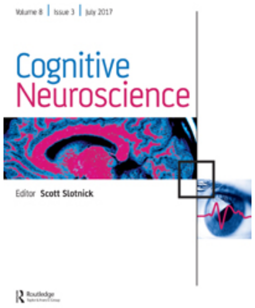

Cognitive Neuroscience

ISSN: 1758-8928 (Print) 1758-8936 (Online) Journal homepage: https://www.tandfonline.com/loi/pcns20

A new fun and robust version of an fMRI localizer for the frontotemporal language system

Terri L. Scott, Jeanne Gallée & Evelina Fedorenko

To cite this article: Terri L. Scott, Jeanne Gallée & Evelina Fedorenko (2017) A new fun and robust version of an fMRI localizer for the frontotemporal language system, Cognitive Neuroscience, 8:3, 167-176, DOI: 10.1080/17588928.2016.1201466

To link to this article: https://doi.org/10.1080/17588928.2016.1201466

COGNITIVE NEUROSCIENCE, 2017
VOL. 8, NO. 3, 167–176
https://doi.org/10.1080/17588928.2016.1201466

EMPIRICAL REPORTS

A new fun and robust version of an fMRI localizer for the frontotemporal language system

Terri L. Scott [a], Jeanne Gallée [b] and Evelina Fedorenko [c,d]

[a]Boston University, Graduate Program for Neuroscience, Boston, MA, USA; [b]Department of Cognitive and Linguistic Sciences, Wellesley College, Wellesley, MA, USA; [c]Massachusetts General Hospital, Harvard Medical School, Department of Psychiatry, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA; [d]Massachusetts Institute of Technology, McGovern Institute for Brain Research, Cambridge, MA, USA

ABSTRACT
A set of brain regions in the frontal, temporal, and parietal lobes supports high-level linguistic processing. These regions can be reliably identified in individual subjects using fMRI, by contrasting neural responses to meaningful and structured language stimuli vs. stimuli matched for low-level properties but lacking meaning and/or structure. We here present a novel version of a language 'localizer,' which should be suitable for diverse populations including children and/or clinical populations who may have difficulty with reading or cognitively demanding tasks. In particular, we contrast responses to auditorily presented excerpts from engaging interviews or stories, and acoustically degraded versions of these materials. This language localizer is appealing because it uses (a) naturalistic and engaging linguistic materials, (b) auditory presentation, (c) a passive listening task, and can be easily adapted to new stimulus materials enabling comparisons of language activation in children and speakers of diverse languages.

ARTICLE HISTORY
Received 2 March 2016
Revised 8 June 2016
Published online
7 July 2016

KEYWORDS
Language; fMRI; individual subject analyses; brain imaging methods; functional localizers

Introduction

The use of functional localizers to define regions of interest (ROIs) in functional magnetic resonance imaging (fMRI) studies enables researchers to more accurately target functionally distinct regions of the brain that are difficult or impossible to delineate using anatomical markers alone. This is because the relationship between functional activations and macroanatomical landmarks (sulci and gyri) is highly variable across individuals, especially in the association cortices (Fischl et al., 2008; Frost & Goebel, 2012; Tahmasebi et al., 2012). Combining anatomical information and functional responses allows researchers to be more confident that they are examining the *same* brain region across individuals, studies, and labs, as well as to delve deeply into the human cognitive architecture within each individual brain (Laumann et al., 2015). This approach—long used in the field of vision research (Julian, Fedorenko, Webster, & Kanwisher, 2012; Kanwisher, McDermott, & Chun, 1997) and more recently extended to other domains, including speech/voice perception (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Hickok, Okada, & Serences, 2009) and higher-level language processing (Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010)—has helped to functionally characterize a large number of brain regions, narrowing down the space of possible hypotheses about the computations they support. For example, using the functional localization approach, it has been shown that there exist language-responsive areas of the brain that are highly selective for linguistic processing, and these areas are distinct from areas that support other high-level cognitive processes, such as working memory, inhibitory cognitive control, arithmetic processing or music perception (Fedorenko, Behr, & Kanwisher, 2011; Fedorenko, Duncan, & Kanwisher, 2012).

The language localizer described in Fedorenko et al. (2010) contrasts sentences and sequences of pronounceable nonwords (e.g., FLORP) presented visually (using rapid serial visual presentation) or auditorily. This contrast targets word-level processing

CONTACT Terri L. Scott tlscott@bu.edu Boston University, Graduate Program for Neuroscience, 635 Commonwealth Avenue, Rm. 352, Boston, MA 02215, USA
The supplemental data for this article can be accessed here.
The underlying research materials for this article can be accessed at https://evlab.mit.edu/papers/Scott_CogNeuro

and higher-level semantic/syntactic processing and is robust to changes in task (Fedorenko, 2014; Fedorenko et al., 2010). The baseline condition (nonword lists) controls for lower-level visual/acoustic processes. This localizer has been successfully used in a number of studies (Axelrod, Bar, Rees, & Yovel, 2015; Blank, Kanwisher, & Fedorenko, 2014; Fedorenko et al., 2011, 2012; Koster-Hale & Saxe, 2011; Maruyama, Pallier, Jobert, Sigman, & Dehaene, 2012; Xu et al., 2015). However, this paradigm is somewhat limited in that it is not well suited for children or certain clinical populations. The visual (reading) version is obviously not suitable for illiterate populations, and more generally, regardless of the presentation modality, processing single sentences is not very engaging, especially for populations that have general difficulties with maintaining attention. Failure to maintain attention is likely to lead to weaker activations, in part due to increased motion. Furthermore, rapid reading is a fairly cognitively demanding task. Comparisons of language activation between clinical populations and typical subjects during rapid serial visual presentation may be confounded due to disproportionate task difficulty unrelated to language processing.

Here we present an alternative version of a language localizer, which (i) uses auditory presentation, (ii) involves engaging naturalistic materials, and (iii) requires only passive listening. Moreover, this localizer is easily adaptable to different age groups or speakers of different languages, by simply replacing the stimuli with a new set (e.g., taken from age- and language-appropriate audio books or based on custom recordings). We validate this new localizer against the more traditional *sentences* > *nonwords* contrast introduced in Fedorenko et al. (2010). To foreshadow the key result, the new localizer quickly (as short as ~6 minutes per run, although we recommend administering two runs, which allows for the estimation of response magnitudes) and robustly identifies the high-level language-processing regions at the individual subject level.

Materials and methods

Participants

Eleven participants (five females) between the ages of 18 and 30—students at MIT and members of the surrounding community—were paid for their participation. Participants were right-handed native speakers of English, naïve to the purposes of the study. All participants gave informed consent in accordance with the requirements of the MIT's Committee on the Use of Humans as Experimental Subjects (COUHES).

Design, materials and procedure

Each participant performed the standard visual language localizer task (Fedorenko et al., 2010) and the novel auditory language localizer task. Five participants completed both localizer tasks in the same scanning session; the remaining participants completed the localizers during separate scanning sessions. Prior work has shown that responses to the standard language localizer are highly stable within individual subjects both within and across scanning sessions (Mahowald & Fedorenko, 2016). Participants also completed a few additional tasks for unrelated studies. Each scanning session lasted approximately 2 hours and typically included approximately 20 minutes of structural data collection (including diffusion tensor imaging), 10 minutes of resting state fMRI, and 40–50 minutes of task-based fMRI.

Standard language localizer task

Participants read sentences (e.g., BARBARA WAS CONCERNED ABOUT THE SURGERY WHICH THE DENTIST HAD RECOMMENDED YESTERDAY) and lists of unconnected pronounceable nonwords (e.g., DEVIVED U SOLLTERCIDED PURQUIFECT CRE SHRY EXTINORE DELUNTOR CRE NIS OLP CRE) in a blocked design. Each stimulus consisted of 12 words/nonwords. For details of how the language materials were constructed, see Fedorenko et al. (2010). The materials are available at http://web.mit.edu/evelina9/www/funcloc/funcloc_localizers.html. Stimuli were presented in the center of the screen, one word/nonword at a time, at the rate of 450 ms per word/nonword. Participants were prompted to press a button at the end of each sequence by a visually displayed hand icon, which stayed on the screen for 400 ms. Additionally, a 100 ms duration blank screen appeared at the beginning of each trial and after the hand image, so that each trial lasted a total of 6 seconds. The button press task was included to confirm that participants were paying attention. Earlier versions of this localizer included a memory probe at the end of each trial or no task at all (passive reading), and it was shown that activations are similar regardless of task (Fedorenko, 2014; Fedorenko et al., 2010).

Condition order was counterbalanced across runs and participants. Experimental blocks lasted 18 seconds (with three trials per block) and fixation blocks lasted 14 seconds. Each run (consisting of 5 fixation blocks and 16 experimental blocks) lasted 358 seconds. Each participant completed two runs.

Novel auditory language localizer task

Participants listened to excerpts from speeches and talks from several sources (e.g., The Moth podcast, TED talks, celebrity interviews) and acoustically degraded versions of those excerpts. All materials for this localizer, as well as the script used to present materials, and our subjects' activation maps are available at https://evlab.mit.edu/papers/Scott_CogNeuro. Thirty-two items were created (where an item is an intact and degraded pair of stimuli) and distributed across two experimental lists following the standard Latin square design, so that each list contained only one version of an item. Any given participant was presented with one of the lists. The texts of the materials are included in the Supplemental Information. Degraded speech clips were created from the intact versions using the procedure described below. The resulting clips sounded like muffled speech, where it is no longer possible to discern the content (see Discussion for a discussion of possible alternative control conditions). Each clip lasted between 16 and 18 seconds. Participants were told that they would listen to some fun audio clips and some clips that have been distorted in a way that makes it impossible to understand what the speaker is saying. They were instructed to listen attentively. Prior to the experiment, it was ensured that the volume level was sufficiently loud yet comfortable.

Condition order was counterbalanced across runs and participants. Each block consisted of a single intact or degraded clip. Clips that were less than 18 seconds long were padded with silence at the end so that each block was exactly 18 seconds in duration. Fixation blocks were 14 seconds long. The structure of runs mirrored that of the visual language localizer, with each run consisting of 5 fixation blocks and 16 experimental blocks, for a total run duration of 358 seconds. All participants completed two runs, except for one participant who completed one run.

Audio degradation procedure

All audio processing was done using MATLAB (The Mathworks, Nattick, MA). The intact and degraded audio clips used here were 16-bit wave files with a sampling rate of 44,100 Hz. To create the degradation effect, we first created a low-pass filtered copy of each of the 32 audio clips, using a pass-band frequency of 500 Hz. We also created a noise track from the intact clip by randomizing the time-points (~0.02 ms long each). The noise track was multiplied by the amplitude envelope of the intact clip to produce variations in the volume of the noise. We then low-pass filtered the noise track in order to 'soften' the highest frequencies using a pass-band frequency of 8,000 Hz and a stop frequency of 10,000 Hz. Artifacts at the beginning and end of the noise tracks were removed and the noise track was added to the low-pass filtered copies of the clips. The level of the noise track was adjusted by listening to a few clips with different levels of noise, choosing a single value for the level of noise which rendered the clips unintelligible and applying the same level of noise to all of the degraded clips. The resulting degraded versions sound like poor radio reception of speech, but are indecipherable.

fMRI data acquisition

Structural and functional data were collected on the whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 128 axial slices with 1.33 mm isotropic voxels (TR = 2,000 ms, TE = 3.39 ms). Functional, blood oxygenation level dependent (BOLD), data were acquired using an EPI sequence (with a 90° flip angle and using GRAPPA with an acceleration factor of 2), with the following acquisition parameters: thirty-one 4 mm thick near-axial slices acquired in the interleaved order (with 10% distance factor), 2.1 mm × 2.1 mm in-plane resolution, FoV in the phase encoding (A ≫ P) direction 200 mm and matrix size 96 × 96, TR = 2,000 ms and TE = 30 ms. The first 10 s of each run were excluded to allow for steady state magnetization. Sound was delivered via scanner-safe headphones (Sensimetrics, Malden, MA).

fMRI data preprocessing

MRI data were analyzed using SPM5 and custom MATLAB scripts (available—in the form of an SPM toolbox—from http://www.nitrc.org/projects/spm_ss). Each participant's data were motion corrected and then normalized into a common brain space (the Montreal Neurological Institute (MNI) template) and resampled into 2 mm isotropic voxels. The data were then smoothed with a 4 mm Gaussian filter and high-pass filtered (at 200 s). For both of the localizer tasks, effects were estimated using a General Linear Model (GLM) in which each experimental condition was modeled with a boxcar function convolved with the canonical hemodynamic response function (HRF).

Group-constrained Subject-Specific (GSS) analysis

For all the analyses, regions of interest were defined functionally in each individual participant using the *sentences* > *nonwords* contrast in the standard visual language localizer and the *intact* > *degraded* contrast in the novel auditory language localizer. To do so, we used the Group-constrained Subject-Specific (GSS) analysis method developed in Fedorenko et al. (2010) and Julian et al. (2012). In particular, functional regions of interest (fROIs) were constrained to fall within a set of functional 'masks' which indicated the expected gross locations of activations for the *sentences* > *nonwords* contrast and which were generated based on a group-level data representation from an independent group of participants (see Fedorenko et al., 2010). These masks were intersected with each individual participant's activation map for the *sentences* > *nonwords* or the *intact* > *degraded* contrast. The voxels falling within each mask were sorted based on their *t*-values for the relevant localizer contrast and the top 10% of voxels were chosen as that participant's fROI. This top *n*% approach ensures that the fROIs can be defined in every participant—thus enabling us to generalize the results to the entire population (Nieto-Castañón & Fedorenko, 2012)—and that fROI sizes are the same across participants. However, qualitatively similar results were obtained in an alternative analysis approach where the fROIs were defined as all the voxels that (i) fell within the relevant mask and (ii) passed a fixed significance threshold ($p < 0.001$, uncorrected at the whole-brain level).

Eight fROIs were defined in each participant, for each of the two localizer contrasts. These included three fROIs on the lateral surface of the left frontal cortex in the inferior frontal gyrus (LIFG) and its orbital part (LIFGorb) as well as in the middle frontal gyrus (LMFG), and five fROIs on the lateral surface of the temporal and parietal cortex, in the anterior temporal cortex (LAntTemp), middle anterior temporal cortex (LMidAntTemp), middle posterior temporal cortex (LMidPostTemp), posterior temporal cortex (LPostTemp) and angular gyrus (LAngG). We chose to focus on these 'core' regions in the left hemisphere, which most robustly and consistently emerge in investigations of the language system.

To estimate (a) the responses of the *sentences* > *nonwords* fROIs to the conditions of the standard visual language localizer or (b) the responses of the *intact* > *degraded* fROIs to the conditions of the new auditory language localizer, we used an across-runs cross-validation procedure. In particular, each participant's activation map was first computed for the relevant (e.g., *sentences* > *nonwords*) contrast using the first run, and the 10% of voxels with the highest *t*-values within a given mask were selected as that subject's fROI. The response of each fROI to the same contrast was then estimated using the second run. This procedure was repeated using the second run to define the fROIs and the first run to estimate the responses. Finally, the responses were averaged across the two runs to derive a single response magnitude for each condition in a given fROI/participant. This cross-validation procedure allows one to use all of the data for defining the ROIs and for estimating their responses (see Nieto-Castañón & Fedorenko, 2012, for discussion), while ensuring the independence of the data used for fROI definition and for response estimation (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). To estimate (a) the responses of the *sentences* > *nonwords* fROIs to the *intact* and *degraded* conditions of the novel auditory language localizer or (b) the responses of the *intact* > *degraded* fROIs to the sentences and nonwords conditions of the standard visual language localizer, data from both runs of the standard visual language localizer or the novel auditory localizer, respectively, were used for defining the fROIs.

To summarize the logic of our approach: the visual language localizer is used as a standard against which to test the efficacy of the novel auditory language localizer for identifying a set of brain regions that respond robustly during language processing. To evaluate the new localizer, we characterize the responses elicited by both localizers qualitatively and quantitatively, with respect to (i) effect sizes and (ii) the similarity of the fine-grained spatial patterns across runs. Statistical tests across subjects were performed on the percent BOLD signal change values extracted from the fROIs as described above.

Results

The responses to the novel auditory language localizer (the *intact* > *degraded* contrast) were highly similar to the responses to the standard visual language localizer (the *sentences* > *nonwords* contrast). We illustrate this similarity in several ways.

Figure 1 shows whole-brain activation maps for the two localizer contrasts and their conjunction in five individual subjects. The maps were thresholded at an FDR-corrected *p*-value of 0.05. Although the intersubject variability in the exact locations and extent of activations is apparent, responses elicited from the two localizer contrasts are remarkably consistent within subjects. Similar whole-brain activation maps for all 10 subjects who completed two runs of each localizer are included in the Supplemental Information as Figure S1.

Figure 2(a,b) shows the responses of the fROIs (defined by the *sentences* > *nonwords* contrast) to the conditions of the two localizers relative to the fixation baseline in eight 'core' language regions. Both, the *sentences* > *nonwords* and the *intact* > *degraded* contrasts were reliable in each of the eight fROIs, and the response magnitudes were similar across localizers, although—consistent with our prior work (Fedorenko et al., 2010)—some regions appear to be overall more responsive to visual and others to auditory language stimuli. Detailed statistics are included in Table 1 (see Supplemental Information for a similar figure (Figure S2) and table (Table S1) for an analysis where the fROIs are defined by the *intact* > *degraded* contrast).

To examine the similarity of the activation patterns for the two localizer contrasts in greater detail, we computed linear correlations within each ROI mask (i) between the two runs of the traditional visual localizer, (ii) between the two runs of the novel auditory localizer, and (iii) between the first run of the visual localizer and the second run of the auditory localizer (to have the same amount of data as the within-localizer comparisons). In this analysis, we included all the voxels within each ROI mask, regardless of whether they showed a reliable effect for either of the localizer contrasts. Only the 10 subjects who completed two runs of each localizer were included in this analysis. The average Fisher-transformed correlation values are presented in Figure 2 (c). High correlations indicate that the same voxels that show a large effect size in one run also show a large effect size in the other run, thus indicating that the fine-grained activation patterns are reproducible. High correlations were observed for the within-localizer comparisons (all *r*s between 1.32 and 0.73) and, critically, for the between-localizer comparison (all *r*s between 1.01 and 0.52) in every ROI, suggesting that the activations are highly similar between the two localizer contrasts. We have included a qualitative comparison of the correlations between the two different localizers in subjects who either completed each localizer in the same session or over the course of different sessions in the Supplemental Information as Figure S3 in order to better understand the effect

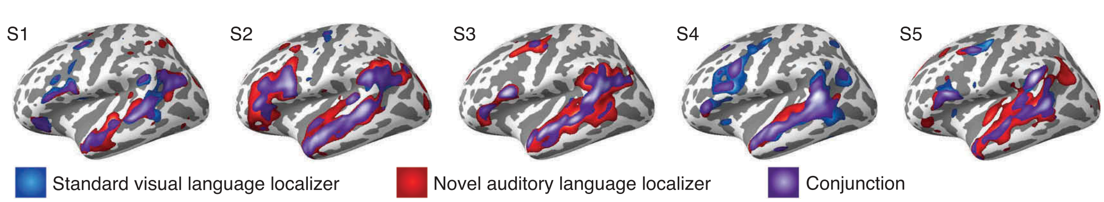

Figure 1. Activation maps from five representative subjects. All maps are thresholded at a corrected *p*-value of 0.05. Blue maps indicate *sentences—nonwords* activations from two runs of the standard visual language localizer, red maps indicate *intact—degraded* activations from two runs of the novel auditory language localizer, and purple maps indicate the conjunction of both localizers.

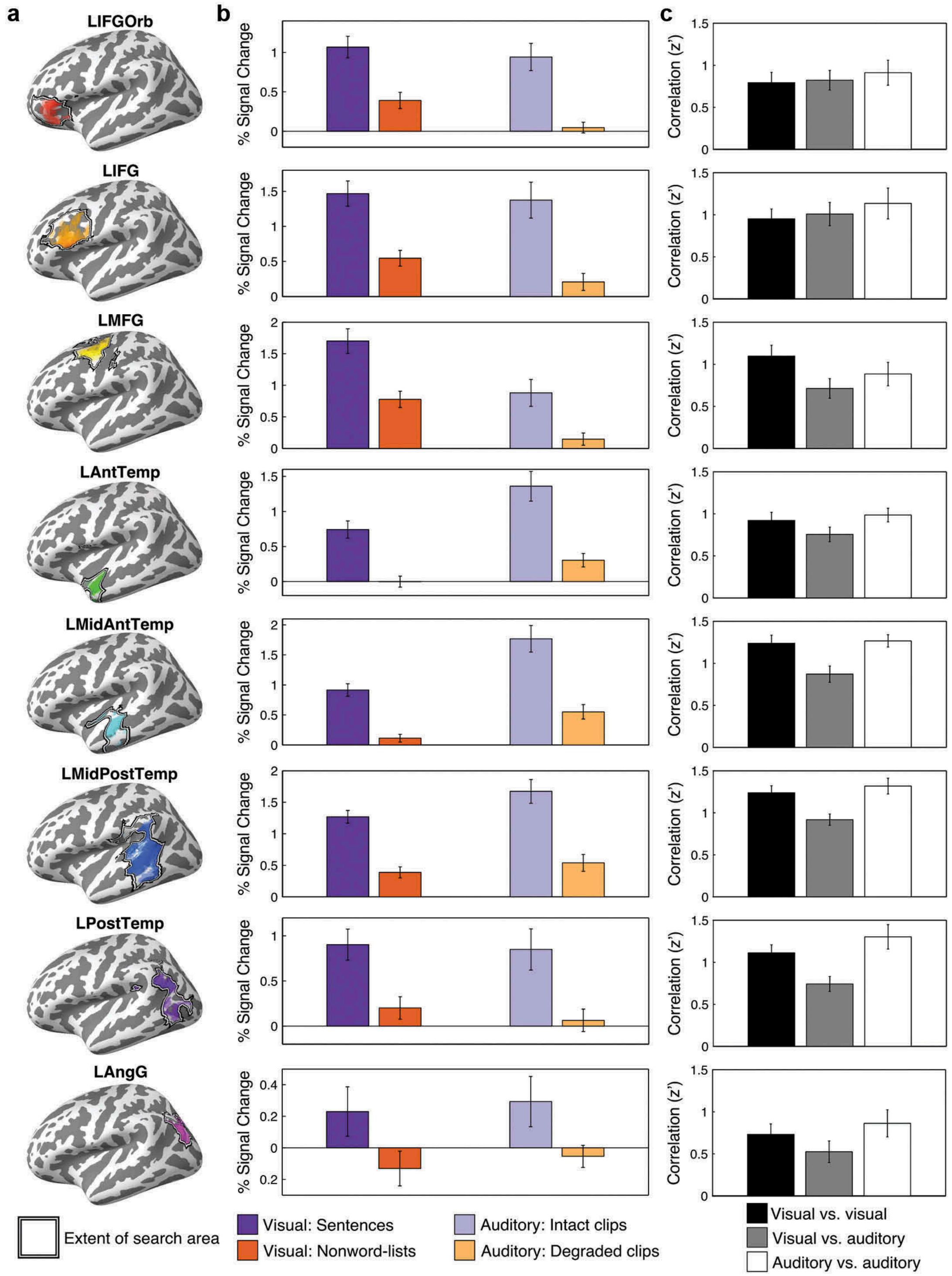

Figure 2. (a) Individual subjects' fROIs are overlaid in color onto a template inflated surface. Each subject's fROI is plotted with equal transparency, so that higher opacity in the map indicates overlap of multiple subjects' fROIs. Contrast lines indicate the extent of the search areas from which the top 10% of voxels were selected for analysis. fROIs defined using two runs of the standard visual language localizer are shown. (b) Percent signal change (PSC) measured from baseline fixation in independent data in each fROI and averaged across subjects is shown. In the case of the visual conditions, PSC is measured using a leave-one-out approach. For the auditory conditions, PSC is measured in the regions defined using the visual language localizer. (c) Average Fisher-transformed correlations are shown between the PSC in each voxel from two independent runs across the entire search area of each ROI, regardless of whether they showed a reliable effect for either of the localizer contrasts. The black bars show the average correlation between two independent runs of the visual language localizer, the white bars show the average correlation between two independent runs of the auditory language localizer, and the gray shows the average correlation between the first run of the visual language localizer and the second run of the auditory language localizer.

Table 1. Response magnitudes to contrasts of interest within ROIs defined using the visual language localizer. Statistics for the *sentences—nonwords* contrast from the standard visual language localizer and the *intact—degraded* contrast from the novel auditory language localizer measured in each ROI. The size of each ROI is fixed across participants due to the decision to include the top 10% of voxels from each broader region of interest. All *p*-values have been FDR-corrected for the number of fROIs. All contrasts are statistically significant to a level of at least 10^{-2}, except the *intact* > *degraded* contrast in the LAngG fROI, which is significant to a level of 0.014.

ROI	Size in voxels	Sentences—nonword lists (Visual)	Intact—degraded clips (Auditory)
LIFGOrb	292	0.68 ± 0.10 $t(10) = 6.65$ p-FDR < 10^{-4}	0.89 ± 0.16 $t(10) = 5.49$ p-FDR < 10^{-3}
LIFG	346	0.92 ± 0.11 $t(10) = 8.11$ p-FDR < 10^{-4}	1.17 ± 0.23 $t(10) = 5.05$ p-FDR < 10^{-3}
LMFG	303	0.93 ± 0.11 $t(10) = 5.40$ p-FDR < 10^{-3}	0.73 ± 0.19 $t(10) = 3.87$ p-FDR < 10^{-2}
LAntTemp	195	0.74 ± 0.11 $t(10) = 6.87$ p-FDR < 10^{-4}	1.06 ± 0.13 $t(10) = 7.95$ p-FDR < 10^{-4}
LMidAntTemp	209	0.80 ± 0.10 $t(10) = 8.14$ p-FDR < 10^{-4}	1.21 ± 0.14 $t(10) = 8.67$ p-FDR < 10^{-4}
LMidPostTemp	480	0.88 ± 0.05 $t(10) = 17.18$ p-FDR < 10^{-4}	1.13 ± 0.09 $t(10) = 11.99$ p-FDR < 10^{-5}
LPostTemp	239	0.70 ± 0.11 $t(10) = 6.42$ p-FDR < 10^{-4}	0.79 ± 0.16 $t(10) = 4.79$ p-FDR < 10^{-3}
LAngG	223	0.36 ± 0.10 $t(10) = 3.50$ p-FDR < 10^{-2}	0.35 ± 0.13 $t(10) = 2.58$ p-FDR = 0.014

of separating the sessions in some participants on our results.

Finally, we repeated our GSS analysis with a separate set of fROIs believed to constitute the multiple demand network, which has been shown to be modulated by task difficulty (Blank et al., 2014; Fedorenko, Duncan, & Kanwisher, 2013). We found that the opposite novel contrast of *degraded* > *intact* speech revealed significant positive signal changes in 14 of the 18 regions tested ($p < 0.05$, FDR-corrected), suggesting that subjects were highly attentive to the degraded speech, even though they were not able to understand it.

Discussion and conclusions

Comparisons of the locations, effect sizes, and the fine-grained spatial patterns of neural activity within left hemisphere language ROIs revealed great similarity between the activations elicited from our standard visual language localizer and the novel auditory language localizer presented here. These results suggest that the new auditory localizer can be used for quickly and reliably identifying the high-level language-processing brain regions at the individual subject level in future work.

There are several changes to our paradigm researchers could implement to better suit their own needs. First, fixation blocks could be omitted in order to shorten task runs. We included fixation blocks in order to provide a low-level baseline of BOLD activity, to which responses to the experimental conditions can be compared. This baseline may be important if one were interested not just in the size of the difference between the two experimental conditions, but also in the magnitude of response to each experimental condition. These rest periods also provide short breaks for participants, which can be helpful for some individuals/populations. One could omit fixation blocks if the primary goal was only to localize language regions as quickly as possible. Second, in order to make sure subjects are awake and attentive, one could add response prompts after each clip. This may be especially desirable when working with clinical populations.

One methodological point is worth a brief discussion. We here used acoustically degraded versions of the clips. However, based on other data collected in our group, as well as some prior literature (Bavelier, Corina, & Neville, 1998; Binder et al., 1997; Diaz & McCarthy, 2009; Fedorenko et al., 2010; Kuperberg et al., 2003; Noppeney & Price, 2004; Petersen, Fox, Snyder, & Raichle, 1990; Glenberg & Robertson, 2000; Snijders et al., 2009), we suspect that the precise nature of the control condition will not matter too much for identifying the high-level language system. In particular, other studies have used control conditions like vocoded speech (Scott, Blank, Rosen, & Wise, 2000), reversed speech (Bedny et al., 2011; Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002), or foreign speech (Perani et al., 1996). Although the nature of the control condition may matter a great deal for defining lower-level speech-processing brain regions (Overath, McDermott, Zarate, & Poeppel, 2015; Stoppelman, Harpaz, & Ben-Shachar, 2013), it appears that for defining the higher-level language regions the key requirement is that the control condition lacks meaning and structure that characterizes language. It is possible, of course, that—depending on which control condition is chosen—speech-responsive regions may or may not be activated *in addition* to the high-level language regions (Fedorenko & Thompson-Schill, 2014), but if we restrict our search to the brain areas that we know support high-level language processing, we are likely to be able to narrow in on the same sets of voxels, as we here show for two contrasts that differ in the materials, modality of presentation and task.

Whereas many of the features of the novel contrast of intact versus acoustically degraded auditory stimuli—such as ease of adaptation to new testing materials and appropriateness for illiterate populations—could also be attributed to similar contrasts, such as that between speech and reversed speech, there are advantages to the current paradigm. Muffled speech is more natural than reversed speech and more frequently encountered. This relative naturalness of muffled speech is likely to encourage participants to *try* to understand the degraded speech, and therefore may be less likely to produce an attentional confound. Furthermore, this extra effort to decipher muffled speech provides an added benefit: the opposite contrast (*degraded* > *intact*) can be used to identify the multiple demand network, which becomes more active as cognitive tasks become more difficult (Fedorenko et al., 2013).

To conclude, using the paradigm described in the current study (a contrast between engaging speech clips and acoustically degraded versions of those clips) researchers can quickly and reliably identify high-level language-processing brain regions in diverse populations, including children and clinical populations, who may have difficulty with cognitively demanding tasks or reading (see Tie et al., 2014, 2015; for similar efforts with different paradigms). Furthermore, as noted above, this paradigm can be easily adapted by simply replacing the materials with a different set, as needed for studying children of various ages or speakers of diverse languages.

Acknowledgments

E.F. was supported by the NICHD Award HD-057522. For technical support during scanning, the authors thank Atsushi Takahashi, Steve Shannon, and Sheeba Arnold. For help with finding some of the clips, we thank Jacovie Rodriguez. For helping to make our data and experimental materials publicly available, we thank Zach Mineroff. We also acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, MIT.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the NICHD: [award HD-057522].

ORCID

Terri L. Scott http://orcid.org/0000-0001-9205-8581

References

Axelrod, V., Bar, M., Rees, G., & Yovel, G. (2015). Neural correlates of subliminal language processing. *Cereb Cortex*, *25*(8), 2160–2169. doi:10.1093/cercor/bhu022

Bavelier, D., Corina, D. P., & Neville, H. J. (1998). Brain and language: A perspective from sign language. *Neuron*, *21*(2), 275–278. doi:10.1016/S0896-6273(00)80536-X

Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., & Saxe, R. (2011). Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences*, *108*(11), 4429–4434. doi:10.1073/pnas.1014818108

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312. doi:10.1038/35002078

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *17*(1), 353–362.

Blank, I. A., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, *112*, 1105–1118. doi:10.1152/jn.00884.2013

Diaz, M. T., & McCarthy, G. (2009). A comparison of brain activity evoked by single content and function words: An fMRI investigation of implicit word processing. *Brain Research*, *1282*, 38–49. doi:10.1016/j.brainres.2009.05.043

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600), 2013–2015. doi:10.1126/science.1077066

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*, 335. doi:10.3389/fpsyg.2014.00335

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*, 16428–16433. doi:10.1073/pnas.1112937108

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*(21), 2059–2062. doi:10.1016/j.cub.2012.09.011

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616–16621. doi:10.1073/pnas.1315235110

Fedorenko, E., Hsieh, P., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. doi:10.1152/jn.00032.2010

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, *18*(3), 120–126. doi:10.1016/j.tics.2013.12.006

Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B. T. T., … Zilles, K. (2008). Cortical folding patterns and predicting cytoarchitecture. *Cerebral Cortex*, *18*(8), 1973–1980. doi:10.1093/cercor/bhm225

Frost, M. A., & Goebel, R. (2012). Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *NeuroImage*, *59*(2), 1369–1381. doi:10.1016/j.neuroimage.2011.08.035

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*(3), 379–401. doi:10.1006/jmla.2000.2714

Hickok, G., Okada, K., & Serences, J. T. (2009). Area spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, *101*(5), 2725–2732. doi:10.1152/jn.91099.2008

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, *60*(4), 2357–2364. doi:10.1016/j.neuroimage.2012.02.055

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, *17*(11), 4302–4311.

Koster-Hale, J., & Saxe, R. (2011). Theory of mind brain regions are sensitive to the content, not the structural complexity, of belief attributes. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Cognitive Science Society Conference* (pp. 3356–3361). Austin, TX: Cognitive Science Society.

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272–293. doi:10.1162/089892903321208204

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. doi:10.1038/nn.2303

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M. … Petersen, S. E. (2015). Functional system and areal organization of a highly sampled individual human brain. *Neuron*, *87*(3), 657–670. doi:10.1016/j.neuron.2015.06.037

Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *NeuroImage*, Advanced online publication. doi:10.1016/j.neuroimage.2016.05.073

Maruyama, M., Pallier, C., Jobert, A., Sigman, M., & Dehaene, S. (2012). The cortical representation of simple mathematical expressions. *NeuroImage*, *61*(4), 1444–1460. doi:10.1016/j.neuroimage.2012.04.020

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, *63*(3), 1646–1669. doi:10.1016/j.neuroimage.2012.06.065

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *NeuroImage*, *22*(1), 164–170. doi:10.1016/j.neuroimage.2003.12.010

Overath, T., McDermott, J. H., Zarate, J. M., & Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nature Neuroscience*, *18*, 903–911. doi:10.1038/nn.4021

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., … Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport*, *7*(15), 2439–2444. doi:10.1097/00001756-199611040-00007

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*(4972), 1041–1044. doi:10.1126/science.2396097

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406. doi:10.1093/brain/123.12.2400

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*(7), 1493–1503. doi:10.1093/cercor/bhn187

Stoppelman, N., Harpaz, T., & Ben-Shachar, M. (2013). Do not throw out the baby with the bath water: Choosing an effective baseline for a functional localizer of speech processing. *Brain and Behavior*, *3*(3), 211–222. doi:10.1002/brb3.129

Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., & Johnsrude, I. S. (2012). Is the link between anatomical structure and function equally strong at all cognitive levels of processing? *Cerebral Cortex*, *22*(7), 1593–1603. doi:10.1093/cercor/bhr205

Tie, Y., Rigolo, L., Norton, I. H., Huang, R. Y., Wu, W., Orringer, D., … Golby, A. J. (2014). Defining language networks from resting-state fMRI for surgical planning- A feasibility study. *Human Brain Mapping*, *35*(3), 1018–1030. doi:10.1002/hbm.22231

Tie, Y., Rigolo, L., Ovalioglu, A. O., Olubiyi, O., Doolin, K. L., Mukundan, S. Jr., & Golby, A. J. (2015). A new paradigm for individual subject language mapping: movie-watching fMRI. *Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging*, *25*(5), 710–720. doi:10.1111/jon.2015.25.issue-5

Xu, M., Beuckelaer, A. D., Wang, X., Liu, L., Song, Y., & Liu, J. (2015). Regional amplitude of the low-frequency fluctuations at rest predicts word-reading skill. *Neuroscience*, *298*, 318–328. doi:10.1016/j.neuroscience.2015.04.025