NeuroImage 56 (2011) 2310–2320

Contents lists available at ScienceDirect

NeuroImage

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

Dynamic assignment of neural resources in auditory comprehension of complex sentences

Jonas Obleser *, Lars Meyer, Angela D. Friederici

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

Article history:
Received 10 December 2010
Revised 8 March 2011
Accepted 11 March 2011
Available online 21 March 2011

Keywords:
Audition
Adverse listening
Intelligibility
Syntax
Language comprehension
Functional MRI
Processing pathways

ABSTRACT

Under real-life adverse listening conditions, the interdependence of the brain's analysis of language structure (syntax) and its analysis of the acoustic signal is unclear. In two fMRI experiments, we first tested the functional neural organization when listening to increasingly complex syntax in fMRI. We then tested parametric combinations of syntactic complexity (argument scrambling in three degrees) with speech signal degradation (noise-band vocoding in three different numbers of bands), to shed light on the mutual dependency of sound and syntax analysis along the neural processing pathways. The left anterior and the posterior superior temporal sulcus (STS) as well as the left inferior frontal cortex (IFG) were linearly more activated as syntactic complexity increased (Experiment 1). In Experiment 2, when syntactic complexity was combined with improving signal quality, this pattern was replicated. However, when syntactic complexity was additive to degrading signal quality, the syntactic complexity effect in the IFG shifted dorsally and medially, and the activation effect in the left posterior STS shifted from posterior toward more middle sections of the sulcus. A distribution analysis of supra- as well as subthreshold data was indicative of this pattern of shifts in the anterior and posterior STS and within the IFG. Results suggest a signal quality gradient within the fronto-temporal language network. More signal-bound processing areas, lower in the processing hierarchy, become relatively more recruited for the analysis of complex language input under more challenging acoustic conditions ("upstream delegation"). This finding provides evidence for dynamic resource assignments along the neural pathways in auditory language comprehension.

Introduction

Research into the functional neuroanatomy of speech and language perception has come far over the last twenty years of neuroimaging (for recent reviews see e.g., Hickok and Poeppel, 2007; Rauschecker and Scott, 2009), and major pathways of acoustic information processing have been identified (Hackett, 2008; Romanski et al., 1999).

However, it is difficult to conclusively separate the auditory analysis of the speech signal – the "carrier" – from the structural analysis of its "content", i.e. the grammatical relations between different elements as conveying who is doing what to whom. Specifically, the processing of syntax is mostly studied apart from the processing of the auditory carrier signal, often even in visual reading experiments. Yet substantial progress beyond Broca's and Wernicke's initial findings has been made: distinct substructures in the inferior frontal gyrus (IFG) as well as in the posterior superior temporal gyrus (STG) and sulcus (STS) and the fiber tracts connecting them are now known to be decisive for one's ability to "parse" the structure of heard (or read) sentences (for recent reviews see e.g., Friederici, 2009; Grodzinsky and Santi, 2008; Hagoort, 2005), especially so as a sentence's structure becomes increasingly complex — a key feature of human grammar (Friederici et al., 2006a, 2006b; Hauser et al., 2002).

The intimate link and interdependence of carrier and content in speech become obvious when the acoustic listening conditions become compromised, such as in noisy environments, in degraded hearing, or most drastically in listeners with a cochlear implant (for simulations see e.g. Scott et al., 2006; Shannon et al., 1995).

Neural correlates of speech intelligibility have been studied extensively over the last decade or so in functional neuroimaging experiments (Davis and Johnsrude, 2003; Obleser and Kotz, 2010; Okada et al., 2010; Scott et al., 2000; Zekveld et al., 2006). The bilateral anterior and the lateral superior temporal cortex (in some studies with an additional posterior extent; e.g., Narain et al., 2003) and also the left inferior frontal cortex (e.g., Davis and Johnsrude, 2003) have been found to increase in hemodynamic activity with more intelligible speech. Mostly, not very much emphasis had been on the specific linguistic processes that are of course directly linked to intelligibility. Only more recently, researchers have directly addressed the interaction of linguistic processing and speech intelligibility on the one hand (e.g., Friederici et al., 2010) and the effort that accompanies speech comprehension under degraded conditions on the other hand (Harris

---

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany.
E-mail address: obleser@cbs.mpg.de (J. Obleser).

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.03.035

et al., 2009; Davis and Johnsrude, 2003; Peelle et al., 2010; Scott et al., 2009).

With respect to the present study, it is unclear how exactly such challenging auditory bottom-up circumstances interact with higher-level language processes during comprehension, and what the consequences within the cortical language network are. In particular, how do the IFG and the posterior STG/STS react as listening (i.e., the acoustic analysis) becomes more taxing, as it is often the case in real-life situations? Here, syntactic complexity is operationalized as so-called argument scrambling, where – in three levels of complexity – the relative order of verbal arguments in German sentences is changed, while keeping the sentences grammatically well-formed (e.g. Friederici et al., 2006b; Grewe et al., 2005).

In the first parametric functional imaging experiment, we test for the functional neuroanatomy of demanding syntactic processes in the auditory domain. We hypothesize that the pars opercularis (PO) of the left IFG (Brodmann's Area [BA] 44) and the posterior third of the STS (pSTS) will be increasingly activated as syntactic complexity of acoustically presented sentences increases (Friederici et al., 2006b; Vigneau et al., 2006).

In the second experiment, this linguistic manipulation is combined with demanding acoustic processing by parametrically degrading the spectral detail of the "carrier" through noise-band vocoding (Shannon et al., 1995). Unlike inference of meaning, which can profit from strong semantic associations and elicits widely distributed extra-auditory brain areas (e.g., Obleser and Kotz, 2010; Obleser et al., 2007; Rodd et al., 2005), complex grammar under such circumstances can only be parsed by analyzing the acoustic details as thoroughly as possible.

Accordingly, in the second experiment we will look into the same parametric variation of syntactic complexity as in the first. However, we hypothesize that additional gradual signal degradations will affect the peak locations of the syntax complexity effect within the fronto-temporal language network, reflecting the additional or relatively stronger recruitment of neural structures most responsive to the acoustic variation that is introduced by speech signal degradation. If so, this would provide evidence for the dynamic allocation of fronto-temporal brain structures to syntactic analysis, depending on further constraints and specifics, such as signal quality.

Methods

The study consists of two functional MRI experiments and one flanking behavioral experiment. fMRI Experiment 1 was designed to establish whether the syntactic complexity (with three levels of argument scrambling) that had been observed with visual presentation earlier (Friederici et al., 2006b) could be also found in the auditory domain, in a setup without any speech degradation. A behavioral pilot established the acoustic degradation levels selected for fMRI Experiment 2. Functional MRI Experiment 2 was the main experiment and combined the three levels of syntactic complexity with three levels of acoustic signal degradation.

Subjects

Sixteen participants (8 females, 26.7 ± 2.7 years of age, M ± SD) took part in fMRI Experiment 1. 14 participants (6 females, 25.7 ± 3.3 years of age, M ± SD) volunteered for the behavioral pilot experiment. Another 14 participants (8 females, 23.4 ± 2.1 years of age, M ± SD) were recruited for fMRI Experiment 2. All participants in all experiments were right-handed and were native speakers of German. They had normal or corrected to normal vision, and did not report any history of neurological, psychiatric, or hearing disorder. No subject had had previous exposure to noise-vocoded speech, and all were naïve as to the purpose of the study. Participants received financial compensation of 14 € for the fMRI experiments and 7 € for the behavioral pilot experiment, respectively. The local ethics committee of the University of Leipzig approved all procedures.

Stimuli

All experiments used auditory recordings of a sentence set developed and tested previously only as visual stimuli (Friederici et al., 2006b). The set consists of 48 ditransitive German sentences in the past tense with the verb participle at the end of the sentence. Each sentence occurs in a triplet of three syntactic conditions that varied the order of the verb's so-called arguments (i.e., the subject, the direct object, and the indirect object; Fig. 2B): condition A uses the German canonical word order, subject–indirect object–direct object (i.e., the arguments are not "scrambled"). Conditions B and C apply the argument scrambling iteratively, with one or two repositioned arguments, respectively (Friederici et al., 2006b). Thus, this design allows for a parametric variation of syntactic complexity. Note, however, that all variations are well within the bounds of grammatical German; see Friederici et al. (2006b) for grammaticality judgments.

All sentences were single-channel recorded by a trained female speaker of German and digitized at a sampling rate of 44.10 kHz. Offline, the sentences were edited into separate audio files, re-sampled to 22.05 kHz, and normalized for root mean squared amplitude. The mean length of the sentences was 2.96 s (±0.3 s SD). This set of 48 × (A,B,C argument order) = 144 stimuli was used in fMRI Experiment 1.

For the main fMRI experiment (Experiment 2) as well as the related behavioral pilot experiment (see below), all sentence recordings were additionally submitted to a Matlab-based (The MathWorks, Inc., Natick, MA, USA) routine for speech degradation in arbitrary levels (noise-band vocoding; Shannon et al., 1995). Noise-band vocoding is an effective manipulation of the amount of spectral detail. It preserves the temporal envelope of the speech signal and renders it more or less intelligible in a graded and controlled fashion, depending on the number of bands used. More bands yield a more intelligible speech signal.

In vocoding, the bands were equally spaced using the Greenwood formula (as implemented in Rosen et al., 1999). The filter cut-offs were adjacent but non-overlapping and were linearly spaced on the log frequency axis. The passband for filtering into channels/bands and envelope extraction was set to 70–9000 Hz; the lowpass filter cutoff for the envelope extraction was set at 256 Hz.

Of each sentence, five different vocoded versions were created with varying numbers of filter bands (2, 4, 8, 16 and 32 bands) to allow for a parametric variation of speech intelligibility. Two bands result in a drastic reduction in spectral resolution and a concomitantly very low intelligibility, whereas 32 bands almost restore the normal spectral distribution of the signal and are thereby almost readily intelligible to naïve listeners.

Procedure

MRI acquisition

Scanning was performed using a Siemens Trio 3-T scanner with a 12-channel SENSE head coil. Participants were comfortably positioned in the bore and wore air-conduction headphones (Resonance Technology). All data were acquired in a sparse temporal sampling setup, where volume acquisition is clustered at the beginning of each TR to allow for auditory sentence presentation in silence and for the hemodynamic changes driven by scanner noise to fade somewhat before the next volume acquisition (Hall et al., 1999). Sentences were, in both experiments, presented 5.5 s prior to the ensuing volume acquisition onset to capture the rising slope or maximum of the hemodynamic response elicited by the sentences in the ensuing volume acquisition (Obleser et al., 2007).

Echo-planar imaging (EPI) scans were acquired in 26 axial slices covering the forebrain for almost all head sizes, ensuring temporal and inferior frontal coverage. Scans had an in-plane resolution of 3×3 mm^2 and a 3-mm slice thickness (TR = 9 s, TA = 2 s, TE = 30 ms, flip angle 90°, field of view 192 mm, matrix size 64×64, interleaved slice acquisition and no gap between slices). For all participants, individual high-resolution 3D T1-weighted MR scans acquired in previous sessions were available for normalization and co-registration (MP-RAGE; TR = 1300 ms; TE = 3.93 ms; FOV = 256 mm × 240 mm; slab thickness = 192 mm; 128 partitions; sagittal orientation; spatial resolution $1 \times 1 \times 1.5$ mm).

The acquired number of volumes varied between both experiments: Experiment 1 had 240 scans, and Experiment 2 only 210 (as no catch trials were used; see below). Two dummy volume acquisitions preceded all acquisitions of fMRI time series to limit initial longitudinal magnetization effects.

Functional MRI Experiment 1

In the first experiment, no speech degradation but a clear audio signal was used in order to establish a reliable estimate of brain regions solely activated by the syntactic complexity manipulation.

After a brief (10-trial) familiarization period, the actual experiment was started. Participants were stimulated with a carefully pseudo-randomized sequence of the 48 sentences in either one of the complexity (argument scrambling) levels A, B, or C. Together with a set of 48 simple filler sentences spoken by the same speaker and recorded previously (e.g., [He builds the house], [Everyone eats]); and an equal amount of silent trials (null events), this yielded a total of 240 trials/volumes of interest. Four different pseudo-randomized stimulus sequences (constrained not to begin with silent trials, to have a roughly evenly-spaced distribution oft the silent trials across time, and not to have two succeeding task trials) were designed and used in a counterbalanced fashion across all participants.

Interspersed we rarely presented catch trials (20% in total; equally drawn from all three syntactic complexity levels), which prompted the participant after a trial's audio playback with a sentence written on-screen. The latter was either the identical sentence or one of the two syntactically non-matching sentences from the same (semantic) triplet (see Fig. 3; for example, the audio in a catch trial could have been sentence B, while the written and visually presented one could have been either sentence B again (match) or sentence A or C; non-match). The participants indicated match/non-match by a (counterbalanced) left index finger — match, and right index finger — non-match button press. In catch trials, there was a response window of 2000 ms just before the next volume was acquired, and participants indicated their response by a (counterbalanced) left index finger — correct, and right index finger — incorrect button press. The task fulfilled a double purpose. It established a moderate control of the participants' attention (but see Experiment 2), but more importantly it ensured that the first auditory experiment on this material would be as closely comparable as possible to the previous fMRI study using these complex sentences in written format (Friederici et al., 2006b).

Behavioral experiment

A sentence–sentence matching task was used to find adequate signal degradation levels to be further used in the ensuing fMRI Experiment 2. To this end, an auditory sentence was presented via headphones first. Then, either the identical – matching – or one of the two syntactically non-matching sentences from the same triplet would appear on-screen. As in the catch trials in the first fMRI study, the combination of a sentence heard in complexity level A and seen in complexity level A would be correctly classified as a match, whereas the combination of a sentence heard in complexity level A, but seen in complexity level B or C would be correctly classified as a non-match.

Stimuli were pseudo-randomized and counterbalanced into seven lists, of which one was presented to each participant using the Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA) on a Windows PC. While a fixation cross was shown, an auditory stimulus was presented, followed by a visual stimulus that stayed on-screen until the participant pushed a button (with a time out after 6 s at most; average response time was 1.34 s ± 0.63 s standard deviation). The total duration of the procedure was 30 min.

The percentage of correct button presses was analyzed using SPSS (SPSS, Inc., Chicago, IL, USA). The two independent variables were syntactic complexity and acoustic degradation, dependent being percentage correct. Since Mauchly's test on sphericity was found significant ($p < 0.001$), a multivariate analysis of variance (ANOVA) statistic was applied and Wilks' Lamda approximated F-values are being reported.

Functional MRI Experiment 2

In the second fMRI experiment, a parametric 3×3 design of complexity and degradation was used with 14 trials in each cell (i.e., 42 for each vocoding level or 42 for each argument scrambling level, respectively). Here, we also used 56 filler sentences (see description above), also in noise-vocoded formats with a varying number of bands to avoid any perceptual pop-out effects or introducing confounding correlations between number of bands and sentence complexity. With 28 silent trials (null events) added, a total amount of 210 volumes/trials of interest was acquired. Three different pseudo-randomized stimulus sequences (constrained only not to begin with silent trials and to have a roughly evenly-spaced distribution oft the silent trials across time) were designed and used in a counterbalanced fashion.

The participants just listened attentively to the sentences. Unlike in Experiment 1, they performed no additional active task for two reasons. Firstly, various previous studies have shown advantageously that passive listening tasks that mimic natural speech comprehension yield strong activations throughout the whole perisylvian region (Obleser et al., 2007; Price et al., 2005; Scott et al., 2000, 2006) including the inferior frontal gyrus. Secondly, an active sentence-matching task (as employed in Experiment 1; see above) could have biased the well-balanced acoustic × syntactic manipulation design in various ways. For example, the syntactic task would have been very likely to be more difficult to perform on more degraded speech. Also, performing a syntactic task, but not a concurrent intelligibility task, would have possibly biased the results toward syntactic mechanisms. Participants were placed in the scanner and, after a few familiarization trials, were instructed to listen attentively.

Data analysis

For data processing and analysis, SPM8 was used (Wellcome Trust Centre for Neuroimaging, University College, London, UK). Volumes of the fMRI time series were resampled to a cubic $2 \times 2 \times 2$ mm^3 voxel size; realigned and corrected for field inhomogeneities ("unwarped"); normalized to a template in the MNI coordinate system (using the unified segmentation-based procedure, which first segments an individual brain into tissue-specific images using according tissue probability maps in MNI space; Ashburner and Friston, 2005); and smoothed using an isotropic 8-mm^3 kernel.

In both MRI experiments, a general linear model (GLM) based on the parametric (i.e., three-level) variation of *syntactic complexity* (Exp. 1) or the three-by-three level variations *syntactic complexity* and *degradation* (Exp. 2) was estimated in each participant, using a finite impulse response basis function (order 1, window length 1). Occurrences of filler sentences (Exps. 1 and 2) or response trials (only applicable in Experiment 1) were modeled as regressors of no interest in the single-subject GLMs. Contrast estimates of the three syntactic complexity levels (compared against the global mean) from all participants were submitted to a second-level within-subject analysis of variance (ANOVA) in SPM in Experiment 1. In Experiment 2, there were of course nine such contrast estimates (three syntactic levels in three acoustic degradation formats), which were accordingly submitted to the second-level within-subject analysis of variance (ANOVA) in SPM. Fig. 4 exemplifies the two SPM contrasts of complexity and degradation, entitled “trade-off” and “additivity”, that were assessed at the second level.

For thresholding of the statistical parametrical maps, a Matlab-implemented Monte Carlo simulation (Slotnick et al., 2003; 1000 iterations, no volume mask) suggested a cluster extent threshold of at least 40 resampled voxels and an uncorrected p-value of 0.001 to ensure a whole-volume type I error probability smaller than 0.05 (code available at http://www2.bc.edu/~slotnics/scripts.htm). Thus, all activation overlay figures presented show only suprathreshold activation.

Distribution analysis

In addition, we pursued the hypothesis that for the two main comparisons (“additivity”, shown in red, versus “trade-off” shown in blue in Fig. 4) Z-scores from the group-level analysis should vary differently as a function of space (e.g., the posterior–anterior axis). To this end, we extracted all (i.e., sub- as well as supra-threshold) voxels' Z-scores along a given axis and from a given brain region, for example, the left mid to posterior STS/STG. This results in two histogram-like distributions (Fig. 6), which show peaks at the voxel locations known from the conventional SPM results yet also allow taking into account the neighboring sub-threshold Z-scores, that is, the distribution.

To give an example in more detail, a plot was created of reasonable MNI coordinates along the axis of interest (e.g., $-20>x>-60$) versus the maximum statistical score (Z-score) observed at each coordinate (effectively projecting down onto this axis and ignoring the other two spatial axes). This was done separately for the Z-scores from the “additivity” and from the “trade-off” contrast (Fig. 4). Treating these two sets of data as observations in a (spatial) histogram, the data could be submitted to a non-parametric test for two-sample distribution inequality, the Cramér–Von Mises Criterion (Anderson, 1962). A significant or trend-level-significant difference in distribution between the two conditions would be taken as indicative of a shift or reconfiguration of the underlying brain activations. For visualization purposes only, probability density estimates were obtained and plotted using the *ksdensity* function in Matlab. Fig. 6 will give a comprehensive overview over the distributions of all Z-scores along the geometric axes of the MNI space.

Results

Functional MRI Experiment 1

Strong and extensive bilateral activation of superior temporal cortex areas by speech (compared to silence) was observed in all participants, and all 16 scanned participants were included in the group statistics.

The principal contrast to be tested in fMRI Experiment 1 was the effect of syntactic complexity, that is, which brain region would show an increase in activation as syntactic complexity in auditorily presented sentences increased.

As shown in Fig. 1 and Table 1, fronto-temporal areas of mostly the left hemisphere were activated when syntactic complexity increased. This involved the very anterior parts of the left STG (temporal pole) as well as the IFG, including the left PO. The right PO was also activated, albeit somewhat smaller in extent. A circumscribed cluster in very posterior aspects of the left STS was also found active; while the Z-score was very strong, the cluster size verged on significance and survived the thorough 41-voxel threshold only in a slightly more lenient contrast (i.e., a direct comparison of most scrambled against canonical argument order).

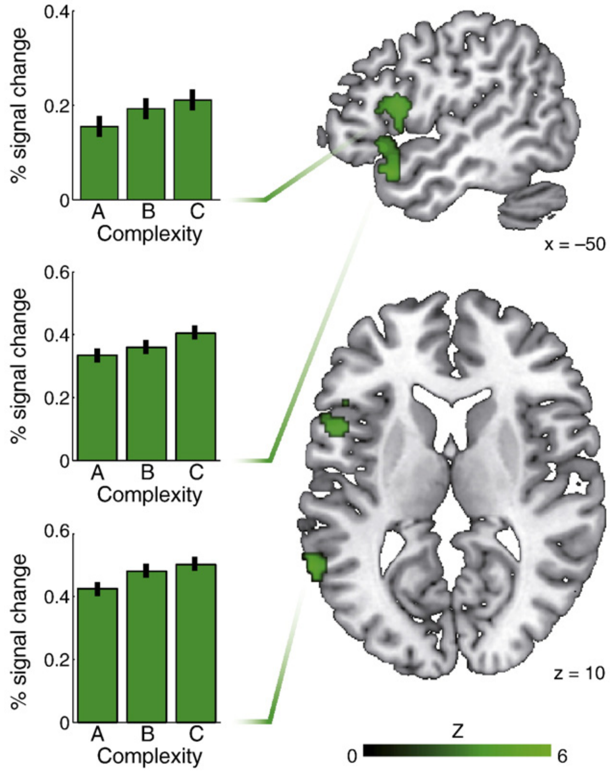

Fig. 1. Effect of increasing syntactic complexity in Experiment 1. All activations are thresholded at $p<0.001$ and a minimum cluster size of 41 suprathreshold voxels to control for family-wise error at $p<0.05$. Bar graph panels show % signal change from the parametric syntactic complexity contrast for the left IFG (henceforth referred to as *pars opercularis*, *PO*) in panel 1 and two regions of the superior temporal cortex (BA 13 in panel 2 and BA 22 in panel 3). Note that error bars have been corrected to reflect within-subject error (Jarmasz and Hollands, 2009).

Bar graphs on extracted percentage of signal change in the three left-hemispheric clusters illustrate the main effect of *syntactic complexity* in the PO region, in the very anterior left STG and in the left posterior STS (Fig. 1).

Behavioral experiment

Next, a sentence–sentence-matching paradigm was tested to identify adequate degradation levels to be used in the parametric complexity × degradation fMRI experiment. Note that this behavioral test was *not* designed to measure speech intelligibility of the vocoded stimuli per se, but to quantify the *lower bound* of spectral information needed to perform a syntactic analysis of a sentence. It was expected that this should be possible at intermediate degradation levels, such as 8-band speech, where enough information is available to distinguish pivotal consonants such as /n/ and /r/ (as in “den” and “der”, marking German case), although not all parts of the sentence might be understood equally well.

Table 1
Overview of significant clusters in Experiment 1, random-effects contrasts, thresholded at $p<0.001$; and cluster extent>41 voxels (~369 µl; equaling whole-brain $p<0.05$). Specifications refer to peak voxels.

Site	MNI coordinate			Cluster size (µl)	Z
	x	y	z		
Syntactic complexity effect[a]					
Left STG/BA 38	−50	16	−20	4320	4.62
Left IFG/BA 44	−64	−54	10		4.12
Left STS/BA 22	−52	12	14	666	5.26
Right insula/BA 13	44	12	18	918	4.57
Right IFG	52	16	2	468	3.48

[a] C>B>A.

The correctness of responses in the sentence–sentence matching task is shown in Fig. 2A. It revealed a main effect of both syntactic complexity (Wilks' Lamda approximated $F(2,12) = 10.53$, $p < 0.001$) and number of bands ($F(5,9) = 7.64$, $p < 0.006$), with no interaction present ($F(10,4) = 1.80$, n.s.). For syntactic complexity, condition B (one argument order scrambling) was found to significantly differ from condition C (two argument order scramblings; $t(13) = -4.36$, $p < 0.001$), no other contrasts reaching significance. For number of bands, 2-band speech was found to significantly differ from 4, 8, 16, and 32 bands and the non-vocoded version ($t = -2.68$; $t = -6.01$; $t = -6.14$; $t = -6.89$; $t = 6.48$, all $p < 0.05$, $df = 13$), respectively, while 4-band speech differed significantly from 2, 8, 16 and 32 bands ($t = -2.68$; $t = -3.64$; $t = -3.17$; $t = -3.47$; all $p < 0.05$, $df = 13$). Importantly, 8-, 16- and 32-band speech did not differ from each other.

We concluded from the pattern of results (Fig. 2A) that the set of 8-, 16- and 32-band speech conditions would be suited best for a parametric fMRI study in a 3 × 3 design with the factors syntactic complexity (degree of scrambling) and acoustic degradation (number of bands): sentences in these three degradation conditions varied in spectral detail and were clearly perceptually different (see spectrograms in Fig. 2B). Crucially, however, they did not yield significant differences in correctly performing the sentence matching task and thus allowing the listener to disambiguate the sentence structure.

Functional MRI Experiment 2

In Experiment 2, we re-applied the syntactic complexity modulation in acoustically presented sentences that effectively drove the activation in the left PO, left temporal pole and left posterior STS in Experiment 1.

Here we additionally applied three levels of acoustic degradation to study the effects of acoustic and syntactic complexity, as well as the effects of a complexity/degradation "trade-off" (Which brain areas are relatively more driven by complexity, as the signal improves in quality?) and a complexity/degradation "additivity" (Which brain areas are relatively more driven by complexity, as the signal

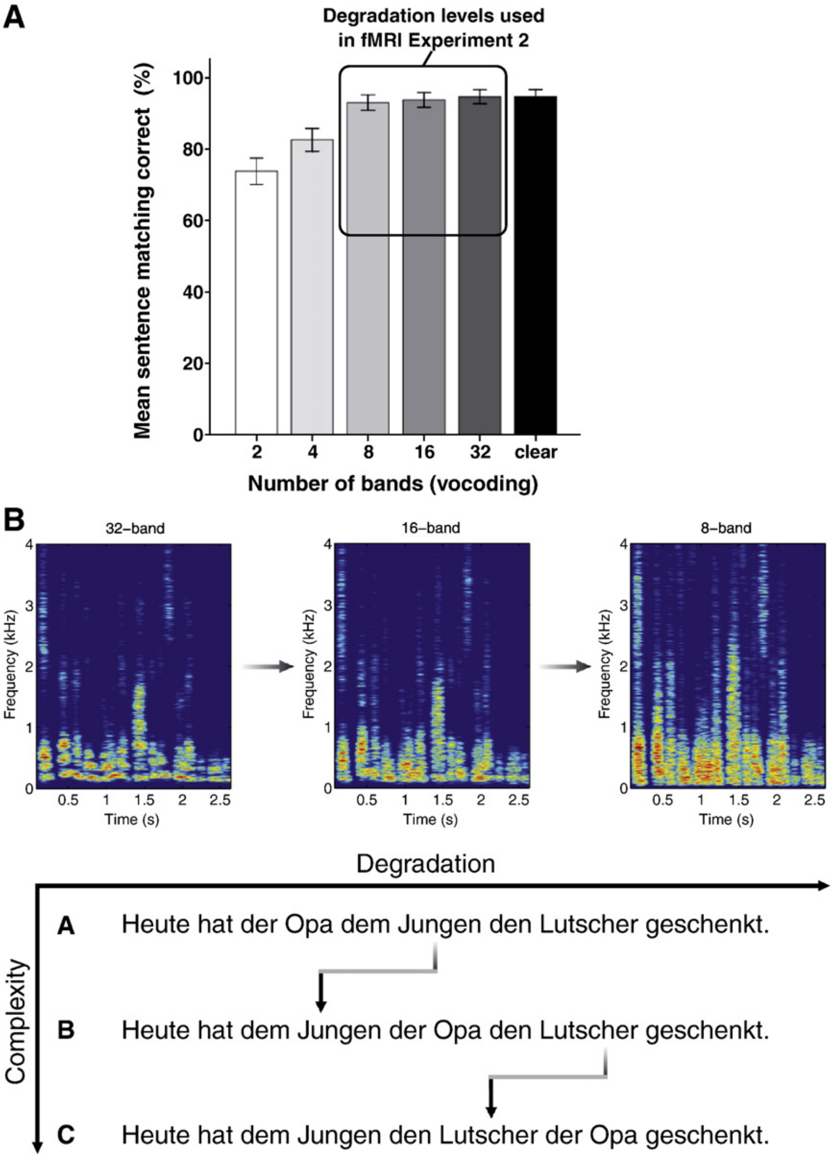

Fig. 2. A. Results of the behavioral pilot experiment to Experiment 2. The mean percentage of correct sentence matching is shown for 2- to 32-band vocoded and clear speech. From 8-band on, scores do not differ significantly from each other (clear speech; rightmost bar). Thus, we chose vocoding with 8, 16, and 32 bands (lower panels) for fMRI Experiment 2. B. Example stimuli for Experiment 2 (increasing in syntactic complexity from top to bottom). All sentences literally translate as "today the grandfather has given the lollipop to the boy." For Experiment 2, acoustically degraded versions were additionally introduced, as exemplified by the spectrograms with spectral detail declining from left (32-band speech, almost clear speech) to right (8-band speech, moderately degraded speech).

additionally degrades? See Fig. 4 for an illustration of the contrast formulation for these two patterns).

As expected, linear increases of activation from 8- to 16- to 32-band speech (improving signal quality) that mirrored the increasing intelligibility showed a monotonic increase of activation in bilateral superior temporal cortex (Fig. 3, Table 2; cf. Obleser and Kotz, 2010; Scott et al., 2006).

The first main finding with respect to syntactic complexity and acoustic degradation was the overall confirmation of the inferior frontal activation for increasing syntactic complexity under degradation: The left PO was activated, as expected from Experiment 1 (Fig. 3; note that a few of those left PO voxels showed a main effect of intelligibility as well). However, the exact peak locations that were observed for a parametric increase of syntactic complexity underwent interesting changes as acoustic degradation was taken into account.

We were interested to see which voxels in the IFG would be activated strongest when syntactic complexity adds up with gradual signal *degradation* (Fig. 4; red arrow in Figs. 5 and 6). Here, a notable change of peak activation to a substantially more medial and superior location was found (Table 1). As Figs. 5 and 6 illustrate, the activation peak (shown in blue) appears now to be shifted into the inferior frontal sulcus.

On the contrary, when syntactic complexity traded off with gradual signal *improvement*, the PO/IFG activation peak observed in Experiment 1 shifted only very marginally, with peak location changes in the 2-mm range (Table 2). This is a strong across-participants confirmation of the syntactic complexity effect for acoustically presented sentences in the IFG, as identified before in Experiment 1. Additionally, it shows that the presence or absence of rare active task trials did not affect this sub-process.

Along the STS, shifts or displacements of peak activation were also evident (Figs. 5 and 6). The syntactic complexity effects did again elicit a peak in pSTS when testing for the trade-off of complexity and degradation; just as had been observed under unhampered acoustic conditions in Experiment 1 (cf. the green peaks along the STS in Fig. 1; and the blue peaks in Fig. 5).

However, when the additivity of syntactic complexity and signal degradation was tested, the posterior STS peaks appeared shifted towards the mid section of the STS and the peak was also more likely to involve STG (i.e., BA 22; shown in detail in Figs. 5 and 6).

In Fig. 6, the distribution of maximum Z-scores per coordinate along the posterior–anterior axis (for the two STS subareas) and the inferior–superior and lateral–medial axes (for the IFG) illustrates this in greater detail. The observed peaks in the distributions of Z-scores and their relative displacement, depending on the two directions of interaction of complexity and signal quality (additivity coded in red and trade-off coded in blue), do further reflect the relative shifts of activation within these areas. Two-sample Cramér–Von Mises tests on the distribution of Z-scores proved significant for the posterior-to-anterior shift in the posterior STS ($p = 0.022$) and the inferior-to-superior shift in the IFG ($p = 0.015$), and bordered on significance for the lateral-to-medial shift in the IFG ($p < 0.07$).

Though clearly discernible as a peak in the distribution of Z-scores, the very anterior STS peak was not as prominent as in Experiment 1; a few voxels surpassed the $Z = 3.09$ — threshold in the additivity contrast, however (red in Fig. 6; bottom left panel). These were less anterior, more in mid STS/STG, than the activation in the trade-off contrast (blue).

Thus, more signal-bound processing regions in the mid sections of the STG/STS seem to become relatively more recruited as the quality of the signal (from which complex syntactic information has to be decoded) drops.

Discussion

The goal of this study was to specify the neural relationship of complex grammar analysis and auditory signal analysis. This was tested

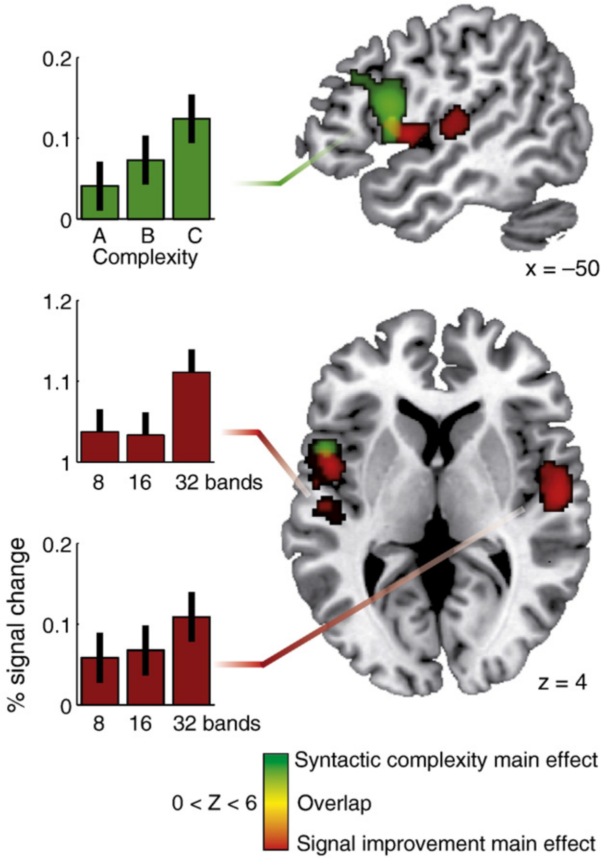

Fig. 3. Activation overlays from Experiment 2 for the main effects of syntactic complexity (shown in green) and signal improvement (shown in red). All activations are thresholded at $p<0.001$ and a minimum cluster size of 41 supra-threshold voxels to control for family-wise error at $p<0.05$. A small overlap of the main effects is seen in the left inferior frontal cortex (shown in yellow). Note that the posterior STS activation "cancels out" in the main effect of complexity, as Figs. 5 and 6 show how activation there is shifted as an interaction of complexity and degradation. Note that error bars have been corrected to reflect within-subject error (Jarmasz and Hollands, 2009).

Table 2
Overview of significant clusters in Experiment 2, random-effects contrasts, thresholded at $p<0.001$; and cluster extent>41 voxels (~369 µl), equaling whole-brain ($p<0.05$). Specifications refer to peak voxels.

Site	MNI coordinate			Cluster size (µl)	Z
	x	y	z		
Syntactic complexity effect[a]					
Left IFG/BA 44	−48	10	18	7173	6.80
Signal improvement effect[b]					
Right STG/BA 22	56	−12	4	1800	4.92
Left STG/BA 22	−50	2	4	1836	4.52
Insula/BA 13	−44	−22	12	1458	4.19
Additivity: syntactic complexity and signal degradation[c]					
Left dorsal IFG	−42	10	22	2610	4.60
Left STG/upper bank of STS	−64	−38	12	603	3.85
Left mid-anterior STS	−64	−12	−2	112[d]	3.40
Trade-off: syntactic complexity and signal improvement[e]					
Left lateral IFG	−50	10	16	4968	6.23
Right mid-anterior STG	56	−12	4	558	3.99
Left posterior MTG/STS	−54	−50	10	405	3.78

[a] C>B>A.
[b] 32>16>8 band.
[c] (C>B>A)+(8>16>32 band), see also Fig. 4.
[d] Does not fulfill the cluster extent criterion, but see Fig. 6.
[e] (C>B>A)+(32>16>8 band), see also Fig. 4.

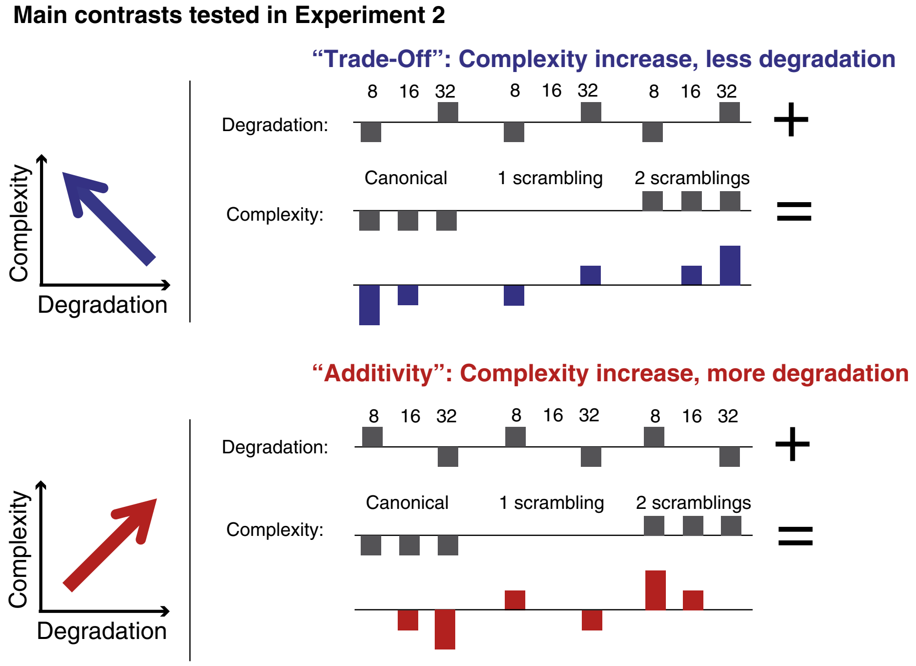

Fig. 4. Contrast design for additivity and trade-off. Schematic display of the linear contrasts formulated for testing "additivity" (i.e., increasing syntactic complexity adds up with signal degradation, leading to the strongest activation at 8-band/two-scramblings speech; shown in red) and "trade-off" (i.e., increasing syntactic complexity trades off with signal degradation, leading to the strongest activation at 32-band/two-scramblings speech; shown in blue).

by transferring a syntactic complexity manipulation to the auditory domain (Experiment 1) and by parametrically varying syntactic complexity and acoustic degradation (Experiment 2). Our results indicate a clear influence of bottom-up auditory processes on higher level syntax comprehension processes, reflected in a topographic shift of the complexity-related activation toward more primary sensory (i.e.

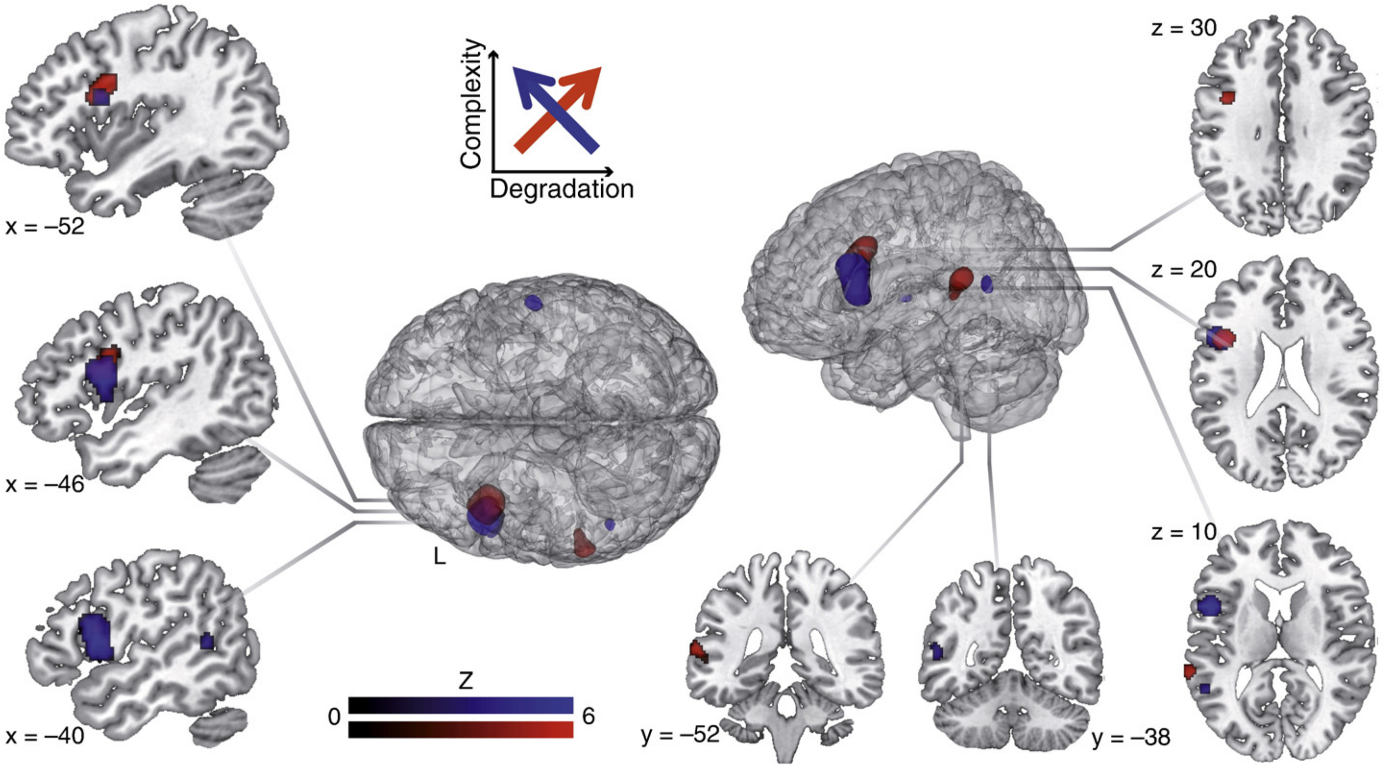

Fig. 5. Effects of increasing syntactic complexity under varying acoustic degradations in Experiment 2. The panels show activation overlays of the two major contrasts: increasing syntactic complexity trading off with signal improvement (blue); and increasing syntactic complexity additive to signal degradation (red). See Table 2 for exact contrast formulation. All activations are thresholded at $p < 0.001$ and a minimum cluster size of 41 supra-threshold voxels to control for family-wise error at $p < 0.05$. Linear contrasts that yielded these patterns are schematized with arrows in the coordinate system.

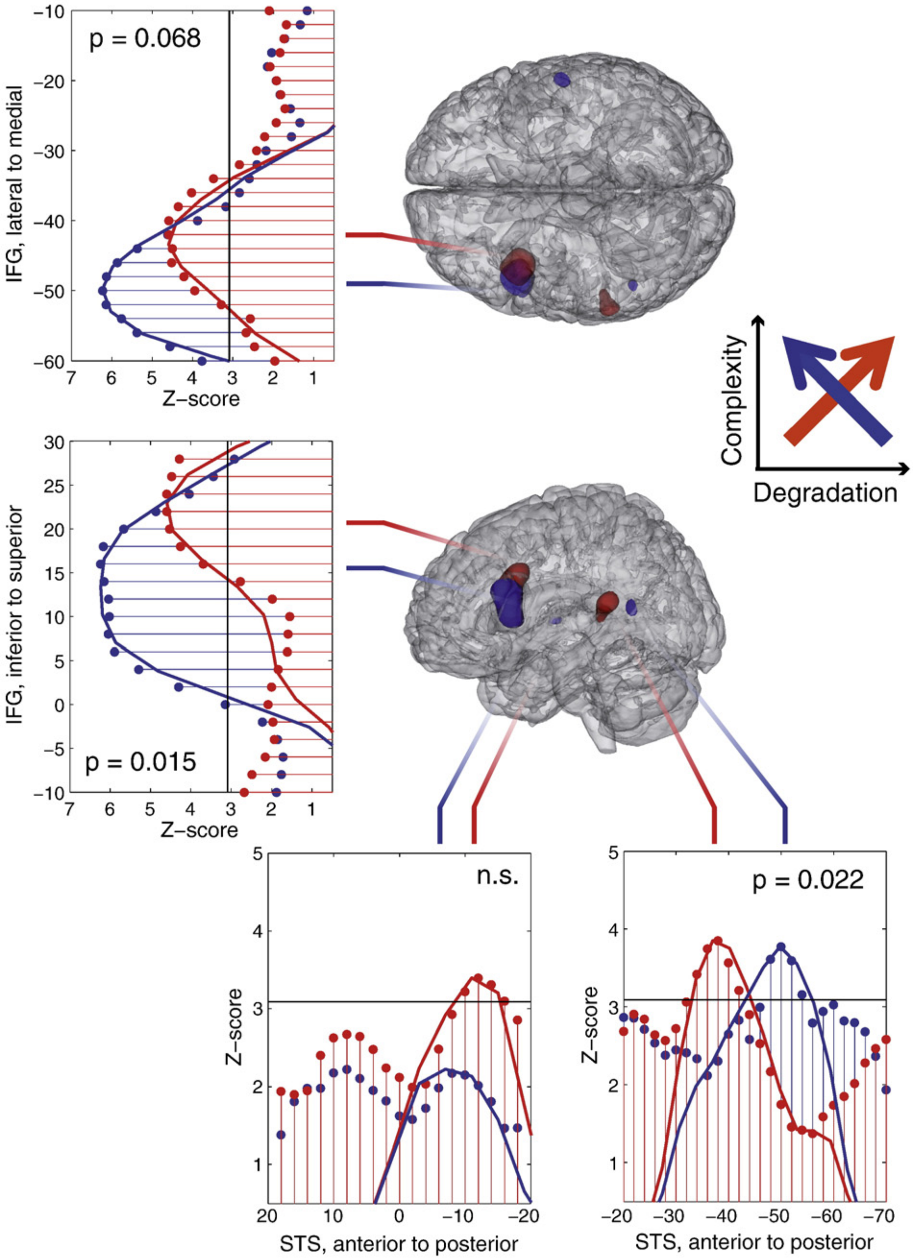

Fig. 6. Coordinate shifts for Experiment 2 ("upstream delegation"). The activation overlays and maximum Z-score per coordinate plot illustrate the syntactic complexity activation cluster in the IFG and STS sites shift from signal improvement toward signal degradation (i.e., blue to red). Parametric trade-off between syntactic complexity and signal improvement (blue) drives more inferior and lateral left IFG (left panels) and more anterior and posterior left STS sites (bottom panels). P-values indicate the significance of two-sample Cramér–Von Mises tests run on the indicated spatial distributions. Parametric additivity of syntactic complexity and signal degradation activates a more medio-dorsal prefrontal site (left panels) and more middle STS sites (bottom panels). Linear contrasts that yielded these patterns are schematized with arrows in the coordinate system.

"upstream") processing regions in the temporal cortex, as well as a shift in the inferior frontal cortex to regions that are likely to support basic processes of sequencing and working memory.

In more detail, the first fMRI experiment without any acoustic manipulation established that comprehending increasingly complex sentences from auditory input involves a left-lateralized array of posterior superior temporal, very anterior superior temporal and inferior frontal cortex (Experiment 1). This corroborates previous studies on such parametric complexity variation in visually presented sentences (Friederici et al., 2006b), with an additional focus on the anterior temporal cortex (anterior of Heschl's; see also a very recent study, Brennan et al., in press). This result is generally in line with the broad functional neuroanatomy for sentence processing as suggested by recent meta-analyses (Price, 2010; Vigneau et al., 2006). It adds weight to these areas' specific relevance to syntax, as our design used parametric variation rather than cognitive subtraction to isolate syntactic processing.

Second, the behavioral experiment showed that listeners would be able to correctly grasp the pivotal parts of complex grammatical sentences despite varying levels of considerable acoustic degradation, and we chose the degradation levels in Experiment 2 accordingly. It is important to keep in mind, though, that the overall intelligibility of these sentences certainly varied; Obleser et al., using structurally much simpler sentences yet focusing on sentence meaning, report comprehension scores of about 55–95% for semantically unpredictable sentences for 8–32-band speech; for highly predictable sentences scores for 8–32 band speech were more homogenous and scores for 8-band sentences were even above 90% (Obleser et al., 2007). The present sets of sentences, while being complex in syntactic structure, were certainly more of high predictability (e.g., keywords could be "mechanic–motor–driver–repair", and "spectator–magician–trick–given away"). Thus, although general intelligibility of the sentences might have varied, the present sentence matching results allow to conclude that intelligibility was sufficient to allow parsing of the complex (and varying) sentence structures. The question arising, however, was *to what cost and by engaging which specific brain regions* listeners would resolve the sentences when they are increasingly difficult to understand.

To this end, the main experiment (Experiment 2) combined the parametric changes in syntactic complexity with parametric changes in signal quality. Results showed that when increasing syntactic complexity added up with signal degradation, the center of gravity in activation shifted, presumably toward more "upstream" processing regions: in the left posterior superior temporal cortex activation shifted from the more medio-posterior sulcus anteriorily to the more lateral-middle sulcus, and in the left inferior frontal cortex from more ventral and lateral parts of the PO (BA 44) to more dorsal and medial parts, involving the inferior frontal sulcus (IFS) and reaching into the precentral gyrus.

These data lead us to argue that all observed changes of activation between syntactic complexity *with signal improvement* (blue activations; cf. also Experiment 1) and syntactic complexity *with signal degradation* (red activations) can be interpreted within a framework of "upstream delegation" of the analysis of complex syntax: when syntactic complexity adds up with signal degradation, regions more proximate to primary sensory and premotor areas become relatively more engaged (Figs. 5 and 6).

The following sections will put these findings into perspective with previous studies on syntactic complexity and then further outline the idea of "upstream delegation" as a tentative principle of neural resource allocation in language comprehension.

Syntactic complexity and the fronto-temporal network

A left-lateralized fronto-temporal network consisting of the IFG (PO) and the posterior STG/STS has been established as supporting syntactic complexity (Bornkessel et al., 2005; Friederici et al., 2010; Just et al., 1996; Makuuchi et al., 2009; Peelle et al., 2010). In addition, the anterior sections of the left STG/STS (aSTG), also responding to syntactic complexity in the present study, have previously been reported increasingly active for syntactically structured compared to non-structured sequences in auditory (Friederici et al., 2000; Humphries et al., 2005; Rogalsky and Hickok, 2009) as well as visual sentence processing (Stowe et al., 1998). Leff et al. (2009) recently presented compelling brain lesion evidence for the role of the pSTG/STS in supporting verbal working memory. Thus, posterior STS involvement in complex syntax processing altogether is a very established structure–function link, with the posterior STS' specific role remaining somewhat unclear and rendering the shifts observed here (see below) the more relevant.

For the left-anterior temporal region, it has been argued that the aSTG functions support on-line local phrase structure building (Friederici et al., 2003), whereas others assume that the aSTG do not differentiate between combinatorial aspects in syntax and semantics when under respective attentional control (Rogalsky and Hickok, 2009). In the present Experiment 1, all activations (PO, pSTG, and aSTG) are taken to reflect "downstream" areas along the auditory pathways that are responsive to abstract, complex language structure.

Note that the anterior temporal activation in Experiment 2 fell just short of conventional significance levels. Too few voxels surpassed the threshold but a clear peak-like activation topography is discernible from the sub- and supra-threshold activations evident in the Z-score-plot in Fig. 6 (lower left panel). The Z-scores exhibit a bimodal shape and it is the more posterior or more "upstream" one of these two peaks that is most strongly activated. This is in line with our main conclusion.

"Upstream delegation" — a framework for resource allocation?

The clusters of activations commonly analyzed and interpreted in BOLD fMRI are statistical peaks of broadly distributed activations (Fig. 6). This is helpful to keep in mind when interpreting the observed shifts in peak activation as relative shifts within interwoven processing networks — some operating on more abstract and complex codes, some more tuned toward detailed acoustic analysis.

Despite considerable differences in the details, most current models assume processing streams or gradients of hierarchical processing. Originally proposed as a powerful heuristic in the visual domain (Ungerleider et al., 1983), dual-stream-models of processing sensory input have become adopted into the auditory domain (Kaas and Hackett, 2000; Rauschecker and Tian, 2000). Many studies in humans and primates have substantiated the idea that two partly segregated processing streams originate from primary auditory areas (in humans most likely located along medial Heschl's gyrus; Humphries et al., in press; Wessinger et al., 2001) and run anterior and lateral (e.g., Binder et al., 2004; Scott et al., 2000) as well as posterior (e.g., Warren et al., 2002, 2005; for a study showing functional as well as anatomical evidence for anterior and posterior connections see Upadhyay et al., 2008). The processing streams target neighboring yet distinct prefrontal cortex areas (Romanski et al., 1999). A variety of modified stream models have been put forward (e.g., Belin and Zatorre, 2000; Hickok and Poeppel, 2007; Scott and Johnsrude, 2003). However, all of these models are consistent in assuming increasingly abstract levels of processing as the information propagates away from primary auditory areas. Within this framework, it is of note that the more middle STG/STS region that we observe to be more activated when the acoustic detail is pivotal (i.e., under degrading signal conditions, red activations in Figs. 5 and 6) matches well with cortical locations indicated in phonemic perception (e.g., Dehaene-Lambertz et al., 2005; Liebenthal et al., 2010; for a review see Obleser and Eisner, 2009) — as pointed out above, the phonetic information of the case-marking German consonants in our stimuli were particularly critical here.

The corresponding shift observed within the inferior frontal cortex toward the precentral gyrus and inferior frontal sulcus, detailed in the distributions of Z-scores in Fig. 6, might be taken to indicate an increased recruitment of working memory resources, as the left IFS has been characterized before as supporting memory-related processing during sentence comprehension (Makuuchi et al., 2009). Closely related to the current observations, a recent study by Peelle and colleagues varied syntactic complexity and speech signal quality (by means of temporal compression) in older adults and found, on the one hand a "classic" left PO activation for increasing syntactic complexity, but on the other hand also a peak clearly dorsal to that (in the precentral gyrus). This dorsal cluster's activation correlated positively with older adults' accuracy in comprehending the complex degraded signal (Peelle et al., 2010). Both our and Peelle et al.'s findings are compatible with the suggestion that premotor cortex/precentral gyrus involvement reflects basic mechanisms of sequence processing gaining importance (Schubotz and von Cramon, 2003). More generally, gradients of functional specialization for increasing complexity or abstractness running posterior–anterior in the prefrontal cortex have also been suggested (e.g., Koechlin and Summerfield, 2007).

How do these previous findings relate to the shifts in peak activation reported here? We argue that the activation differences observed between syntactic complexity in ideal acoustics on the one hand and syntactic complexity accompanied by declining acoustic quality on the other hand shed light on the processing streams in the temporal and the frontal lobe. They speak to a more dynamic assignment of neural resources for complex cognitive functions (i.e., understanding a sentence) than one might have concluded from comparing visual and auditory experiments under ideal stimulus conditions alone.

By adding acoustic degradation to the equation, our data imply that a given processing stage (i.e., comprehending a certain level of syntactic complexity) is not hard-wired to a certain stage along the functional neuroanatomical pathways. Instead, certain requirements such as a greater demand for thorough acoustic analysis relatively emphasize more signal-bound, less abstract upstream processes. Figs. 5 and 6 illustrate this in detail. This upstream shift observed is likely to reflect the reverse direction of shifts reported for intelligibility per se. The current syntactic complexity manipulation is able to additionally show that language activation "downstream" in the intelligibility pathways (Okada et al., 2010; Rauschecker and Scott, 2009) is not an all-or-nothing result of intelligibility per se. It rather reflects a weighted mixture of successful phonological, semantic and syntactic inferences (for a review on inference and interpretation in intelligibility, see e.g. Davis and Johnsrude, 2007).

Comparable phenomena have been described in the domain of deploying attention to different aspects of an acoustic signal (e.g., voice content versus voice identity, Formisano et al., 2008; Obleser et al., 2004; von Kriegstein et al., 2003). In the present data, a more parsimonious explanation would be that the listening system automatically allocates more resources to upstream processes of acoustic analysis. Our behavioral data show that listeners were sufficiently able to parse the increasingly complex sentences under all degradation conditions used here (Fig. 2A), despite the compromised intelligibility. At the same time, it is also known that increasingly complex sentences are more taxing to the system (e.g. response times to judge grammaticality rise accordingly, Friederici et al., 2006b).

Thus, the additionally imposed acoustic degradation does not corrupt the sentence comprehension process as such. Rather, comprehension is achieved by engaging hierarchically lower processing resources (in more central and superior parts of the temporal cortex) relatively more. In other words, areas subserving the acoustic analysis are required to a greater extent. Given that the crucial information indicating the sentence's underlying syntactic structure, which is necessary to understand the sentence, is encoded in single phonemes (e.g. *der* versus *den*, *den* versus *dem*), additional efforts in the acoustic analysis are well-placed. The areas supporting this analysis are auditory areas, most likely to be counted as human homologues of the parabelt cortex (Hackett, 2008). These regions are not differentially activated when the acoustic domain is either bypassed entirely (as in the majority of studies on sentence processing using the visual domain; e.g., Bornkessel et al., 2005; Hagoort et al., 2004; Indefrey et al., 2001; Makuuchi et al., 2009) or when acoustic signal quality is no pivotal precondition in sentence comprehension (present Experiment 1).

In sum, our results speak to a dynamic "upstream delegation" in auditory sentence comprehension under acoustic degradation: more signal-bound, less abstract processing areas not only in the left temporal but also in the left frontal cortex become relatively more recruited as the quality of the signal (from which syntactic structure has to be decoded) drops.

These data also show that the addition of parametric acoustic manipulations to studies of sentence comprehension is critical to isolate syntax-specific from more unspecific, signal-dominated processing stages. Parametric variation of syntactic complexity alone could not have yielded this result (Friederici et al., 2006b), neither did previous parametric studies on acoustic degradation alone (e.g., Obleser et al., 2008) allow for such conclusions. The data provide evidence for a dynamic assignment of neural processing resources to sentence processing and encourage further studies on dynamic neural pathways in auditory language comprehension.

Acknowledgments

This research is supported by the Max Planck Society. The research reported here was also partly supported by The German Ministry of Education and Research (BMBF; Grant Nr. 01GW0773). The authors are grateful to Christian Fiebach for providing the set of written sentences that formed the basis for our auditory stimuli. Simone Wipper, Annett Wiedemann and Anke Kummer helped acquiring the MR data. Two anonymous reviewers were particularly helpful in improving this manuscript.

References

Anderson, T.W., 1962. On the distribution of the two-sample Cramér–von Mises criterion. Ann. Math. Stat. 33, 1148–1159.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851.

Belin, P., Zatorre, R.J., 2000. 'What', 'where' and 'how' in auditory cortex. Nat. Neurosci. 3, 965–966.

Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A., Ward, B.D., 2004. Neural correlates of sensory and decision processes in auditory object identification. Nat. Neurosci. 7, 295–301.

Bornkessel, I., Zysset, S., Friederici, A.D., von Cramon, D.Y., Schlesewsky, M., 2005. Who did what to whom? The neural basis of argument hierarchies during language comprehension. Neuroimage 26, 221–233.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D.J., Pylkkanen, L., in press. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. Epub ahead of print. doi:10.1016/j.bandl.2010.04.002.

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

Davis, M.H., Johnsrude, I.S., 2007. Hearing speech sounds: top-down influences on the interface between audition and speech perception. Hear. Res. 229, 132–147.

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., Dehaene, S., 2005. Neural correlates of switching from auditory to speech perception. Neuroimage 24, 21–33.

Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. Science 322, 970–973.

Friederici, A.D., 2009. Pathways to language: fiber tracts in the human brain. Trends Cogn. Sci. 13, 175–181.

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 Lang. 74, 289–300.

Friederici, A.D., Ruschemeyer, S.A., Hahne, A., Fiebach, C.J., 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb. Cortex 13, 170–177.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., Anwander, A., 2006a. The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc. Natl. Acad. Sci. U. S. A. 103, 2458–2463.

Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., von Cramon, D.Y., 2006b. Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb. Cortex 16, 1709–1717.

Friederici, A.D., Kotz, S.A., Scott, S.K., Obleser, J., 2010. Disentangling syntax and intelligibility in auditory language comprehension. Hum. Brain Mapp. 31, 448–457.

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D.Y., Schlesewsky, M., 2005. The emergence of the unmarked: a new perspective on the language-specific function of Broca's area. Hum. Brain Mapp. 26, 178–190.

Grodzinsky, Y., Santi, A., 2008. The battle for Broca's region. Trends Cogn. Sci. 12, 474–480.

Hackett, T.A., 2008. Anatomical organization of the auditory cortex. J. Am. Acad. Audiol. 19, 774–779.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304, 438–441.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W., 1999. "Sparse" temporal sampling in auditory fMRI. Hum. Brain Mapp. 7, 213–223.

Harris, K.C., Dubno, J.R., Keren, N.I., Ahlstrom, J.B., Eckert, M.A., 2009. Speech recognition in younger and older adults: a dependency on low-level auditory cortex. J. Neurosci. 29, 6078–6087.

Hauser, M.D., Chomsky, N., Fitch, W.T., 2002. The faculty of language: what is it, who has it, and how did it evolve? Science 298, 1569–1579.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Humphries, C., Love, T., Swinney, D., Hickok, G., 2005. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Hum. Brain Mapp. 26, 128–138.

Humphries, C., Liebenthal, E., Binder, J.R., in press. Tonotopic organization of human auditory cortex. Neuroimage 50, 1202–1211.

Indefrey, P., Hagoort, P., Herzog, H., Seitz, R.J., Brown, C.M., 2001. Syntactic processing in left prefrontal cortex is independent of lexical meaning. Neuroimage 14, 546–555.

Jarmasz, J., Hollands, J.G., 2009. Confidence intervals in repeated-measures designs: the number of observations principle. Can. J. Exp. Psychol. 63, 124–138.

Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, K.R., 1996. Brain activation modulated by sentence comprehension. Science 274, 114–116.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. U. S. A. 97, 11793–11799.

Koechlin, E., Summerfield, C., 2007. An information theoretical approach to prefrontal executive function. Trends Cogn. Sci. 11, 229–235.

Leff, A.P., Schofield, T.M., Crinion, J.T., Seghier, M.L., Grogan, A., Green, D.W., Price, C.J., 2009. The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke. Brain 132, 3401–3410.

Liebenthal, E., Desai, R., Ellingson, M.M., Ramachandran, B., Desai, A., Binder, J.R., 2010. Specialization along the left superior temporal sulcus for auditory categorization. Cereb. Cortex 20, 2958–2970.

Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A.D., 2009. Segregating the core computational faculty of human language from working memory. Proc. Natl. Acad. Sci. U. S. A. 106, 8362–8367.

Narain, C., Scott, S.K., Wise, R.J., Rosen, S., Leff, A., Iversen, S.D., Matthews, P.M., 2003. Defining a left-lateralized response specific to intelligible speech using fMRI. Cereb. Cortex 13, 1362–1368.

Obleser, J., Eisner, F., 2009. Pre-lexical abstraction of speech in the auditory cortex. Trends Cogn. Sci. 13, 14–19.

Obleser, J., Kotz, S.A., 2010. Expectancy constraints in degraded speech modulate the speech comprehension network. Cereb. Cortex 20, 633–640.

Obleser, J., Elbert, T., Eulitz, C., 2004. Attentional influences on functional mapping of speech sounds in human auditory cortex. BMC Neurosci. 5, 24.

Obleser, J., Wise, R.J., Dresner, M.A., Scott, S.K., 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. J. Neurosci. 27, 2283–2289.

Obleser, J., Eisner, F., Kotz, S.A., 2008. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. J. Neurosci. 28, 8116–8123.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.H., Saberi, K., Serences, J.T., Hickok, G., 2010. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cereb. Cortex 20, 2486–2495.

Peelle, J.E., Troiani, V., Wingfield, A., Grossman, M., 2010. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. Cereb. Cortex 20, 773–782.

Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N. Y. Acad. Sci. 1191, 62–88.

Price, C.J., Thierry, G., Griffiths, T.D., 2005. Speech-specific auditory processing: where is it? Trends Cogn. Sci. 9, 271–276.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12, 718–724.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11800–11806.

Rodd, J.M., Davis, M.H., Johnsrude, I.S., 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb. Cortex 15, 1261–1269.

Rogalsky, C., Hickok, G., 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb. Cortex 19, 786–796.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat. Neurosci. 2, 1131–1136.

Rosen, S., Faulkner, A., Wilkinson, L., 1999. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. J. Acoust. Soc. Am. 106, 3629–3636.

Schubotz, R.I., von Cramon, D.Y., 2003. Functional–anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. Neuroimage 20 (Suppl 1), S120–S131.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci. 26, 100–107.

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.

Scott, S.K., Rosen, S., Lang, H., Wise, R.J., 2006. Neural correlates of intelligibility in speech investigated with noise vocoded speech—a positron emission tomography study. J. Acoust. Soc. Am. 120, 1075–1083.

Scott, S.K., McGettigan, C., Eisner, F., 2009. A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. Nat. Rev. Neurosci. 10, 295–302.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270, 303–304.

Slotnick, S.D., Moo, L.R., Segal, J.B., Hart Jr., J., 2003. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. Brain Res. Cogn. Brain Res. 17, 75–82.

Stowe, L.A., Broere, C.A., Paans, A.M., Wijers, A.A., Mulder, G., Vaalburg, W., Zwarts, F., 1998. Localizing components of a complex task: sentence processing and working memory. Neuroreport 9, 2995–2999.

Ungerleider, L.G., Mishkin, M., Macko, K.A., 1983. Object vision and spatial vision: two cortical pathways. Trends Neurosci. 6, 414–417.

Upadhyay, J., Silver, A., Knaus, T.A., Lindgren, K.A., Ducros, M., Kim, D.S., Tager-Flusberg, H., 2008. Effective and structural connectivity in the human auditory cortex. J. Neurosci. 28, 3341–3349.

Vigneau, M., Beaucousin, V., Herve, P.Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. Neuroimage 30, 1414–1432.

von Kriegstein, K., Eger, E., Kleinschmidt, A., Giraud, A.L., 2003. Modulation of neural responses to speech by directing attention to voices or verbal content. Brain Res. Cogn. Brain Res. 17, 48–55.

Warren, J.D., Zielinski, B.A., Green, G.G., Rauschecker, J.P., Griffiths, T.D., 2002. Perception of sound-source motion by the human brain. Neuron 34, 139–148.

Warren, J.E., Wise, R.J., Warren, J.D., 2005. Sounds do-able: auditory–motor transformations and the posterior temporal plane. Trends Neurosci. 28, 636–643.

Wessinger, C.M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., Rauschecker, J.P., 2001. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J. Cogn. Neurosci. 13, 1–7.

Zekveld, A.A., Heslenfeld, D.J., Festen, J.M., Schoonhoven, R., 2006. Top-down and bottom-up processes in speech comprehension. Neuroimage 32, 1826–1836.