the other hand, aid in speech production, motor planning, and some online processing of spoken language (Hickok and Poeppel, 2004; Rauschecker, 2012). However, spoken language constantly relies on both ventral and dorsal stream processing, often alternating or working harmoniously (partially or wholly) to perceive and produce speech and language (Indefrey and Levelt, 2004).

2.1.1. Syntactic processing

In behavioral research studies contrasting clause structures, slower reaction times and reduced accuracy were observed for object-relative (OS; e.g., *houses that men build are strong*) compared to subject-relative (SO; e.g., *men that build houses are strong*) clause structures when adults either read (Just and Carpenter, 1992; Vos et al., 2001) or listened to (Wingfield et al., 2003) sentences and were asked to make decisions about them. In a series of functional positron emission tomography (PET) and magnetic resonance imaging (MRI) studies, the left IFG has been identified as playing a key role in the processing of syntactically complex (OS) sentences compared to syntactically simple (SO) sentences (Alain et al., 2018; Bilenko et al., 2008; Caplan, 2001, 2008; Caplan et al., 1998, 2000, 2003; Caplan et al., 2008; Friederici et al., 2006; Stromswold et al., 1996). Although the left IFG appears to be the primary processing area for syntactic complexity, Caplan et al. (2002) and Stromswold et al. (1996) observe greater regional cerebral blood flow in the left perisylvian language regions when individuals were asked to decide whether sentences were possible compared to when individuals had to decide whether syntactically identical sentences contained a nonsense word. These scientific findings suggest that the left IFG, which is recruited for processing syntactically complex sentences, plays a role in mediating greater memory loads associated with processing these kinds of sentence structures (Caplan et al., 2002; Stromswold et al., 1996). The perisylvian association cortex, which is recruited for general sentence processing, is suggested to relate to language-general processes.

2.1.2. Information processing and listening effort

As the acoustic environment becomes more challenging, it becomes increasingly more difficult to perceive the target signal, and thus, the neural networks for listening, language, and attention processes are modulated. Specifically, listeners rely on cognitive systems to a greater extent for successful extraction of meaning from distorted speech signals. These neural network changes are posited to reflect manipulations of mechanisms that "(i) establish and maintain behavioral goals and listening to targets, (ii) prioritize target signals (i.e., sensory gain), and (iii) suppress interfering sound sources and distracting information" (Obleser and Erb, 2020, p. 2). Distinct types of listening challenges result in distinct neural network patterns.

In a classic study, Rabbitt (1968) found that understanding speech in adverse listening conditions interfered with other cognitive processes, such as those used for short-term verbal working memory encoding. In subsequent studies with hearing adults, accuracy in speech sound identification has been inversely related to the amount of background noise present (i.e., accuracy increases as noise decreases; Gosselin and Gagné, 2011; Zekveld et al., 2011). Similarly, comprehension decreases with decreasing amounts of spectral information in studies manipulating the spectral features or quality of speech sounds (i.e., noise-vocoded speech; Shannon et al. 1995). Even when speech is understood, Heinrich et al. (2008) and Surprenant (1999) observed that acoustically degraded words and syllables are more difficult to remember. In the presence of adverse listening conditions, DeCaro, Peelle, Grossman, and Wingfield (2016) and Wingfield et al. (2006) found that sentence processing is also impacted.

Listening under adverse conditions modulates a widely distributed network with major contributions from the prefrontal cortex, suggesting that listening effort relies in part on cognitive control, attention, and working memory processes (Du et al., 2016; Erb and Obleser, 2013; Love et al., 2006; Mattys et al., 2012; Peelle, 2018, 2019; Peelle et al., 2009; Rodd et al., 2010; Vaden et al., 2015; Zekveld et al., 2012). Functional MRI and PET studies investigating the effect of listening to spectrally degraded speech (e.g., noise vocoded speech) observed increased activation in the left IFG, bilateral STG, and bilateral middle temporal gyrus (MTG; Davis and Johnsrude, 2003; Evans and Davis, 2015; Meyer et al., 2004; Scott et al., 2006; Vaden et al., 2011; Wild et al., 2012). These findings suggest that accurate processing of challenging speech relies on a widely distributed language processing network in the brain.

The orbitofrontal cortex, specifically the right middle frontal gyrus (MFG), has also been implicated in mediating auditory attentional, memory, and decision-making processes (Euston et al., 2012; Vigneau et al., 2011). In a meta-analysis of right-hemisphere contribution to linguistic processing, Vigneau et al. (2011) observed fMRI activation clusters in the right anterior MFG that are involved with selective auditory attention processing. The right MFG has been implicated in mediating auditory attention processes for both linguistic and non-linguistic tasks (e.g., Bavelier et al., 2000; Jäncke and Shah, 2002; Lipschutz et al., 2002; Poldrack et al., 1999; Vigneau et al., 2011). A large body of neurocognitive research studies have also suggested that the orbital right MFG and adjacent brain areas supplement the mediation of hard, but not easy, working and episodic memory processes, including online monitoring of information and the ability to temporarily access, reason with, and manipulate or make decisions about stored information (Aron et al., 2004; Frey and Petrides, 2002; Henson et al., 1999; Lipschutz et al., 2002; Nyberg et al., 1995; Okuda et al., 1998; Olesen et al., 2004; Petrides et al., 2002; Rypma et al., 1999).

2.2. Listening with hearing aids and cochlear implants

Hearing aids uniquely compress the speech signal and limit spectral range (wide-bandwidth), and cochlear implants transpose and reduce the amount of spectral information (narrow-bandwidth). Both types of signal degradation impact behavioral performance and cognitive processing for speech comprehension. These listening challenges are the result of a bottom-up or signal-driven processing failure when attempting to construct lexical units from mapping acoustic-phonetic features to segmental representations. The consequences of these failures result in uncertainty (e.g., weak lexical activation of a larger lexical neighborhood of a single lexical item). These challenges, to some extent, are ameliorated by top-down or knowledge-driven processing demands that rely on contextual information for lexical selection. Taken together, listeners with reduced hearing acuity experience a systems-level challenge that persistently requires the allocation of executive cognitive resources for speech understanding (Peelle, 2018). Thus, by using these signal types (wide-band hearing aid and narrow-band cochlear implant), our findings can inform and advance scientific inquiry about cortical speech processing and cognitive neuroplasticity in hearing aid and cochlear implant users using fNIRS brain imaging.

2.3. Hypotheses and predictions

Auditory degradation and linguistic complexity in the brain and behavior have been well investigated. The work here provides a novel test of the hypotheses that listening effort is partially reliant on higher cognitive auditory attention and working memory mechanisms in the frontal lobe (Pichora-Fuller et al., 2016; Rabbitt, 1968; Rönnberg et al., 2008, 2013), and partially reliant on hierarchical linguistic computation in the brain's left hemisphere (Friederici and Gierhan, 2013; Hickok and Poeppel, 2007; Peelle, 2018; Rauschecker and Scott, 2009). Here, we specifically hypothesize that these models are robust and can be applied in ecologically relevant and coarse-grain contexts developed here that rigorously control for acoustic and linguistic listening challenges. To our knowledge, this is the first investigation of the computational organization of sustained and motivated speech comprehension in hierarchically challenging conditions. Notably, to our knowledge, this is also the first investigation of the relationship between the brain, task performance, and internal perceptions of mental workload. We predict that neural tissues in the antero-ventral information processing stream will be most active during everyday listening with low cognitive demand, whereas tissues in the postero-dorsal stream will be most active during difficult listening with high cognitive demand. In response to increasing levels of cognitive demand, we predict (i) poorer comprehension, (ii) slower reaction times, (iii) increasing levels of perceived mental effort, (iv) increasing levels of brain activity in the prefrontal cortex, (v) hierarchical modulation of core language processing regions that reflect increasingly higher-order auditory-linguistic processing, and (vi) recruitment of right hemisphere auditory and homologous language regions.

3. Experimental design

3.1. Participants

Behavioral and optical brain imaging data were analyzed from 29 young, right-handed, healthy, monolingual adults with clinically defined typical hearing, characterized by audiometric three-frequency pure-tone averages ≤ 25 dB HL at 500 Hz, 1 and 2 kHz (Hall and Mueller, 1997). All participants demonstrated excellent speech recognition abilities and were native speakers of Standard American English with little to no experience with a second language. Additionally, all participants were matched based on psychometric measures of nonverbal intelligence and English language understanding (see Table 1 for participant demographics). Participants in this study comprise a subset of data from a larger research project with multiple groups (White, 2019; White and Langdon, 2019).

Table 1
Participant demographics.

Demographic	*M* (*SD*)
N	29
Sex	18 F, 11 M
Age	30.88 (6.12)
Speech Recognition[a]	99.66 (1.20)
Physical Health[b]	54.85 (7.09)
Mental Health[b]	53.18 (9.66)
Nonverbal Intelligence[c]	104.4 (13.88)
English Comprehension[d]	96.93 (13.33)

[a] CID Everyday Sentences Test.
[b] PROMIS Global Health Short Form.
[c] Kaufman Brief Intelligence Test II: Matrices.
[d] Woodcock-Johnson IV Tests of Achievement: Passage Comprehension.

Our rigorous screening and assessment battery consisted of six components: (i) screening of audiometric pure-tone thresholds, (ii) the CID Everyday Sentences Test (Davis and Silverman, 1970; Silverman and Hirsh, 1955), (iii) the PROMIS Global Health Short Form (Hays et al., 2009), (iv) the Language Background and Use Questionnaire (e.g., Kovelman et al., 2008), (v) the Kaufman Brief Intelligence Test II: Matrices (Kaufman and Kaufman, 2004), and (vi) the Woodcock-Johnson IV Test of Achievement: Passage Comprehension (Schrank et al., 2014).

3.2. Materials

Participants were presented with a classic Plausibility Judgment Task (for a review, see Caplan, 2007). We used a total of 300 unique sentences (12 practice, and 288 experimental). All sentences were constructed to vary in four features:

(i) Semantic plausibility: plausible and implausible (e.g., *people feeding birds* and *birds feeding people*).
(ii) Syntactic complexity: subject-relative and object-relative clauses (Just et al., 1996).
(iii) Acoustic clarity: undegraded, hearing aid simulation, and cochlear implant simulation.
(iv) Speech rate: typical (205 wpm) and fast (273–410 wpm).

The semantic plausibility feature presents participants with a basic linguistic task to attend to. The other features allow us to construct stimuli that are increasingly difficult to perceive and comprehend. For example, we expect that a subject-relative sentence presented at a typical speech rate and without degradation would be the easiest listening condition, whereas an object-relative sentence presented at a fast speech rate and with a cochlear implant simulation would be the most difficult.

3.2.1. Syntactic structure

The stimuli were constructed from 72 plausible and 72 implausible, syntactically appropriate, subject-relative (SO) clause structures. For each of these sentences, a counterpart sentence was constructed that had the same content words and characters performing the action as the original, but with an object-relative (OS) clause structure. All sentences were matched for the number of total syllables (ranging from seven to nine syllables), length (six words), and subject and object animacy (i.e., human, animal, object). Sentences containing a human subject or object were balanced for male and female nouns. This procedure resulted in a total of 288 sentences, 72 of each of the following types:

(i) Subject-relative clause, plausible: *Women that feed birds are nice.*
(ii) Object-relative clause, plausible: *Birds that women feed are nice.*
(iii) Subject-relative clause, implausible: *Baseballs that eat keys are nasty.*
(iv) Object-relative, implausible: *Keys that baseballs eat are nasty.*

All the sentences were recorded by a female native speaker of Standard American English at an average speech rate of 205 words per minute (wpm) and at a sampling rate of 44.1 kHz in an acoustically treated, professional audio recording studio (e.g., Fig. A2.a). Prior to manipulating the sentences to create the experimental stimuli, all sentences were trimmed to remove silence at the beginning and end of the recording.

3.2.2. Hearing aid simulation

We created a hearing aid simulation from each of the 288 sentences using the Hearing Loss and Prosthesis Simulator (HeLPS) v2 (Sensimetrics Corp., Malden, MA). To do this, we programmed a simulated three-band digital omnidirectional HA with automatic gain and compression (i.e., automatically selected by the HeLPS v2 software) to an audiogram reflecting a symmetric, flat 65 dB HL hearing loss at all frequencies (moderately-severe hearing loss; see Fig. A1). This procedure resulted in a compressed-speech signal with limited spectral range, resembling speech through a hearing aid (see Fig. A2.b.; Zurek and Desloge, 2007).

3.2.3. Cochlear implant simulation

We also created a cochlear implant simulation from each of the 288 sentences using MATLAB (MathWorks, Natick, MA). To do this, we applied an eight-channel noise-vocoding algorithm with logarithmic spacing between frequency bands to each of the audio files (Peelle, 2016; Shannon et al., 1995). This noise-vocoding procedure reduces the amount of spectral information and speech clarity while simultaneously preserving the temporal information in the speech envelope (see Fig. A2.c). We chose eight-channel noise-vocoded speech because it is significantly more intelligible than lower numbers of channels and comparably as intelligible as higher numbers of channels (Chen and Loizou, 2011; Fishman et al., 1997; Friesen et al., 2001; Obleser et al., 2007; Scott et al., 2006). Therefore, eight-channel noise-vocoded speech is challenging, but it is not so challenging that it makes speech completely unintelligible or reduces the participants' motivation to listen; it is an ideal condition to investigate listening effort.

3.2.4. Tempo alterations

For each of the 288 sentences across all three listening conditions (undegraded, hearing aid simulation, and cochlear implant simulation), we created a version of the sentence that was faster than the original. For the undegraded speech condition, we created a version that was 50% of the original recording time (i.e., twice as fast, or 410 wpm). For the hearing aid and cochlear implant simulations, we created a version that was 66% of the originally recorded time (i.e., one-third as fast, or 273 wpm). In-lab testing of speeded acoustically-degraded stimuli suggested that hearing aid and cochlear implant simulations were too difficult to process at 50% of the original time, thus reducing motivation to try and perform well on the task. However, degraded signals at 66% of the original time did not exhibit this effect. All tempo alterations were made using Audacity (Audacity Team, Boston, MA) and the software's built-in "Change Tempo" effect with high quality stretching. Audacity uses sub-band sinusoidal modeling synthesis and pitch scaling to speed-up or slow-down audio files while simultaneously preserving the pitch of the audio recording (Audacity Team, 2018).

3.2.5. Equalization and calibration

In addition to the sentence stimuli, our speech model recorded 12 unique practice sentences and a version of "The Rainbow Passage" (Fairbanks, 1960, pp. 124–139). All 301 recordings were equalized based on the average root mean square, and the passage was then used for establishing the participants' most comfortable listening levels and for calibrating the volume for the entire experiment.

3.3. Plausibility judgment task

Participants were presented with a total of 300 audio-recorded sentences, 12 unique practice sentences and 288 experimental sentences (see Fig. A3). Experimental trials were divided into three runs of 96 sentences each: (i) undegraded, (ii) hearing aid simulated, and (iii) cochlear implant simulated speech. Each run contained four blocks: (i) 24 subject-relative sentences presented at a typical rate (205 wpm), (ii) 24 object-relative sentences presented at a typical rate (205 wpm), (iii) 24 subject-relative sentences presented at a fast rate (273–410 wpm), and (iv) 24 object-relative sentences presented at a fast rate (273–410 wpm). Trials in blocks were pseudorandomized. Blocks across all 3 runs were randomized. The task lasted approximately 25 min, and no sentences were repeated.

Participants were asked to carefully listen to sentences at their most comfortable listening level and make judgments about their plausibility or implausibility by pressing a button on a Serial Response Box (SR Box; Psychology Software Tools, Sharpsburg, PA). The leftmost button on the SR Box was green and labeled with a black "Y", and the button to the right of the green button was red and labeled with a black "N". Participants were asked to indicate their responses by pressing the green button for "Yes" (i.e., plausible) or a red button for "No" (i.e., implausible). After each block, participants were administered the Subjective Mental Effort Question (SMEQ; see 3.4.).

3.4. Subjective mental effort question

The Subjective Mental Effort Question (SMEQ) is a single-item, post-task questionnaire for measuring perceived cognitive load (Sauro and Dumas, 2009; Zijlstra, 1993; Zijlstra and van Doorn, 1985). The SMEQ consists of a vertical number line with "0″ labeled at the bottom-most end and "150″ labeled at the top-most end. Nine descriptive, scaled labels are marked on the line. These labels were chosen based on psychometrically calibrating them against tasks (Zijlstra, 1993; Zijlstra and van Doorn, 1985). Although the original scale was developed in Dutch as a paper-based task, it has since been translated into English and converted into a computer-based rating scale (Sauro and Dumas, 2009). Previous studies have found the SMEQ to be a reliable and easy to use tool for measuring perceived cognitive load (Sauro and Dumas, 2009; Zijlstra, 1993; Zijlstra and van Doorn, 1985). We administered a computer-based version of the SMEQ. Participants were prompted with "On average, this task was: (select any point on the line)" and instructed to use a computer mouse to indicate their response. The mouse cursor appeared in the top right of the screen only after the participant had moved the mouse to reduce priming of responses based on initial cursor location. The mouse cursor disappeared after a response was recorded. Participant responses were rounded to the nearest integer.

3.5. Functional NIRS brain imaging

Functional near-infrared spectroscopy (fNIRS) is a non-invasive optical brain imaging technique that uses red to near-infrared light (650–900 nm) to measure changes in the optical properties of cortical tissues, specifically hemoglobin oxygen saturation and total hemoglobin volume (Ferrari and Quaresima, 2012; Scholkmann et al., 2014; Wolf et al., 2007). A series of light-emitting sources and light-measuring detectors (collectively referred to as optodes) are arranged into a probe array and placed on the surface of the scalp in strict adherence to the international 10–20 system (Klem et al., 1999). Positioning the optodes in this way allows the fNIRS system to measure changes in the absorption of light as it passes along a diffuse path between adjacent source and detector optodes, thus inferences about underlying brain activity in different brain regions can be made based on the localized hemodynamic response (Ferrari and Quaresima, 2012). Furthermore, fNIRS has been shown to be especially well suited for language and cognition research (Quaresima et al., 2012; Rossi et al., 2012), especially in people who use listening devices (Basura et al., 2018). One primary benefit of using fNIRS brain imaging is that the equipment is virtually silent, allowing for a truer quiet listening baseline and less confounding equipment sound during the experiment.

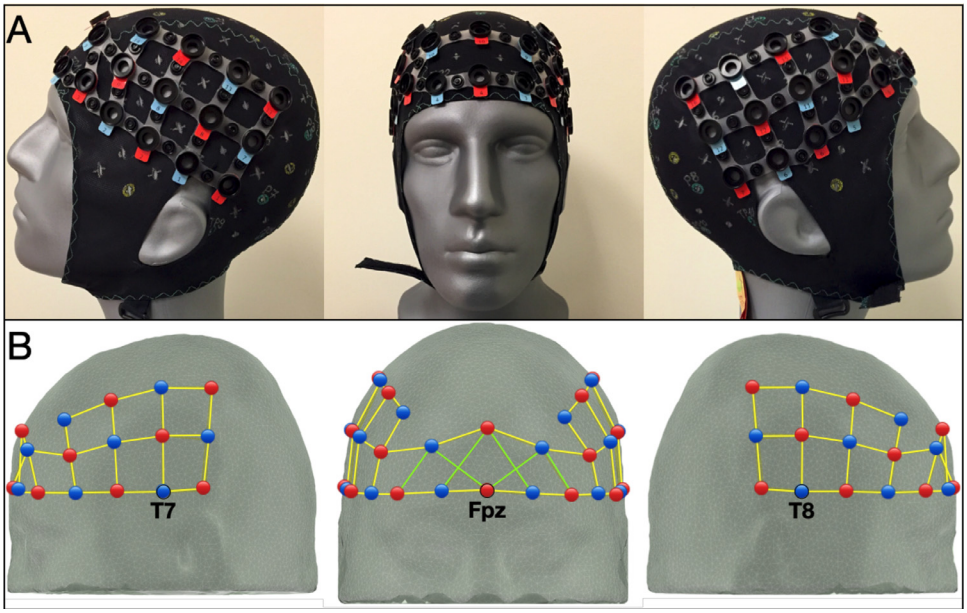


Fig. 1. Images depicting the fNIRS probe: (a) situated on the scalp in adherence with the international 10–20 system and (b) mapped in standard sensor space using AtlasViewer (Aasted et al., 2015). Source optodes are indicated by red circles, detector optodes are indicated by blue circles, and channels are indicated by yellow (3 cm) and green (3–4 cm) lines between adjacent source and detector optodes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We used a continuous-wave NIRScout fNIRS system (NIRx Medizintechnik GmbH, Berlin, Germany). A total of 16 LED light-emitting sources and 16 avalanche photodiode light-measuring detectors were arranged 3 cm apart (3–4 cm apart in the frontal cortex) to collect data from 50 measurement channels at 7.812 Hz (referred to as the probe; see Fig. 1). The light sources emitted near-infrared light at two factory-default wavelengths, 760 and 850 nm. Oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (HbR) have distinct absorption properties; thus, it is possible to determine concentration changes in HbO_2 and HbR by using these two separate wavelengths. With each measurement, a ratio of light emitted from the source optode to light received by the detector optode for each wavelength is recorded. These raw signal intensity ratios give insight into the probable absorption of light by the tissue (optical density) and are related to HbO_2 and HbR concentration changes (Jacques, 2013).

4. Procedures

4.1. Experimental data collection

This study was approved by and carried out in accordance with the policies and procedures of the Institutional Review Board at Gallaudet University. After obtaining written informed consent, participants were administered the first four components of the screening and assessment battery (see 3.1.): (i) screening of audiometric pure-tone thresholds, (ii) the CID Everyday Sentences Test, (iii) the PROMIS Global Health Short Form, (iv) the Language Background and Use Questionnaire. Only a portion of the assessment battery was administered prior to the experiment to control for fatigue. The pure-tone hearing screening and the CID Everyday Sentences Test were administered before the experiment to ensure that participants met the inclusion criteria and were able to complete the Plausibility Judgment Task. The PROMIS Global Health Scale was administered before the experiment so that participants were more likely to respond to items about their overall life and health without factoring in their participation in the research study. Finally, the Language Background and Use Questionnaire was administered before the experiment, to keep the participant occupied while the researchers set up the fNIRS apparatus. Unless noted otherwise, all screening and assessment battery components were administered according to their standard administration protocols (White, 2019).

Participants were seated in a stationary chair facing a 68.58 cm monitor with speakers positioned on both sides. The participants' heads were positioned approximately 65 cm away from the monitor. The volume for the experiment was adjusted by the participant to match their most comfortable listening level using a calibrated recording of "The Rainbow Passage," which had been previously equalized to match the intensity of the experimental sentences.

An fNIRS EasyCap (Brain Products GmbH, Gilching, Germany) with the probe configuration was then placed on the participants' scalps strictly according to the international 10–20 system (See Fig. 1; Klem et al., 1999) and according to standard fNIRS protocols (Shalinsky et al., 2009). The light sources and detectors were held in the cap using spring-loaded grommets, and a black cloth was placed over the probe to mitigate interference from ambient light sources, including the experimental monitor. Prior to recording, the room's overhead lights were turned off and the signal quality of every channel was assessed for optimal connectivity (i.e., signal-to-noise ratio) using the NIRx NIRStar native computer software for data acquisition. The experiment was administered and controlled using a Python-based stimulus presentation software, PsychoPy v1.85.6 (Peirce, 2009). Instructions for the Plausibility Judgment Task and SMEQ were given verbally and in writing on the experiment monitor. The neuroimaging task lasted approximately 25–30 min. Once finished, the researcher removed the fNIRS EasyCap and light sources and detectors.

Finally, the researcher administered the last two components of the screening and assessment battery (see 3.1.): (v) the Kaufman Brief Intelligence Test II: Matrices, and (vi) the Woodcock-Johnson IV Test of Achievement: Passage Comprehension. The administration order of these assessments was counterbalanced across all participants but was always administered after the experiment had completed. This was done to limit the number of activities before the experiment, which may tire or fatigue the participant before the experiment begins. All participants were compensated for their time at the rate of $20.00 per hour.

All original behavioral and fNIRS brain imaging data are part of a larger study and dataset. All data presented in this paper are available as supplemental documents. All original source code for behavioral and brain-behavioral analyses are available in the appendices. Functional NIRS brain imaging data were converted, analyzed, and verified using the NIRS Brain AnalyzIR Toolbox (April 26, 2019) with MATLAB R2017b: https://github.com/huppertt/nirs-toolbox/tree/1c1d70b52f261ad7353d7f5bd893f4011b40cbb2.

4.2. Data analyses

4.2.1. Behavioral analyses

We performed restricted maximum likelihood (ReML) linear mixed-effects statistical modeling (Bates et al., 2015) with Welch-Satterthwaite approximation Kenward and Roger (1997) in R (R Core Team, 2017). Speech clarity, speed, and syntax were modeled as fixed effects to predict participants' average accuracy, reaction times, and SMEQ ratings, where each participant was treated as a random effect. We used a full factorial combination model to demonstrate the impact of each hierarchical level of listening challenge (see Appendix A for Models and Appendix B for Behavioral Analysis R Source Code).

4.2.2. Functional NIRS source localization

Detector optodes one and six and source optode four were anchored at T7, Fpz, and T8 international 10–20 scalp positions, respectively (see Fig. 1). We virtually registered the probe in standard sensor space using AtlasViewer (Aasted et al., 2015). Channel locations were identified by calculating the centroid between adjacent source and detector optode pairs using custom MATLAB (Mathworks, Natick, MA) scripts. The estimated positions of the channels were projected from the scalp to the brain and registered to the Montreal Neurological Institute (MNI) probabilistic reference system (Mazziotta et al., 2001; Okamoto et al., 2004; Okamoto and Dan, 2005; Tsuzuki et al., 2012). Brain cytoarchitectural labels were assigned to a spherical region around each projected point with a radius of 5 mm using Automated Anatomical Labeling (AAL; Tzourio-Mazoyer et al., 2002). Probabilistic registration of fNIRS probes to standard coordinate systems is essential for comparing channel-wise results to other fNIRS and fMRI studies (see Appendix C for fNIRS Source Localization Results).

4.2.3. Functional NIRS channel-wise analysis

Functional NIRS data were analyzed using the NIRS Brain AnalyzIR Toolbox (Santosa et al., 2018). Raw fNIRS signal intensity was first converted to optical density and then to oxygenated and deoxygenated hemoglobin concentration using the Modified Beer-Lambert Law (Jacques, 2013; Santosa et al., 2018). We employed a novel analysis approach that renders the minimum number of manipulations to the data possible (Santosa et al., 2018, p. 29). No channels were pruned, and we did not attempt to remove or pre-process physiological or motion artifacts prior to analysis. Rather, we used statistical models that are more robust to the effects of these artifacts and are better at controlling false-positive rates. At the subject-level, we used general linear regression modeling (GLM) with an autoregressive iterative reweighted least squares (AR-IRLS) pre-whitening approach (Barker et al., 2013). We used a canonical hemodynamic response function model (double gamma function) with default parameters (peak, 4 s; dispersion times constants, 1 s^{-1}; Santosa et al., 2018). At the group-level, we used linear mixed-effects statistical modeling to calculate the group mean for each condition, where each participant was treated as a random effect (see Appendix A for Models). The channel-wise estimated regression coefficients (i.e., the absolute estimated "strength" of hemodynamic activity, β) for each condition were calculated and compared using Student's *t*-statistic estimates. These statistical tests allow us to make inferences about how brain areas respond to different listening conditions. We performed six comparisons to demonstrate brain activity during increasing levels of cognitive demand:

(A) Undegraded, OS (205 wpm) > SO (205 wpm).
(B) Undegraded, OS (410 wpm) > SO (205 wpm).
(C) HA Simulation, OS (205 wpm) > SO (205 wpm).
(D) HA Simulation, OS (273 wpm) > SO (205 wpm).
(E) CI Simulation, OS (205 wpm) > SO (205 wpm).
(F) CI Simulation, OS (273 wpm) > SO (205 wpm).

We performed the easiest comparison (A) as a control and in order to verify brain activation reported in previous fMRI studies. The exploratory comparisons (B–F) increased incrementally in listening challenge in order to test the hypotheses. Multiple comparisons and false discovery rates were controlled for using the Benjamini-Hochberg correction (*pFDR*). Due to the interdependent nature of HbO_2 and HbR, we performed statistical modeling with both hemoglobin chromophores. HbO_2 results are available in full in the supplemental documents. However, HbO_2 has been shown to be more sensitive to global components and physiological noise than HbR (Kirilina et al., 2012; Strangman et al., 2002; Tachtsidis and Scholkmann, 2016; Zhang et al., 2016). Additionally, fNIRS HbR measurements most closely correlate with and resemble the blood oxygen level dependent (BOLD) measurements acquired by fMRI and are thus reported here (Huppert et al., 2006; Ogawa et al., 1990; Sato et al., 2013). fNIRS studies have found discrepancies between HbO_2 and HbR measurements and further support reporting HbR for speech and listening studies (Hirsch et al., 2018, 2020). Combined with fNIRS Source Localization and a control comparison (A), this allows us to most reliably make comparisons to the fMRI BOLD literature we reviewed.

4.2.4. Brain-behavior correlation analysis

Lastly, we used Pearson's Product Moment Correlation to test the relationship (i.e., correlation coefficient, *r*) between brain activity and the behavioral response variables (i.e., accuracy, reaction time, and SMEQ rating). We computed subjects' brain activity using a *post hoc* cerebral region of interest, namely the orbital right MFG where we identified a pattern of metabolic expenditure indicative of effort in the previous channel-wise analysis (see Appendix B for Brain-Behavior Correlation Analysis R Source Code). Subject-level unweighted HbR beta values were first extracted from the NIRS Brain AnalyzIR Toolbox in MATLAB, then outliers were removed before correlating the values with the subject-level unweighted response variables in R. We were particularly interested in relationships with the SMEQ Rating Scale, which could provide important insight into the psychophysiological nature of self-reported effort measures (McGarrigle et al., 2014).

5. Results

5.1. Behavioral results

We performed ReML linear mixed effects statistical modeling (see 4.2.1.). Speech clarity and speed significantly predicted task accuracy, reaction time, and SMEQ Rating (see Table 2 and Fig. 2). Additionally, there were 2-way (i.e., $Clarity^{CISim}$:$Syntax^{OS}$) and 3-way (i.e., $Clarity^{CISim}$:$Speed^{Fast}$:$Syntax^{OS}$) interactions that significantly predicted task accuracy.

5.2. Functional NIRS channel-wise results

We performed channel-wise analysis of fNIRS HbR brain imaging data in MATLAB (see 4.2.3.). Specifically, we made six comparisons to demonstrate brain activity during increasing levels of cognitive demand (see Fig. 3). We observed significant activation in the left IFG (BA44) and right STG (BA22) for the classic contrast A (OS > SO; Caplan, 2007; Just et al., 1996). We also observed increasing levels of activation in the orbital right MFG from contrasts A–E, with a decrease in activation for the most challenging conditions in contrast F (see Fig. 4.a). Additionally, we observed a pattern of activation in the left hemisphere that follows the ventral and then dorsal auditory processing streams as the level of cognitive demand increases (see Fig. 4.b). This pattern starts with significant recruitment of the left IFG (BA44) in contrasts A–B, then the left supramarginal gyrus (SMG, BA40) in contrast C, the left postcentral gyrus (BA1,2,3) in contrast D, the left MFG (BA9) in contrast E, and finally the orbital right MFG (BA10) in contrast F.

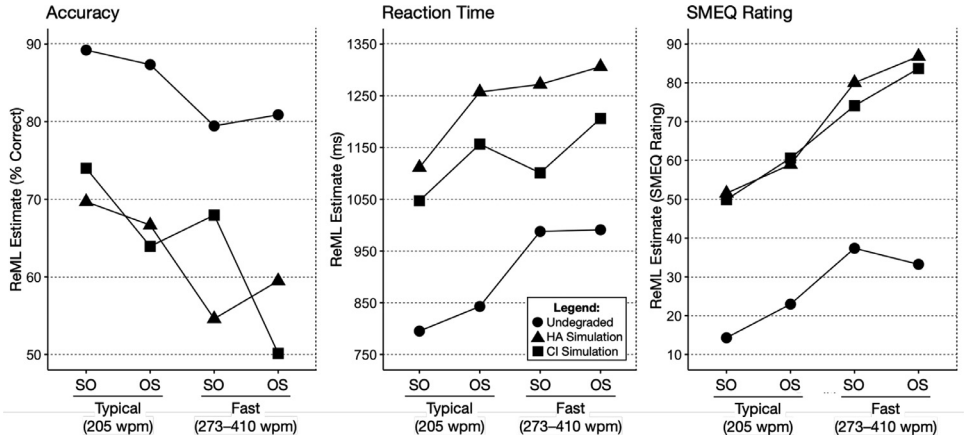


Fig. 2. The impact of clarity, speed, and syntax on task accuracy, reaction time, and SMEQ rating (see Table 2). Index: SO, subject-relative clause structure; OS, object-relative clause structure.

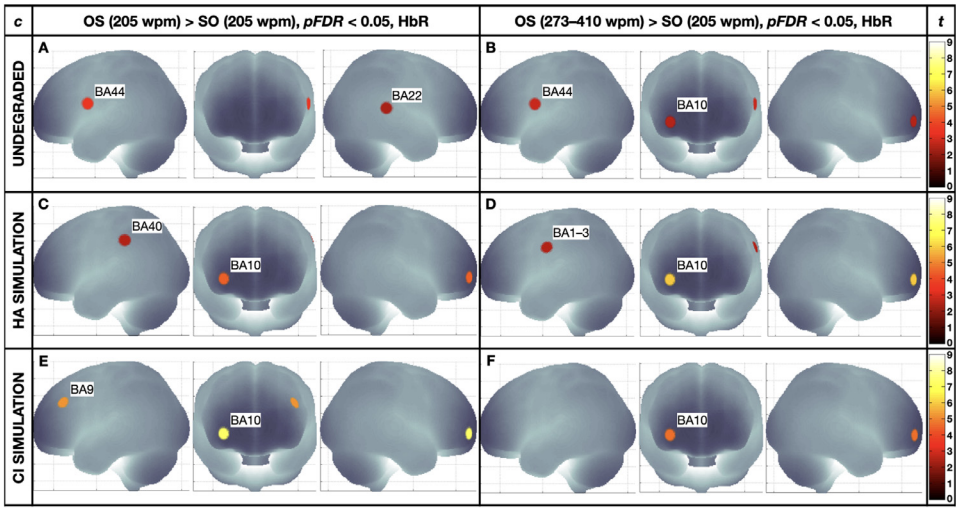


Fig. 3. Channel-wise HbR regression coefficient maps depicting the *t*-score (*t*) for each contrast (*c*). Index. HA, hearing aid; CI, cochlear implant; OS, object-relative clause structure; SO, subject-relative clause structure; BA, Brodmann area(s).

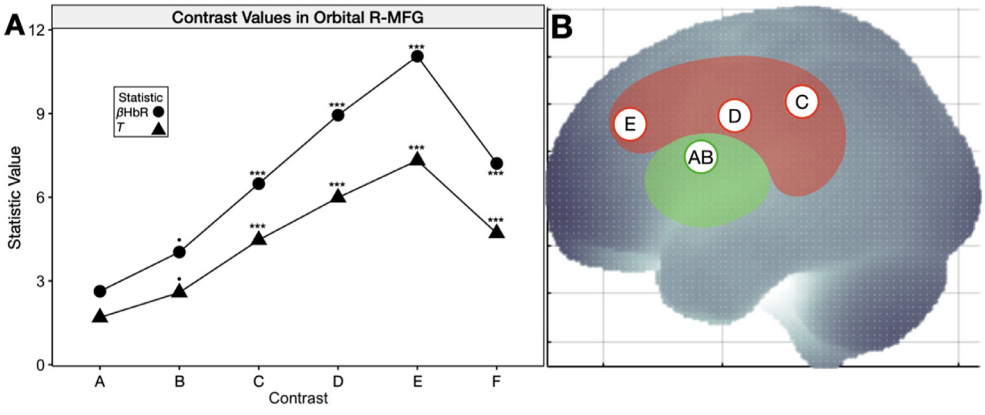


Fig. 4. (a) Channel-wise HbR regression coefficient (β, 'brain activity strength') and *t*-statistic values (*T*) in the orbital right MFG for each contrast. Index: ***, $p < 0.001$; •, $p < 0.05$. (b) Left hemisphere pattern of significant brain activation ($pFDR < 0.05$) for contrasts A–E within ventral (green) and dorsal (red) processing streams. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.3. Brain-behavior correlation results

We performed correlation analysis between brain activity and the behavioral response variables (i.e., accuracy, reaction time, and SMEQ rating; see 4.2.4.). We observed significant correlations between the behavioral response variables (see Fig. 5). There were significant moderate negative correlations between accuracy and reaction time ($r = -0.452$, $p < 0.001$) and between accuracy and SMEQ rating ($r = -0.475$, $p < 0.001$), as well as a significant weak positive correlation between reaction time and SMEQ rating ($r = 0.309$, $p < 0.001$). Lastly, we observed no significant correlations between brain activity (β HbR) in the orbital right MFG and responses.

6. Discussion

Using fNIRS brain imaging during a plausibility judgment task, we observed two distinct patterns of neural activity while hearing listeners attended to increasingly challenging speech signals. The first pattern appeared in the frontopolar cortex, specifically the right orbital MFG (BA10), and appears to reflect the manipulation of auditory attentional resources expended by the participant during the task. The second pat-

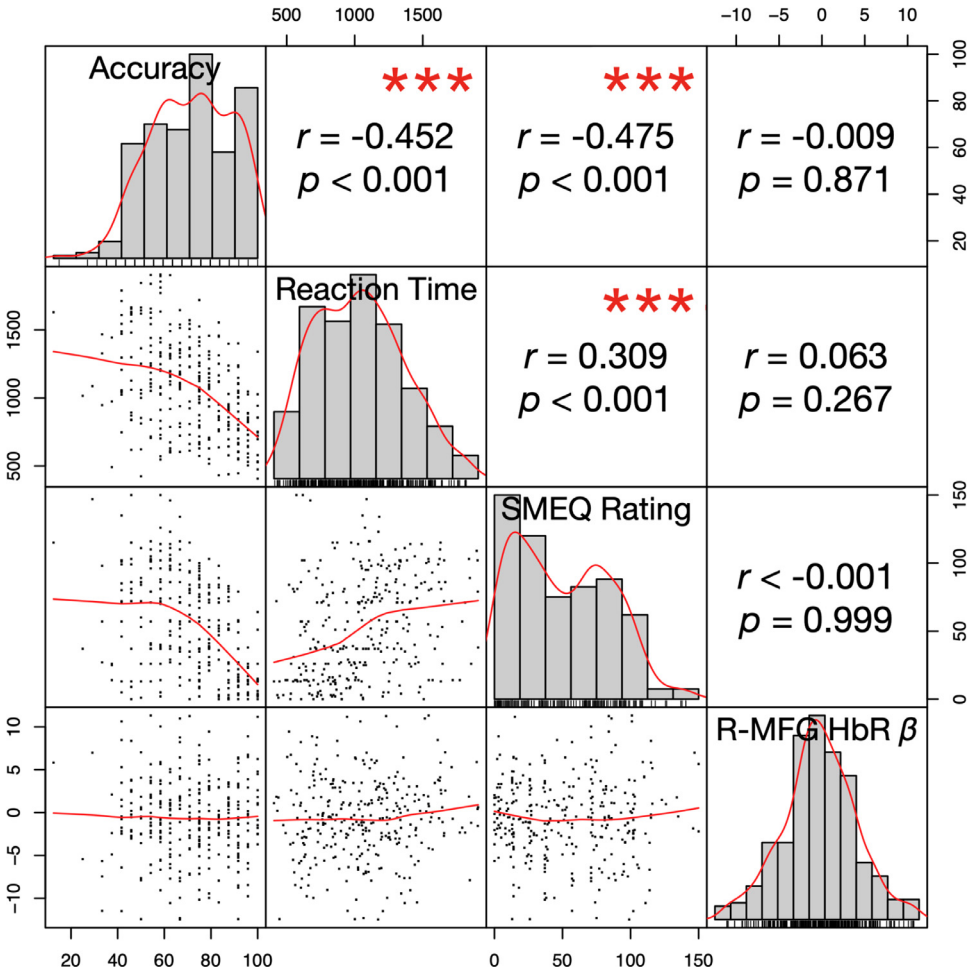


Fig. 5. Matrix depicting the Pearson Product Moment Correlation Coefficients (r) and significance values (p) between brain activity (β HbR) in the orbital right MFG and the behavioral response variables (accuracy, reaction time, and SMEQ rating). Index: ***, $p < 0.001$.

tern appears in the left hemisphere's core speech and language network and reflects a clear hierarchical organization for processing linguistic information in the presence of increasing listening challenge. Fascinatingly, this pattern can be easily traced along the ventral and dorsal auditory pathways, yielding new insight into the computational organization of the human cortex for speech and language. Lastly, participant's ratings of perceived mental effort significantly correlated with their performance on the task, highlighting the validity of using single item post-task questionnaires to assess listening effort and listening performance.

6.1. Auditory attention processing

The first distinct pattern of neural activity that we observed was in the frontopolar cortex, specifically the right orbital MFG (BA10). This cortical site has been well documented as the primary relay area for complex auditory information between the frontal lobe and auditory cortices in the STG in both humans and non-human primates (Medalla and Barbas, 2014). Further, BA10 is thought to be the core functional node responsible for delegating attentional resources and maintaining information for decision making (see 2.1.3.; Medalla and Barbas, 2014). As listening became more difficult during the task, activity in BA10 became stronger. However, during the most difficult condition, we observed a steep decrease in activation. Taken together, these findings suggest that participants were investing a greater amount of attentional resources for processing complex auditory and linguistic information but started to lose motivation as the task became too difficult to perform. One possible explanation for this is that the cognitive demand in the speech signal exceeded the available cognitive resources necessary for successful comprehension. This interpretation is supported by our behavioral findings, which reflect a similar pattern, as well as current models of auditory cognition and listening effort (e.g., Peelle, 2018; Pichora-Fuller et al., 1995). To the best of our knowledge, this is the first study to identify a possible neural signature for sustained auditory attention, or listening effort.

6.2. Auditory linguistic processing

We also observed a second distinct pattern of neural activity in the left hemisphere's core speech and language processing areas. The pattern begins with expected neural activation in the left IFG (BA44) and right STG (BA22) for the classic contrast between complex and simple syntax constructions (Caplan, 2007; Stromswold et al., 1996). These brain regions are functionally specialized for mediating longer memory traces associated with processing more complex OS syntactic structures. As listening became more difficult, neural activation pivoted posteriorly to the left SMG (BA40), to the left postcentral gyrus (BA1,2,3), to the left MFG of the dorsolateral prefrontal cortex (DLPFC; BA9), and finally terminating in the orbital right MFG (BA10). In relation to classic computational models for speech and language processing (e.g., Rauschecker and Scott, 2009), this neural pattern of activation can be traced along the ventral (easier conditions) to increasingly more dorsal (more difficult conditions) processing streams. One possible explanation for observing this pattern is that easier listening conditions allow for easier mapping of speech sounds to meaning, whereas more difficult listening conditions interfere with this mapping process and increas-

Table 2
Linear mixed effects coefficients for task accuracy, reaction time, and SMEQ rating.

Fixed Effect	ReML Est.	*SE*	*t*	*p*
Accuracy				
Clarity$^{Clear(Intercept)}$	**89.22**	**2.48**	**36.04**	**< 2e-16**
Clarity^{HASim}	**−19.54**	**2.67**	**−7.33**	**2.03e-12**
Clarity^{CISim}	**−15.23**	**2.67**	**−5.71**	**2.61e-8**
Speed^{Fast}	**−9.77**	**2.67**	**−3.67**	**2.91e-4**
Syntax^{OS}	−1.87	2.67	−0.70	0.484
Speed^{Fast}:Syntax^{OS}	3.31	3.77	0.88	0.381
Clarity^{HASim}:Speed^{Fast}	−5.32	3.77	−1.41	0.159
Clarity^{HASim}:Syntax^{OS}	−1.15	3.77	−0.31	0.761
Clarity^{CISim}:Speed^{Fast}	3.74	3.77	0.99	0.322
Clarity^{CISim}:Syntax^{OS}	**−8.19**	**3.77**	**−2.17**	**0.031**
Clarity^{HASim}:Speed^{Fast}:Syntax^{OS}	4.60	5.33	0.86	0.389
Clarity^{CISim}:Speed^{Fast}:Syntax^{OS}	**−11.06**	**5.33**	**−2.08**	**0.039**
Reaction Time				
Clarity$^{Clear(Intercept)}$	**795.24**	**58.38**	**13.62**	**< 2e-16**
Clarity^{HASim}	**316.15**	**58.47**	**5.41**	**1.29e-7**
Clarity^{CISim}	**251.91**	**58.47**	**4.31**	**2.22e-5**
Speed^{Fast}	**192.61**	**58.47**	**3.29**	**0.001**
Syntax^{OS}	47.59	58.47	0.81	0.416
Speed^{Fast}:Syntax^{OS}	−44.39	82.70	−0.54	0.592
Clarity^{HASim}:Speed^{Fast}	−31.87	82.70	−0.39	0.700
Clarity^{HASim}:Syntax^{OS}	98.31	82.70	1.19	0.235
Clarity^{CISim}:Speed^{Fast}	−138.76	82.70	−1.68	0.094
Clarity^{CISim}:Syntax^{OS}	61.84	82.70	0.75	0.455
Clarity^{HASim}:Speed^{Fast}:Syntax^{OS}	−67.40	116.95	−0.576	0.565
Clarity^{CISim}:Speed^{Fast}:Syntax^{OS}	40.05	116.95	0.342	0.732
SMEQ Rating				
Clarity$^{Clear(Intercept)}$	**14.31**	**5.41**	**2.65**	**0.0097**
Clarity^{HASim}	**37.24**	**5.33**	**6.99**	**1.76e-11**
Clarity^{CISim}	**35.62**	**5.33**	**6.68**	**1.11e-10**
Speed^{Fast}	**23.07**	**5.33**	**4.33**	**2.05e-5**
Syntax^{OS}	8.66	5.33	1.62	0.106
Speed^{Fast}:Syntax^{OS}	−12.79	7.54	−1.70	0.091
Clarity^{HASim}:Speed^{Fast}	5.41	7.54	0.72	0.473
Clarity^{HASim}:Syntax^{OS}	−1.24	7.54	−0.17	0.869
Clarity^{CISim}:Speed^{Fast}	1.10	7.54	0.15	0.884
Clarity^{CISim}:Syntax^{OS}	2.03	7.54	0.27	0.787
Clarity^{HASim}:Speed^{Fast}:Syntax^{OS}	12.17	10.66	1.14	0.255
Clarity^{CISim}:Speed^{Fast}:Syntax^{OS}	11.69	10.66	1.10	0.274

Index. Clear, undegraded; HASim, hearing aid simulation; CISim, cochlear implant simulation; OS, object-relative clause structure.

ingly rely on the brain's ability to derive meaning and make decisions from predictive and controlling contexts (Hickok and Poeppel, 2004). The left SMG (BA40) of the IPL is also associated with linguistic functions (Caplan et al., 1992), such as the phonological-articulatory loop (Baddeley et al., 1984). However, some evidence suggests that activity in this area is driven by higher-order linguistic processes for improving speech comprehension as opposed to acoustic processing to recover intelligibility of the degraded speech signal (Obleser et al., 2007). The postcentral gyrus (BA1,2,3) has also been uniquely implicated in linguistic processing, specifically phonological (Schwartz et al., 2012), semantic (Sabri et al., 2008), and syntactic processing (Lee et al., 2016), with a strong interaction effect with attention (e.g., Sabri et al., 2008) and Wild et al., 2012). Lastly, the left MFG of the DLPFC (BA9) is a region that has been typically associated with short-term verbal working memory for linguistic processing (Baddeley, 2003; Curtis and D'Esposito, 2003; D'Esposito, 2007). Taken together, we conclude that motivated processing of complex speech signals relies on a distributed network of cortical brain regions. Crucially, this network is hierarchical and can be traced in terms of current computational models of speech and language processing. Easy listening conditions with easy mapping of speech sounds to meaning are processed ventrally and closer to the auditory cortex, whereas difficult listening conditions rely on increasingly dorsal higher-order linguistic mechanisms that are further in proximity from the auditory cortex (i.e., from phonological-articulatory buffering in the IPL, to syntactic-attention recovery in the postcentral parietal lobe, to short-term-verbal working memory in the left DLPFC). Crucially, this network is hierarchical and can be mapped in terms of current computational models of speech and language processing. To the best of our knowledge, this is the first study to demonstrate the hierarchical composition of neural activation patterns for motivated and effortful speech processing in these contexts.

6.3. Brain-behavior correlations

Participants' mental effort ratings were significantly correlated with their accuracy and reaction times on the task. Voluntary attention affects performance in both task accuracy and reaction time, whereas involuntary attention affects task reaction time but not task accuracy (Prinzmetal et al., 2005, 2009). Taken together, our findings support the view that perceived mental effort is largely dependent on motivated, voluntary allocation of attentional resources. Additionally, this further validates the use of off-line single item usability questionnaires for assessing listening effort.

Of important note, we also found no significant correlations between neural activity in BA10 and the response variables. There are three possible reasons why we did not observe stronger brain-behavior correlations. First, the patterns of task performance and activation in BA10 appear to be nonlinear. Our brain-behavior correlation analysis only tests for linear relationships between variables. Indeed, this is supported by current models of auditory cognition and listening effort that are also nonlinear in nature (e.g., Peelle, 2018; Pichora-Fuller et al., 2016). Second, there exists an inherent inter-subject variability in cognitive perception and cognitive performance abilities (e.g., Peelle, 2018), as well as participant ratings for mental effort. Psychometrically, mental effort is also very difficult to interpret based on absolute or individual ratings and measurements. Between-group analyses may be more appropriate and statistically robust (Zijlstra, 1993, p. 59). This is partly due to an ill-defined nil-point (i.e., baseline), where even the simplest baseline measures require some degree of effort and are not rated at "0" on the scale across participants. Lastly, we used subject-level unweighted brain activation beta values and response variables. It is possible that weighted brain activation beta and response variables could provide more reliable population estimates. Further investigation is warranted to explore 'why' a relationship between brain and behavior was not found in our analyses.

Interestingly, hearing participants rated hearing aid and cochlear implant simulations to be comparably difficult. However, we observed distinct cortical processing patterns for each simulation. In BA10, we observed greater activation when listening to cochlear implant simulations than hearing aid simulations. In the left hemisphere, we observed recruitment of posterior temporal-parietal areas for processing hearing aid simulations (BA1,2,3,40) and anterior temporal-parietal areas for cochlear implant simulations (BA9,10). These findings suggest that the relationship between the brain and human behavior for listening effort is highly complex, and there may exist some similarities and differences between naïve listening of different types of acoustic distortions (e.g., hearing aid and cochlear implant simulations; Alain et al., 2018; Mattys et al., 2012).

6.4. Limitations

While our findings provide new evidence for the cortical and cognitive processing of complex speech signals, some limitations of this study must be noted. First, fNIRS brain imaging is a region-of-interest technique that is limited to the outer surface of cortical tissue. Therefore, we are not able to investigate additional, subcortical brain regions that might play an important role in auditory cognitive processes. The cingulo-opercular network, including the anterior cingulate and bilateral, but primarily the right, insula, are suggested to play an impor-

tant role in attention-based performance monitoring and the flexible allocation of neurocognitive resources for online information processing (Alain et al., 2018; Eckert et al., 2009; Erb and Obleser, 2013; Peelle, 2018; Vaden et al., 2013; Wild et al., 2012). Studies that integrate fNIRS and fMRI brain imaging techniques are promising to investigate the relationship between the cingulo-opercular network and listening effort. However, introducing magnetic resonance or another electromagnetic-based imaging technique would diminish the compatibility to investigate this line of research questioning with hearing aid and cochlear implant users. Future studies are planned to integrate other optical imaging approaches, such as fNIRS with functional infrared thermography (fIRT), which is sensitive for measuring different emotional-attentional states (Ioannou et al., 2014) but not affected by the use of hearing aids or cochlear implants during the experiment. Second, this study utilizes variations in spectral (i.e., wide-band vs. narrow-band speech) and temporal (i.e., normal vs. fast) domains. There are also other listening challenges that impact auditory attention that this study does not address, such as many different kinds of noise and multi-talker babble. The goal of this investigation was to isolate the cognitive effects of spectral degradation experienced everyday by hearing aid and cochlear implant users in the absence of additional acoustic challenge (i.e., competing acoustic signals). Further, cognitive processing in both hearing aid and cochlear implant users is uniquely challenged by speech rate (Piquado et al., 2012). However, typical listening often occurs in the presence of competing acoustic signals, and future studies would benefit from exploring the cognitive impacts of overlapping challenges within the acoustic domain. Lastly, all participants in this study were naïve listeners of distorted speech, and some research has documented a training effect in naïve listeners of noise-vocoded speech (e.g., Davis et al., 2005). This raises interesting theoretical questions about the role of sensorimotor feedback in the acquisition of speech and language by early-deafened hearing aid and cochlear implant users. Modern computational models suggest that adequate access to afferent inputs is necessary to complete the internal models in dorsal stream auditory processing and to facilitate neuroplasticity (e.g., Rauschecker, 1995). Future studies of early, long-term hearing aid and cochlear implant users are planned to further investigate the impact of sensory feedback in auditory linguistic processing.

7. Conclusion

In the present study, participants carried out a plausibility judgment task while undergoing fNIRS brain imaging. Sentences included a range of acoustic and linguistic challenges in order to increase cognitive demand (listening effort) and test the hypothesis that the ventral and dorsal auditory processing streams, with recruitment of the prefrontal cortex, have distinct contributions to motivated, effortful listening. We assessed the behavioral and neural impacts of listening to clear speech spoken at a normal rate with low syntactic complexity in comparison to degraded speech (i.e., wide-bandwidth hearing aid simulation, narrow-bandwidth cochlear implant simulation, temporal compression) spoken at a fast rate with high syntactic complexity (i.e., object-relative clauses). This experimental design of hierarchically challenging conditions permits a novel investigation with fNIRS brain imaging into the computational organization of sustained and motivated speech comprehension. Notably, to our knowledge, this is also the first investigation of the relationship between the brain, task performance, and internal perceptions of mental workload.

Here, we found that when mapping speech sounds to meaning is unimpeded, listening processes are carried out ventrally and closer to the auditory cortex. In comparison, difficult listening conditions are increasingly processed in dorsal higher-order linguistic mechanisms that are further in proximity from the auditory cortex. Specifically, neural networks for speech perception are modulated to reflect changing demands in the speech signal, from phonological-articulatory buffering in the IPL, to syntactic-attention recovery in the postcentral parietal lobe, to short-term-verbal working memory in the left DLPFC. Taken together with the "inverted-V" pattern of activation in BA10 that reflects increasing demands on auditory attention processing, these findings contribute towards a better understanding of the hierarchical cortical organization for cognitive processes that underlie not only clear, easily perceived speech, but also motivated, effortful listening. Our findings can further inform and advance scientific inquiry about cortical speech processing and cognitive neuroplasticity in hearing aid and cochlear implant users.

Data/code availability statement

All original behavioral and fNIRS brain imaging data are part of a larger study and dataset. All data presented in this paper are available as supplemental documents. All original source code for behavioral and brain-behavioral analyses are available in the appendices. Functional NIRS brain imaging data were converted, analyzed, and verified using the NIRS Brain AnalyzIR Toolbox (April 26, 2019) with MATLAB R2017b: https://github.com/huppertt/nirstoolbox/tree/1c1d70b52f261ad7353d7f5bd893f4011b40cbb2.

Declaration of Competing Interest

The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Bradley E. White: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Clifton Langdon:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision.

Acknowledgments

The work presented here was directly funded by a Small Research Grant (White, PI; Langdon, Supervisor) from the Office of Research Support and International Affairs, Gallaudet University. This work was also made possible in part by a Florence R. Foerderer Fellowship (White) and scholarship support (White) from the National Science Foundation grant (SBE-1041725) awarded to the Science of Learning Center, Visual Language and Visual Learning, Gallaudet University. We are greatly appreciative for the assistance of Patrick Harris and Suzanne Scheuermann for their professional audio recording expertise and services and for Diana Andriola and Lauren Berger for their day-to-day contributions and support of this work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118324.

Appendix A. Models

Models are expressed in Wilkinson-Rogers notation.

1 **Behavioral Analysis** estimate ~ 1 + clarity*syntax*speed + (1|subject)
2 **Group-Level fNIRS Analysis** beta ~ −1 + condition + (1|subject)

Appendix B. Behavioral and brain-behavior correlation analysis, R source code

```
# ReML LME Statistics with WS Approximation
library(tidyverse)
library(lme4)
library(lmerTest)
# Import data
statsData <- read_csv("~/statsData.csv", col_names = TRUE, na = c("", "NA"), quoted_na = TRUE)
# Define data class
statsData$clarity <- factor(statsData$clarity, levels = c("C", "D", "V"), ordered = FALSE)
statsData$speed <- factor(statsData$speed, levels = c("N", "F"), ordered = FALSE)
statsData$syntax <- factor(statsData$syntax, levels = c("S", "O"), ordered = FALSE)
statsData$subject <- factor(statsData$subject)
# LME Models
acc.lmer <- lmer(acc ~ 1 + clarity*syntax*speed + (1|subject), data = statsData)
summary(acc.lmer)
rt.lmer <- lmer(rt ~ 1 + clarity*syntax*speed + (1|subject), data = statsData)
summary(rt.lmer)
smeq.lmer <- lmer(smeq ~ 1 + clarity*syntax*speed + (1|subject), data = statsData)
summary(smeq.lmer)
# Pearson Product Moment Correlation Matrix
library(dplyr)
outliers <- boxplot(statsData$ch13hbrB, plot=FALSE)$out
statsData2 <- statsData[-which(statsData$ch13hbrB %in% outliers),]
chart.Correlation(select(statsData2, acc, rt, smeq, ch13hbrB), method = "pearson", pch = 46)
```

Appendix C. Functional NIRS source localization results

Table A1 reports the results from the fNIRS source localization analyzis presented in §4.2.2.

Table A1
Channel probabilistic brain MNI coordinates and AAL cytoarchitectural labels.

Ch.	Src.	Det.	MNI Coordinates			Cytoarchitecture[a]	Prob.
			X	*Y*	*X*		
1	1	1	−70.00	−33.67	8.67	Temporal Mid *L*	0.83
						Temporal Sup *L*	0.18
2	1	7	−67.00	−46.33	26.67	SupraMarginal *L*	0.83
						Temporal Sup *L*	0.18
3	2	1	−66.00	−7.67	5.33	Temporal Sup *L*	0.79
						Heschl *L*	0.08
						Temporal Mid *L*	0.08
						Rolandic Oper *L*	0.05
4	2	2	−58.00	20.67	3.67	Frontal Inf Tri *L*	0.90
						Frontal Inf Oper *L*	0.11
5	2	8	−63.00	9.00	18.00	Frontal Inf Oper *L*	0.70
						Precentral *L*	0.23
						Postcentral *L*	0.07
6	3	2	−52.00	44.67	−3.33	Frontal Inf Orb *L*	0.67
						Frontal Inf Tri *L*	0.29
						Frontal Mid Orb *L*	0.04
7	3	3	−34.33	65.00	−4.67	Frontal Mid Orb *L*	0.72
						Frontal Sup *L*	0.23
						Frontal Sup Orb *L*	0.06
8	3	9	−37.67	62.00	7.00	Frontal Mid *L*	0.95
						Frontal Sup *L*	0.05
9	4	3	−11.00	72.00	−7.00	Frontal Sup Orb *L*	0.62
						Frontal Mid Orb *L*	0.38
10	4	4	14.33	73.00	−5.67	Frontal Sup Orb *R*	0.69
						Frontal Mid Orb *R*	0.30
						Frontal Sup *R*	0.01
11	4	9	−15.00	73.00	5.00	Frontal Sup *L*	0.90
						Frontal Sup Medial *L*	0.10
12	4	10	17.67	73.00	5.67	Frontal Sup *R*	0.82
						Frontal Sup Medial *R*	0.18
13	5	4	37.67	65.33	−3.33	Frontal Mid Orb *R*	0.54
						Frontal Sup Orb *R*	0.32
						Frontal Mid *R*	0.13
						Frontal Sup *R*	0.01

(*continued on next page*)

Table A1 (*continued*)

Ch.	Src.	Det.	MNI Coordinates			Cytoarchitecture[a]	Prob.
			X	*Y*	*X*		
14	5	5	53.33	46.67	−1.67	Frontal Inf Orb *R*	0.48
						Frontal Inf Tri *R*	0.47
						Frontal Mid *R*	0.03
						Frontal Mid Orb *R*	0.03
15	5	10	40.67	62.33	8.33	Frontal Mid *R*	0.92
						Frontal Sup *R*	0.08
16	6	5	60.33	22.67	5.33	Frontal Inf Tri *R*	0.82
						Frontal Inf Oper *R*	0.18
17	6	6	69.00	−7.33	8.33	Temporal Sup *R*	0.82
						Rolandic Oper *R*	0.17
						Heschl *R*	0.01
18	6	11	65.33	10.67	20.67	Precentral *R*	0.84
						Frontal Inf Oper *R*	0.16
19	7	6	72.00	−33.33	11.33	Temporal Sup *R*	0.98
						Temporal Mid *R*	0.02
20	7	12	68.00	−44.67	28.67	SupraMarginal *R*	0.94
						Temporal Sup *R*	0.06
21	8	1	−68.67	−18.67	22.67	Postcentral *L*	0.69
						SupraMarginal *L*	0.31
22	8	7	−66.00	−33.00	41.00	SupraMarginal *L*	1.00
23	8	8	−63.67	−4.67	34.67	Postcentral *L*	1.00
24	8	13	−59.00	−17.33	50.67	Postcentral *L*	0.62
						Parietal Inf *L*	0.36
						SupraMarginal *L*	0.02
25	9	2	−55.00	35.00	10.00	Frontal Inf Tri *L*	1.00
26	9	8	−56.67	22.67	26.33	Frontal Inf Tri *L*	0.99
						Frontal Inf Oper *L*	0.01
27	9	9	−42.67	52.33	20.67	Frontal Mid *L*	1.00
28	9	14	−45.67	39.00	32.00	Frontal Mid *L*	1.00
29	10	3	−10.67	73.00	11.67	Frontal Sup *L*	0.58
						Frontal Sup Medial *L*	0.42
30	10	4	14.67	73.00	12.33	Frontal Sup Medial *R*	0.79
						Frontal Sup *R*	0.21
31	10	9	−14.00	68.00	23.67	Frontal Sup *L*	0.96
						Frontal Sup Medial *L*	0.04
32	10	10	17.00	68.67	24.00	Frontal Sup *R*	1.00
33	11	5	57.33	36.67	11.67	Frontal Inf Tri *R*	1.00
34	11	10	44.33	52.67	22.33	Frontal Mid *R*	1.00
35	11	11	57.33	25.33	28.33	Frontal Inf Tri *R*	0.78
						Frontal Inf Oper *R*	0.22
36	11	15	47.33	39.33	33.33	Frontal Mid *R*	1.00
37	12	6	70.00	−18.33	26.33	SupraMarginal *R*	0.99
						Postcentral *R*	0.01
38	12	11	65.00	−2.33	37.67	Postcentral *R*	0.97
						Precentral *R*	0.03
39	12	12	68.00	−31.33	42.67	SupraMarginal *R*	1.00
40	12	16	59.00	−16.00	53.67	Postcentral *R*	1.00
41	13	7	−58.00	−46.33	52.33	Parietal Inf *L*	1.00
42	13	13	−49.67	−30.67	63.33	Postcentral *L*	1.00
43	14	8	−52.67	11.67	43.33	Precentral *L*	0.76
						Frontal Mid *L*	0.24
44	14	13	−47.67	−1.67	57.33	Precentral *L*	1.00
45	14	14	−41.67	26.67	48.67	Frontal Mid *L*	1.00
46	15	11	54.00	13.33	44.33	Precentral *R*	0.55
						Frontal Mid *R*	0.42
						Frontal Inf Oper *R*	0.03
47	15	15	43.67	27.67	48.67	Frontal Mid *R*	1.00
48	15	16	48.33	−0.67	58.67	Frontal Mid R	0.93
						Precentral R	0.07
49	16	12	58.33	−44.33	54.67	Parietal Inf R	1.00
50	16	16	49.67	−29.67	64.33	Postcentral R	1.00

Index. Ch., Channel; Src., Source Optode; Det., Detector Optode; Prob., Coverage Probability.

[a] For a reference to AAL cytoarchitectural labels, see Tzourio-Mazoyer et al. (2002).

Appendix D. Functional NIRS channel-wise results

Table A2 reports the fNIRS channel-wise HbR regression coefficient contrast results presented in Figs. 3 and 4.

Table A2
Channel-wise HbR regression coefficient contrast results.

C[a]	Ch.[b]	MNI Coord. *X*	*Y*	*X*	*β*	*t*	*pFDR*	Prob.	Cytoarchitecture[c]
A	5	−63	9	18	11.54	3.58	0.003	0.70	Frontal Inf Oper *L*
								0.23	Precentral *L*
								0.07	Postcentral *L*
	19	72	−33	11	5.33	2.44	0.046	0.98	Temporal Sup *R*
								0.02	Temporal Mid *R*
B	5	−63	9	18	9.20	2.86	0.020	0.70	Frontal Inf Oper *L*
								0.23	Precentral *L*
								0.07	Postcentral *L*
	13	38	65	−3	4.03	2.59	0.033	0.54	Frontal Mid Orb *R*
								0.32	Frontal Sup Orb *R*
								0.13	Frontal Mid *R*
								0.01	Frontal Sup *R*
C	13	38	65	−3	6.49	4.47	1.44e-4	0.54	Frontal Mid Orb *R*
								0.32	Frontal Sup Orb *R*
								0.13	Frontal Mid *R*
								0.01	Frontal Sup *R*
	22	−66	−33	41	4.90	2.57	0.034	1.00	SupraMarginal *L*
D	13	38	65	−3	8.94	5.99	1.24e-7	0.54	Frontal Mid Orb *R*
								0.32	Frontal Sup Orb *R*
								0.13	Frontal Mid *R*
								0.01	Frontal Sup *R*
	23	−64	−5	35	5.98	2.56	0.035	1.00	Postcentral *L*
E	13	38	65	−3	11.05	7.32	6.84e-11	0.54	Frontal Mid Orb *R*
								0.32	Frontal Sup Orb *R*
								0.13	Frontal Mid *R*
								0.01	Frontal Sup *R*
	28	−46	39	32	6.18	5.27	4.08e-6	1.00	Frontal Mid *L*
F	13	38	65	−3	7.21	4.71	5.33e-5	0.54	Frontal Mid Orb *R*
								0.32	Frontal Sup Orb *R*
								0.13	Frontal Mid *R*
								0.01	Frontal Sup *R*

Index. C, Contrast; Ch., Channel; MNI Coord., MNI Coordinates; Prob., Coverage Probability.
[a] For a reference to contrast labels, see 4.2.3. and Fig. 3.
[b] For a reference to channel labels, see Table A1.
[c] For a reference to AAL cytoarchitectural labels, see Tzourio-Mazoyer et al. (2002).

Appendix E. Figures

Fig. A1 depicts the parameters for the simulated digital hearing aid used to make the hearing aid simulation. Fig. A2 depicts an example of the undegraded, hearing aid simulated, and cochlear implant simulated speech conditions. Fig. A3 depicts the trial and block designs.

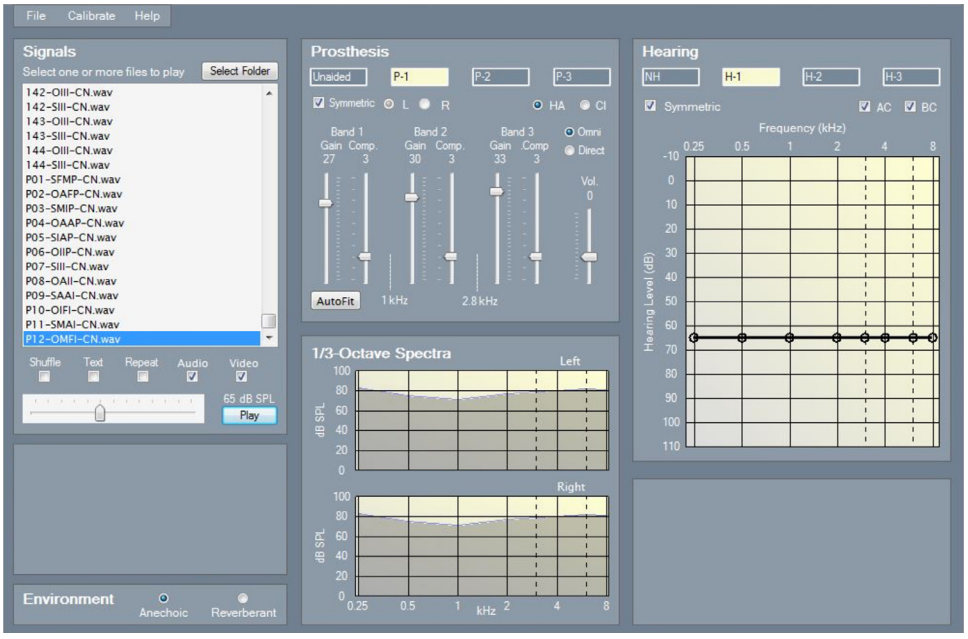


Fig. A1. A screen capture image depicting the parameters for the programmed simulated digital hearing aid using the "AutoFit" feature within the HeLPS v2 computer software.

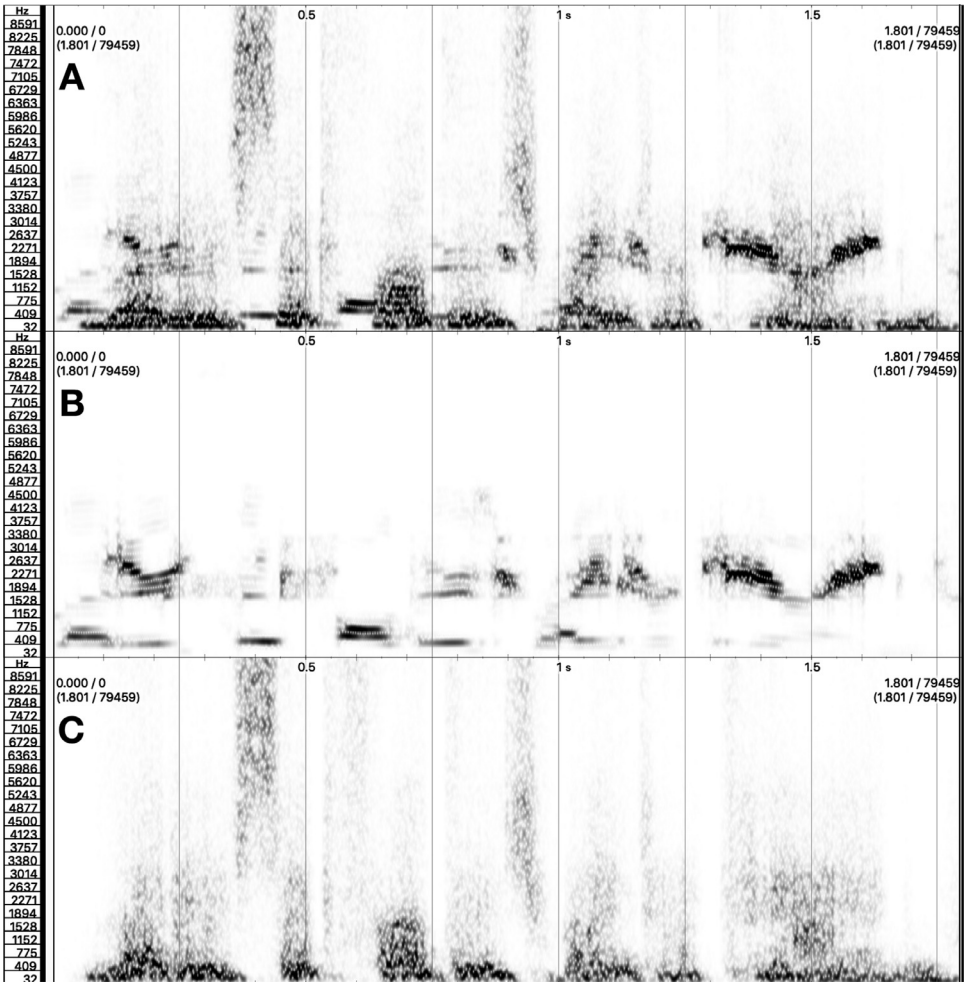


Fig. A2. Spectrograms depicting the: (a) original speech signal, (b) speech signal processed through a simulated hearing aid, and (c) speech signal processed through an eight-channel noise-vocoder algorithm (simulated cochlear implant). All images depict the sentence, "Letters that fathers write are caring.".

.

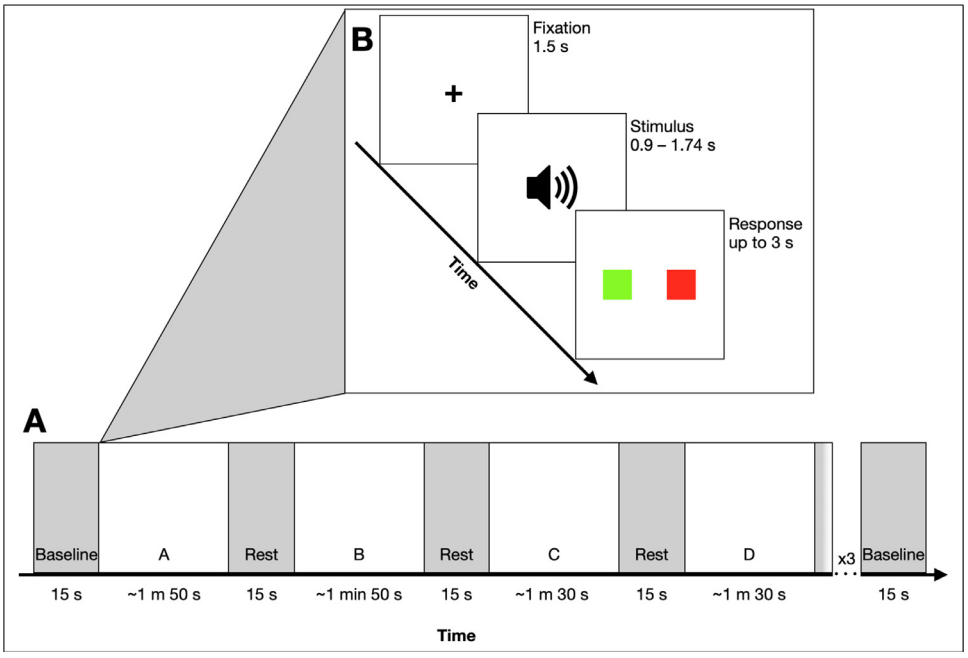


Fig. A3. A schematic of (a) the block design and (b) the trial design. There were 4 main conditions in the block design: (a) 24 subject-relative sentences presented at a typical rate (205 wpm), (b) 24 object-relative sentences presented at a typical rate (205 wpm), (c) 24 subject-relative sentences presented at a fast rate (273–410 wpm), and (d) 24 object-relative sentences presented at a fast rate (273–410 wpm). All conditions were presented in 3 unique runs: (i) undegraded, (ii) hearing aid simulated, and (iii) cochlear implant simulated speech. Trials in blocks were pseudorandomized. Blocks across all 3 runs were randomized.

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. Neurophotonics 2 (2), 020801.

Alain, C., Du, Y., Bernstein, L.J., Barten, T., Banai, K., 2018. Listening under difficult conditions: an activation likelihood estimation meta-analysis. Hum. Brain Mapp. 39 (7), 2695–2709.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8 (4), 170–177 (Regul. Ed.).

Audacity Team, 2018. Change Tempo. Audacity Team https://manual.audacityteam.org/man/change_tempo.html.

Baddeley, A., 2003. Working memory: looking back and looking forward. Nat. Rev. Neurosci. 4 (10), 829–839.

Baddeley, A., Lewis, V., Vallar, G., 1984. Exploring the articulatory loop. Q. J. Exp. Psychol. Sect. A 36 (2), 233–252.

Barker, J.W., Aarabi, A., Huppert, T.J., 2013. Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS. Biomed. Opt. Express. 4 (8), 1366–1379.

Basura, G.J., Hu, X.S., Juan, J.S., Tessier, A.M., Kovelman, I., 2018. Human central auditory plasticity: a review of functional near-infrared spectroscopy (fNIRS) to measure cochlear implant performance and tinnitus perception. Laryngoscope Investig. Otolaryngol. 3 (6), 463–472.

Bates, D., Maechler, M., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Grothendieck, G., Green, P., Bolker, M.B., 2015. Package “lme4.”. Convergence 12 (1).

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., Neville, H.J., 2000. Visual attention to the periphery is enhanced in congenitally deaf individuals. J. Neurosci. 20 (17), RC931–RC936.

Bilenko, N.Y., Grindrod, C.M., Myers, E.B., Blumstein, S.E., 2008. Neural correlates of semantic competition during processing of ambiguous words. J. Cogn. Neurosci. 21 (5), 960–975.

Blanco-Elorrieta, E., Ding, N., Pylkkänen, L., Poeppel, D., 2020. Understanding requires tracking: noise and knowledge interact in bilingual comprehension. J. Cogn. Neurosci. 32 (10), 1975–1983.

Caplan, D., 2001. Functional neuroimaging studies of syntactic processing. J. Psycholinguist. Res. 30 (3), 297–320.

D. Caplan (2007). Functional neuroimaging studies of syntactic processing in sentence comprehension: a critical selective review. In Language and linguistics compass (Vol. 1, Issues 1-2, pp. 32–47). https://doi.org/ 10.1111/j.1749-818x.2007.00005.x

Caplan, D., Alpert, N., Waters, G., 1998. Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. J. Cogn. Neurosci. 10 (4), 541–552.

Caplan, D., Alpert, N., Waters, G., Olivieri, A., 2000. Activation of Broca’s area by syntactic processing under conditions of concurrent articulation. Hum. Brain Mapp. 9 (2), 65–71.

Caplan, D., Chen, E., Waters, G., 2008a. Task-dependent and task-independent neurovascular responses to syntactic processing. Cortex 44 (3), 257–275.

Caplan, D., Rochon, E., Waters, G.S., 1992. Articulatory and phonological determinants of word length effects in span tasks. Q. J. Exp. Psychol. A 45 (2), 177–192.

Caplan, D., Stanczak, L., Waters, G., 2008b. Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. J. Cogn. Neurosci. 20 (4), 643–656.

Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., Dale, A.M., 2002. Vascular responses to syntactic processing: event-related fMRI study of relative clauses. Hum. Brain Mapp. 15 (1), 26–38.

Caplan, D., Waters, G., Alpert, N., 2003. Effects of age and speed of processing on rCBF correlates of syntactic processing in sentence comprehension. Hum. Brain Mapp. 19 (2), 112–131.

Chen, F., Loizou, P.C., 2011. Predicting the intelligibility of vocoded speech. Ear. Hear. 32 (3), 331.

Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D., Wise, R.J.S., 2003. Temporal lobe regions engaged during normal speech comprehension. Brain A J. Neurol. 126 (5), 1193–1201.

Curtis, C.E., D’Esposito, M., 2003. Persistent activity in the prefrontal cortex during working memory. Trends Cogn. Sci. 7 (9), 415–423 (Regul. Ed.).

Davis, H., Silverman, S.R., 1970. Hearing and Deafness. Holt, Rinehart & Winston of Canada Ltd.

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. J. Cogn. Neurosci. 23 (12), 3914–3932.

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

Davis, M.H., Johnsrude, I.S., Hervais-Adelman, A., Taylor, K., McGettigan, C., 2005. Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. J. Exp. Psychol. Gen. 134 (2), 222–241.

D’Esposito, M., 2007. From cognitive to neural models of working memory. Philos. Trans. R. Soc. Lond., B Biol. Sci. 362 (1481), 761–772.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2016. Increased activity in frontal motor cortex compensates impaired speech perception in older adults. Nat. Commun. 7, 12241.

Eckert, M.A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., Dubno, J.R., 2009. At the heart of the ventral attention system: the right anterior insula. Hum. Brain Mapp. 30 (8), 2530–2541.

Erb, J., Obleser, J., 2013. Upregulation of cognitive control networks in older adults’ speech comprehension. Front Syst. Neurosci. 7, 116.

Euston, D.R., Gruber, A.J., McNaughton, B.L., 2012. The role of medial prefrontal cortex in memory and decision making. Neuron 76 (6), 1057–1070.

Evans, S., Davis, M.H., 2015. Hierarchical organization of auditory and motor representations in speech perception: evidence from searchlight similarity analysis. Cereb. Cortex 25 (12), 4772–4788.

Evans, S., McGettigan, C., Agnew, Z.K., Rosen, S., Scott, S.K., 2016. Getting the cocktail party started: masking effects in speech perception. J. Cogn. Neurosci. 28 (3), 483–500.

Fairbanks, G., 1960. Voice and Articulation Drillbook, 2nd ed. Harper & Row, New York.

Ferrari, M., Quaresima, V., 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 63 (2), 921–935.

Fishman, K.E., Shannon, R.V., Slattery, W.H., 1997. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J. Speech Lang. Hearing Res. JSLHR 40 (5), 1201–1215.

Frey, S., Petrides, M., 2002. Orbitofrontal cortex and memory formation. Neuron 36 (1), 171–176.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392.
Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., Von Cramon, D.Y., 2006. Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb. Cortex 16 (12), 1709–1717.
Friederici, A.D., Gierhan, S.M.E., 2013. The language network. Curr. Opin. Neurobiol. 23 (2), 250–254.
Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J. Acoust. Soc. Am. 110 (2), 1150–1163.
Gosselin, P.A., Gagné, J.P., 2011. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. Int. J. Audiol. 50 (11), 786–792.
Hall, J.W., Mueller, H.G., 1997. Audiologists' Desk reference: diagnostic audiology-principles, Procedures and Practices (Vol. 1). Singular Publishing Group, San Diego, CA.
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.
Hays, R.D., Bjorner, J.B., Revicki, D.A., Spritzer, K.L., Cella, D., 2009. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil. 18 (7), 873–880.
Heinrich, A., Schneider, B.A., Craik, F.I.M., 2008. Investigating the influence of continuous babble on auditory short-term memory performance. Q. J. Exp. Psychol. 61 (5), 735–751.
Henson, R.N.A., Shallice, T., Dolan, R.J., 1999. Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring hypothesis. Brain A J. Neurol. 122 (7), 1367–1381.
Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92 (1–2), 67–99.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402.
Hirsch, J., Adam Noah, J., Zhang, X., Dravida, S., Ono, Y., 2018. A cross-brain neural mechanism for human-to-human verbal communication. Soc. Cogn. Affect. Neurosci. 13 (9), 907–920.
Hirsch, J., Tiede, M., Zhang, X., Noah, J.A., Salama-Manteau, A., Biriotti, M., 2020. Interpersonal agreement and disagreement during face-to-face dialogue: an fNIRS investigation. Front Hum. Neurosci. 14, 606397.
Huppert, T.J., Hoge, R.D., Diamond, S.G., Franceschini, M.A., Boas, D.A., 2006. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. Neuroimage 29 (2), 368–382.
Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. Cognition 92 (1–2), 101–144.
Ioannou, S., Gallese, V., Merla, A., 2014. Thermal infrared imaging in psychophysiology: potentialities and limits. Psychophysiology 51 (10), 951–963.
Jacques, S.L., 2013. Optical properties of biological tissues: a review. Phys. Med. Biol. 58 (11), R37.
Jäncke, L., Shah, N.J., 2002. Does dichotic listening probe temporal lobe functions? Neurology 58 (5), 736–743.
Just, M.A., Carpenter, P.A., 1992. A capacity theory of comprehension: individual differences in working memory. Psychol. Rev. 99 (1), 122.
Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, K.R., 1996. Brain activation modulated by sentence comprehension. Science 274 (5284), 114–116.
Kaufman, A.S., Kaufman, N.L., 2004. Kaufman Brief Intelligence Test. Wiley Online Library.
Kenward, M.G., Roger, J.H., 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 983–997.
Kirilina, E., Jelzow, A., Heine, A., Niessing, M., Wabnitz, H., Brühl, R., Ittermann, B., Jacobs, A.M., Tachtsidis, I., 2012. The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. Neuroimage 61 (1), 70–81.
Klem, G.H., Lüders, H.O., Jasper, H.H., Elger, C., 1999. The ten-twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 52 (3), 3–6.
Kovelman, I., Baker, S.A., Petitto, L.A., 2008. Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible neural signature" of bilingualism. J. Cogn. Neurosci. 20 (1), 153–169.
Lee, Y.S., Min, N.E., Wingfield, A., Grossman, M., Peelle, J.E., 2016. Acoustic richness modulates the neural networks supporting intelligible speech processing. Hear. Res. 333, 108–117.
Lipschutz, B., Kolinsky, R., Damhaut, P., Wikler, D., Goldman, S., 2002. Attention-dependent changes of activation and connectivity in dichotic listening. Neuroimage 17 (2), 643–656.
Love, T., Haist, F., Nicol, J., Swinney, D., 2006. A functional neuroimaging investigation of the roles of structural complexity and task-demand during auditory sentence processing. Cortex 42 (4), 577–590.
Mattys, S.L., Davis, M.H., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: a review. Lang. Cogn. Process. 27 (7–8), 953–978.
Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: international Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond. B Biol. Sci. 356 (1412), 1293–1322.
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 (7), 433–445.
McGettigan, C., Evans, S., Rosen, S., Agnew, Z.K., Shah, P., Scott, S.K., 2012. An application of univariate and multivariate approaches in fMRI to quantifying the hemispheric lateralization of acoustic and linguistic processes. J. Cogn. Neurosci. 24 (3), 636–652.
Medalla, M., Barbas, H., 2014. Specialized prefrontal "auditory fields": organization of primate prefrontal-temporal pathways. Front Neurosci. 8, 77.
Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D., von Cramon, D.Y., 2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang. 89 (2), 277–289.
Nyberg, L., Tulving, E., Habib, R., Nilsson, L.G., Kapur, S., Houle, S., Cabeza, R., McIntosh, A.R., 1995. Functional brain maps of retrieval mode and recovery of episodic information. Neuroreport 7 (1), 249–252.
Obleser, J., Erb, J., 2020. Neural filters for challenging listening situations. In: Poeppel, D., Mangun, G.R., Gazzaniga, M. (Eds.), The Cognitive Neurosciences. MIT Press, pp. 167–176.
Obleser, J., Wise, R.J.S., Dresner, M.A., Scott, S.K., 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. J. Neurosci. 27 (9), 2283–2289.
Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. U S A. 87 (24), 9868–9872.
Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.H., Saberi, K., Serences, J.T., Hickok, G., 2010. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cereb. Cortex 20 (10), 2486–2495.
Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., Oda, I., Isobe, S., Suzuki, T., Kohyama, K., Dan, I., 2004. Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. Neuroimage 21 (1), 99–111.
Okamoto, M., Dan, I., 2005. Automated cortical projection of head-surface locations for transcranial functional brain mapping. Neuroimage 26 (1), 18–28.
Okuda, J., Fujii, T., Yamadori, A., Kawashima, R., Tsukiura, T., Fukatsu, R., Suzuki, K., Ito, M., Fukuda, H., 1998. Participation of the prefrontal cortices in prospective memory: evidence from a PET study in humans. Neurosci. Lett. 253 (2), 127–130.
Olesen, P.J., Westerberg, H., Klingberg, T., 2004. Increased prefrontal and parietal activity after training of working memory. Nat. Neurosci. 7 (1), 75.
J.E. Peelle (2016). JP Vocode [MATLAB code]. GitHub. https://github.com/jpeelle/jp_matlab/blob/master/jp_vocode.m
Peelle, J.E., 2018. Listening effort: how the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear. Hear. 39 (2), 204–214.
J.E. Peelle (2019). The Neural Basis For Auditory and Audiovisual Speech Perception.
Peelle, J.E., Eason, R.J., Schmitter, S., Schwarzbauer, C., Davis, M.H., 2010. Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. Neuroimage 52 (4), 1410–1419.
Peelle, J.E., Troiani, B., Wingfield, A., Grossman, M., 2009. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. Cereb. Cortex 20 (4), 773–782.
Peirce, J.W., 2009. Generating stimuli for neuroscience using PsychoPy. Front Neuroinform. 2, 10.
Petrides, M., Alivisatos, B., Frey, S., 2002. Differential activation of the human orbital, mid-ventrolateral, and mid-dorsolateral prefrontal cortex during the processing of visual stimuli. Proc. Natl. Acad. Sci. 99 (8), 5649–5654.
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., Richter, M., Rudner, M., Sommers, M., Tremblay, K., Wingfield, A., 2016. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). Ear. Hear. 37, 5S–27S.
Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97 (1), 593–608.
Piquado, T., Benichov, J.I., Brownell, H., Wingfield, A., 2012. The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. Int. J. Audiol. 51 (8), 576–583.
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 (1), 15–35.
Prinzmetal, W., McCool, C., Park, S., 2005. Attention: reaction time and accuracy reveal different mechanisms. J. Exp. Psychol. Gen. 134 (1), 73.
Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., Dilem, L., 2009. Voluntary and involuntary attention have different consequences: the effect of perceptual difficulty. Q. J. Exp. Psychol. 62 (2), 352–369.
Quaresima, V., Bisconti, S., Ferrari, M., 2012. A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. Brain Lang. 121 (2), 79–89.
Rabbitt, P.M.A., 1968. Channel-capacity, intelligibility and immediate memory. Q. J. Exp. Psychol. 20 (3), 241–248.
Rauschecker, J.P., 1995. Compensatory plasticity and sensory substitution in the cerebral cortex. Trends Neurosci. 18 (1), 36–43.
Rauschecker, J.P., 2012. Ventral and dorsal streams in the evolution of speech and language. Front Evol. Neurosci. 4, 7.
Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12 (6), 718–724.
Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of what" and where" in auditory cortex. Proc. Natl Acad. Sci. 97 (22), 11800–11806.
R Core Team, 2017. R: a language and Environment For Statistical Computing. R Foundation for Statistical Computing.
Rodd, J.M., Davis, M.H., Johnsrude, I.S., 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb. Cortex 15 (8), 1261–1269.

Rodd, J.M., Johnsrude, I.S., Davis, M.H., 2010a. The role of domain-general frontal systems in language comprehension: evidence from dual-task interference and semantic ambiguity. Brain Lang. 115 (3), 182–188.
Rodd, J.M., Longe, O.A., Randall, B., Tyler, L.K., 2010b. The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. Neuropsychologia 48 (5), 1324–1335.
Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M.K., Rudner, M., 2013. The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. Front Syst. Neurosci. 7.
Rönnberg, J., Rudner, M., Foo, C., Lunner, T., 2008. Cognition counts: a working memory system for ease of language understanding (ELU). Int. J. Audiol. 47 (2), 99–105.
Rossi, S., Telkemeyer, S., Wartenburger, I., Obrig, H., 2012. Shedding light on words and sentences: near-infrared spectroscopy in language research. Brain Lang. 121 (2), 152–163.
Rypma, B., Prabhakaran, V., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E., 1999. Load--dependent roles of frontal brain regions in the maintenance of working memory. Neuroimage 9 (2), 216–226.
Sabri, M., Binder, J.R., Desai, R., Medler, D.A., Leitl, M.D., Liebenthal, E., 2008. Attentional and linguistic interactions in speech perception. Neuroimage 39 (3), 1444–1456.
Santosa, H., Zhai, X., Fishburn, F., Huppert, T.J., 2018. The NIRS brain AnalyzIR toolbox. Algorithms 11 (5), 73.
Sato, H., Yahata, N., Funane, T., Takizawa, R., Katura, T., Atsumori, H., Nishimura, Y., Kinoshita, A., Kiguchi, M., Koizumi, H., Fukuda, M., Kasai, K., 2013. A NIRS-fMRI investigation of prefrontal cortex activity during a working memory task. Neuroimage 83, 158–173.
Sauro, J., Dumas, J.S., 2009. Comparison of three one-question, post-task usability questionnaires. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1599–1608.
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.
Schrank, F.A., McGrew, K.S., Mather, N., Wendling, B.J., LaForte, E.M., 2014. Woodcock--Johnson IV Tests of Achievement: Woodcock-Johnson IV Tests of Cognitive Abilities. Riverside Publishing Company.
Schwartz, M.F., Faseyitan, O., Kim, J., Coslett, H.B., 2012. The dorsal stream contribution to phonological retrieval in object naming. Brain A J. Neurol. 135 (Pt 12), 3799–3814.
Scott, S.K., Rosen, S., Lang, H., Wise, R.J.S., 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.
Shalinsky, M.H., Kovelman, I., Berens, M.S., Petitto, L.A., 2009. Exploring cognitive functions in babies, children, and adults with near infrared spectroscopy. J. Vis. Exp. JoVE 29.
Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270 (5234), 303–304.
Silverman, S.R., Hirsh, I.J., 1955. Problems related to the use of speech in clinical audiometry. Ann. Otol. Rhinol. Laryngol. 64 (4), 1234–1244.
Strangman, G., Culver, J.P., Thompson, J.H., Boas, D.A., 2002. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. Neuroimage 17 (2), 719–731.
Stromswold, K., Caplan, D., Alpert, N., Rauch, S., 1996. Localization of syntactic comprehension by positron emission tomography. Brain Lang. 52 (3), 452–473.
Surprenant, A.M., 1999. The effect of noise on memory for spoken syllables. Int. J. Psychol. J. Int. de Psychol. 34 (5–6), 328–333.
Tachtsidis, I., Scholkmann, F., 2016. Erratum: publisher's note: false positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. Neurophotonics 3 (3), 039801.
Tsuzuki, D., Cai, D., Dan, H., Kyutoku, Y., Fujita, A., Watanabe, E., Dan, I., 2012. Stable and convenient spatial registration of stand-alone NIRS data through anchor-based probabilistic registration. Neurosci. Res. 72 (2), 163–171.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.
Vaden, K.I., Kuchinsky, S.E., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A., 2015. Cortical activity predicts which older adults recognize speech in noise and when. J. Neurosci. 35 (9), 3929–3937.
Vaden, K.I., Kuchinsky, S.E., Cute, S.L., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A., 2013. The cingulo-opercular network provides word-recognition benefit. J. Neurosci. 33 (48), 18979–18986.
Vaden, K.I., Kuchinsky, S.E., Keren, N.I., Harris, K.C., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A., 2011. Inferior frontal sensitivity to common speech sounds is amplified by increasing word intelligibility. Neuropsychologia 49 (13), 3563–3572.
Van Engen, K.J., Peelle, J.E., 2014. Listening effort and accented speech. Front Hum. Neurosci. 8, 577.
Vigneau, M., Beaucousin, V., Hervé, P.Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., Tzourio-Mazoyer, N., 2011. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: insights from a meta-analysis. Neuroimage 54 (1), 577–593.
Vos, S.H., Gunter, T.C., Schriefers, H., Friederici, A.D., 2001. Syntactic parsing and working memory: the effects of syntactic complexity, reading span, and concurrent load. Lang. Cogn. Process. 16 (1), 65–103.
White, B.E., 2019. The Role of Auditory Experience in the Neurocognitive Systems For Everyday and Effortful Listening. Gallaudet University (C. Langdon (ed.)) [Ph.D.].
White, B.E., Langdon, C., 2019. Listening Experience and Syntactic Complexity Modulate Neural Networks For Speech and language: A functional Near-Infrared Spectroscopy Study. Society of the Neurobiology of Language, Helsinki, Finland https://www.researchgate.net/profile/Bradley_White11/publication/335635990_Listening_experience_and_syntactic_complexity_modulate_neural_networks_for_speech_and_language_A_functional_near-infrared_spectroscopy_study/links/5d713dea92851cacdb23ac52/Listening-experience-and-syntactic-complexity-modulate-neural-networks-for-speech-and-language-A-functional-near-infrared-spectroscopy-study.pdf.
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 (40), 14010–14021.
Wingfield, A., 2016. Evolution of models of working memory and cognitive resources. Ear. Hear. 37, 35S–43S.
Wingfield, A., McCoy, S.L., Peelle, J.E., Tun, P.A., Cox, C.L., 2006. Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. J. Am. Acad. Audiol. 17 (7), 487–497.
Wingfield, A., Peelle, J.E., Grossman, M., 2003. Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn. 10 (4), 310–322.
Wolf, M., Ferrari, M., Quaresima, V., 2007. Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. J. Biomed. Opt. 12 (6), 062104.
Zekveld, A.A., Kramer, S.E., Festen, J.M., 2011. Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. Ear. Hear. 32 (4), 498–510.
Zekveld, A.A., Rudner, M., Johnsrude, I.S., Heslenfeld, D.J., Rönnberg, J., 2012. Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. Brain Lang. 122 (2), 103–113.
Zhang, X., Noah, J.A., Hirsch, J., 2016. Separation of the global and local components in functional near-infrared spectroscopy signals using principal component spatial filtering. Neurophotonics 3 (1), 015004.
Zijlstra, F.R.H., 1993. Efficiency in Work Behaviour: A Design Approach For Modern Tools. Delft University of Technology.
F.R.H. Zijlstra, & L.V. van Doorn (1985). The construction of a scale to measure subjective effort. Delft, Netherlands, 43, 124–139.
Zurek, P.M., Desloge, J.G., 2007. Hearing loss and prosthesis simulation in audiology. Hear. J. 60 (7), 32.