NEURAL CORRELATES OF PROCESSING LEXICALLY AND GRAMMATICALLY DEGRADED LINGUISTIC STIMULI IN A FAMILIAR NARRATIVE CONTEXT

By

ALEXA GRACE BAUTISTA

A Thesis Submitted to the Honors College

In Partial Fulfillment of the Bachelor's degree
With Honors in

Speech, Language, and Hearing Sciences

THE UNIVERSITY OF ARIZONA

M A Y 2 0 1 5

Approved by:

Stephen Wilson

Dr. Stephen M. Wilson
Department of Speech, Language, and Hearing Sciences

Abstract

Functional activation for language processing in left hemisphere language regions has been shown to be correlated with intelligibility of the speech signal, with more intelligible stimuli yielding greater activation. We created stimuli in which intelligibility was reduced in two distinct ways. First, we created lexically degraded stimuli in which lexical items were spectrally rotated. Second, we created grammatically degraded stimuli in which function words and morphemes were spectrally rotated and scrambled in order. We hypothesized that contrasting these conditions would reveal brain areas differentially involved in processing lexical or grammatical information when matched for intelligibility. We used functional MRI to investigate differences in the neural response to normal and degraded stimuli. Stimuli were presented such that normal segments alternated with degraded ones in the context of a familiar narrative. Our results showed no statistically significant differences in activation between lexically and grammatically degraded stimuli. Surprisingly, we also found only a small difference in the anterior middle temporal gyrus between either kind of degraded stimuli and normal stimuli. A possible explanation is that the context of a familiar narrative may increase attention and intelligibility, reducing degradation effects.

Introduction

Prior neuroimaging studies that have contrasted activation levels between grammatical sentences and their semantically degraded counterparts provide evidence for the role of the anterior temporal lobe in syntactic processing (Pallier et al., 2011; Humphries et al., 2006). In order to isolate the effects of syntax, the literature has removed semantic information by using matched word lists (Pallier et al., 2011; Humphries et al., 2006; Vandenberghe et al., 2002), semantically random sentences (Humphries et al. 2006), and by replacing words with pseudowords (Pallier et al., 2011). Of these, the most compelling evidence that the ATL may be involved in syntactic processing is the fact that greater activation has been observed in this region for spoken pseudoword sentences (Humphries et al., 2006) when compared to pseudoword lists. Pseudowords are an effective control to isolate syntactic processing because pseudowords do not provide lexicosemantic information.

However, stimuli that are more intelligible have also shown a positive correlation with activation in the anterior temporal lobe (Davis & Johnsrude, 2003). Intelligibility, or meaningfulness, is defined in the present study as the amount of speech that a listener can understand. Because previous experiments that argued for role of the anterior temporal lobe in syntactic processing did not match their stimulus conditions for intelligibility, it is possible that the observed effects in this region were not specific to the presence or absence of syntactic structure but were instead associated with varying levels of meaningfulness. Therefore, to determine whether the activation in this area is due to syntactic processing or a result of more meaningful stimuli, it is necessary to match for intelligibility between conditions.

The goal of the present study was to use functional MRI to observe whether anterior temporal activation for syntactic construction is observed when compared to semantically degraded stimuli with matched intelligibility. We did this by removing either grammatical or lexical information and then matching intelligibility between the two conditions. Grammatical information was removed from speech segments by spectrally rotating and pseudorandomizing the order of grammatical morphemes while maintaining the relative word order of lexical items (nouns, verbs, adjectives. To degrade lexical information, lexical items (nouns, verbs, adjectives) were spectrally rotated, but mixed with normal speech in a proportion that matched the intelligibility of the condition where grammatical information was removed. All stimuli were presented within a narrative context with segments alternating between clear and degraded speech to provide a context for assessing meaningfulness.

We hypothesized that reducing intelligibility would reduce anterior temporal activation regardless of whether grammatical or semantic information was removed. The absence of contrasting activation between intelligibility-matched stimuli could suggest that previous effects attributed to syntactic processing in the anterior temporal lobe may instead have reflected varying levels of intelligibility.

Experiment 1 Methods

The aim of this experiment was to create stimuli in which grammatical or lexical information was selectively removed and to match intelligibility between the two. This was done to provide a framework for our second experiment that investigated areas specific to grammatical and lexical processing, as well as to ensure that one condition was not harder than the other.

Participants

Seven native English speakers (3 female; mean age 21.4 years; range 20-28 years) participated in a perceptual experiment to rate the intelligibility of auditory stimuli in which either grammatical or lexical elements were removed. All participants were healthy individuals who reported no hearing or cognitive impairments and provided written informed consent. The study was approved by the Institutional Review Board at the University of Arizona.

Stimuli

The stimuli were created from Rosemary Kingston's (2001) audiobook adaptations of "Cinderella" and "Sleeping Beauty," read by Marguerite Gavin. In both stories, the onset and offset of each morpheme were marked. Next, the morphemes were categorized by their part of speech as lexical or functional items. Lexical items were nouns, verbs, and adjectives. Functional items were determiners, pronouns, auxiliary verbs, copulae, conjunctions, prepositions, bound inflectional and bound derivational morphemes, and the possessive clitic. From the transcribed morphological time signatures, the stories were broken into segments. Each segment was as long as possible without separating morphemes of the same word and without exceeding 7 seconds. Given these parameters, four different conditions were constructed: normal, delexicalized, degrammaticalized, and scrambled stimuli.

Normal stimuli were low pass filtered at 4 kHz. Performing a low pass filter was necessary because parts of the normal stimuli would be spectrally rotated to create the delexicalized, degrammaticalized, and scrambled stimuli.

To create the delexicalized stimuli, we preserved word order and the acoustic integrity of functional words, but replaced the lexical items with a blend of spectrally rotated and normal speech. The methods of using spectrally rotated speech as an acoustically matched control have been previously described by Scott et al. (2000). This involved using a rotated speech algorithm devised by Blesser (1972) in which the signal is first low pass filtered at 4 kHz then inverted around 2 kHz as its central axis. In our study, the degree of delexicalization could be controlled by changing the proportion of rotated to non-rotated speech.

Degrammaticalized stimuli were created by randomly ordering all morphemes in a segment then rearranging the functional items in such a way that they paralleled their original positions. By keeping the lexical items in their original order relative to one another, we preserved some level of intelligibility. However, the functional items remained fixed in their randomized position and were completely spectrally rotated.

Lastly, scrambled stimuli were created to control for auditory activation. These segments were entirely spectrally rotated and every morpheme was randomly ordered.

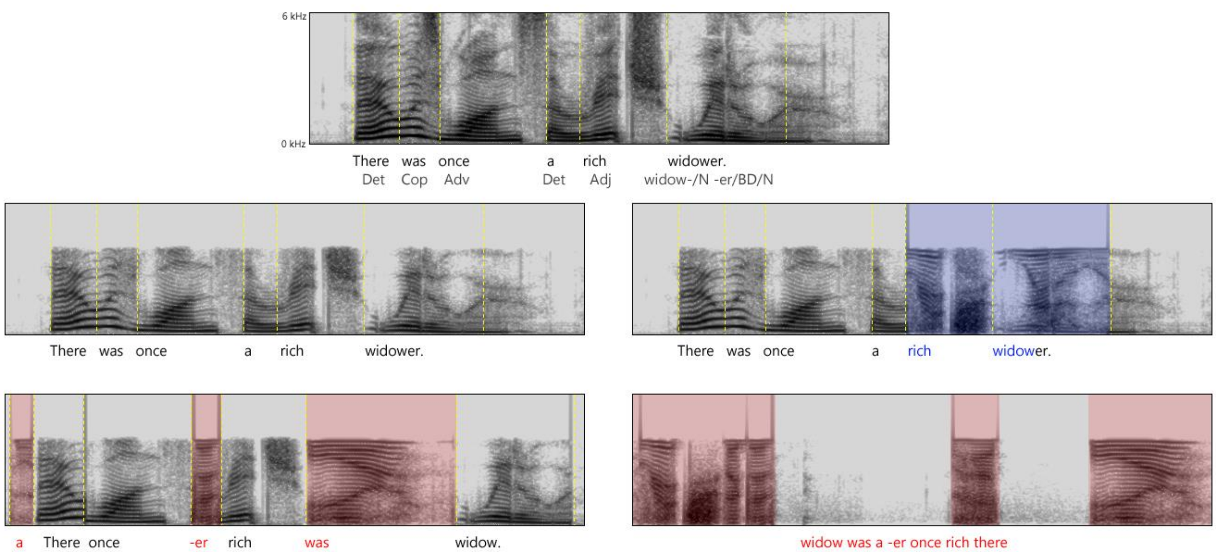

Figure 1. Spectrograms of the phrase "There was once a rich widower" and creation of stimuli.
A) Original stimuli. B) Normal stimuli after low pass filtering at 4 kHz. C) Delexicalized stimuli, words highlighted in blue have been mixed with spectrally rotated speech by 83.5%. D) Degrammaticalized stimuli, all words in the sentence are randomized, then lexical words are rearranged to correlate with their original order. Function words remain in their scrambled order and are spectrally rotated by 100%, indicated in red. E) Scrambled stimuli, all words are placed in a randomized order and subsequently spectrally rotated by 100%.

Experimental Design

"Cinderella" was clipped into 86 segments in total. Forty-four of the segments were normal, 18 were degrammaticalized, and 24 were delexicalized. Segments alternated between normal audio and degrammaticalized or delexicalized audio to ensure that participants could still comprehend the story despite the presence of degraded stimuli. It was expected that participants would be able to easily follow the progression of each story, because the two narratives are familiar children's stories and every second segment is intact. The 24 delexicalized segments were broken into four groups representing various weighted averages between normal and spectrally rotated speech in ten percent intervals from 70 to 100 percent.

The seven participants listened to "Cinderella" and were instructed to follow the story as much as possible and rate intelligibility on a six point scale (1 = not at all; 2 = just a bit; 3 = about half, maybe less; 4 = about half, maybe more; 5 = mostly; 6 = perfectly). Using a six point scale required the participant to decide whether more or less of the signal was understood. The norming experiment was presented by a MATLAB program that prompted the participant to input a response before playing the next segment.

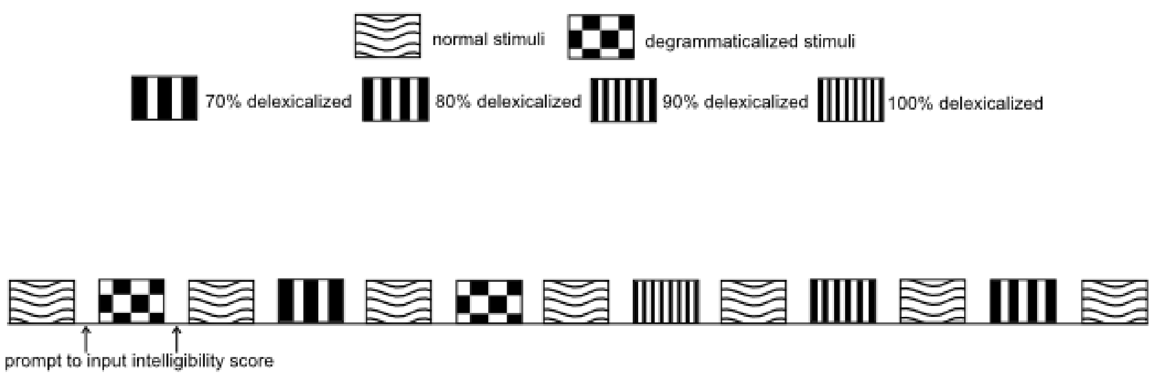

Figure 2. Example of the order in which stimuli was presented for Experiment 1. Normal and degraded stimuli alternated between each other. Delexicalized and degrammaticalized stimuli were presented in a pseudorandom order.

Results

We calculated the average of the four delexicalized intervals and the degrammaticalized segments across participants. The average results for each condition can be found in Figure 3. To match intelligibility between the two conditions, the degrammatical intelligibility score was linearly interpolated using the delexicalized values as reference points. From this, we derived that delexicalized segments degraded by 83.5% were as equally intelligible as the degrammaticalized segments. This percentage was used in our imaging study.

	Average Intelligibility Score
Normal Stimuli	6.00
Degrammaticalized	3.09
70% Delexicalized	4.02
80% Delexicalized	3.36
90% Delexicalized	2.61
100% Delexicalized	2.18

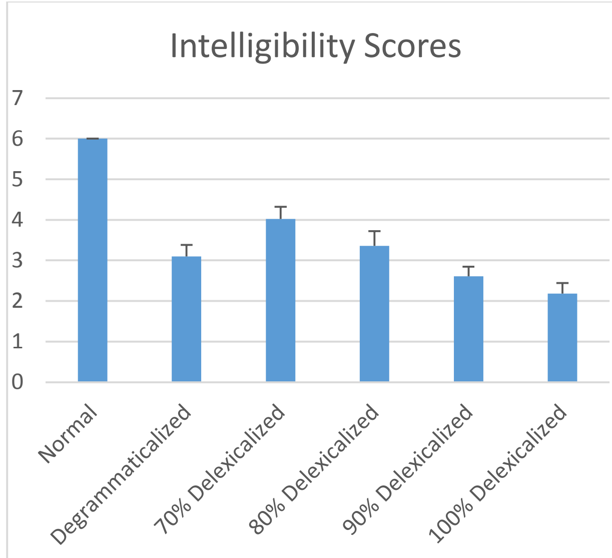

Figure 3. Results from Experiment 1 demonstrating the intelligibility score for each condition. Error bars indicate the standard error of the mean.

Experiment 2 Methods

The purpose of this experiment was to use functional magnetic resonance imaging (fMRI) to determine whether there are differences in the brain regions involved in processing delexicalized and delexicalized stimuli that are equally intelligible.

Participants

Twelve healthy, native English speakers were successfully scanned with fMRI (9 female; mean age 22.8 years; range 20-33 years; 1 left-handed). One additional participant was unable to complete the task due to claustrophobia. All participants were healthy individuals who reported no hearing or cognitive impairments and provided written informed consent. The study was approved by the Institutional Review Board at the University of Arizona.

Experimental Design

Participants were asked to lie supine in the MR scanner and listen to the stories using insert earphones (S14, Sensimetrics, Malden, MA). Prior to the functional runs, the presentation volume was adjusted to a comfortable level for each participant. There was no task and participants were simply instructed to listen and follow the story as much as possible, paying attention to both normal and degraded parts. After the scans, participants were asked what they had heard and if they were able to follow the narrative.

Stimuli were created for “Cinderella” and “Sleeping Beauty” using the methods found in the first experiment. The stimuli were controlled with the Psychophysics Toolbox version 3.0.10 (Brainard, 1997; Pelli, 1997) running under MATLAB R2012b (Mathworks, Natick, MA) on a Lenovo S30 workstation. Unlike the first experiment, the story progressed without the prompting of the participant, and delexicalized stimuli were fixed at a weighted average of 85.3% spectrally rotated speech and normal speech. The first two or three stimuli were always normal. Afterwards, the stimuli alternated between degraded (either delexicalized or degrammaticalized) and normal speech. Interspersed pseudorandomly among this sequence were 10 scrambled trials and 10 silent trials. Although the scrambled trials were completely unintelligible, the progression of the narrative was not interrupted because they were not based on the next segment in the sequence but rather on random segments from other parts of the narrative.

Half of the participants were presented with “Cinderella” first, and half were presented with “Sleeping Beauty” first. The “Cinderella” runs were 16.8 minutes long and 107 functional images (44 normal; 21 delexicalized; 21 degrammaticalized; 10 silent; 10 scrambled; 1 initial discarded image) were collected. The “Sleeping Beauty” runs were 14.3 minutes long and 91 functional images (36 normal; 17 delexicalized; 17 degrammaticalized; 10 silent; 10 scrambled; 1 initial discarded image) were collected.

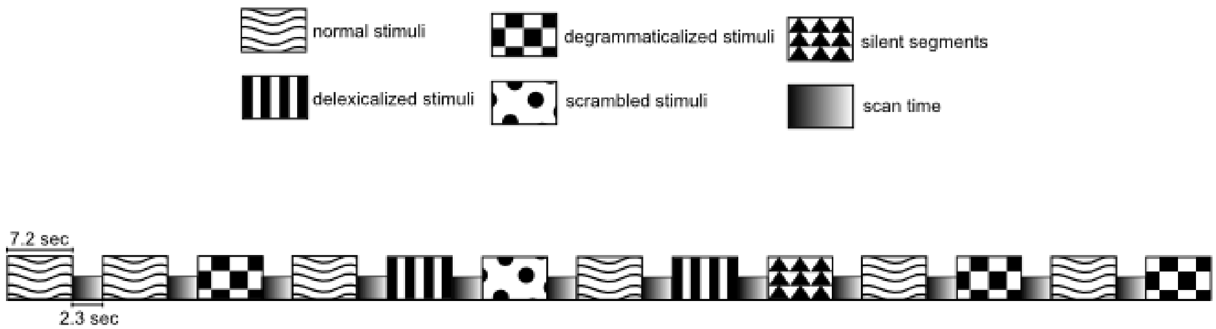

Figure 4. Example of the order in which stimuli were presented. Stimuli were not presented while acquiring scans. The length of each stimulus was 7 seconds or less and the acquisition time was around 2.3 seconds, summing up to the 9.5 second repetition time.

Neuroimaging Protocol

Functional and structural images were acquired on a Siemens Skyra 3T scanner with a 32-channel head coil at the University of Arizona.

T2*-weighted BOLD echo planar images were acquired in two sparse sampling runs with the following parameters: 30 axial slices in ascending order; slice thickness = 3.5 mm with a 0.9 mm skip; field of view = 240 x 240 mm; matrix = 86 x 96 mm; TR = 9500 ms; TA = 2288 ms; TE = 30 ms; flip angle = 90°; voxel size = 2.5 x 2.5 x 3.5 mm.

For each participant, we also collected a T1 MPRAGE (192 sagittal slices; slice thickness = 0.9 mm; field of view = 240 x 240 mm; matrix = 256 x 256 mm; TR = 2300 ms; TI = 900 ms; TE = 2.32 ms; flip angle = 8°; GRAPPA acceleration factor = 2; voxel size = 0.9 x 0.9 x 0.9 mm) to provide an anatomical reference for the functional scans.

Analysis

The T1-weighted anatomical image was segmented into gray matter, white matter, and cerebrospinal fluid, then warped into MNI space using SPM5 (Friston et al., 2007). All functional images were coregistered to their participant's corresponding anatomical image, warped into MNI space, and smoothed to minimize the effects of noise on the data. Data were thresholded at a p-value of 0.005, corresponding to a t-score of 3.11. Statistical significance was determined based on cluster size according to Gaussian random field theory.

We masked the contrast “normal vs delexdegram” with “normal vs scrambled” ($p < 0.05$) in order to restrict the regions of analyses to areas of activation present in both contrasts. This isolated regions of neural activation for the normal condition, but only within a subset of voxels where the normal condition had greater activation than in the scrambled condition.

Results

Delexicalized vs. Degrammaticalized Contrast

There were no significant differences in activation between the delexicalized and degrammaticalized conditions or vice versa ($p < 0.05$, corrected). Because we did not find distinct areas of the brain involved in processing syntactic or lexical information, we conducted further analyses to observe the effects of degradation in general on neural activation.

Delexicalized vs. Scrambled

Delexicalized stimuli demonstrated higher levels of activation along the left superior temporal gyrus than scrambled stimuli.

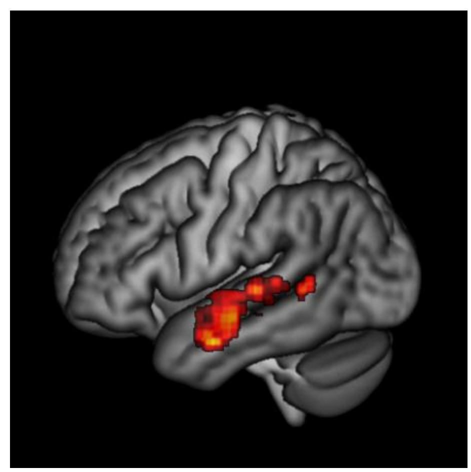

Figure 5. Delexicalized vs. Scrambled.

Degrammaticalized vs. Scrambled

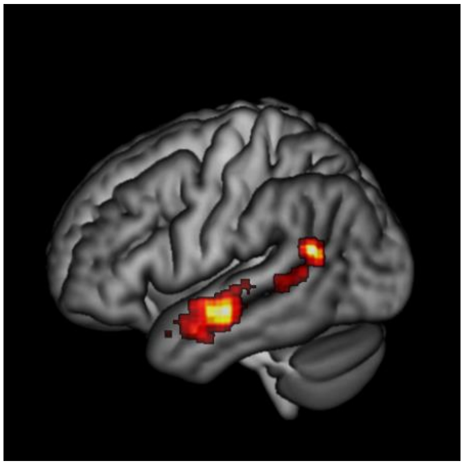

Figure 6. Degrammaticalized vs. Scrambled.

Similar to the contrast between delexicalized and scrambled stimuli, this contrast showed greater activation along the left superior temporal gyrus than scrambled stimuli.

"Delexdegram" vs. Scrambled

Because there were no significant differences between delexicalized and degrammaticalized, we combined the two conditions into "delexdegram" to increase the reliability of the contrast with scrambled stimuli. "Delexdegram" showed greater activation along the left superior temporal gyrus and superior temporal sulcus than did scrambled stimuli.

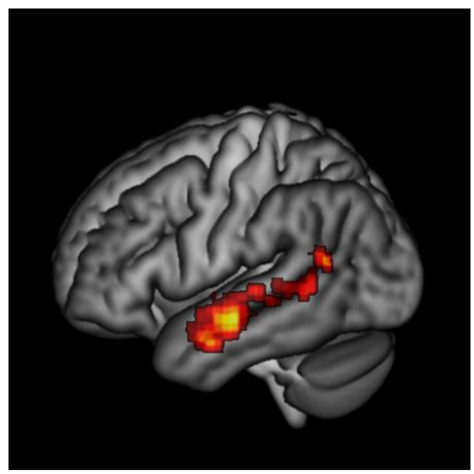

Figure 7. "Delexdegram" vs. Scrambled.

Normal vs. Scrambled

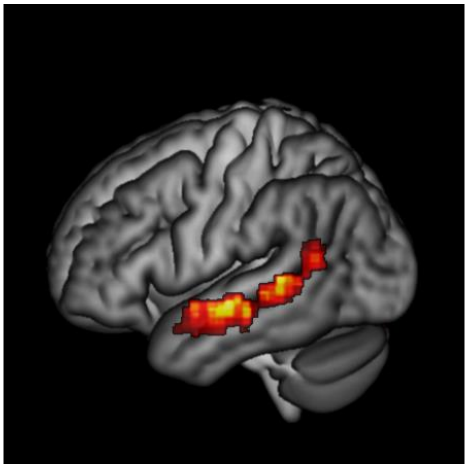

Figure 8. Normal vs. Scrambled

Comparing the normal and scrambled conditions demonstrates greater activation along the left superior temporal gyrus and superior temporal sulcus for normal stimuli than for scrambled ones. The results from this contrast resemble those of "delexdegram" vs. scrambled.

Normal vs. "Delexdegram"

Our fourth contrast was between the normal and "delexdegram" conditions, masked with the "normal vs. scrambled" contrast as described in our methods. The contrast yielded no statistically significant results, but we did observe a small area of greater activation (cluster size: 584 mm^3, *a priori* p-value $p = 0.017$ for single cluster closest to expected location, Friston et al., 1997) in the middle temporal gyrus for normal stimuli. This cluster reflects an intelligibility effect between the two conditions.

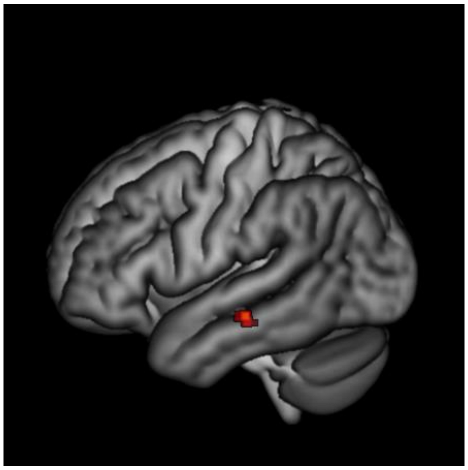

Figure 9. Normal vs. "Delexdegram."

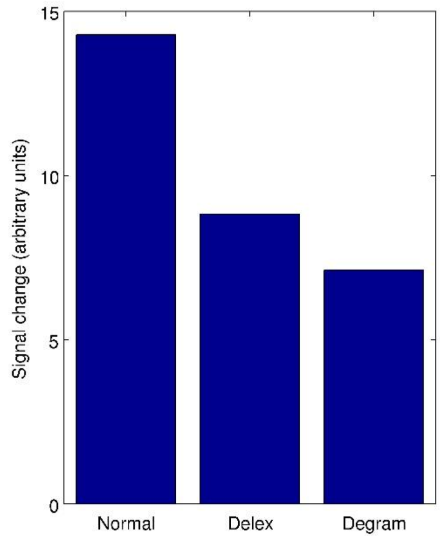

Figure 10. ROI analysis showing the amount of signal change in the MTG for 3 conditions. There were no significant differences between delexicalized and degrammaticalized stimuli.

Though smaller than anticipated, we performed a region-of-interest (ROI) analysis to determine whether this area was sensitive to how intelligibility had been degraded. Results from the ROI analysis showed that this region of the middle temporal gyrus were insensitive to how speech was degraded since there are no significant differences in signal between delexicalized and degrammaticalized stimuli.

Discussion

The selective removal of lexical and grammatical information, controlling for intelligibility, did not reveal distinct brain areas for lexical and grammatical processing. This contrasts with previous findings in the literature that have implicated regions such as the anterior temporal pole and the posterior superior temporal gyrus for syntactic processing.

When comparing our two partially intelligible stimuli (delexicalized and degrammaticalized) to our normal condition, we found a region in the middle temporal gyrus that show increased activation for intelligibility. This intelligibility effect in our results is small, did not reach significance outside of ROI analysis ($p = 0.84$, corrected), and provided no evidence for differential degradation effects as has been previously seen in the literature with low-level auditory differences (Davis & Johnsrude, 2003). Because prior studies that had used isolated sentences found more robust effects, it is possible that our narrative context played a role in the absence of differential degradation effects due to our overall small degradation effects.

Our interpretation of our findings is that placing the sentences within a familiar, narrative context minimized the effects of degradation due to increased attention. We hypothesize that attention to the narrative would increase the intelligibility of the speech signal such that lexical or grammatical degradation would not significantly affect the overall level of activation in left language regions. This is supported by a recent study by Wild et al. (2012) in which they investigated how attention to auditory stimuli affects the level of activation in left hemisphere language regions. Similar to our study, their pilot participants perceptually rated stimuli that were degraded to differing levels and found that intelligibility scores decreased as the stimuli became more degraded. Despite this, their imaging results demonstrated no significant differences in activation between clear speech and noise-vocoded speech when participants attended to the auditory stimuli.

The use of a narrative context is a potential limitation for our study. Given that we did not find strong degradation effects in our analyses, a follow-up study using isolated sentences might be necessary to see if differential effects of various types of degradation would arise.

References

Blesser, B. (1972). Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *J Speech Hear Res, 15*(1), 5-41.

Brainard, D. (1997). The psychophysics toolbox [computer software]. *Spat Vis* 10:433-436.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*(8), 3423-3431.

Friston KJ. (2007). Statistical parametric mapping the analysis of functional brain images. Amsterdam: Elsevier/Academic Press.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665-679.

Kingston, R., (Editor) & Gavin, M. (Speaker). (2001). *Fifty Famous Fairy Tales* [CD]. Blackstone Audio.

Pallier, C., Devauchelle, A., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *PNAS*, *108*(6), 2522-2527.

Scott, S., Blank, C., Rosen, S. & Wise, R. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123*(12), 2400-2406.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550-560.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience, 32*(40), 14010-14021.