8-1-2023

Mapping cortical activations underlying covert and overt language production using high-density diffuse optical tomography

Mariel L Schroeder
Washington University School of Medicine in St. Louis
Arefeh Sherafati
Washington University School of Medicine in St. Louis
Rachel L Ulbrich
Washington University School of Medicine in St. Louis
Muriah D Wheelock
Washington University School of Medicine in St. Louis
Alexandra M Svoboda
Washington University School of Medicine in St. Louis

See next page for additional authors

NeuroImage 276 (2023) 120190

Contents lists available at ScienceDirect

NeuroImage

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

Mapping cortical activations underlying covert and overt language production using high-density diffuse optical tomography

Mariel L. Schroeder[a,b,1], Arefeh Sherafati[a,c,1], Rachel L. Ulbrich[a,d], Muriah D. Wheelock[a], Alexandra M. Svoboda[a,e], Emma D. Klein[a,f], Tessa G. George[a], Kalyan Tripathy[a,g], Joseph P. Culver[a,h,i,j], Adam T. Eggebrecht[a,h,j,*]

[a] Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, MO, USA
[b] Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN, USA
[c] Department of Neurology, University of California, San Francisco, San Francisco, CA, USA
[d] University of Missouri School of Medicine, Columbia, MO, USA
[e] University of Cincinnati Medical Center, Cincinnati, Oh, USA
[f] Icahn School of Medicine at Mount Sinai, New York, NY, USA
[g] Washington University School of Medicine, St Louis, MO, USA
[h] Division of Biology & Biomedical Sciences, Washington University School of Medicine, St Louis, MO, USA
[i] Department of Physics, Washington University in St. Louis, St Louis, MO, USA
[j] Department of Biomedical Engineering, Washington University in St. Louis, St Louis, MO, USA

ARTICLE INFO

Keywords:
High-density diffuse optical tomography
Naturalistic language
Overt
Covert
Speech
Articulation

ABSTRACT

Gold standard neuroimaging modalities such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and more recently electrocorticography (ECoG) have provided profound insights regarding the neural mechanisms underlying the processing of language, but they are limited in applications involving naturalistic language production especially in developing brains, during face-to-face dialogues, or as a brain-computer interface. High-density diffuse optical tomography (HD-DOT) provides high-fidelity mapping of human brain function with comparable spatial resolution to that of fMRI but in a silent and open scanning environment similar to real-life social scenarios. Therefore, HD-DOT has potential to be used in naturalistic settings where other neuroimaging modalities are limited. While HD-DOT has been previously validated against fMRI for mapping the neural correlates underlying language comprehension and covert (i.e., "silent") language production, HD-DOT has not yet been established for mapping the cortical responses to overt (i.e., "out loud") language production. In this study, we assessed the brain regions supporting a simple hierarchy of language tasks: silent reading of single words, covert production of verbs, and overt production of verbs in normal hearing right-handed native English speakers (n = 33). First, we found that HD-DOT brain mapping is resilient to movement associated with overt speaking. Second, we observed that HD-DOT is sensitive to key activations and deactivations in brain function underlying the perception and naturalistic production of language. Specifically, statistically significant results were observed that show recruitment of regions in occipital, temporal, motor, and prefrontal cortices across all three tasks after performing stringent cluster-extent based thresholding. Our findings lay the foundation for future HD-DOT studies of imaging naturalistic language comprehension and production during real-life social interactions and for broader applications such as presurgical language assessment and brain-machine interfaces.

1. Introduction

Robust mapping of cortical brain function underlying naturalistic language production has broad potential impacts from tracking childhood development of language (Kuhl et al., 2014) to diagnosis and treatment of language impairments such as post-stroke aphasia (Allendorfer et al., 2012), apraxia of speech (Dronkers, 1996; Dronkers et al., 2004a; Richardson et al., 2012), or for presurgical language assessment (Benjamin et al., 2018; Tyndall et al., 2017). Despite extensive neuroimaging studies investigating different aspects of the human language system, neural computations underlying language production in real-life settings especially in the early years of development

* Corresponding author at: Mallinckrodt Institute of Radiology; Washington University School of Medicine; 660 S. Euclid Ave, St Louis, MO 63110; USA.
E-mail address: aeggebre@wustl.edu (A.T. Eggebrecht).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neuroimage.2023.120190.
Received 27 December 2022; Received in revised form 5 May 2023; Accepted 23 May 2023
Available online 26 May 2023.

remains challenging, mainly due to the constraints of the current gold-standard neuroimaging modalities.

Recent studies using invasive methods such as electrocorticography (ECoG) have provided unparalleled access to the neural substrates of language comprehension and production (Anumanchipalli et al., 2019b; Flinker et al., 2015; Li et al., 2021; Yi et al., 2019). However, ECoG not only provides a limited spatial sampling of cortical tissue but also is limited to a small group of clinical patients who are already undergoing neurosurgery (Metzger et al., 2022; Moses et al., 2019). By contrast, non-invasive methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have been used to map the hierarchical and spatially distributed nature of the language system in typical and various atypical populations (Bohland and Guenther, 2006; Brown et al., 2005; Church et al., 2011, 2008; Eickhoff et al., 2009; Fedorenko et al., 2015; Ihnen et al., 2009; Petersen et al., 1988; Wise et al., 1999). However, aside from high costs, PET uses ionizing radiation and has poor temporal resolution and fMRI entails a noisy ambient environment of 110 dB (human speech typically ranges between 50 and 65 dB). These limitations aside, none of the current neuroimaging modalities provide naturalistic studies of language processing in real-life settings as they either rely on certain patient populations in surgery settings (ECoG) or they require the participants to lie alone in an isolated scanner (PET or fMRI).

Functional near infrared spectroscopy (fNIRS) methods similar to fMRI, measure the changes in the cerebral blood flow following neuronal activity, but offer a silent (<30 dB), non-invasive, portable, and potentially wearable alternative that enable mapping of human brain function in open and naturalistic social environments (Pinti et al., 2018, 2020). Furthermore, the flexibility of the fNIRS-based scanning environments makes it more feasible for use in developing brains (Billing et al., 2021; Blanco et al., 2021; Ferradal et al., 2016; Frijia et al., 2021), a critical phase for studying the development of language. Despite all these advantages, fNIRS has not been widely used among neurolinguistics or cognitive psychologists due to its low spatial resolution, lack of depth profiling, and limited field of view.

High-density diffuse optical tomography (HD-DOT) is a fast-growing fNIRS-based modality with a dense arrangement of optical sources and detectors that leads to a significantly higher spatial resolution and depth profiling compared to that of the standard fNIRS (Eggebrecht et al., 2014, 2012; White and Culver, 2010). This higher spatial resolution, paired with the potential for portability and wearability, not only makes HD-DOT a reliable surrogate for fMRI but also enables a naturalistic scanning environment for use in real-life settings such as during face-to-face dialogues (Hirsch et al., 2021) or as a brain-computer interface (BCI) (Anumanchipalli et al., 2019a; Makin et al., 2020) that could restore communication capabilities in patients with locked in syndrome (Biasiucci et al., 2018; Collinger et al., 2013). Despite these possibilities, while HD-DOT has been used to map the hierarchical structure of processing single words (Eggebrecht et al., 2014), and the functional activation (Hassanpour et al., 2015) and connectivity (Hassanpour et al., 2017) underlying syntactic complexity, it has not yet been leveraged to investigate naturalistic production of language. This topic is particularly important, as the resulting motion artifacts from overt speaking have been reported to be a key limitation in mapping language production using traditional neuroimaging modalities such as PET and fMRI. Thus, covert tasks are typically used instead of overt tasks to avoid introducing motion (Black et al., 2017). This underscores the importance of demonstrating HD-DOT's resilience to the level of motion artifacts during overt speaking and its feasibility for mapping the differential response to covert vs. overt language production. This may lay the foundation for future applications in more naturalistic real-life social interactions and presurgical language assessments.

Herein, we imaged 41 healthy adults with HD-DOT as they performed a simple hierarchy of language tasks: silent reading of single words and covert and overt production of verbs. We chose this hierarchy based on the hypothesis that successful language production can be modeled as the covert (silent and internalized) generation of language followed by the overt (out loud and externalized) articulation of the planned content and phonemes. As all stimuli for the covert and overt verb production tasks were presented orthographically, the silent word reading task could be used as a baseline for mapping the brain response to the visual language comprehension task.

We aimed to quantify the level of motion artifacts during covert and overt verb production using a previously defined motion detection index called the global variance of the temporal derivatives (GVTD) (Sherafati et al., 2020). We hypothesized that moderate jaw movements (as quantified by GVTD motion index) during the overt verb production task would not result in significant artifacts in HD-DOT data relative to control tasks that require no jaw movement: silent reading of words and covert production of verbs. We also hypothesized that our sample size and spatial resolution would be sufficient for mapping the neural response to language comprehension through reading words and the differential responses to covert vs. overt verb production.

2. Materials and methods

2.1. Data and code availability

Data and analysis scripts will be publicly available upon acceptance at our NeuroDOT repository, https://www.nitrc.org/projects/neurodot, and on OpenfNIRS, https://openfnirs.org/.

2.2. Participants

Adult participants were recruited from the Washington University in St. Louis community and surrounding area. The research was approved by the Human Research Protection Office at Washington University in St. Louis School of Medicine, and informed consent was obtained for all subjects. In total, 41 native English-speaking participants (21 females) 20–44 years old participated in the study. To help control for variance beyond the scope of this study, exclusion criteria included being a native speaker of an additional language to English (n = 2), being left-handed or strongly ambidextrous (n = 4), and not complying with task instructions (n = 2). The final dataset was comprised of 33 native English speaking, right-handed participants ages 20–44 years (17 females). To establish handedness, participants completed the short version of the Edinburgh Handedness Inventory (Oldfield, 1971). Based on this questionnaire, we ascertained the handedness of the participants based on the Laterality Quotient (Left-handedness = Less than −40, Ambidexterity = Between −40 and +40, and Right-handedness = More than +40) (Table 1). To assess general cognitive ability, the number of years of education was used (Matarazzo, 1972; Matarazzo and Herman, 1984) (Table 1). All participants had no known history of speech disorder or language-related learning disabilities (e.g., childhood lisping, dysgraphia, dyslexia).

2.3. HD-DOT system and data acquisition

The custom-built continuous wave HD-DOT instrument used in this experiment has been previously described in detail (Eggebrecht et al., 2014). The instrument contained a dense array of 96 LED sources and 92 avalanche photodiode detectors (188 total optodes) spatially arranged in a regular grid with first through fourth nearest neighbor distances of 1.3, 3.0, 3.9 and 4.7 cm, respectively (Fig. 1A). This dense grid of overlapping source-detector measurements provided sensitivity in a field of view over about a third of optically-accessible cortex, including multiple cortical regions previously implicated in processing of language (Fig. 1B). The 96 sources illuminated the head at two wavelengths (750 nm and 850 nm), resulting in ~1200 source-detector measurements per wavelength sampled at a frame rate of 10 Hz. Optodes were coupled to the scalp at a perpendicular angle via a custom-built imaging cap and connected to the console supporting the optoelectronics with 4.2 m long

Table 1
Participant demographics.

Mean	*Age in Years (+/-STD)*	*Laterality Quotient (+/-STD)*	*Years of Education (+/-STD)*	*Estimated Full Scale IQ (+/-STD)*
Male $n = 16$	28.89 (5.10)	78.28 (23.89)	19 (2.99)	113.43 (1.34)
Female $n = 17$	27.39 (6.36)	80.59 (23.16)	17.5 (2.80)	112.54 (2.25)
Full Sample	28.12 (5.75)	79.47 (23.17)	18.23 (2.95)	112.96 (1.91)

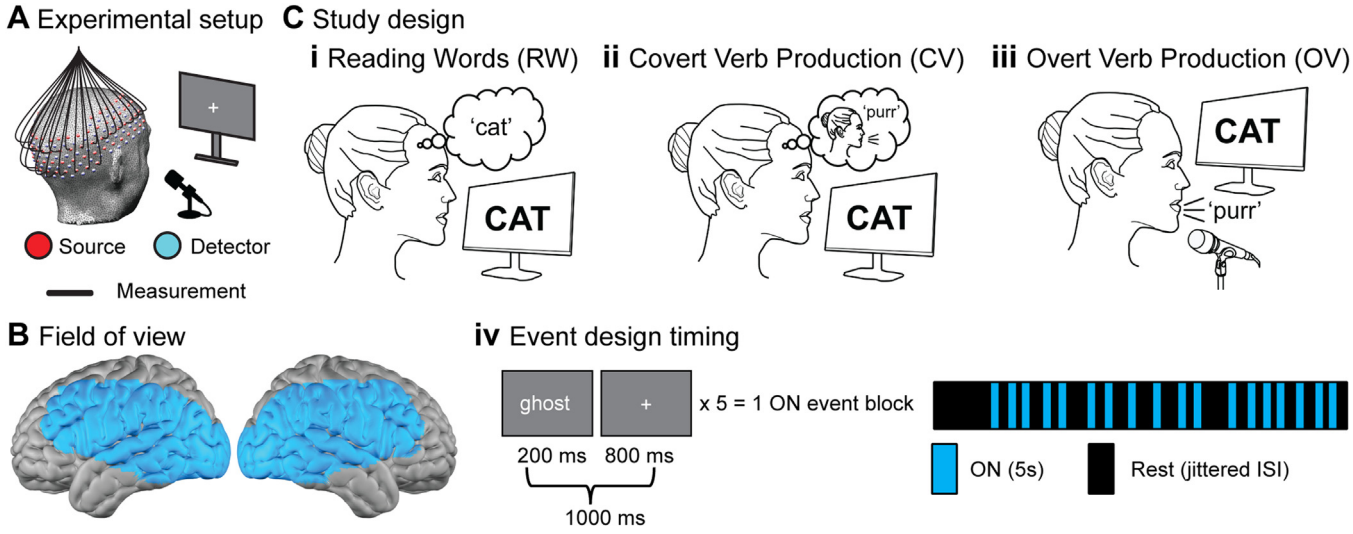

Fig. 1. HD-DOT system and study design. (A) Schematic of a participant wearing the imaging cap (subset of fibers shown for clarity). (B) The field of view of the HD-DOT array on a model of the lateral perspectives of the cortex. (C) Stimulus designs and event design timing. All three task paradigms used visually presented nouns, with only OV requiring a spoken response.

fiber-optic bundles. The weight of the cap and the fibers were supported by an extruded aluminum frame and a "double halo" of two, collinear rings to prevent participants from bearing any of the imaging hardware weight while allowing gross head motion of a few centimeters. Participants were seated in an adjustable chair in a quiet, sound-insulated room (<30 dB during data acquisition) and underwent a roughly 10 min cap-fit procedure. The cap position, comfort, and fiber-to-scalp optical coupling were optimized with the aid of real time visual feedback on data quality including measurement-specific light levels and signal-to-noise as well as optode-specific assessments of coupling coefficients and temporal variability.

2.4. Stimuli and procedure

Participants completed an approximately 35 min-long HD-DOT session. Before beginning the experiment, participants were briefed on task instructions. During the experiment, participants completed three different tasks (Fig. 1C): (i) reading single words (RW), (ii) covert verb production (CV), and (iii) overt verb production (OV). For the RW task, participants were instructed to read each word silently. For the CV task, participants were instructed to covertly (silently) generate a verb associated with each noun. For the OV task, participants were instructed to speak aloud a verb associated with each noun. In order to minimize the effect of motion artifacts due to overt speaking, we instructed our participants to avoid any unnecessary jaw or head movements during overt speaking. Each imaging session consisted of two RW runs, three CV runs, and three OV runs. All runs were approximately 3 min and 45 s in length. Additionally, participants completed a practice run at the beginning of each session to ensure they understood task instructions and to confirm that audio responses could be clearly captured. All three tasks used visual presentation of single nouns in horizontally and vertically centered white text (Arial font, 80 pt) on a 50% gray background on a computer monitor positioned 90 cm from the nasion. Nouns were presented using an event design paradigm. Each event block (ON) consisted of five nouns presented one second apart (presentation time of 200 ms; inter-word display interval of 800 ms; Fig. 1C (iv). A white fixation crosshair was presented on the monitor between word presentations (Rest) with pseudorandomized exponentially distributed inter-stimulus intervals (2–13 s). Participants were instructed to fixate on the crosshair whenever it appeared, including during the initial 30 s resting fixation period at the beginning of each run.

High-frequency, concrete nouns ($n = 765$) from Standard American English were selected from the SUBTLEXus corpus (Brysbaert and New, 2009). Only nouns with Zipfian frequencies (Van Heuven et al., 2014) of 3.86 and higher were included. Nine lists of words were created (one practice run list plus a word list for each of the other eight runs). No noun was presented more than once per imaging session, and the order of words presented within each list was counterbalanced across participants. Four pseudorandomized orders of nouns were created for each of the eight lists of words. All stimuli were delivered using MATLAB (2010b) and Psychophysics Toolbox 3 via custom scripts.

2.5. Behavioral response collection

Audio responses (i.e., verbs generated during OV runs) were captured using a Sennheiser e835 Live Vocal Microphone positioned two to four inches laterally from and slightly posterior to participants' mouths and recorded using the Audacity software. Audio responses were monitored live (via visualizations of the waveforms in Audacity) during the scan to ensure that participants were not speaking during the RW or CV tasks and that they were speaking during the OV task.

2.6. HD-DOT data processing

HD-DOT data were processed using the NeuroDOT (https://www.nitrc.org/projects/neurodot/) processing pipeline in MATLAB (Eggebrecht and Culver, 2019). Individual source-detector pair light level measurements were converted into 5D-movies (oxygena-

tion species concentrations by space by time) of relative hemodynamic changes at the voxel-level through five steps: light-level measurement pre-processing, anatomical light modeling, image reconstruction, spectroscopy, and spatial normalization. Details regarding these procedures have been previously published (Eggebrecht et al., 2014) and are described in brief below.

The baseline of the raw light-level measurement for each source-detector (SD) channel for an individual run was calculated as the mean value across the entire run. This baseline was then used to convert the SD measurements to time-series of log-ratio optical density data by taking the logarithm of the ratio of the instantaneous light level and the mean value (baseline) over the entire run. Any measurements with temporal standard deviation > 7.5% were excluded from image reconstruction because they were deemed overly contaminated by non-physiological variance such as motion artifact. In summary, only 3.2 ± 3.5% of measurements were excluded for each subject across all their RW runs, and 3.5 ± 3.7% and 4.0 ± 4.6% for CV and OV runs, respectively. The density of the SD grid ensured that each voxel is sampled by multiple measurements, thereby minimizing data fidelity loss when individual channel measurements are excluded due to high temporal variance. Long term drift was removed from the remaining measurements via high-pass filtering at a cutoff frequency of 0.02 Hz with a 5-pole Butterworth filter. The mean across all included first nearest neighbor measurements (source-detector pairs with a 13 mm separation) was used as an approximation of systemic superficial signals and regressed out of all measurements (Gregg et al., 2010). Measurements were then low-pass filtered at a cutoff frequency of 0.5 Hz and resampled to 1 Hz. The resampling to 1 Hz was performed primarily because the aim of this study is to evaluate the efficacy of HD-DOT in localizing the language responsive cortical regions as compared to previous fMRI results (which typically samples at or below ~1 Hz) rather than exploring new temporal dynamics of the speech related areas.

As described in prior work (Eggebrecht et al., 2014; Ferradal et al., 2014), a finite-element, wavelength-dependent forward light-model was used to calculate the sensitivity for the full array of SD measurements using a heterogeneous head model composed of five tissue types, each with specific optical properties: scalp, skull, gray matter, white matter, and cerebrospinal fluid (Eggebrecht et al., 2012). The non-linear ICBM152 atlas from the Montreal Neurological Institute (Fonov et al., 2009) was used as the anatomical model (Ferradal et al., 2014).

Image reconstruction was performed to convert from optical density to relative changes in absorption in a voxelated space using spatially variant regularization parameter of 0.1, Tikhonov regularization parameter of 0.01, and a Gaussian 3D smoothing kernel of 3 mm full width at half maximum was used to remove speckly spatial noise. These data were then spectroscopically unmixed to obtain volumetric time-series data of relative changes in oxyhemoglobin (ΔHbO_2), deoxyhemoglobin (ΔHbR), and total hemoglobin (ΔHbT) concentrations (Bluestone et al., 2001). For brevity, while the present analyses focused on the ΔHbO_2 contrast, as we have consistently found higher contrast-to-noise ratios with ΔHbO_2 compared to ΔHbR or ΔHbT, we are also presenting figures and tables of significant clusters for ΔHbR to highlight the complementary results obtainable with ΔHbO_2 and ΔHbR.

2.7. Detection and removal of motion artifacts

To ensure that motion artifacts (especially during the OV task) do not result in spurious effects in our results, we quantified motion using the global variance of the temporal derivatives (GVTD). The GVTD, much like the DVARS measure of fMRI (Smyser et al., 2010), is calculated by taking the root mean square of the first temporal derivative of all first nearest neighbor measurements (Sherafati et al., 2017). GVTD as a data-driven index of motion for HD-DOT has been previously validated to outperform direct accelerometer measures and other motion correction methods for HD-DOT data (Sherafati et al., 2020). Lower mean GVTD values have been shown to correlate with high spatial similarity of the HD-DOT maps with corresponding gold standard fMRI maps (Sherafati et al., 2020). Here, we used GVTD to quantify the levels of motion across three paradigms and to remove the runs of data with high median levels of motion using an empirical GVTD threshold.

2.8. Response modeling and statistical analysis

A general linear model was used with an HD-DOT-derived adult hemodynamic response function (Hassanpour et al., 2014) to generate within-participant beta values of stimulus contrasted against rest for each run. Mean beta value maps for each participant were generated by averaging beta values from all runs for a given task type. Cluster-extent based thresholding was performed in SPM12 (Wellcome Trust centre for Neuroimaging, London, UK) using the mean beta value maps for each participant as input. Cluster-extent based thresholds (k) were calculated with the Gaussian random field (GRF) method. The residual images were used to estimate intrinsic smoothness for the task vs. rest effects and the task contrast effects (e.g., OV vs. CV) using a primary voxel-level $p < 1 \times 10^{-4}$. One sample t-tests were calculated for the task vs. rest effects and paired t-tests were calculated for the task contrast effects, with $p < 0.05$ for each tail. As a total, 24 tails were assessed (positive and negative for each of the twelve comparisons for each of ΔHbO_2 and ΔHbR). Only clusters with $p < 0.0021$ (family wise error critical value, FWEc) were included as significant ($0.05 \div 24 = 2.1 \times 10^{-3}$).

To describe anatomical regions and Brodmann areas for each cluster, SPM outputs were entered into the WFU PickAtlas 3.0.5b (Maldjian et al., 2004, 2003) (Tables 2–5). For each region, we then also calculated the Cohen's-d to assess the effect size of the task response relative to rest, as well as for the contrasts between the paradigms. Finally, we estimated the temporal profile of the ΔHbO_2 contrast hemodynamic activity for each stimulus type within a variety of implicated regions of interest using a finite impulse response model (Glover, 1999; Hassanpour et al., 2014; Miezin et al., 2000). This procedure allowed us to both evaluate the time course of the evoked responses to each task type to ensure they were physiologically plausible and to highlight spatio-temporal characteristics differentiating brain function underlying these language tasks.

3. Results

3.1. Motion evaluation across paradigms

Motion was quantified using the time-point by time-point GVTD index. Fig. 2A-C demonstrates two GVTD time-traces and their corresponding histograms for a low motion (clean)run and a high motion (noisy) run for each of the three language paradigms. Additionally shown are the cumulative distribution for all runs for all subjects for each paradigm, with a different color line for each run. Insets show mean +/- STD across all clean runs for each paradigm with additional dashed traces for noisy runs that were excluded from further analyses. Red boxes show median GVTD values (horizontal line intersecting 50% normalized values, equivalent to the median value) for each paradigm and empirical threshold for the study (7.04×10^{-4}). To ensure equal levels of motion across three tasks, an empirically derived threshold of the median GVTD value per run of 7.04×10^{-4} was used to exclude runs with too much motion contamination from analysis (Fig. 2). In all, 11 runs were excluded due to high motion as quantified by median of the GVTD value of the run ($n = 4$ RW; $n = 3$ CV runs; $n = 4$ OV runs) (Fig. 2D). As assessed by the GVTD metric, though small but significant differences were observed between OV and each of RW and CV, the majority of data collected survived our empirically-derived threshold of 7.04×10^{-4} for median GVTD within a run (Fig. 2D). In order to determine whether GVTD changed as a function of task type, GVTD was averaged across scan sessions (2 RW, 3 CV, 3 OV) and across days for all subjects ($n = 33$; Fig. 2E). A repeated measures ANOVA

Table 2
ΔHbO_2 Task vs. rest effects.

Significant Clusters	*Brodmann Area*	*k (cm³)*	*T peak, mean*	*Z*	*X (mm)*	*Y (mm)*	*Z (mm)*
RW > Rest							
Middle Occipital Gyrus (L)	18	23.01	7.12, 5.51	5.47	−30	−94	−4
Middle Occipital Gyrus (R)	18	11.61	6.72, 5.07	5.27	35	−95	−20
Superior Temporal Gyrus (L)	22	1.217	6.49, 4.53	5.15	−70	−52	9
CV > Rest							
Middle Occipital Gyrus (L)	18	20.51	7.89, 5.59	5.84	−53	−87	6
Inferior Occipital Gyrus (R)	18	16.37	8.04, 5.59	5.90	33	−101	3
Middle Temporal Gyrus (L)	21	9.24	5.84, 4.99	4.78	−56	−51	−5
Inferior Frontal Gyrus (L)	45	8.91	6.83, 4.77	5.33	−43	24	13
CV < Rest							
Superior Temporal Gyrus (R)	39	1.57	4.95, 4.35	4.24	58	−67	25
OV > Rest							
Inferior Frontal Gyrus (L)	9	24.23	6.24, 5.00	5.01	−53	32	29
Inferior Occipital Gyrus (L)	18	10.72	6.96, 4.76	5.39	−48	−89	−22
Inferior Occipital Gyrus (R)	18	4.65	5.80, 4.73	4.76	29	−98	−14
Middle Temporal Gyrus (L)	37	4.21	5.68, 4.77	4.69	−53	−48	−17
Postcentral Gyrus (R)	43	3.34	5.35, 4.57	4.49	57	−10	17
Precentral Gyrus (L)	4	1.87	5.50, 4.65	4.58	−55	−13	51
Inferior Frontal Gyrus (R)	44	1.61	5.29, 4.56	4.45	63	10	14
Middle Frontal Gyrus (R)	6	1.59	5.60, 4.52	4.64	46	1	56
OV < Rest							
Superior Temporal Gyrus (R)	40	1.91	5.38, 4.41	4.50	63	−66	26

RW, Reading Words; CV, Covert Verb Generation; OV, Overt Verb Generation; L, left; R, right; *k,* Cluster-extent based threshold P Town.

Table 3
ΔHbR Task vs. rest effects.

Significant Clusters	*Brodmann Area*	*k (cm³)*	*T peak, mean*	*Z*	*X (mm)*	*Y (mm)*	*Z (mm)*
RW < REST							
Inferior Occipital Gyrus (L)	18	17.76	7.79, 5.29	5.79	−31	−95	−7
Inferior Occipital Gyrus (R)	18	6.92	6.35, 4.88	5.07	33	−80	−17
CV > REST							
Supramarginal Gyrus (R)	40	5.12	6.17, 4.80	4.97	64	−64	23
Inferior Parietal Lobule (R)	40	3.10	4.77, 5.56	4.62	59	−31	45
CV < REST							
Middle Occipital Gyrus (L)	18	23.19	8.73, 5.53	6.2	−32	−100	−8
Middle Occipital Gyrus (R)	18	18.35	7.38, 5.30	5.6	32	−100	18
Middle Frontal Gyrus (L)	9	20.34	6.53, 4.93	5.17	−48	37	30
Inferior Temporal Gyrus (L)	20	3.28	5.82, 4.66	4.77	−65	−60	−16
Superior Temporal Gyrus (L)	22	2.20	5.56, 4.64	4.62	−49	−40	5
OV < REST							
Middle Frontal Gyrus	9	42.05	11.48, 5.60	7.18	−70	−8	4
Precentral Gyrus (R)	6	19.51	8.14, 5.19	5.95	73	−17	5
Middle Occipital Gyrus (L)	18	18.57	6.91, 4.94	5.37	−48	−88	−22
Inferior Occipital Gyrus (R)	18	7.65	6.10, 4.68	4.93	35	−100	2
Inferior Temporal Gyrus (L)	37	4.67	6.15, 5.03	4.96	−64	−59	−11
Precentral Gyrus (L)	6	3.53	6.67, 4.95	5.24	−54	−1	52
Lingual Gyrus (R)	18	0.68	4.92, 4.44	4.22	6	−97	−19

RW, Reading Words; CV, Covert Verb Generation; OV, Overt Verb Generation; L, left; R, right; *k*, Cluster-extent based threshold.

Table 4
ΔHbO_2 Task vs. task effects.

Cluster-level			Peak-level					
Cluster	*Region*	*Brodmann Area*	*k (cm³)*	*T peak, mean*	*Z*	*X (mm)*	*Y (mm)*	*Z (mm)*
CV>RW	**Inferior Frontal Gyrus (L)**	**45**	5.97	6.90, 4.97	5.36	−51	25	13
CV>RW	**Middle Frontal Gyrus (L)**	**9**	0.71	5.43, 4.43	4.54	−65	16	26
CV<RW	**Supramarginal Gyrus (R)**	**40**	3.93	5.34, 4.56	4.48	48	−47	29
OV>RW	**Inferior Frontal Gyrus (L)**	**46**	9.15	5.91, 4.68	4.83	−60	25	21
OV>RW	**Precentral Gyrus (L)**	**44**	3.20	6.03, 4.54	4.89	−64	13	13
OV<RW	**Supramarginal Gyrus (R)**	**40**	3.38	4.94, 4.40	4.23	51	−43	24
OV>CV	**Postcentral Gyrus (R)**	**43**	1.04	4.88, 4.46	4.19	54	−17	19

RW, Reading Words; CV, Covert Verb Generation; OV, Overt Verb Generation; L, left; R, right; *k,* Cluster-extent based threshold.

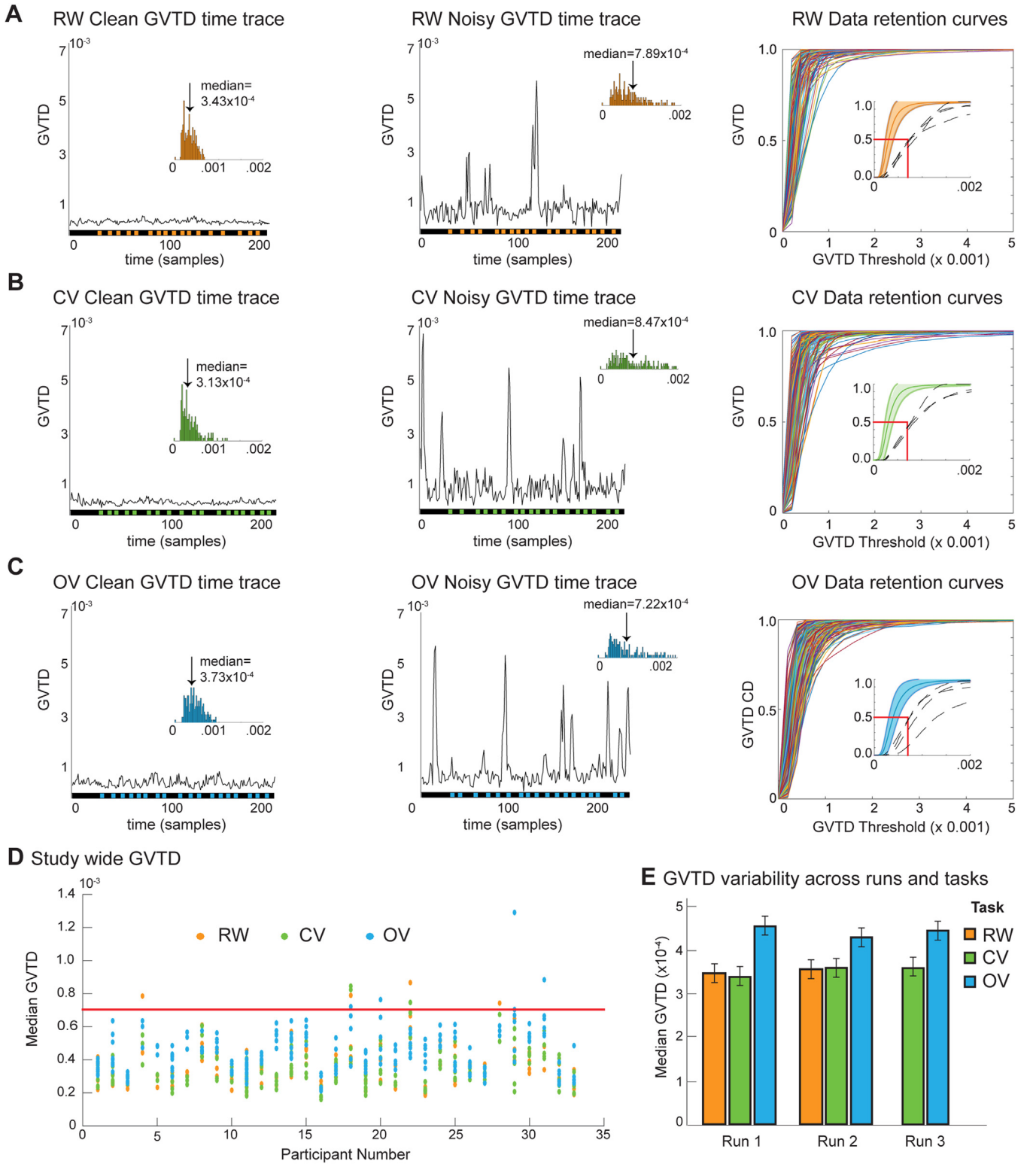

Fig. 2. Tracking motion artifacts across language paradigms. Representative individual time-traces and histograms (insets) of GVTD values for clean and noisy data acquired for representative (A) RW, (B) CV, and (C) OV runs. Black and colored boxes on x-axis denote the stimulus paradigm timings (10 Hz). Insets show mean +/- STD across all clean runs with additional dashed traces for noisy runs. Additionally, cumulative distributions for all runs for all subjects for each paradigm are shown, with a different color for each run to distinguish the overlapping lines. Red horizontal lines show median GVTD values for each paradigm (as it intersects 0.5 on the cumulative distributions) and red vertical lines show the empirical threshold for the study (7.04×10^{-4}) that was applied to the median values for each run. (D) Head motion was quantified by median GVTD for each participant and run. Red line represents empirically derived threshold (same as in A-C) to remove data from analyses. (E) GVTD variability across scan runs and tasks after exclusion of high motion runs. Collapsing task types across days revealed a main effect of task type. GVTD did not differ by scan run or as an interaction of scan run and task type. Error bars represent standard error of the mean [+/−1SE]).

Table 5
ΔHbR Task vs. task effects.

Cluster-level			Peak-level					
Cluster	*Region*	*Brodmann Area*	*k (cm³)*	*T peak, mean*	*Z*	*X (mm)*	*Y (mm)*	*Z (mm)*
CV>OV	**Precentral Gyrus (R)**	6	17.74	8.39, 5.33	6.06	71	−11	5
CV>OV	**Superior Temporal Gyrus (L)**	22	8.66	7.21, 5.10	5.52	−61	0	−3
CV<RW	**Middle Frontal Gyrus (L)**	9	13.00	6.38, 4.79	5.09	−58	28	21
CV>RW	**Supramarginal Gyrus (R)**	40	8.78	6.70, 5.09	5.26	48	−47	28
OV<RW	**Inferior Frontal Gyrus (L)**	9	28.35	9.22, 5.25	6.39	−70	−10	4
OV<RW	**Precentral Gyrus (R)**	6	12.58	7.89, 5.21	5.84	74	−18	5
OV<RW	**Precentral gyrus (L)**	3	0.94	5.11, 4.52	4.34	−51	−17	61
OV<RW	**Precentral Gyrus (R)**	4	0.55	5.20, 4.55	4.40	64	−8	45

RW, Reading Words; CV, Covert Verb Generation; OV, Overt Verb Generation; L, left; R, right; *k*, Cluster-extent based threshold.

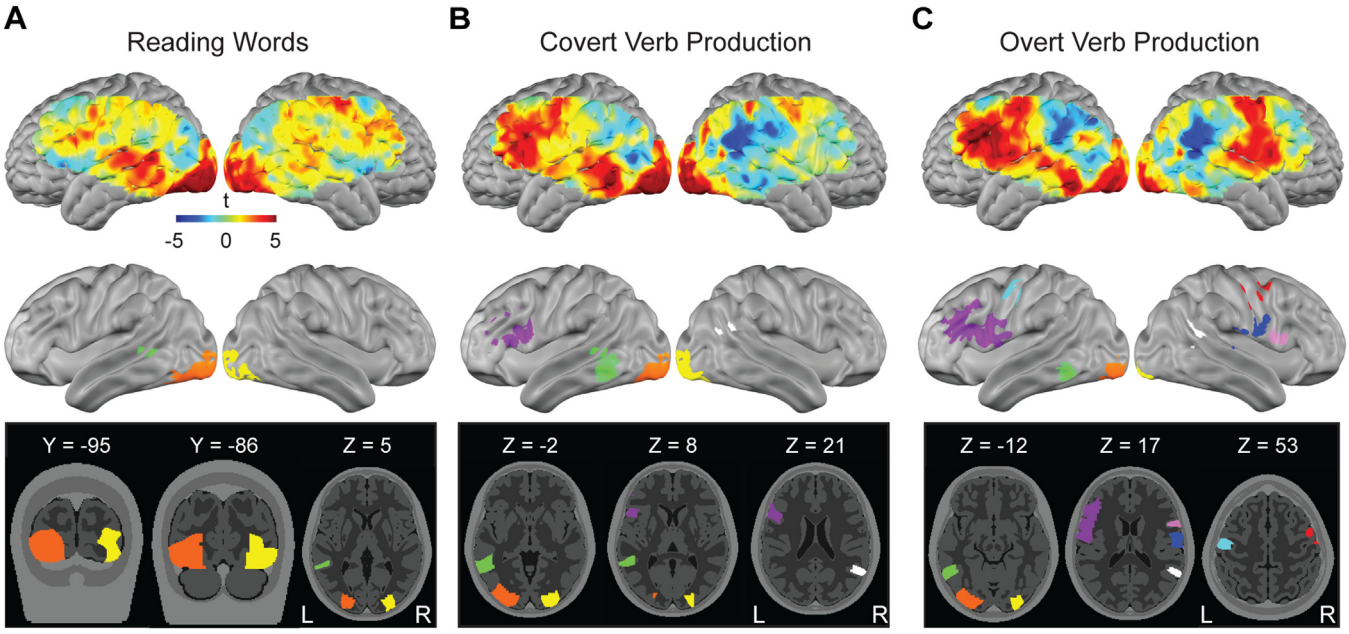

Fig. 3. Task-based responses for HbO_2. (A) Group-wise t-maps highlight spatially distributed regions recruited for responses to reading words. Statistically significant clusters shown on an inflated cortical surface and in selected slices through the volume with MNI-based coordinates for each slice. (B-C) Corresponding maps for covert verb production (B) and overt verb production (C).

with Greenhouse-Geisser correction determined that mean GVTD significantly varied across tasks $F(1.26,39.9) = 24.4$ ($p < 0.001$). *Post hoc* paired samples tests with Bonferroni correction revealed that CV (mean = 3.5×10^{-4}, STD = 1.0×10^{-4}) and RW (mean = 3.6×10^{-4}, STD = 1.0×10^{-4}) did not differ in GVTD $t(32) = 0.14$ ($p = 0.894$). However OV GVTD (mean = 4.4×10^{-4}, STD = 1.0×10^{-4}) was significantly greater than CV $t(32) = 6.49$ ($p < 0.001$) and RW $t(32) = 5.78$ ($p < 0.001$). (FWE $p < 0.05$) Additional *Post hoc* analyses confirmed that OV GVTD was greater than both RW and CV GVTD and this difference was greater than variability in GVTD across repeated scan runs within a session and across days (Fig. 2E, Supplemental Fig. S1).

3.2. Silent reading of words

Silently reading single nouns caused strong increases in oxygenated hemoglobin (ΔHbO_2) (Fig. 3A) and decreases in oxygenated hemoglobin (ΔHbR) (Fig. 4A) throughout bilateral occipital cortex as well as spatially distributed regions including left posterior temporal lobe and bilateral regions around dorsolateral prefrontal cortex. Stringent cluster-extent based thresholding at voxel-level $p < 1 \times 10^{-4}$ and cluster-level FWEc $p < 2.1 \times 10^{-3}$ yielded three clusters for RW > rest, with peaks located in left middle occipital gyrus (orange), right middle occipital gyrus (yellow), and left superior temporal gyrus (green) (Fig. 3A). No clusters passed significance for RW < rest for ΔHbO_2. Cluster-extent based thresholding results are given in Tables 2,3 with one anatomical region named for each significant cluster, corresponding with the peak-level coordinate for each cluster. Clusters are listed in order of decreasing size (*k*) for each task type.

3.3. Covert verb production

In response to covertly producing a verb for each noun, robust responses are again apparent in bilateral occipital regions along with a strong unilateral recruitment of the left temporal lobe and left prefrontal cortex (Figs. 3B and 4B). For ΔHbO_2 responses to CV > rest, statistically significant activations were located in four total clusters localized with peaks in left middle occipital gyrus (orange), right inferior occipital gyrus (yellow), left middle temporal gyrus (green), and left inferior frontal gyrus (purple) after the same stringent cluster-extent based thresholding as the RW task (Fig. 3B, Table 2). For CV < rest, one statistically significant cluster was found with a peak in right superior temporal gyrus (white) for ΔHbO_2 (Fig. 3B, Table 2). Complementary results are presented for the ΔHbR contrast in Fig. 4, Table 3.

3.4. Overt verb production

Successful responses to this paradigm required overt language generation or speaking aloud. In response to overtly producing a verb for

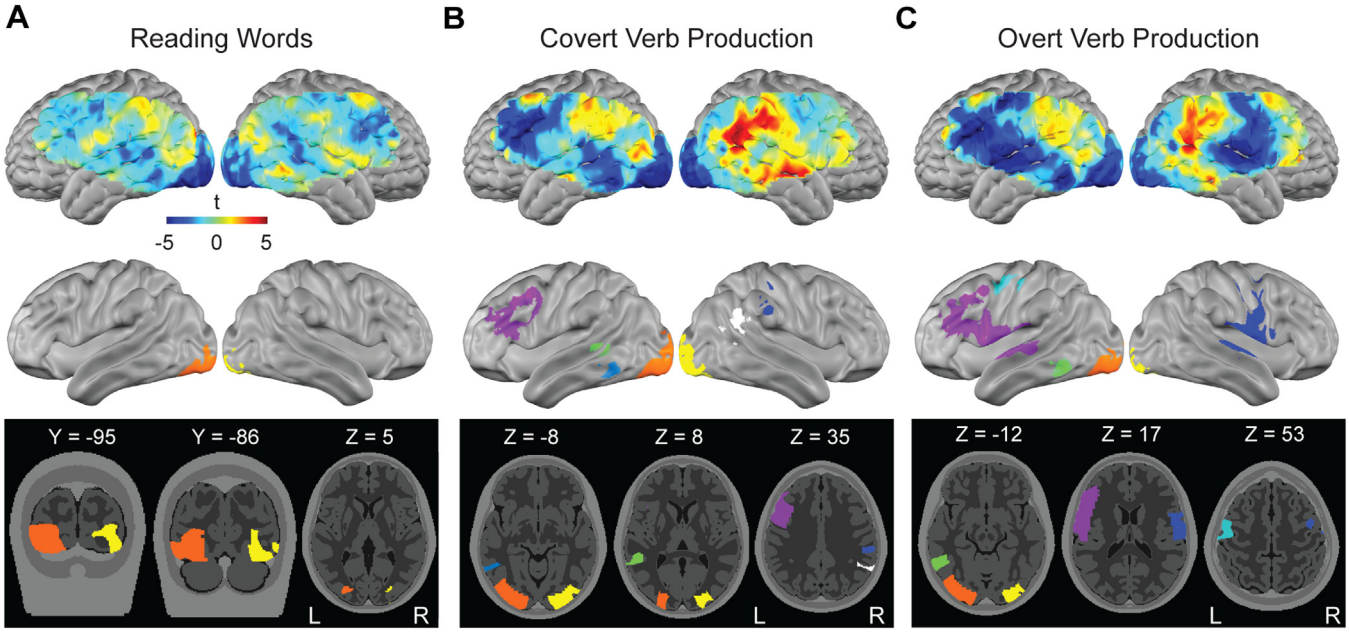

Fig. 4. Task-based responses for HbR. (A) Group-wise t-maps highlight spatially distributed regions recruited for responses to reading words. Statistically significant clusters shown on an inflated cortical surface and in selected slices through the volume with MNI-based coordinates for each slice. (B-C) Corresponding maps for covert verb production (B) and overt verb production (C).

each noun, robust responses are again apparent in bilateral occipital regions, left temporal lobe, and left prefrontal cortex, along with notable recruitment of bilateral motor-mouth regions (Figs. 3C, 4C). For ΔHbO_2 responses to OV > rest, eight statistically significant clusters were found after the same stringent cluster-extent based thresholding as the RW and CV tasks with peaks in left inferior frontal gyrus (purple), left inferior occipital gyrus (orange), right inferior occipital gyrus (yellow), left middle temporal gyrus (green), right postcentral gyrus (dark blue), left precentral gyrus (cyan), right inferior frontal gyrus (pink), and right middle frontal gyrus (red) (Fig. 3C, Table 2). The OV < rest contrast localized one statistically significant cluster in right superior temporal gyrus (white) (Fig. 3C, Table 2). Complementary results are presented for the ΔHbR contrast in Fig. 4, Table 3.

3.5. Contrasts between task paradigms

We identified regions showing significantly different activity between two given task types using cluster-extent based thresholding with the same parameters as the task vs. rest effects (primary $p < 1 \times 10^{-4}$ and cluster level FWEc $p < 2.1 \times 10^{-3}$). The comparison between activity underlying production of language relative to passive word reading allows for differentiation of higher-level language processing regions from general visual processing regions. Compared to passive reading, both covert and overt production of words led to significant increases in oxygenated hemoglobin (ΔHbO_2) (Fig. 5A, Table 4) and decreases in oxygenated hemoglobin (ΔHbR) (Fig. 6A, Table 5) . Specifically, for the ΔHbO_2 CV > RW contrast, two clusters were localized with peaks located in the left inferior frontal gyrus (purple, Cohen's $d = 0.93$) and left middle frontal gyrus (dark pink, Cohen's $d = 0.77$) (Fig. 5A). Additionally, the OV > RW contrast yielded two clusters, with peaks in left inferior frontal gyrus (purple, Cohen's $d = 0.98$) and left precentral gyrus (brown, Cohen's $d = 1.09$) (Fig. 5B). Further, we found significant decreases in activity relative to passive reading while generating language either overtly or covertly in right supramarginal gyrus (Fig. 5A-B). Strong effect sizes were observed for both CV < RW (white cluster in Fig. 5A, Cohen's $d = 1.38$) and for OV < RW (white cluster in Fig. 5B, Cohen's $d = 1.09$). We then identified regions showing significantly different activity between covert and overt production of language (Fig. 5C, Table 4). We found one region with a significant increase in activity for OV > CV (dark blue cluster Fig. 5C, Cohen's $d = 0.96$) with the peak in right postcentral gyrus (Fig. 5C). No clusters survived cluster-extent based thresholding at $p > 0.0001$ for OV < CV. Complementary results are presented for the ΔHbR contrast in Fig. 6, Table 5.

3.6. Temporal characteristics of responses

To examine the temporal profile of the ΔHbO_2 responses across the group of $n = 33$ participants, we extracted the mean hemodynamic responses evoked for each task type at selected clusters of interest from Figs. 3 and 5 (Fig. 7). The hemodynamic response starts with a delay between 2 and 5 s from the start of the event block and peaks around 4–6 s after the end of the stimulus block (5 s). These responses serve as a face-valid quality control to check and verify that the responses recorded with HD-DOT are valid hemodynamic patterns and not just spurious results due to noise.

3.7. Motion artifact evaluation in OV responses

In addition to quantifying the expected higher motion levels in the OV runs compared to RW and CV runs (Fig. 2E), we evaluated the Pearson correlation between the median GVTD of each OV run and the magnitude of the mean beta value in each of the nine significant clusters in the ΔHbO_2 responses to the OV task across $n = 33$ participants. Our results indicated no significant correlations in the magnitudes of the mean beta value in these nine clusters and the median GVTD value of that run (Table 6).

4. Discussion

Language is a complex cognitive process exclusive to humans that integrates several levels of linguistic analysis (phonemes, semantics, morphemes, clauses, etc.), social cues, and processing modes (reading, writing, listening, and speaking). Despite this complexity, the stimuli most often chosen for language experiments largely lack similarity to how language is used naturally. Those that have used naturalistic stimuli have found that the neural response to language is much more

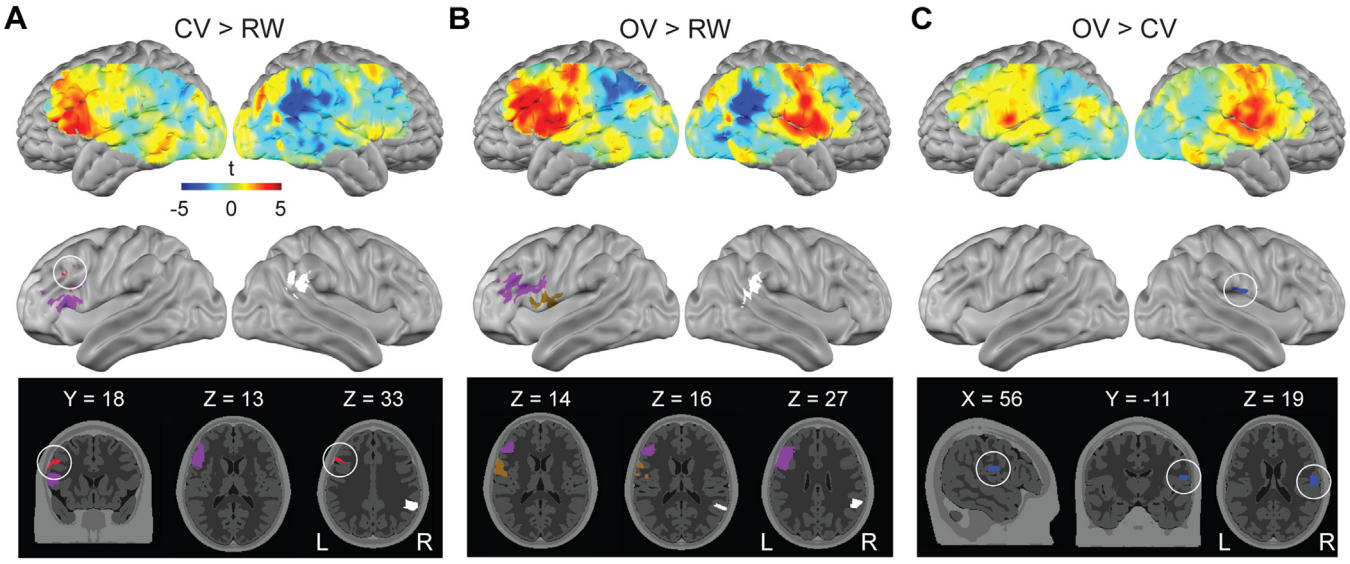

Fig. 5. Task contrast maps for ΔHbO_2. (A) Group-wise t-maps highlight spatially distributed regions recruited for responses to covert production of verbs compared to passive reading of words. Statistically significant clusters are shown on an inflated cortical surface and in selected slices through the volume with MNI-based coordinates for each slice. (B-C) Corresponding maps for overt verb production compared to passive reading of words (B) and overt vs. covert verb production (C). Spatially smaller clusters are highlighted with a white circle.

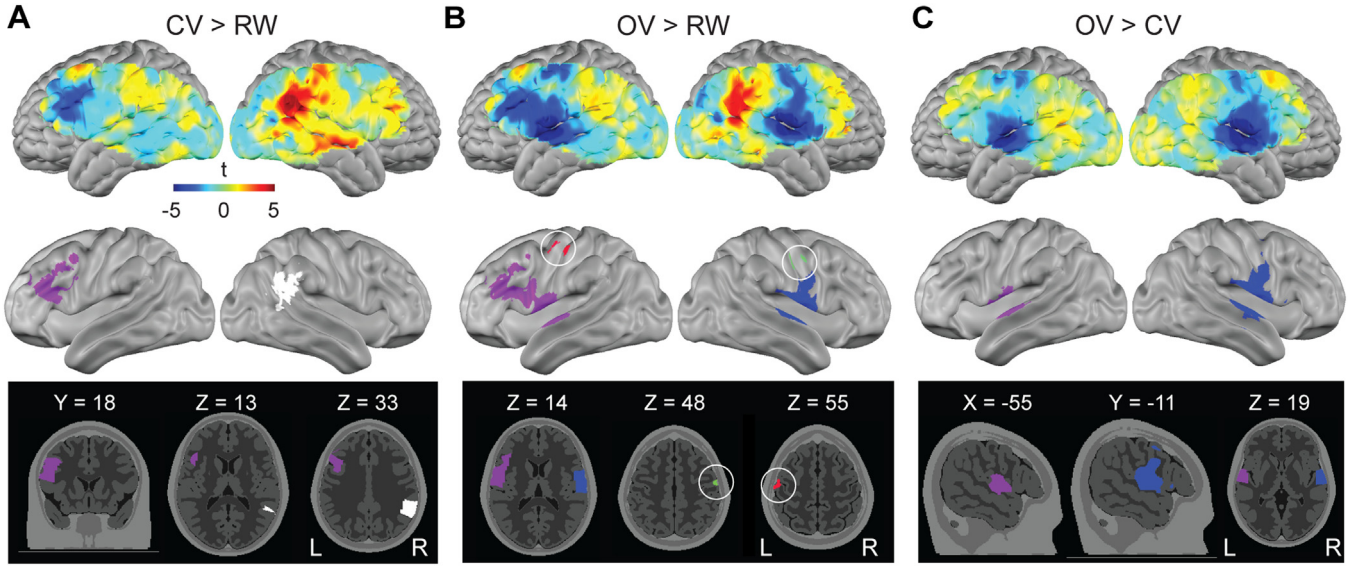

Fig. 6. Task contrast maps for ΔHbR. (A) Group-wise t-maps highlight spatially distributed regions recruited for responses to covert production of verbs compared to passive reading of words. Statistically significant clusters are shown on an inflated cortical surface and in selected slices through the volume with MNI-based coordinates for each slice. (B-C) Corresponding maps for overt verb production compared to passive reading of words (B) and overt vs. covert verb production (C). Spatially smaller clusters are highlighted with a white circle.

Table 6
Correlations between median GVTD and mean beta in each OV ΔHbO_2 significant cluster.

Cluster	*r-value*	*p-value*
Inferior Frontal Gyrus (L)	0.06	0.70
Inferior Occipital Gyrus (L)	−0.13	0.46
Inferior Occipital Gyrus (R)	−0.28	0.11
Middle Temporal Gyrus (L)	−0.03	0.83
Postcentral Gyrus (R)	−0.11	0.51
Precentral Gyrus (L)	0.10	0.54
Inferior Frontal Gyrus (R)	−0.15	0.37
Middle Frontal Gyrus (R)	−0.23	0.19
Superior Temporal Gyrus (R)	0.11	0.53

widespread and less left-lateralized than studies of non-naturalistic language brain function have suggested (Cogan et al., 2014; Hamilton and Huth, 2018; Lerner et al., 2011). While naturalistic language experiments can be very efficient for testing multiple hypotheses at once without having to specify categories ahead of time (e.g., syntactic, semantic, spectrotemporal, phonemic, etc.) (de Heer et al., 2017; Hamilton and Huth, 2018; Wehbe et al., 2014), they are not commonly used in neurolinguistics studies mainly due to the limitations of the current gold standard neuroimaging modalities and their incompatibility with social environments. In this study, we aimed to assess the feasibility of HD-DOT, a silent, non-invasive, portable, and potentially wearable modality for mapping the differential brain responses to three tasks: reading nouns, covert verb production, and overt verb production. We first quan-

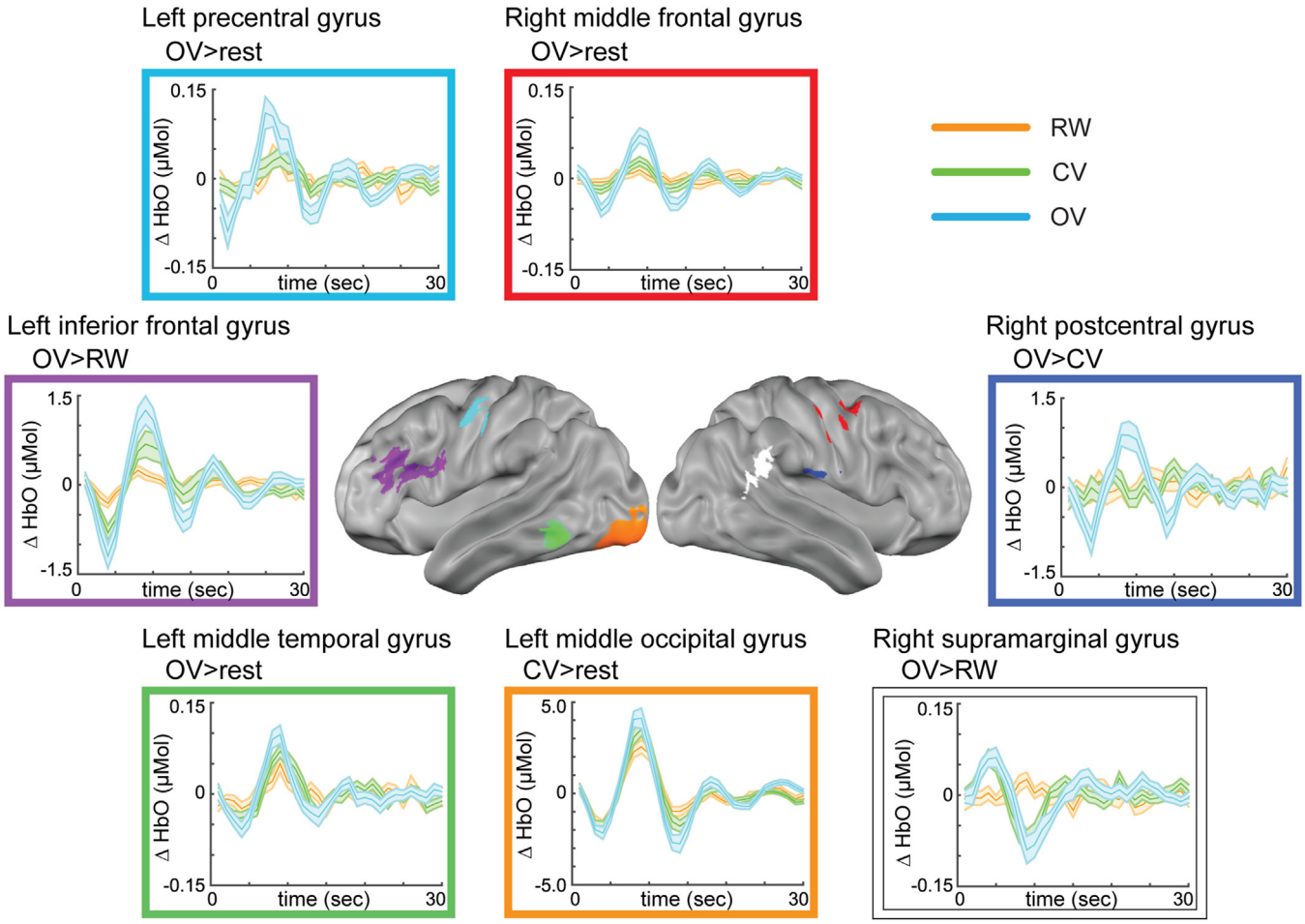

Fig. 7. Temporal profiles of the ΔHbO_2 hemodynamic response (oxygenated hemoglobin) for each task contrasted with rest for a selection of clusters, with shaded areas representing the standard error of the mean across participants. The color of the box outline for a given set of time traces corresponds to the color of the associated cluster on the cortical surface.

tified the levels of motion in all 3 task responses using a previously validated motion detection index for HD-DOT called GVTD (Sherafati et al., 2017, 2020). We also showed that HD-DOT successfully localizes the brain regions that respond during single word perception as well as during covert and overt verb production. Our findings are in general agreement with a set of previously localized regions responding to the perception and production of language based on previous PET and fMRI studies (Bohland and Guenther, 2006; Eickhoff et al., 2009; Fedorenko et al., 2015; Petersen et al., 1988, 1990; Wise et al., 1999) that include portions of the superior temporal cortex, inferior and middle frontal cortex, supramarginal gyrus and precentral and postcentral gyri. These regions showed statistically significant responses either in task vs. rest effects or task vs. task effects after cluster-extent based thresholding (see Tables 2–5 for details). Finally, we demonstrated the resilience of HD-DOT in tolerating the levels of motion artifacts during overt language production by performing a correlation analysis between the magnitudes of the median GVTD value of each OV run and the mean beta value in each significant cluster in the OV response.

Below we discuss these results in more detail, note their limitations, and explore some future potential applications of HD-DOT for mapping naturalistic real-life language processing and pre-surgical language assessment in light of our present findings.

4.1. Resilience of HD-DOT to motion due to overt speaking

Currently, fMRI is the gold standard neuroimaging tool for localizing the components of language network specially prior to neurosurgeries (Black et al., 2017). Overt speaking is known to induce artifacts in fMRI data due to the involvement of the respiratory system and the moving articulatory organs such as tongue, lips, and jaw (Barch et al., 1999; Birn et al., 1998; Friston et al., 1996; Huang et al., 2008; Kemeny et al., 2005). Thus, *covert* language production tasks are commonly used to avoid motion artifacts (Black et al., 2017; Hesling et al., 2019; Tzourio-Mazoyer et al., 2016). However, using covert paradigms does not allow localizing of specific cortical regions that are recruited during overt language production. Here, we demonstrated that HD-DOT is resilient to the level of motion artifacts generated during overt speaking. We quantified the level of motion across three tasks (RW, CV, and OV) using median GVTD values for each subject. We found that, although the OV task, results in a slightly higher median compared to RW and CV tasks (Fig. 2E), this difference was below the threshold for spurious motion-induced artifact as reflected by a lack of significant correlations between the median GVTD values of each OV run and the magnitudes of the ΔHbO_2 response activations in each significant cluster (Table 6). More importantly, the number of the OV runs excluded due to motion was comparable to the RW and CV tasks (four RW runs, three CV runs, and four OV runs were removed due to high median GVTD values for the run) (Fig. 2D).

4.2. Cortical regions responding to language comprehension

All three tasks required language comprehension through the reading of single words. The impetus for including the RW task was to distinguish between the cortical responses to language comprehension

and production that were present in the CV and OV tasks. We successfully localized bilateral activations spanning visual cortices, especially around the left lingual gyrus (Petersen et al., 1988, 1990), an expected result from reading words across three tasks. Beside the activations in the visual cortex due to reading of the words, receptive language processing is typically localized around left superior temporal gyrus (STG), middle temporal gyrus (MTG), or even supramarginal or angular gyri (Barch et al., 1999; Demonet et al., 2005; Dronkers et al., 2004b; Geschwind, 1970; Shuster and Lemieux, 2005; Tremblay and Dick, 2016; Wernicke, 1881) although it is worth noting that the anatomical regions identified as Wernicke's area are variable across the literature (Bogen and Bogen, 1976; Tremblay and Dick, 2016). As expected, our results show a cluster in these areas (Figs. 3,4) across all task > rest contrasts, with magnitudes peaking in left STG for RW and left MTG for CV and OV.

4.3. Cortical activations in response to covert language production

Beside the expected left and right visual cortex activations due to reading words, we found a significant cluster in response to CV in left MTG. We also found a cluster in the right STG which did not survive our stringent cluster-extent based thresholding (Table 2). This result is in agreement with the hypothesis that the areas around the STG (especially left STG) are recruited when we prepare to produce speech output (Basilakos et al., 2018) or during phonological planning (Hickok, 2009). An additional interesting point is that the location of the left STG cluster putatively associated with language comprehension descends inferiorly in CV compared to the RW task, approaching the visual word form area in the left fusiform gyrus as verb production is added to the word perception (Cohen et al., 2000; Dehaene et al., 2002) (c.f. (Howard et al., 1992; Moore and Price, 1999; Price and Devlin, 2003). This is not surprising given that the processing required by the two tasks are different. The RW task required participants to simply read the words, while CV required participants to form a concrete semantic representation in order to generate a related verb simultaneously with reading words. This result is also consistent with the lemma model of lexical selection and form encoding (Levelt, 1999), which posits phonological code retrieval in posterior STG (consistent with our RW task results) (Acheson et al., 2011; Pillay et al., 2014; Hamberger et al., 2016; Binder, 2017) and lemma retrieval in more inferior areas of the temporal lobe (consistent with our CV task results) (Indefrey and Levelt, 2004; Baldo et al., 2013).

We also found a statistically significant cluster localized in the left IFG in CV > rest that supports previous findings that this region supports language production (Basilakos et al., 2018; Flinker et al., 2015). The exact role of the left IFG in language production is still debated. However, it is most often associated with resolving conflicts between multiple co-activated lemmas (Schnur et al., 2009; Nozari and Pinet, 2020), phonological processing such as syllabification (Indefrey and Levelt, 2004), the translation of phonological representations into an articulatory code (Flinker et al., 2015; Hickok, 2012; Hickok and Poeppel, 2007), and the implementation of the articulatory code via manipulation of the articulators (Caplan, 1992; Dronkers, 1996; Trupe et al., 2013). Our results highlight the fact that left IFG is recruited even when overt speaking or articulation is not performed. It also emphasizes that this region is involved in language production rather than just mediating the information flow from the temporal region (phonological planning) to motor regions (overt articulation) (Hickok, 2009).

Beside the exact *role* of the IFG in language production, there is debate in the literature about the exact *location* and *extent* of this region across different individuals (Fedorenko et al., 2013, 2010). Therefore, depending on the sample or type of analysis this region could be found at slightly different locations or even disappear in stringent thresholding of group results (Basilakos et al., 2018). However, we did identify significant clusters in the IFG in both CV > rest and CV > RW effects as well as a significant cluster in the MFG in CV > RW effect, which underscores the recruitment of this region in the differential response between covert verb production and reading words.

Another interesting observation is the bilateral activations in the motor cortex in response to the CV task (evident in the unthresholded CV vs. rest t-maps in Figs. 3B, 4B). However, these activations did not survive cluster-extent based thresholding (Figs. 3B, 4B). This could be speculated to be related to imagined speaking as the participants were instructed to imagine saying verbs out loud for the CV task without actually moving their articulators.

4.4. Cortical activations in response to overt language production

As we hypothesized, overt production of verbs (OV > rest) recruited all cortical regions that responded to RW and CV tasks (left and right occipital cortex, left MTG, right STG, and left IFG) in addition to four new regions (right IFG, right MFG, left precentral gyrus, and right postcentral gyrus) (Figs. 3,4). The successful localization of the left MTG underscores its role in mapping sounds to meanings for overt speech (Hickok, 2009).

Interestingly, unlike CV > rest, we found the recruitment of the right IFG and MFG in response to the OV task. This is in agreement with previous fMRI studies identifying the role of both left and right prefrontal cortex (especially IFG) in response to articulation (Basilakos et al., 2018).

It is also worth noting that the cluster around the left and right IFG and MFG (especially the left IFG identified in OV > rest) might seem to extend into the dorsolateral prefrontal cortex, an area known to be activated during cognitive demand (Duncan, 2010). While the left IFG is primarily associated with domain-specific language processing (Bogen and Bogen, 1976; Burton et al., 2000; Friederici, 2002; Hagoort, 2005; Mesulam et al., 2015; Palmer et al., 2001; Sahin et al., 2009), dorsolateral prefrontal cortex is associated with several domain-general functions (such as working memory) that facilitate novel problem solving (Fedorenko et al., 2013, 2012). Although our cluster-extent based thresholding accurately shows that the peaks of these clusters are located in IFG and MFG and are not overlapping with the dorsolateral prefrontal cortex (Basilakos et al., 2018), we did not explore the exact overlap between the single-subject language and multiple demand ROIs. Further analyses of these data may explore such possibilities.

We also found a strong bilateral activation in the left and right pre- and post-central cortices in the un-thresholded OV t-map (Figs. 3C,4C). After cluster-extent based thresholding we found two clusters in these regions with their peaks located at the left precentral gyrus and right postcentral gyrus, suggesting their role in the motor aspects of speech production. This result is in accordance with the recent findings that regions in the precentral cortex are exclusively contributing to the motor aspects of articulation, as they are shown to be engaged during both speech and non-speech oral movements (Basilakos et al., 2018).

As for the differential response between OV and RW, our results localized the left IFG and MFG and areas around the left precentral and right postcentral gyri (Figs. 5,6). However, the only two ΔHbO_2 response clusters that survived the OV vs. RW effect after cluster-extent based thresholding were left IFG and left precentral gyrus (Fig. 5).

4.5. Differential responses underlying covert vs. overt language production

One of the main aims of this study was to demonstrate the feasibility of HD-DOT in localizing the differential cortical responses to covert vs. overt verb production. As explained in the previous section, we identified that similar to CV, left STG and left IFG were recruited during the OV task. Crucially, we found four additional clusters that contributed exclusively to OV (i.e., right IFG, right MFG, left precentral gyrus and right postcentral gyrus) (Fig. 3C). Out of these four regions, right postcentral gyrus was the only region that showed a significant effect in OV > CV analysis after stringent cluster-extent based thresholding (Fig. 5C). This cluster located in Brodmann area 43, has been found to respond to tympanic membrane pressure changes (Job et al., 2011). This finding

suggests that while it is tempting to hypothesize that the largest differences between CV and OV language tasks might be related to higher order cognition and executive control, we cannot ignore the embodied nature of language and its integration with somatosensory responses. Another interpretation of this result is that this cluster unique to OV but not CV is localized in cortical regions that might be associated with the auditory and/or somatosensory state, target, and/or error maps outlined in the DIVA model (Directions Into Velocities of Articulators; an adaptive neural network that describes the sensorimotor interactions involved in articulator control during speech production) of speech motor control (Guenther, 2016). These results highlight the fact that while the other three regions beside the right postcentral gyrus (i.e., right IFG, right MFG, and left precentral gyrus) are playing an exclusive role during OV > rest (and not CV > rest), they might not be linked to the key differences distinguishing overt and covert language production.

Our results also confirm the validity of the original hypothesis that there are significant differences between the brain's response to CV and OV tasks that extend through motor and frontal cortices, and thus these two tasks cannot be used interchangeably.

4.6. Cortical deactivations in response to language production

There was an additional and remarkably robust ΔHbO_2 response deactivation pattern located around the right posterior STG in the CV > rest and OV > rest contrasts (Fig. 3) and in the neighboring right supramarginal gyrus in the CV > RW and OV > RW contrasts (Fig. 5). These two regions are parts of the default mode network and are known to be suppressed during cognitively challenging tasks (Palmer et al., 2001; Shulman et al., 1997). It is interesting to note that participants in general reported finding the CV task more challenging than the RW task, and the OV task more challenging than the CV task. While this is only anecdotal evidence, our results support this hypothesis reflecting that these regions are indeed being suppressed during higher cognitive loads (CV and OV) but not suppressed during lower cognitive loads (RW).

4.7. Temporal profile of the hemodynamic response

Studying the temporal aspects of the brain's response to different paradigms is arguably as important as its spatial aspects. Although, investigating the temporal dynamics of the cortical responses to language was not the focus of the current study, we plotted the ΔHbO_2 hemodynamic responses across all three tasks (Fig. 7) to confirm that the hemodynamic response function pattern was evident in all the significant clusters and they were not just false positive results due to noise (Aguirre et al., 1998). Also important to note is that the temporal dynamics of both fMRI and HD-DOT hemodynamic responses are an order of magnitude slower than the ECoG responses that have recently been shown to distinguish processing stages of language generation (Anumanchipalli et al., 2019a; Flinker et al., 2015). However, future studies leveraging the broad spatial coverage of HD-DOT combined with other potential advancements in optical sampling of brain function such as frequency-modulated HD-DOT systems (Doulgerakis et al., 2019; Fan et al., 2021) show exciting promise in non-invasive measurement of brain function.

4.8. Limitations

This study explored the neural correlates of naturalistic language using overt single verb production, which is more naturalistic than common covert language tasks but still not representative of language in the real world aside from social word-association games like Taboo or Catchphrase. Although single word production is arguably a key aspect of naturalistic learning of language especially in the developing brains, future studies are needed to map the neural responses to higher levels of linguistic analysis in more naturalistic real-life settings or in different age groups and social contexts. In this study, we used the GVTD motion index to quantify and exclude high motion runs. Future studies can be done to evaluate the effect of motion censoring or motion correction on covert and overt language production HD-DOT data. The field of view of the HD-DOT instrument used in this study is limited to a depth of about 15–20 mm and does not include the entirety of the motor, frontal, and subcortical areas. It is important to note that areas in the cerebellum and supplementary motor area are also known to contribute to speech production (Basilakos et al., 2018); however, they were outside of our field of view. As the HD-DOT hardware continues to develop (e.g., denser and expanded arrays (Tripathy et al., 2021b), frequency domain systems (Blaney et al., 2020), wearable wireless caps (Uchitel et al., 2022; Vidal-Rosas et al., 2021), etc.), we expect that imaging truly naturalistic language tasks with a more expansive field of view and temporal resolution will be possible. Further, we analyzed our data using voxel-wise significance threshold of $p < 0.0001$ that errs on the stringent side. The unthresholded t-maps make it apparent that there is signal in some regions that did not survive cluster-extent based thresholding. Although our group results showed sufficient significance in all the expected cortical regions known to contribute to language perception and production that were in our field of view, we should note that a voxel-wise functional correspondence across individuals was assumed. Future studies could adopt a functional localization approach by identifying the single-subject language regions of interest (ROIs) prior to group analysis (Fedorenko et al., 2010; Saxe et al., 2006). As this is an incredibly rich dataset, we are planning on releasing the full set of raw optical measurement data for the community to further analyze to explore additional observables present in the total hemoglobin contrasts. Additionally, while the t-maps present some level of variability across the group, the data collection schedule of multiple runs per session over two sessions provides opportunities for deeper within and across-participant analyses of these data in a future study. Further, it is possible that brain activations due to producing language, either covertly or overtly, that are action words in response to concrete nouns may lead to differential responses in parts of somato-motor networks or other locations due to varying levels of embodied language. While beyond the scope of the current study, the uniquely quiet and natural environment of HD-DOT may provide a compelling opportunity for future investigations on brain function underlying embodied language. Last, the relationship between the variations in cortical activity as a function of behavioral performance (e.g., number of verbs generated on the OV task) is another interesting analysis beyond the scope of this paper that may be investigated in future studies.

5. Conclusions

In this study, we not only established the resilience of HD-DOT to the level of motion due to overt language production, but we also replicated the cortical activations underlying language comprehension (covert word reading task) and language production (covert and overt verb generation tasks). Our results are consistent with previous fMRI, PET, and ECoG studies of the language network. Our sample size of 33 subjects and localization of all key components of language comprehension and production in our field of view, despite a stringent cluster-extent based thresholding method, confirms that HD-DOT has a comparable signal-to-noise ratio to fMRI. These findings suggest that HD-DOT can be used as a method of choice for presurgical language assessment even with overt language production tasks instead of the gold-standard fMRI. Furthermore, HD-DOT's use of an open scanning environment with seated participants allows for studying language and communication in a more naturalistic setting than with supine participants alone inside the bore of an MRI or PET scanner or in a surgery room. Given recent demonstrations of optical brain mapping techniques for decoding activity in visual cortex (Tripathy et al., 2021a), and functional imaging approaches for decoding language throughout superficial cortex (Tang et al., 2023), the sensitivity and specificity to overt and covert language production demonstrated in the current study points to

exciting future directions in non-invasive measurement and use of human language processing, including for brain-machine interfaces. HD-DOT is more conducive to imaging various populations for whom other neuroimaging modalities are challenging or contraindicated, e.g., small children, individuals with implanted metal, and those who are claustrophobic. Thus, HD-DOT is well-positioned to aid researchers to capitalize on using naturalistic language tasks and scanning environments to explore brain function in real-life settings such as during face-to-face dialogues, as a BCI, or during early childhood development.

Funding

This work was supported by the National Institutes of Health (grant number K01-MH103594, R01-MH122751, and R21-MH109775 to A.T.E.; R01NS090874, U01EB02700501, and R01NS109486 to J.P.C.).

Declaration of Competing Interest

The authors declare no competing interests.

Credit authorship contribution statement

Mariel L. Schroeder: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration. **Arefeh Sherafati:** Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Rachel L. Ulbrich:** Investigation, Data curation, Writing – original draft, Project administration. **Muriah D. Wheelock:** Formal analysis, Writing – original draft. **Alexandra M. Svoboda:** Investigation, Data curation, Writing – original draft. **Emma D. Klein:** Conceptualization, Writing – original draft. **Tessa G. George:** Formal analysis, Visualization. **Kalyan Tripathy:** Investigation, Writing – original draft. **Joseph P. Culver:** Resources, Supervision, Funding acquisition. **Adam T. Eggebrecht:** Conceptualization, Methodology, Software, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank Sean Rafferty for assistance with cap fits, Katie Mansfield, Libby Forsen, and Meghana Bhimasani for their efforts on audio response transcription, Anna Abbacchi for her consultation on quantifying interrater reliability, and Jessica Church and Tim Brown for guidance on using number of years of education to estimate full scale IQ. We are also extremely grateful for the time and effort expended by all those who participated in this study. Last, the first author would like to thank all the friends and family that provided support during the development of this manuscript, especially in the form of childcare, as the bulk of the analysis and writing was performed during the COVD-19 pandemic.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120190.

References

Acheson, D.J., Hamidi, M., Binder, J.R., Postle, B.R., 2011. A common neural substrate for language production and verbal working memory. J. Cogn. Neurosci. 23 (6), 1358–1367.

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360–369.

Allendorfer, J.B., Kissela, B.M., Holland, S.K., Szaflarski, J.P., 2012. Different patterns of language activation in post-stroke aphasia are detected by overt and covert versions of the verb generation fMRI task. Med. Sci. Monit. 18, CR135–CR137.

Anumanchipalli, G.K., Chartier, J., Chang, E.F., 2019a. Speech synthesis from neural decoding of spoken sentences. Nature 568, 493–498.

Anumanchipalli, G.K., Chartier, J., Chang, E.F., 2019b. Speech synthesis from neural decoding of spoken sentences. Nature 568, 493–500.

Baldo, J.V., Arévalo, A., Patterson, J.P., Dronkers, N.F., 2013. Grey and white matter correlates of picture naming: Evidence from a voxel-based lesion analysis of the Boston Naming Test. Cortex 49, 658–667.

Barch, D.M., Sabb, F.W., Carter, C.S., Braver, T.S., Noll, D.C., Cohen, J.D., 1999. Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. Neuroimage 10, 642–657.

Basilakos, A., Smith, K.G., Fillmore, P., Fridriksson, J., Fedorenko, E., 2018. Functional characterization of the human speech articulation network. Cereb. Cortex 28, 1816–1830.

Benjamin, C.F., Dhingra, I., Li, A.X., Blumenfeld, H., Alkawadri, R., Bickel, S., Helmstaedter, C., Meletti, S., Bronen, R.A., Warfield, S.K., 2018. Presurgical language fMRI: technical practices in epilepsy surgical planning. Hum. Brain Mapp. 39, 4032–4042.

Biasiucci, A., Leeb, R., Iturrate, I., Perdikis, S., Al-Khodairy, A., Corbet, T., Schnider, A., Schmidlin, T., Zhang, H., Bassolino, M., Viceic, D., Vuadens, P., Guggisberg, A.G., Millan, J.D.R., 2018. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. Nat. Commun. 9, 2421.

Billing, A.D., Cooper, R.J., Scott, S.K., 2021. Pre-SMA activation and the perception of contagiousness and authenticity in laughter sounds. Cortex 143, 57–68.

Binder, J.R., 2017. Current controversies on Wernicke's area and its role in language. Curr. Neurol. Neurosci. Rep. 17, 58.

Birn, R.M., Bandettini, P.A., Cox, R.W., Jesmanowicz, A., Shaker, R., 1998. Magnetic field changes in the human brain due to swallowing or speaking. Magn. Reson. Med. 40, 55–60.

Black, D., Vachha, B., Mian, A., Faro, S., Maheshwari, M., Sair, H., Petrella, J., Pillai, J., Welker, K., 2017. American society of functional neuroradiology–recommended fMRI paradigm algorithms for presurgical language assessment. Am. J. Neuroradiol. 38, E65–E73.

Blanco, B., Molnar, M., Carreiras, M., Collins-Jones, L.H., Rosas, E.E.V., Cooper, R.J., Caballero-Gaudes, C.J.N., 2021. Group-level cortical functional connectivity patterns using fNIRS: assessing the effect of bilingualism in young infants. Neurophotonics 8, 025011.

Blaney, G., Sassaroli, A., Fantini, S.J.O.l., 2020. Dual-slope imaging in highly scattering media with frequency-domain near-infrared spectroscopy. 45, 4464–4467.

Bluestone, A., Abdoulaev, G., Schmitz, C., Barbour, R., Hielscher, A., 2001. Three-dimensional optical tomography of hemodynamics in the human head. Opt. Express 9, 272–286.

Bogen, J.E., Bogen, G.M., 1976. Wernicke's region–Where is it? Ann. N. Y. Acad. Sci. 280, 834–843.

Bohland, J.W., Guenther, F.H., 2006. An fMRI investigation of syllable sequence production. NeuroImage 32, 821–841.

Brown, T.T., Lugar, H.M., Coalson, R.S., Miezin, F.M., Petersen, S.E., Schlaggar, B.L., 2005. Developmental changes in human cerebral functional organization for word generation. Cereb. Cortex 15, 275–290.

Brysbaert, M., New, B., 2009. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav. Res. Methods 41, 977–990.

Burton, M.W., Small, S.L., Blumstein, S.E., 2000. The role of segmentation in phonological processing: an fMRI investigation. J. Cogn. Neurosci. 12, 679–690.

Caplan, D., 1992. Language: Structure, processing, and Disorders. The MIT Press, Cambridge, MA.

Church, J.A., Balota, D.A., Petersen, S.E., Schlaggar, B.L., 2011. Manipulation of length and lexicality localizes the functional neuroanatomy of phonological processing in adult readers. J. Cogn. Neurosci. 23, 1475–1493.

Church, J.A., Coalson, R.S., Lugar, H.M., Petersen, S.E., Schlaggar, B.L., 2008. A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. Cereb. Cortex 18, 2054–2065.

Cogan, G.B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., Pesaran, B., 2014. Sensory–motor transformations for speech occur bilaterally. Nature 507, 94–98.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., Michel, F., 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123 (Pt 2), 291–307.

Collinger, J.L., Wodlinger, B., Downey, J.E., Wang, W., Tyler-Kabara, E.C., Weber, D.J., McMorland, A.J., Velliste, M., Boninger, M.L., Schwartz, A.B., 2013. High-performance neuroprosthetic control by an individual with tetraplegia. Lancet 381, 557–564.

de Heer, W.A., Huth, A.G., Griffiths, T.L., Gallant, J.L., Theunissen, F.E., 2017. The hierarchical cortical organization of human speech processing. J. Neurosci. 37, 6539–6557.

Dehaene, S., Le Clec, H.G., Poline, J.B., Le Bihan, D., Cohen, L., 2002. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. Neuroreport 13, 321–325.

Demonet, J.F., Thierry, G., Cardebat, D., 2005. Renewal of the neurophysiology of language: functional neuroimaging. Physiol. Rev. 85, 49–95.

Doulgerakis, M., Eggebrecht, A.T., Dehghani, H., 2019. High-density functional diffuse optical tomography based on frequency-domain measurements improves image quality and spatial resolution. Neurophotonics 6, 035007.

Dronkers, N.F., 1996. A new brain region for coordinating speech articulation. Nature 384, 159–161.
Dronkers, N.F., Ogar, J., Willock, S., Wilkins, D.P., 2004a. Confirming the role of the insula in coordinating complex but not simple articulatory movements. Brain Lang. 91, 23–24.
Dronkers, N.F., Wilkins, D.P., Van Valin Jr., R.D., Redfern, B.B., Jaeger, J.J., 2004b. Lesion analysis of the brain areas involved in language comprehension. Cognition 92, 145–177.
Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. 14, 172–179.
Eggebrecht, A.T., Culver, J.P., 2019. NeuroDOT: An extensible Matlab toolbox for streamlined optical functional mapping. Clinical and Preclinical Optical Diagnostics II. Optical Society of America, Munich.
Eggebrecht, A.T., Ferradal, S.L., Robichaux-Viehoever, A., Hassanpour, M.S., Dehghani, H., Snyder, A.Z., Hershey, T., Culver, J.P., 2014. Mapping distributed brain function and networks with diffuse optical tomography. Nat. Photonics 8, 448–454.
Eggebrecht, A.T., White, B.R., Ferradal, S.L., Chen, C., Zhan, Y., Snyder, A.Z., Dehghani, H., Culver, J.P., 2012. A quantitative spatial comparison of high-density diffuse optical tomography and fMRI cortical mapping. Neuroimage 61, 1120–1128.
Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K.J.P.T.o.t.R.S.A.M., Physical, Sciences, E., 2009. A systems perspective on the effective connectivity of overt speech production. 367, 2399–2421.
Fan, W., Dehghani, H., Eggebrecht, A.T., 2021. Investigation of effect of modulation frequency on high-density diffuse optical tomography image quality. Neurophotonics 8, 045002.
Fedorenko, E., Duncan, J., Kanwisher, N., 2013. Broad domain generality in focal regions of frontal and parietal cortex. Proc. Natl. Acad. Sci. U. S. A. 110, 16616–16621.
Fedorenko, E., Duncan, J., Kanwisher, N., 2012. Language-selective and domain-general regions lie side by side within Broca's area. Curr. Biol. 22, 2059–2062.
Fedorenko, E., Fillmore, P., Smith, K., Bonilha, L., Fridriksson, J., 2015. The superior precentral gyrus of the insula does not appear to be functionally specialized for articulation. J. Neurophysiol. 113, 2376–2382.
Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., Kanwisher, N., 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. 104, 1177–1194.
Ferradal, S.L., Eggebrecht, A.T., Hassanpour, M., Snyder, A.Z., Culver, J.P., 2014. Atlas-based head modeling and spatial normalization for high-density diffuse optical tomography: *in vivo* validation against fMRI. Neuroimage 85 (Pt 1), 117–126.
Ferradal, S.L., Liao, S.M., Eggebrecht, A.T., Shimony, J.S., Inder, T.E., Culver, J.P., Smyser, C.D., 2016. Functional imaging of the developing brain at the bedside using diffuse optical tomography. Cereb. Cortex 26, 1558–1568.
Flinker, A., Korzeniewska, A., Shestyuk, A.Y., Franaszczuk, P.J., Dronkers, N.F., Knight, R.T., Crone, N.E., 2015. Redefining the role of Broca's area in speech. Proc. Natl. Acad. Sci. U. S. A. 112, 2871–2875.
Fonov, V.S., Evans, A.C., McKinstry, R.C., Almli, C., Collins, D., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage S102.
Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.
Frijia, E.M., Billing, A., Lloyd-Fox, S., Rosas, E.V., Collins-Jones, L., Crespo-Llado, M.M., Amadó, M.P., Austin, T., Edwards, A., Dunne, L., 2021. Functional imaging of the developing brain with wearable high-density diffuse optical tomography: a new benchmark for infant neuroimaging outside the scanner environment. NeuroImage 225, 117490.
Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.
Geschwind, N., 1970. The organization of language and the brain. Science 170, 940–944.
Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9, 416–429.
Gregg, N.M., White, B.R., Zeff, B.W., Berger, A.J., Culver, J.P., 2010. Brain specificity of diffuse optical imaging: improvements from superficial signal regression and tomography. Front Neuroenergetics 2.
Guenther, F.H., 2016. Neural control of speech. MIT Press, Cambridge, MA.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.
Hamberger, M.J., Miozzo, M., Schevon, C.A., Morrison, C., Carlson, C., Mehta, A.D., Klein, G.E., McKhann 2nd, G.M., Williams, A.C., 2016. Functional differences among stimulation-identified cortical naming sites in the temporal region. Epilepsy Behav. 60, 124–129.
Hamilton, L.S., Huth, A.G., 2018. The revolution will not be controlled: natural stimuli in speech neuroscience. Lang. Cogn. Neurosci. 35, 573–582.
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.
Hassanpour, M.S., Eggebrecht, A.T., Peelle, J.E., Culver, J.P., 2017. Mapping effective connectivity within cortical networks with diffuse optical tomography. Neurophotonics 4, 041402.
Hassanpour, M.S., White, B.R., Eggebrecht, A.T., Ferradal, S.L., Snyder, A.Z., Culver, J.P., 2014. Statistical analysis of high density diffuse optical tomography. Neuroimage 85 (Pt 1), 104–116.
Hesling, I., Labache, L., Joliot, M., Tzourio-Mazoyer, N., 2019. Large-scale plurimodal networks common to listening to, producing and reading word lists: an fMRI study combining task-induced activation and intrinsic connectivity in 144 right-handers. Brain Struct. Funct. 224, 3075–3094.
Hickok, G., 2012. Computational neuroanatomy of speech production. Nat. Rev. Neurosci. 13, 135–145.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.
Hickok, G., 2009. The functional neuroanatomy of language. Phys. Life Rev. 6, 121–143.
Hirsch, J., Tiede, M., Zhang, X., Noah, J.A., Salama-Manteau, A., Biriotti, M., 2021. Interpersonal agreement and disagreement during face-to-face dialogue: an fNIRS investigation. Front. Hum. Neurosci. 14, 606397.
Howard, D., Patterson, K., Wise, R., Brown, W.D., Friston, K., Weiller, C., Frackowiak, R., 1992. The cortical localization of the lexicons. Positron emission tomography evidence. Brain 115 (Pt 6), 1769–1782.
Huang, J., Francis, A.P., Carr, T.H., 2008. Studying overt word reading and speech production with event-related fMRI: a method for detecting, assessing, and correcting articulation-induced signal changes and for measuring onset time and duration of articulation. Brain Lang. 104, 10–23.
Ihnen, S.K., Church, J.A., Petersen, S.E., Schlaggar, B.L., 2009. Lack of generalizability of sex differences in the fMRI BOLD activity associated with language processing in adults. Neuroimage 45, 1020–1032.
Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. Cognition 92, 101–144.
Job, A., Paucod, J.C., O'Beirne, G.A., Delon-Martin, C., 2011. Cortical representation of tympanic membrane movements due to pressure variation: an fMRI study. Hum. Brain Mapp. 32, 744–749.
Kemeny, S., Ye, F.Q., Birn, R., Braun, A.R., 2005. Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. Hum. Brain Mapp. 24, 173–183.
Kuhl, P.K., Ramirez, R.R., Bosseler, A., Lin, J.F., Imada, T., 2014. Infants' brain responses to speech suggest analysis by synthesis. Proc. Natl. Acad. Sci. U. S. A. 111, 11238–11245.
Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U., 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J. Neurosci. 31, 2906–2915.
Levelt, W.J.M., 1999. Producing spoken language: A blueprint of the speaker. In: Brown, C.M., Hagoort, P. (Eds.), *The neurocognition of language*. Oxford University Press, New York, pp. 83–122.
Li, Y., Tang, C., Lu, J., Wu, J., Chang, E.F., 2021. Human cortical encoding of pitch in tonal and non-tonal languages. Nat. Commun. 12, 1161.
Makin, J.G., Moses, D.A., Chang, E.F., 2020. Machine translation of cortical activity to text with an encoder-decoder framework. Nat. Neurosci. 23, 575–582.
Maldjian, J.A., Laurienti, P.J., Burdette, J.H., 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. Neuroimage 21, 450–455.
Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239.
Matarazzo, J.D., 1972. Wechsler's Measurement and Appraisal of Adult Intelligence. Williams & Wilkins, Baltimore.
Matarazzo, J.D., Herman, D.O., 1984. Relationship of education and IQ in the WAIS—R standardization sample. J. Consult. Clin. Psychol. 52, 631.
Mesulam, M.M., Thompson, C.K., Weintraub, S., Rogalski, E.J., 2015. The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. Brain 138, 2423–2437.
Metzger, S.L., Liu, J.R., Moses, D.A., Dougherty, M.E., Seaton, M.P., Littlejohn, K.T., Chartier, J., Anumanchipalli, G.K., Tu-Chan, A., Ganguly, K., 2022. Generalizable spelling using a speech neuroprosthesis in an individual with severe limb and vocal paralysis. Nat. Commun. 13, 1–15.
Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E., Buckner, R.L., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11, 735–759.
Moore, C.J., Price, C.J., 1999. Three distinct ventral occipitotemporal regions for reading and object naming. Neuroimage 10, 181–192.
Moses, D.A., Leonard, M.K., Makin, J.G., Chang, E.F., 2019. Real-time decoding of question-and-answer speech dialogue using human cortical activity. Nat. Commun. 10, 1–14.
Nozari, N., Pinet, S., 2020. A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. J. Neurolinguistics 53, 100875.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.
Palmer, E.D., Rosen, H.J., Ojemann, J.G., Buckner, R.L., Kelley, W.M., Petersen, S.E., 2001. An event-related fMRI study of overt and covert word stem completion. Neuroimage 14, 182–193.
Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., Raichle, M.E., 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331, 585–589.
Petersen, S.E., Fox, P.T., Snyder, A.Z., Raichle, M.E., 1990. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. Science 249, 1041–1044.
Pillay, S.B., Stengel, B.C., Humphries, C., Book, D.S., Binder, J.R., 2014. Cerebral localization of impaired phonological retrieval during rhyme judgment. J. Neurol. 76, 738–746.
Pinti, P., Aichelburg, C., Gilbert, S., Hamilton, A., Hirsch, J., Burgess, P., Tachtsidis, I., 2018. A review on the use of wearable functional near-infrared spectroscopy in naturalistic environments. Jpn. Psychol. Res. 60, 347–373.
Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., Burgess, P.W., 2020. The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. Ann. N. Y. Acad. Sci. 1464, 5–29.
Price, C.J., Devlin, J.T., 2003. The myth of the visual word form area. Neuroimage 19, 473–481.
Richardson, J.D., Fillmore, P., Rorden, C., LaPointe, L.L., Fridriksson, J., 2012. Re-establishing Broca's initial findings. Brain Lang. 123, 125–130.

Sahin, N.T., Pinker, S., Cash, S.S., Schomer, D., Halgren, E., 2009. Sequential processing of lexical, grammatical, and phonological information within Broca's area. Science 326, 445–449.

Saxe, R., Brett, M., Kanwisher, N.J.N., 2006. Divide and conquer: a defense of functional localizers. 30, 1088–1096.

Sherafati, A., Eggebrecht, A.T., Burns-Yocum, T.M., Culver, J.P., 2017. A global metric to detect motion artifacts in optical neuroimaging data (Conference Presentation). Neural Imaging and Sensing. SPIE, pp. 108.

Schnur, T.T., Schwartz, M.F., Kimberg, D.Y., Hirshorn, E., Coslett, H.B., Thompson-Schill, S.L., 2009. Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. Proc. Natl. Acad. Sci. 106, 322–327.

Sherafati, A., Snyder, A.Z., Eggebrecht, A.T., Bergonzi, K.M., Burns-Yocum, T.M., Lugar, H.M., Ferradal, S.L., Robichaux-Viehoever, A., Smyser, C.D., Palanca, B.J., Hershey, T., Culver, J.P., 2020. Global motion detection and censoring in high-density diffuse optical tomography. Hum. Brain Mapp. 41, 4093–4112.

Shulman, G.L., Corbetta, M., Buckner, R.L., Fiez, J.A., Miezin, F.M., Raichle, M.E., Petersen, S.E., 1997. Common blood flow changes across visual tasks: I. Increases in subcortical structures and cerebellum but not in nonvisual cortex. J. Cogn. Neurosci. 9, 624–647.

Shuster, L.I., Lemieux, S.K., 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang. 93, 20–31.

Smyser, C.D., Inder, T.E., Shimony, J.S., Hill, J.E., Degnan, A.J., Snyder, A.Z., Neil, J.J., 2010. Longitudinal analysis of neural network development in preterm infants. Cereb. Cortex 20, 2852–2862.

Tang, J., LeBel, A., Jain, S., Huth, A.G., 2023. Semantic reconstruction of continuous language from non-invasive brain recordings. Nat. Neurosci. 26, 858–866.

Tremblay, P., Dick, A.S., 2016. Broca and Wernicke are dead, or moving past the classic model of language neurobiology. Brain Lang. 162, 60–71.

Tripathy, K., Markow, Z.E., Fishell, A.K., Sherafati, A., Burns-Yocum, T.M., Schroeder, M.L., Svoboda, A.M., Eggebrecht, A.T., Anastasio, M.A., Schlaggar, B.L., Culver, J.P., 2021a. Decoding visual information from high-density diffuse optical tomography neuroimaging data. Neuroimage 226, 117516.

Tripathy, K., Markow, Z.E., Fishell, A.K., Sherafati, A., Burns-Yocum, T.M., Schroeder, M.L., Svoboda, A.M., Eggebrecht, A.T., Anastasio, M.A., Schlaggar, B.L., 2021b. Decoding visual information from high-density diffuse optical tomography neuroimaging data. NeuroImage 226, 117516.

Trupe, L.A., Varma, D.D., Gomez, Y., Race, D., Leigh, R., Hillis, A.E., Gottesman, R.F., 2013. Chronic apraxia of speech and Broca's area. Stroke 44, 740–744.

Tyndall, A.J., Reinhardt, J., Tronnier, V., Mariani, L., Stippich, C., 2017. Presurgical motor, somatosensory and language fMRI: technical feasibility and limitations in 491 patients over 13 years. Eur. Radiol. 27, 267–278.

Tzourio-Mazoyer, N., Joliot, M., Marie, D., Mazoyer, B., 2016. Variation in homotopic areas' activity and inter-hemispheric intrinsic connectivity with type of language lateralization: an FMRI study of covert sentence generation in 297 healthy volunteers. Brain Struct. Funct. 221, 2735–2753.

Uchitel, J., Blanco, B., Vidal-Rosas, E., Collins-Jones, L., Cooper, R.J., 2022. Reliability and similarity of resting state functional connectivity networks imaged using wearable, high-density diffuse optical tomography in the home setting. NeuroImage 263, 119663.

Van Heuven, W.J., Mandera, P., Keuleers, E., Brysbaert, M., 2014. SUBTLEX-UK: a new and improved word frequency database for British English. Q. J. Exp. Psychol. 67, 1176–1190.

Vidal-Rosas, E.E., Zhao, H., Nixon-Hill, R.W., Smith, G., Dunne, L., Powell, S., Cooper, R.J., Everdell, N.L.J.N., 2021. Evaluating a new generation of wearable high-density diffuse optical tomography technology via retinotopic mapping of the adult visual cortex. 8, 025002.

Wehbe, L., Murphy, B., Talukdar, P., Fyshe, A., Ramdas, A., Mitchell, T., 2014. Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. PLoS ONE 9, e112575.

Wernicke, C., 1881. Lehrbuch Der Gehirnkrankheiten Für Aerzte Und Studirende. Fischer.

White, B.R., Culver, J.P., 2010. Phase-encoded retinotopy as an evaluation of diffuse optical neuroimaging. NeuroImage 49, 568–577.

Wise, R.J., Greene, J., Büchel, C., Scott, S.K., 1999. Brain regions involved in articulation. Lancet 353, 1057–1061.

Yi, H.G., Leonard, M.K., Chang, E.F., 2019. The encoding of speech sounds in the superior temporal gyrus. Neuron 102, 1096–1110.