Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

Neural responses to grammatically and lexically degraded speech

Alexa Bautista & Stephen M. Wilson

LANGUAGE, COGNITION AND NEUROSCIENCE, 2016
http://dx.doi.org/10.1080/23273798.2015.1123281

Neural responses to grammatically and lexically degraded speech

Alexa Bautista and Stephen M. Wilson

Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ, USA

ABSTRACT
Linguistic stimuli that are degraded in various ways have been used in neuroimaging studies to uncover distinct roles in language processing for different brain regions. To identify regions differentially involved in grammatical and lexical processing, we spectrally rotated specific morphemes and manipulated morpheme order to create speech stimuli that were degraded either grammatically or lexically, yet were matched in intelligibility. Twelve participants were scanned with functional magnetic resonance imaging (fMRI) as they listened to the grammatically and lexically degraded stimuli, interspersed with clear stimuli in the context of a familiar narrative. Contrary to our expectations, we did not find any brain regions that were selectively sensitive to grammatical or lexical degradation. However, there was less signal reduction than anticipated in response to degradation of either type. These findings may reflect increased attention to the degraded stimuli due to the narrative context, attenuating the signal decreases typically associated with reduced intelligibility.

ARTICLE HISTORY
Received 1 June 2015
Accepted 6 November 2015

KEYWORDS
Language comprehension; degraded speech; intelligibility; narrative; fMRI

Introduction

Linguistic stimuli that are degraded in various ways have been used in many functional neuroimaging studies of language processing (Scott, Blank, Rosen, & Wise, 2000). In a seminal study, Davis and Johnsrude (2003) reduced the intelligibility of sentence stimuli to different extents in three ways: by noise vocoding, by presenting speech in noise, and by alternating real speech with noise bursts. They found that regions in the bilateral superior temporal gyri were sensitive not only to the degree of intelligibility, but also to the specific manner of degradation, whereas the middle temporal gyrus was sensitive to the degree of intelligibility, but not to the acoustic form of the degraded speech. Degrading different aspects of the speech signal can reveal distinct roles for different brain regions. For example, Obleser, Eisner, and Kotz (2008) demonstrated differential responses in the left and right auditory cortices to speech that was degraded temporally or spectrally, suggesting preferential processing of different types of acoustic information in the two hemispheres.

Beyond the acoustic level, manipulations of word order and lexicality can also be thought of as means of degrading linguistic stimuli. In particular, scrambling the word order of sentences can be considered a means of degrading grammatical structure, and a number of studies have shown that scrambled sentences yield reduced activation relative to intact sentences in left anterior temporal cortex (Humphries, Binder, Medler, & Liebenthal, 2006; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene, 2011; Vandenberghe, Nobre, & Price, 2002). These findings have been interpreted as suggesting a role for this region in syntactic processing (Humphries et al., 2006) or combinatorial semantic processing (Vandenberghe et al., 2002; Wilson et al., 2014). Another frequently used manipulation is to replace content words with pseudowords ("jabberwocky" prose) (Friederici, Meyer, & von Cramon, 2000; Mazoyer et al., 1993). This can be considered a means of lexical degradation, and has been shown to result in reduced signal throughout the language network (Fedorenko, Nieto-Castañón, & Kanwisher, 2012).

An important open question is the extent to which lexical and grammatical processing are segregated in the brain (Fedorenko et al., 2012; Humphries et al., 2006; Pallier et al., 2011; Wilson et al., 2011). Some evidence suggests that grammatical processing at the sentence level depends on frontal and posterior superior temporal regions and their connections via dorsal fibre pathways, whereas lexical processing is more dependent on anterior and middle temporal and temporo-parietal regions, and ventral fibre pathways (Binder, Desai, Graves, & Conant, 2009; Friederici, 2012; Pallier et al., 2011; Saur et al., 2008; Wilson et al., 2010, 2011, 2014; Wilson, Galantucci, Tartaglia, & Gorno-Tempini, 2012). A

CONTACT Stephen M. Wilson smwilson@u.arizona.edu
Supplemental data for this article can be accessed at doi:10.1080/23273798.2015.1123281

possible approach to identifying regions associated with grammatical and lexical processing might be to compare the brain regions that are relatively sensitive to lexical or grammatical degradation. However, with the manipulations typically employed (scrambling word order, pseudowords), direct comparisons of neural activity are not feasible, because scrambled sentences and jabberwocky sentences differ significantly in "meaningfulness" or intelligibility, with the former judged more meaningful (Humphries et al., 2006). Moreover, the subjective experiences of listening to scrambled sentences and listening to jabberwocky sentences are very different. It would be difficult to draw any firm conclusions from differences in brain activity under these two conditions, given the meaningfulness confound, and the likely metalinguistic consequences of the experiential difference.

In this study, our aim was to address these barriers to comparing neural responses to grammatically and lexically degraded speech stimuli, by creating grammatically and lexically degraded stimuli for which intelligibility is matched, and for which the subjective experience of listening to them is similar. Grammatical degradation was achieved by spectrally rotating and scrambling the order of function words and morphemes. Lexical degradation was achieved by spectrally rotating open class lexical items, and mixing the spectrally rotated signal with the original signal to an empirically determined extent so as to match intelligibility with the grammatically degraded stimuli. We hypothesised that brain regions differentially involved in the processing of grammatical and lexical information would show reduced activity when processing grammatically and lexically degraded stimuli, respectively.

Methods

Participants

Seven native English speakers (3 female; mean age 21.4 years; range 20–28 years) participated in a perceptual norming experiment. Twelve native English speakers were successfully scanned with functional MRI (9 female; mean age 22.8 years; range 20–33 years; 1 left-handed). In the one left-handed participant, language was lateralised to the left hemisphere as revealed by the contrast of clear speech to scrambled speech (described below). One additional participant was unable to complete the imaging study 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.

Stimuli and norming study

The stimuli were created from audiobook recordings of the classic fairy tales "Cinderella" and "Sleeping Beauty" (Kingston & Gavin, 2001). The original durations of the narratives were 9:49 and 8:00, respectively. In each narrative, the onset and offset of each morpheme was marked, and each morpheme was classified according to its part of speech. Morphemes were then classified as lexical (nouns, proper nouns, verbs, and adjectives), grammatical (determiners, pronouns, auxiliary verbs, copulas, conjunctions, prepositions, bound inflectional and bound derivational morphemes, and the possessive clitic) or neither (adverbs, numbers, and interjections). Each narrative was then broken into consecutive segments, such that each segment was as long as possible without separating morphemes of the same word and without exceeding 7 seconds. "Cinderella" was divided into 86 segments (mean duration 6.80 ± 0.51 s) and "Sleeping Beauty" into 70 segments (mean duration 6.80 ± 0.52 s). Each segment was then processed in four different ways, in order to derive speech segments that were clear, grammatically degraded, lexically degraded, and scrambled (Figure 1).

Clear speech segments (Figure 1(A)) were created simply by low-pass filtering the recordings at 4000 Hz. A low-pass filter was required in order to acoustically match the clear condition with the other conditions, which all involved spectrally rotated speech and thus depended on an initial low-pass filter (Blesser, 1972).

Grammatically degraded speech segments (Figure 1(B)) were created by (1) low-pass filtering the recordings at 4000 Hz; (2) scrambling the order of all morphemes, subject to the constraint that lexical morphemes remained in their original order relative to one another, and (3) spectrally rotating grammatical morphemes around the 2000 Hz axis (Blesser, 1972). Keeping the lexical items in their original order relative to one another preserved some degree of intelligibility.

Lexically degraded speech segments (Figure 1(C)) were created by (1) low-pass filtering the recordings at 4000 Hz; (2) spectrally rotating lexical morphemes around the 2000 Hz axis; and (3) blending each spectrally rotated lexical morpheme with its unrotated form. By controlling the proportions of spectrally rotated and unrotated speech in these blends, the degree of lexical degradation, and thus intelligibility, could be manipulated on a continuum.

Scrambled speech segments (Figure 1(D)) were created by (1) low-pass filtering the recordings at 4000 Hz; (2) scrambling the order of all morphemes; and (3) spectrally rotating all morphemes. These stimuli were completely unintelligible.

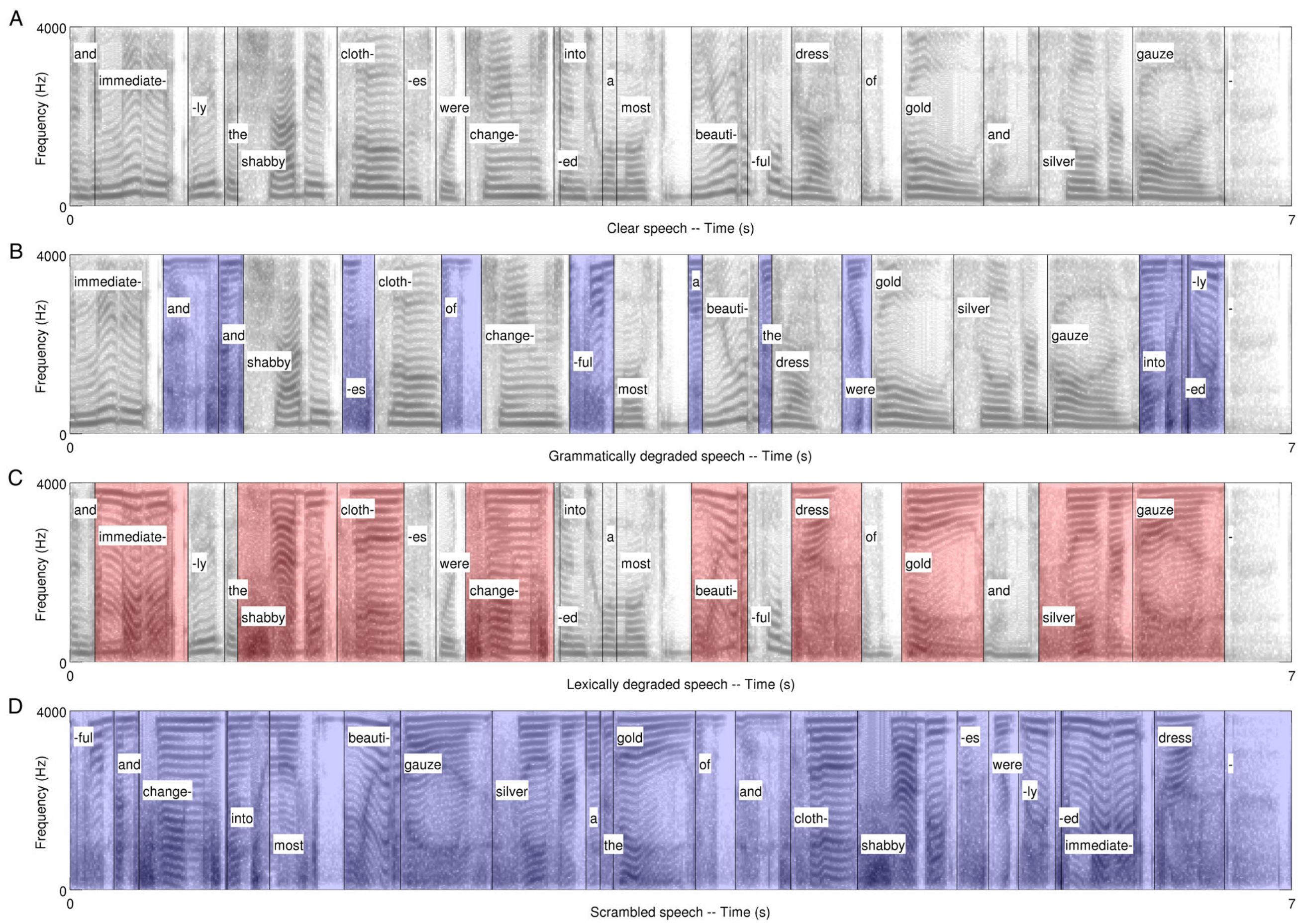

Figure 1. Four types of speech segments. (A) Clear speech. (B) Grammatically degraded speech (blue = spectrally rotated). (C) Lexically degraded speech (red = blend of spectrally rotated and unrotated). (D) Scrambled speech (blue = spectrally rotated). Audio files corresponding to these four examples are included as Supplementary material online.

In order to match the intelligibility of the grammatically and lexically degraded conditions, a perceptual norming study was carried out in which 7 native English speakers listened to the 86 segments of the "Cinderella" narrative in their original order. Clear stimuli alternated with degraded stimuli so that participants could still follow the story despite the presence of degraded stimuli. The familiar narrative was intended to provide a context in which intelligibility could be meaningfully rated for partially intelligible speech segments. The participants were instructed to rate the intelligibility of each segment, based on the following instructions: "You will hear a familiar fairy tale where some segments have been altered to make them more difficult to understand. Please try to follow the story as best you can, and indicate how well you understand each segment, using the following scale: 1 = not at all; 2 = just a bit; 3 = about half, maybe less; 4 = about half, maybe more; 5 = mostly; 6 = perfectly".

There were 44 clear segments, 18 grammatically degraded segments, and 24 lexically degraded segments. The 24 lexically degraded segments were comprised of 6 segments each of 4 different blends of spectrally rotated and unrotated lexical items: 70%/30%, 80%/20%, 90%/10%, and 100%/0%.

The experiment took place in a quiet room. Stimuli were presented and responses recorded using MATLAB R2011a (Mathworks, Natick, MA, USA) on a Lenovo S20 workstation. Participants listened to the stimuli with noise-cancelling headphones (ATH-ANC7, Audio-technica, Tokyo, Japan). After the participant provided a response for each segment, the next segment was presented.

The results of the perceptual experiment are shown in Figure 2. All participants rated all clear stimuli as perfectly intelligible without exception. The mean intelligibility of the grammatically degraded stimuli was 3.09 out of 6, that is, slightly better than "about half, maybe less". The mean intelligibility of the lexically degraded stimuli ranged from 2.18 to 4.02, as the proportion of spectrally rotated signal in the lexical morphemes ranged from 100% to 70%. By linearly interpolating between the

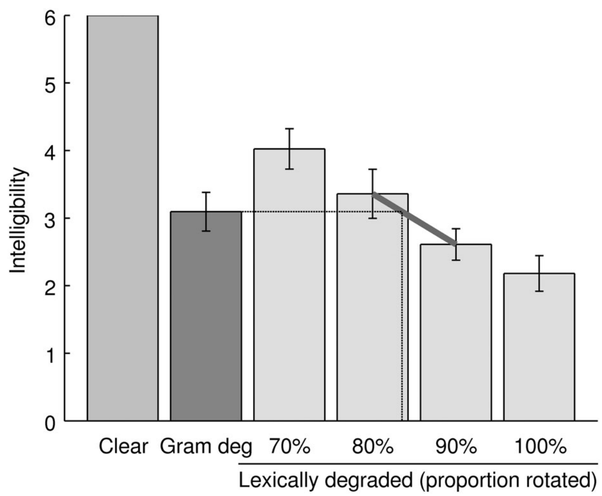

Figure 2. Matching intelligibility between grammatically and lexically degraded stimuli. The mean intelligibility of grammatically degraded stimuli was 3.09 on a 6-point scale. The mean intelligibility of lexically degraded stimuli varied according to the blend of spectrally rotated and unrotated speech in the lexical items. Linear interpolation suggests that a blend of 83.5% rotated to 16.5% unrotated would match intelligibility between the two conditions. Gram deg = grammatically degraded.

intelligibility of the 80% and 90% rotated conditions, we calculated that lexically degraded stimuli with a blend of 83.5% rotated signal and 16.5% unrotated in the lexical morphemes would be equally intelligible to the grammatically degraded stimuli (Figure 2). These proportions were used in the subsequent imaging study.

Neuroimaging experiment

Twelve healthy participants were scanned with functional MRI. Each participant completed two runs. The number of participants, number of runs, and approximate number of trials were based on Davis and Johnsrude (2003), since that study was sufficiently powered to demonstrate differential effects of different forms of acoustic degradation.

The experimental design was similar to the norming study in that the speech segments comprising a narrative were presented in order, with grammatically and lexically degraded stimuli interspersed with clear speech. However unlike the norming study, there was no task: participants were instructed simply to listen to the narratives and follow them as best as possible, paying attention to both clear and degraded parts.

Functional images were acquired with a sparse sampling paradigm so that there were silent intervals (7212 ms) between successive acquisitions. Speech segments were presented temporally centred in these silent intervals between acquisitions. Each run began with 2 or 3 clear speech segments, and then segments alternated between clear or degraded, either grammatically or lexically (pseudorandomised). The regular alternation between clear and degraded stimuli, along with the familiarity of the narratives, made it possible to follow the story despite the fact that the degraded segments were only partially intelligible. Additionally, interspersed pseudorandomly were 10 silent trials and 10 scrambled speech segments. Unlike the grammatically and lexically degraded speech segments, the scrambled segments came from elsewhere in the narratives, so they did not disrupt the progression of the narrative.

In total, "Cinderella" runs included 44 clear speech segments, 21 grammatically degraded segments, 21 lexically degraded segments (with lexical items 83.5% spectrally rotated), 10 scrambled segments, 10 silent trials, and 1 initially discarded image. "Sleeping Beauty" runs included 36 clear speech segments, 17 grammatically degraded segments, 17 lexically degraded segments (with lexical items 83.5% spectrally rotated), 10 scrambled segments, 10 silent trials, and 1 initially discarded image. An example of the first 15 volumes of one run of "Cinderella" is shown in Table 1.

Half of the participants were presented with "Cinderella" first, and half were presented with "Sleeping Beauty" first, and the stimuli were further counterbalanced such that the segments that were clear for one half of the participants were degraded for the other half of the participants, and vice versa, and the particular segments that were grammatically or lexically degraded varied pseudorandomly across participants.

Participants were scanned on a Siemens Skyra 3 T scanner with a 32-channel head coil at the University of Arizona. The stimuli were controlled with the Psychophysics Toolbox version 3.0.10 (Brainard, 1997; Pelli, 1997) running under MATLAB R2012b on a Lenovo S30 workstation. Participants listened to the narratives using insert earphones (S14, Sensimetrics, Malden, MA, USA). The presentation volume was adjusted to a comfortable level for each participant.

T2*-weighted blood oxygenation level dependent (BOLD) echo planar images were collected 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 × 240 mm; matrix = 86 × 96 mm; TR = 9500 ms; TA = 2288 ms; TE = 30 ms; flip angle = 90°; voxel size = 2.5 × 2.5 × 3.5 mm. In the "Cinderella" runs, 107 volumes were acquired (duration 16:57), and in the "Sleeping Beauty" runs, 91 volumes were acquired (duration 14:25). For anatomical reference, T1-weighted MPRAGE structural images were also acquired (voxel size = 0.9 × 0.9 × 0.9 mm).

Table 1. Example first 15 volumes of a functional run.

Volume	Condition	Segment
1	Clear	1
2	Clear	2
3	Lexically degraded	3
4	Clear	4
5	Grammatically degraded	5
6	Clear	6
7	Silence	N/A
8	Grammatically degraded	7
9	Clear	8
10	Lexically degraded	9
11	Clear	10
12	Scrambled	47
13	Grammatically degraded	11
14	Clear	12
15	Lexically degraded	13

Note that silence did not interrupt the sequence of segments, and that the scrambled trial was based on an arbitrary segment from elsewhere in the narrative, and thus also did not interrupt the sequence.

Analysis of neuroimaging data

The data were first preprocessed with tools from AFNI version 2011-06-22 (Cox, 1996). Head motion was corrected, with six translation and rotation parameters saved for use as covariates, then the data were detrended with a Legendre polynomial of degree 2, and smoothed with a Gaussian kernel (FWHM = 6 mm). Next, independent component analysis was performed using the *fsl* tool *melodic* version 3.13 (Beckmann & Smith, 2004). Noise components were manually identified with reference to the criteria of Kelly et al. (2010) and removed using *fsl_regfilt*. A general linear model was fit with the program *fmrilm* from the FMRISTAT package (Worsley et al., 2002). The six motion parameters were included as covariates, as were time-series from white matter and cerebro-spinal fluid (CSF) regions (means of voxels segmented as white matter or CSF in the vicinity of the lateral ventricles) and three cubic spline temporal trends. No hemodynamic response function was modelled; instead, each volume was assumed to reflect the BOLD response to neural activity relating to the immediately preceding segment. The two runs for each participant were combined in a fixed effects model using the FMRISTAT program *multistat*. The T1-weighted anatomical images were warped to Montreal Neurological Institute (MNI) space using unified segmentation in SPM5 (Ashburner & Friston, 2005). Functional images were coregistered with structural images and warped to MNI space.

Group analyses were carried out with random effects models in SPM5. The following contrasts were computed: (1) grammatically degraded speech versus lexically degraded speech; (2) lexically degraded speech versus grammatically degraded speech; (3) clear speech versus scrambled speech; (4) grammatically or lexically degraded speech (i.e. the average of these conditions) versus scrambled speech; (5) clear speech versus grammatically or lexically degraded speech (i.e. the average of these). The last of these contrasts was masked with clear speech versus silence (inclusive mask, $p < .05$, uncorrected).

All contrasts were thresholded at voxelwise $p < .005$, then corrected for multiple comparisons at $p < .05$ based on cluster size based on Gaussian random field theory as implemented in SPM5 (Worsley et al., 1996). A single cluster for the contrast of clear speech versus grammatically or lexically degraded speech that occurred closest to the a priori anticipated location in the left anterior temporal lobe (Davis & Johnsrude, 2003) was thresholded based on the method described by Friston (1997).

We also carried out a region of interest (ROI) analysis to compare neural responses to grammatically and lexically degraded speech in six left hemisphere ROIs identified in a previous study that investigated grammatical and lexical processing (Pallier et al., 2011). Three of these regions showed activation as a function of constituent size for both words and pseudowords, suggesting a role in grammatical processing: the pars triangularis of the inferior frontal gyrus, pars orbitalis of the inferior frontal gyrus, and posterior superior temporal sulcus (STS). The other three regions showed this pattern only with real words, and not with pseudowords, suggesting a role in lexical processing: the temporal pole, anterior STS, and temporo-parietal junction. ROIs were defined as spheres with 10 mm radius around the coordinates reported (see Pallier et al., 2011, Table S5). Paired *t*-tests were performed between estimates of signal change for grammatically degraded and lexically degraded speech.

Results

To test our hypothesis that neural responses would differ between grammatically and lexically degraded speech, we first compared the grammatically degraded and lexically degraded conditions. These contrasts revealed no significant differences in either direction. The six a priori ROIs also showed no differences between responses to grammatically and lexically degraded speech. There was a trend in the anticipated direction in the left temporo-parietal junction, where there was more activation for grammatically degraded speech than lexically degraded speech ($p = .052$, two tailed), but no other regions showed any indication of an effect (all $p \geq .59$).

After failing to find differences between grammatically and lexically degraded speech, we next examined the overall effect of degradation. The contrast of clear speech to scrambled speech revealed activation along the left STS (Figure 3(A) and Table 2), and similar but less extensive activation of the right STS. The contrast of grammatically or lexically degraded speech to scrambled speech yielded a similar pattern of activation (Figure 3(B) and Table 2). When clear speech was compared directly to grammatically or lexically degraded speech, we found that only a small region in the left anterior STS was more active for clear speech than for grammatically or lexically degraded speech (Figure 3(C) and Table 2). This region showed similarly reduced signal for both grammatically or lexically degraded speech (Figure 3(D)).

Discussion

We created grammatically and lexically degraded stimuli in which intelligibility was matched, and the subjective experience of listening to the two types of degraded stimuli was similar. Therefore, it was feasible to contrast neural responses to grammatical and lexical degradation. However, contrary to our hypothesis, we did not observe any brain regions that showed differential activity for processing grammatically or lexically degraded speech.

One possible explanation for this null result is that grammatical and lexical processing may be closely intertwined in the brain (Bates & Goodman, 1997). If the same brain regions are similarly involved in grammatical and lexical processing, then it would follow that these regions would show similar signal reductions when grammatical or lexical information is selectively degraded. However, there is considerable evidence that grammatical and lexical processing are segregated, in particular the selective grammatical and lexical impairments that are often observed in primary progressive aphasia (Gorno-Tempini et al., 2004; Hodges & Patterson, 1996; Schwartz, Marin, & Saffran, 1979; Wilson et al., 2010, 2011, 2012, 2014) and to a lesser extent in other forms of aphasia (Caplan, 1987; Ullman et al., 2005). Functional imaging studies have also suggested distinct neural correlates of grammatical and lexical processing (Binder et al., 2009; Friederici, 2012; Pallier et al., 2011; Saur et al., 2008; Wilson et al., 2014). Given this body of evidence, we should not be too hasty to interpret our null result as evidence against neural segregation of grammatical and lexical processing.

Indeed, the potential of our study to reveal differential effects of grammatical and lexical degradation was severely limited by the small magnitude of the neural effects of degradation in general (Figure 3(C)). Many studies have shown that left temporal regions in particular are robustly sensitive to intelligibility (Davis & Johnsrude, 2003; Erb, Henry, Eisner, & Obleser, 2013; Kyong et al., 2014; Obleser & Kotz, 2010), yet our study showed only a modest effect in this region, which was statistically significant only given its a priori expected location.

Why was our intelligibility effect so modest? Intelligibility was certainly sufficiently reduced in the degraded stimuli, as the results of the perceptual norming study showed. And while we scanned only 12 participants, the design of our study was comparable to a previous study that showed a strong effect of intelligibility (Davis & Johnsrude, 2003). We think that the explanation for our modest intelligibility effect lies in the familiar narrative context in which the degraded stimuli were presented. We chose this experimental design in order to create a relatively naturalistic listening context which would maximise the experiential similarity of perceiving the two kinds of degraded speech. Yet presenting the degraded stimuli in this way created a context in which the degraded stimuli may have been processed with greater effort and attention than is typical in similar experiments, most of which involve isolated sentences (e.g. Davis & Johnsrude, 2003).

A recent neuroimaging study showed that the typical modulation of neural signal in the STS by intelligibility disappeared when participants actively attended to clear or degraded speech in the presence of auditory and visual distracters (Wild et al., 2012). Wild et al. interpreted their results as suggesting that the STS is sensitive simply to intelligibility, regardless of how that intelligibility was achieved. This was based on the fact that

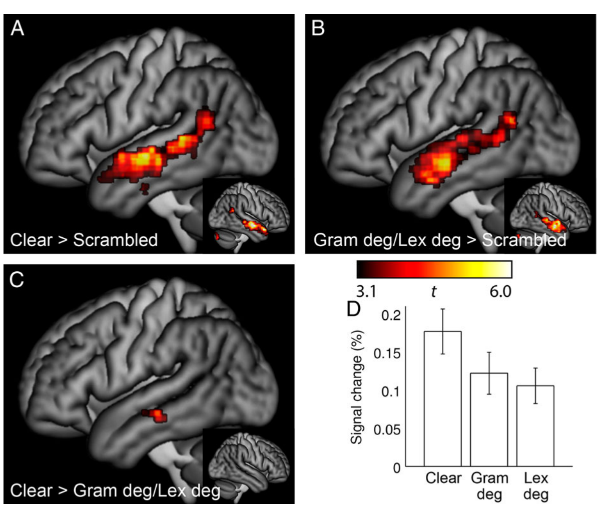

Figure 3. Neuroimaging results. (A) Brain regions that were activated by clear speech relative to scrambled speech. (B) Brain regions that were activated by grammatically or lexically degraded speech relative to scrambled speech. (C) A brain region that was activated by clear speech relative to grammatically or lexically degraded speech, significant due only to its a priori expected location. (D) Signal change in the region shown in (C). This region showed similarly reduced responses for grammatically and lexically degraded speech. Gram deg = grammatically degraded; Lex deg = lexically degraded.

Table 2. Brain activations for key contrasts.

Brain region	MNI coordinates x	y	z	Extent (mm^3)	Max *t*	*p*
Grammatically degraded > Lexically degraded						
(None)						
Lexically degraded > Grammatically degraded						
(None)						
Clear > Scrambled						
Left STS	−55.3	−25.4	−4.5	15,408	6.41	<0.001
Right anterior STS	55.1	−6.9	−13.7	9400	7.68	<0.001
Right cerebellum	20.3	−79.8	−33.7	1688	4.42	0.018
Right posterior STS	47.4	−52.5	21.5	1448	5.18	0.042
Grammatically or lexically degraded > Scrambled						
Right STS	56	−13.2	−9	12,896	6.85	<0.001
Left STS	−57.1	−21.1	−4.8	12,408	5.63	<0.001
Right cerebellum	17.5	−76.5	−31.7	1488	5.17	0.031
Clear > Grammatically or lexically degraded						
Left anterior STS	−55.5	−22.8	−12.2	576	4.57	0.016*

MNI coordinates show centres of mass.
STS, superior temporal sulcus.
*Corrected *p* value based on a priori expected location.

post-scan recall was equivalent for attended clear and attended degraded sentences. However, this interpretation is not entirely consistent with their findings, since although post-scan recall was equivalent for the two stimulus types, response-based measures of intelligibility were not. An alternative account is that STS activity reflects not intelligibility per se, but rather the extent of linguistic processing taking place. In most experiments involving degraded speech, less linguistic processing takes place when sentences are degraded, because some words and phrases cannot be recognised, and thus cannot be processed further. However in certain situations, this reduction in processing may be counterbalanced by an increase in other types of linguistic processing. For instance, in the present study, integrating a partially understood sentence into the narrative context may be more demanding than integrating a completely intelligible sentence. In the Wild et al. (2012) study, extracting degraded sentences from mixed auditory stimuli including distracters may have required more linguistic processing than extracting clear sentences. Our account could be tested in future work by systematically investigating the dependence of STS signal on degradation under different processing conditions, such as context, task, attention, and distracters.

In summary, we did not observe any brain regions that showed differential activity for processing grammatically or lexically degraded speech. However, we found only modest degradation effects in general, which may have been due to the narrative context in which all stimuli were presented. Our findings raise the question of whether STS signal in language comprehension is driven by intelligibility or by extent of linguistic processing.

Acknowledgements

We thank Scott Squire and Angelica McCarron for technical assistance, the individuals who participated in the study, and an anonymous reviewer for his/her constructive comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders (NIH R01 DC013270), the State of Arizona (ADHS14-052688), and the Western Alliance to Expand Student Opportunities (WAESO).

ORCID

Alexa Bautista http://orcid.org/0000-0002-4487-9593
Stephen Wilson http://orcid.org/0000-0001-9884-2852

References

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. doi:10.1016/j.neuroimage.2005.02.018

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*(5–6), 507–584. doi:10.1080/016909697386628

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*(2), 137–152. doi:10.1109/tmi.2003.822821

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. doi:10.1093/cercor/bhp055

Blesser, B. (1972). Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *Journal of Speech and Hearing Research*, *15*(1), 5–41. doi:10.1044/jshr.1501.05

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. doi:10.1163/156856897X00357

Caplan, D. (1987). *Neurolinguistics and linguistic aphasiology: An introduction*. Cambridge: Cambridge University Press.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. doi:10.1006/cbmr.1996.0014

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

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *Journal of Neuroscience*, *33*(26), 10688–10697. doi:10.1523/jneurosci.4596-12.2013

Fedorenko, E., Nieto-Castañón, A., & Kanwisher, N. (2012). Syntactic processing in the human brain: What we know, what we don't know, and a suggestion for how to proceed. *Brain and Language*, *120*(2), 187–207. doi:10.1016/j.bandl.2011.01.001

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. doi:10.1016/j.tics.2012.04.001

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, *75*(2), 289–300. doi:10.1006/brln.2000.2313

Friston, K. J. (1997). Testing for anatomically specified regional effects. *Human Brain Mapping*, *5*(2), 133–136. doi:10.1002/(sici)1097-0193(1997)5:2%3C133::aid-hbm7%3E3.0.co;2-4

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., … Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, *55*(3), 335–346. doi:10.1002/ana.10825

Hodges, J. R., & Patterson, K. (1996). Nonfluent progressive aphasia and semantic dementia: A comparative neuropsychological study. *Journal of the International Neuropsychological Society*, *2*(06), 511–524. doi:10.1017/S1355617700001685

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. doi:10.1162/jocn.2006.18.4.665

Kelly, R. E.Jr, Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murphy, C. F., Morimoto, S. S., … Hoptman, M. J. (2010). Visual inspection of independent components: Defining a procedure for artifact removal from fMRI data. *Journal of Neuroscience Methods*, *189*(2), 233–245. doi:10.1016/j.jneumeth.2010.03.028

Kingston, R. (Author), & Gavin, M. (Narrator). (2001). *Fifty famous fairy tales [Audio book]*. Ashland, OR: Blackstone Audio.

Kyong, J. S., Scott, S. K., Rosen, S., Howe, T. B., Agnew, Z. K., & McGettigan, C. (2014). Exploring the roles of spectral detail and intonation contour in speech intelligibility: An FMRI study. *Journal of Cognitive Neuroscience*, *26*(8), 1748–1763. doi:10.1162/jocn_a_00583

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., … Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467–479. doi:10.1162/jocn.1993.5.4.467

Obleser, J., Eisner, F., & Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *Journal of Neuroscience*, *28*(32), 8116–8123. doi:10.1523/jneurosci.1290-08.2008

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex*, *20*(3), 633–640. doi:10.1093/cercor/bhp128

Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(6), 2522–2527. doi:10.1073/pnas.1018711108

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. doi:10.1163/156856897X00366

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., … Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 18035–18040. doi:10.1073/pnas.0805234105

Schwartz, M. F., Marin, O. S. M., & Saffran, E. M. (1979). Dissociations of language function in dementia: A case study. *Brain and Language*, *7*(3), 277–306. doi:10.1016/0093-934X(79)90024-5

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*(12), 2400–2406. doi:10.1093/brain/123.12.2400

Ullman, M. T., Pancheva, R., Love, T., Yee, E., Swinney, D., & Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language*, *93*(2), 185–238. doi:10.1016/j.bandl.2004.10.001

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. doi:10.1162/08989290260045800

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. *Journal of Neuroscience*, *32*(40), 14010–14021. doi:10.1523/jneurosci.1528-12.2012

Wilson, S. M., DeMarco, A. T., Henry, M. L., Gesierich, B., Babiak, M., Mandelli, M. L., … Gorno-Tempini, M. L. (2014). What role does the anterior temporal lobe play in sentence-level processing? Neural correlates of syntactic processing in semantic variant primary progressive aphasia. *Journal of Cognitive Neuroscience*, *26*(5), 970–985. doi:10.1162/jocn_a_00550

Wilson, S. M., Galantucci, S., Tartaglia, M. C., & Gorno-Tempini, M. L. (2012). The neural basis of syntactic deficits in primary progressive aphasia. *Brain and Language*, *122*(3), 190–198. doi:10.1016/j.bandl.2012.04.005

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., … Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron*, *72*(2), 397–403. doi:10.1016/j.neuron.2011.09.014

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., … Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069–2088. doi:10.1093/brain/awq129

Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., & Evans, A. C. (2002). A general statistical analysis for fMRI data. *NeuroImage*, *15*(1), 1–15. doi:10.1006/nimg.2001.0933

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73. 10.1002/(sici)1097-0193(1996)4:1%3C58::aid-hbm4%3E3.0.co;2-o