

Expectancy Constraints in Degraded Speech Modulate the Language Comprehension Network

Jonas Obleser¹ and Sonja A. Kotz²

¹Department of Neuropsychology and ²Research Group “Neurocognition of Rhythm in Communication,” Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

In speech comprehension, the processing of auditory information and linguistic context are mutually dependent. This functional magnetic resonance imaging study examines how semantic expectancy (“cloze probability”) in variably intelligible sentences (“noise vocoding”) modulates the brain bases of comprehension. First, intelligibility-modulated activation along the superior temporal sulci (STS) was extended anteriorly and posteriorly in low-cloze sentences (e.g., “she weighs the flour”) but restricted to a mid-superior temporal gyrus/STS area in more predictable high-cloze sentences (e.g., “she sifts the flour”). Second, the degree of left inferior frontal gyrus (IFG) (Brodmann’s area 44) involvement in processing low-cloze constructions was proportional to increasing intelligibility. Left inferior parietal cortex (IPC; angular gyrus) activation accompanied successful speech comprehension that derived either from increased signal quality or from semantic facilitation. The results show that successful decoding of speech in auditory cortex areas regulates language-specific computation (left IFG and IPC). In return, semantic expectancy can constrain these speech-decoding processes, with fewer neural resources being allocated to highly predictable sentences. These findings offer an important contribution toward the understanding of the functional neuroanatomy in speech comprehension.

Keywords: angular gyrus, context, degraded speech, fMRI, inferior frontal gyrus, inferior parietal cortex, semantics, speech, superior temporal sulcus

Introduction

It is a common phenomenon in communication not to understand a sentence when you hear it for the first time, and you will, for example, find yourself asking a stranger to repeat his or her utterance. To a great extent, this is due to the lack of any prior common semantic information in such situations. Incidents like this also demonstrate how important the influence of semantic context is in analyzing and comprehending the speech signal.

Recent psycholinguistic (van Linden and Vroomen 2007; Shinn-Cunningham and Wang 2008) and neuroscientific (Sivonen et al. 2006; Davis and Johnsruide 2007; Davis et al. 2007; Hannemann et al. 2007; Obleser et al. 2007; van Atteveldt et al. 2007) approaches to speech are increasingly devoted to the factors in speech communication that transcend the signal analysis mechanisms of the auditory pathways’ (bottom-up) processing. Especially under acoustically compromised conditions, a whole thread of influences facilitates and disambiguates speech comprehension, among them speaker familiarity, prior experience and pragmatic knowledge, prosodic and emotional cues, syntactic expectancies, and constraints. A major source of disambiguation, however, and the focus of the

current study are semantic context and the expectancies built up by it (e.g., Obleser et al. 2007).

With the current experiment, we aim to close a gap in understanding the mechanisms that may explain how semantic context can facilitate speech comprehension. A preceding study by Obleser et al. (2007) identified a heteromodal array of left-hemispheric brain structures (angular gyrus/Brodmann’s area [BA] 39, ventrolateral inferior frontal gyrus [IFG]/BA 47, dorsolateral and medial prefrontal/BA 8,9, and posterior cingulate cortices/BA 30), which became active only when the semantic predictability of a sentence allowed speech comprehension despite compromised speech signal quality. Interestingly, no contextual facilitation was seen in antero-lateral superior temporal gyrus and sulcus (STG/STS), where only the very robust intelligibility effect (Binder et al. 2000; Scott et al. 2000; Davis and Johnsruide 2003) was replicated.

The sentence material used in that study (Speech in Noise [SPIN] sentences corpus, Kalikow et al. 1977) manipulated the predictability of the sentence-final keyword very effectively, although at the cost of subtle semantic confounds such as abstractness, use of sentence-initial names versus objects, reflexive pronouns, possessive adjectives, etc. This complicates the interpretation of the functions the brain regions identified may have in speech comprehension, as word class effects and a variety of semantic features may have confounded the activation pattern. We were therefore interested in further narrowing down possible top-down influences, that is, to restrict the sentence context to a stable minimum. We employed the well-studied semantic phenomenon of cloze probability (here shortened to “cloze”). Originally developed as a measure of text readability (Taylor 1953), it is a solid measure of lexical expectancy of a given word in a context. Variations of it have been adopted in studies of semantic processing, mostly in N400 electroencephalography designs (Kutas and Hillyard 1984; Connolly et al. 1995; Gunter et al. 2000; Federmeier et al. 2007). A stronger N400 in response to a sentence-final keyword is elicited by a low-cloze sentence such as “she weighs the flour” compared with a (otherwise well matched) high-cloze version such as “she sifts the flour,” interpreted as reflecting a higher cognitive effort to integrate the low-cloze (higher lexical competition) elements of the sentence in order to obtain coherent meaning.

The current experiment exploits this simple, yet effective, manipulation in a factorial design: First, simple sentences vary in the strength of the semantic expectancy coupling within that sentence (cloze probability). Second, if we independently subject these sentences to multiple levels of speech degradation (noise vocoding, Shannon et al. 1995; for application in neuroimaging, see Scott et al. 2000, 2006; Davis and Johnsruide 2003; Warren et al. 2006; Obleser et al. 2007), we can

disentangle separate influences as well as interactive processing mechanisms of “cloze probability” (a well-controlled cognitive parameter) and spectral degradation of speech (an equally well-controlled acoustic parameter) on measures of brain activity. Specifically, we expect to examine the relationship between regions previously linked to speech intelligibility, in particular the anterior and posterior STS, and regions repeatedly found in studies of lexico-semantic and top-down-mediated language processing, notably the IFG (BA 44/45; Kan and Thompson-Schill 2004; Hoen et al. 2006; Zempleni et al. 2007) and the inferior parietal cortex (IPC) (BA 39/40; Kotz et al. 2002; Obleser et al. 2007; Raettig and Kotz 2008).

Materials and Methods

Participants

Sixteen participants (9 females; age range 22–32 years) took part in the functional magnetic resonance imaging (fMRI) experiment. All were native speakers of German, right handed, and had normal hearing and no history of neurological or language-related problems. They were also unfamiliar with noise-vocoded speech and had not taken part in any of the pilot experiments. They were paid €15 for their participation. Another 18 participants (9 females; age range 18–31 years) took part in the behavioral pilot experiment prior to the magnetic resonance (MR) experiment. They fulfilled the same criteria as the MR participant sample and had no experience with noise-vocoded speech. These participants received €7.

Stimulus Material

Stimuli were recordings of spoken German sentences, all consisting of pronoun (“Er or Sie”), verb, and object, in the present tense. Every sentence with the framing constituents pronoun and object was used in 2 versions, incorporating either a verb low in cloze probability (e.g., “Er sieht das Bier” [he sees the beer]) or high in cloze probability (e.g., “Er trinkt das Bier” [he drinks the beer]). The materials were previously developed, tested, and used in a reading experiment on cloze probability (Gunter et al. 2000). The final set consisted of 40 pairs of sentences in a low- and a high-cloze probability version (80 sentences in total) and yielded cloze probabilities of $15.3 \pm 4.9\%$ (mean \pm standard deviation; maximum 24) for the low category and $74.2 \pm 12.8\%$ (minimum 56) for the high category, respectively.

Sentences were recorded by 1 male and 1 female speaker (both trained speakers of German) and digitized at 44.1 kHz. Off-line editing included resampling to 22.05 kHz, cutting at zero crossings before and after each sentence, and root mean square normalization of amplitude. From each final audio file (160 recordings), various spectrally degraded versions were created using a Matlab™-based noise-band vocoding algorithm. Noise vocoding is an effective technique of manipulating the spectral detail while preserving the temporal envelope of the speech

signal and rendering it more or less intelligible in a graded and controlled fashion depending on the number of bands used, with more bands yielding a more intelligible speech signal (Shannon et al. 1995).

Experimental Procedures

Pilot Study

The behavioral pilot experiment consisted of 160 sentences, each trial randomly drawn from either high- or low-cloze probability and from 1 of the 5 noise-vocoding level (2, 4, 8, 16, or 32 bands; Fig. 1). Participants (different from the fMRI participants) were seated in front of an LCD screen with comfortable font size and keyboard layout. They wore headphones (Sennheiser HD-202) and were instructed to listen to the stimuli and type in the sentence that they had just heard. The first stimulus was always a sentence drawn from the least degraded condition. On-screen written feedback of the sentence content was supplied for the first 10 trials. Participants could pause at their own discretion. Scoring of responses as correct or incorrect was based on a match of the last word of each sentence. A mean percentage of correct score was calculated for each participant and condition and submitted to a 2×5 repeated measures analysis with factors cloze probability (low and high) and signal degradation (2, 4, 8, 16, and 32 bands).

Results showed the expected and previously reported interaction of signal quality and cloze probability of the sentence ($F_{4,68} = 4.8, P < 0.001$; no sphericity violation indicated), with a significant advantage in sentence comprehension through high-cloze probability at 4-band speech (66% vs. 79%; $t_{17} = -4.7, P < 0.001$). Notably, this is a qualitative replication of earlier results using semantic predictability in less-restricted sentence contexts (Obleser et al. 2007). There was a marked facilitation of speech comprehension through semantic influence at intermediate signal quality. However, the effect is strongest at more degraded (4- instead of 8 band) signals, which is most likely owing to the very simple and predictable sentence structure in the current study, whereas the facilitation is less pronounced (by a factor of 1.2 compared with a factor of 1.8 in the previously mentioned semantic predictability study of Obleser et al. 2007). Consequently, the fMRI study included 3 conditions: 4-band speech, a 1-band condition that was hardly intelligible, and an only moderately degraded (16 band) condition.

fMRI Study

Scanning was performed using a Siemens 3-T scanner with birdcage head coil. Participants were comfortably positioned in the bore and wore air-conduction headphones (Resonance Technology, Northridge, CA). After a brief (10 trial) familiarization period, the actual experiment was started. Participants were required to listen attentively to noise-vocoded sentences, which were pseudorandomly drawn from 1 of the 6 conditions (high- or low-cloze probability, see above, in 1-, 4-, or 16-band noise-vocoded format).

In the fMRI study, we used the 40 pairs of “base sentences” (i.e., pairs of sentences ending in the same keyword) and randomly reused 5 of these (so that repetition of sentences was not a systematic factor of influence), yielding 45 pairs of sentences. These were then used in

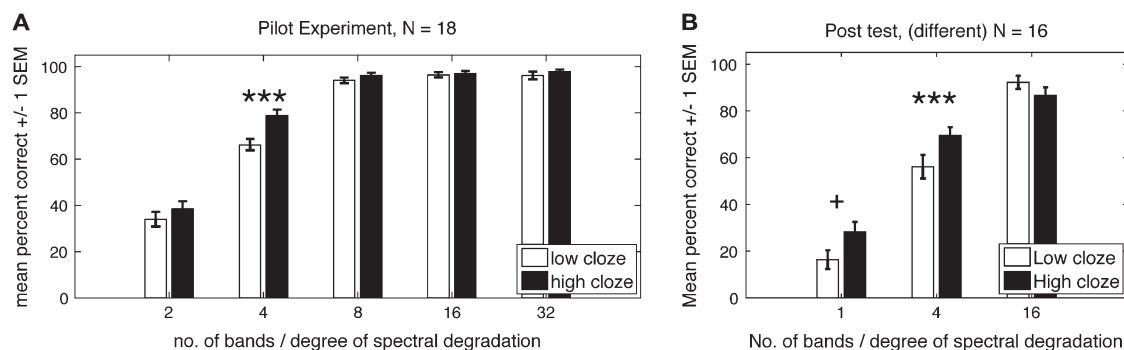


Figure 1. Behavioral data from pilot testing (different sample; left) and from the postscanning test (scanned participant sample; right). Note that the locus of the effect is clearly at 4-band noise-vocoded speech in both studies. Interaction of signal degradation (x axis) and context (cloze probability; white—low, black—high) is significant in both data sets. See text for details.

low- and high-cloze versions, spoken by both a male and a female speaker (=180 sentences) and were pseudorandomly assigned to a level of degradation (1-, 4-, or 16-band noise vocoding). Pseudorandomization was achieved with a Matlab script, such that no direct repetition of degradation levels or of base sentences occurred. Lastly, we inserted 30 trials of no stimulus/silence. Thus, 30 trials of each experimental condition plus 30 trials of no stimulus/silence amounted to a total of 210 trials and MR volume scans, lasting approximately 35 min.

Functional scans were acquired every 10 s, with a stimulus being presented 5.5 s before each scan (sparse temporal sampling; cf., Edmister et al. 1999; Hall et al. 1999). Echo-planar imaging scans were acquired in 22 axial slices covering the entire brain with an in-plane resolution of $3 \times 3 \text{ mm}^2$, a 4-mm slice thickness, and a 0.5-mm gap (repetition time = 10 s, acquisition time = 2000 s, time echo = 30 ms, flip angle 90° , field of view 192 mm, matrix size 64). For each participant, the individual high-resolution 3D T_1 -weighted MR scans acquired in a previous session were available for normalization, coregistration, and data visualization.

Behavioral Posttest

After the MR data acquisition, participants completed a comprehension test, which was similar to the pilot experiments described above but limited to 60 trials and to the conditions actually used in fMRI. Stimuli were also drawn from a subset of sentences not used during fMRI acquisition. They were naive to the task, as they had not taken part in the behavioral pilot study.

Data Analysis

Functional data were motion corrected off-line with the Siemens motion correction protocol (Siemens, Erlangen, Germany). Further analyses were performed using SPM5 (Wellcome imaging Department, University College, London, UK). fMRI time series were resampled to a cubic $2 \times 2 \times 2 \text{ mm}^3$ voxel size, corrected for field inhomogeneities (unwarped), normalized to a standard MR template (using a gray-matter segmentation-based procedure), and smoothed using an isotropic 8-mm^3 kernel.

In each participant, a general linear model using 6 regressors of interest (2 cloze probability levels \times 3 noise-vocoding levels) was estimated with a finite impulse response basis function (order 1 and window length 1). The no-stimulus (silent) trials in the time series formed an implicit baseline. Contrasts of all 6 conditions against the baseline from all 16 participants were then submitted to the SPM5 module for second-level within-subject analysis of variance, in which main effects, interactions, and specific contrasts could be assessed.

For statistical thresholding of the second-level (group) activation maps, we chose an uncorrected threshold of $P < 0.001$ combined with a minimal cluster extent of 87 voxels, yielding a whole-brain alpha of $P < 0.05$, as determined using a Matlab-implemented Monte Carlo Simulation (Slotnick et al. 2003).

Results

All participants showed extensive bilateral activation of the temporal cortices in response to sound and were all included in the ensuing group analyses.

As a first important finding, an increase in activation to more intelligible sentences was observed in the left and right STS (Table 1). Notably, however, both amplitude and topographical extent of this increase in responsiveness were vastly different for low- and high-cloze probability sentences. The anterior STS activation was almost exclusively driven by low-cloze sentences, extending along the entire STG/STS and reaching into the posterior middle temporal gyrus (MTG). In contrast, for high-cloze sentences, the activation appeared less extended and was restricted to middle STG and STS regions, not involving anterior and posterior STS.

As shown in Figure 2, the pattern can be best described as a modulation of the intelligibility-dependent increase: The peak voxel contrast estimates in the bar graphs show the blood oxygen level-dependent increase with increasing spectral

Table 1

Overview of significant clusters in random-effects contrasts ($P < 0.001$; cluster extent > 87 voxels/696 μL , equalling whole brain $P < 0.05$)

Site	MNI coordinate	Z
<i>Main effect of intelligibility (1 < 4 < 16 bands)</i>		
R anterior STS	62, -6, -4	>8
L anterior STS	-60, -8, -6	7.26
L angular gyrus (BA 39)	-46, -64, 38	4.11
<i>Main effect of cloze (low > high)</i>		
L IFG (BA 44)	-60, 12, 16	4.26
L posterior STG/STS	-50, -42, 2	3.82
R posterior STS	40, -32, 8	4.35
<i>Main effect of cloze (high > low)</i>		
Posterior cingulate (BA 23)	2, -36, 20	4.36
L lingual gyrus (BA 18)	-20, -80, -18	4.20
L cuneus (BA 18)	-16, -100, 4	3.80
<i>Intelligibility_{low cloze} > intelligibility effect_{high cloze}</i>		
L anterior STS	-52, -6, -14	4.50
R anterior STS	50, 8, -16	3.75
<i>Low > high cloze, depending on increasing intelligibility</i>		
L IFG (BA 44)	-60, 14, 14	4.88
L anterior STS	-52, -8, -16	4.47
R mid-STS	46, -18, -6	4.33
<i>Low cloze > high cloze, 1 > 4 > 16 bands</i>		
R posterior STG/SMG	64, -28, 18	4.65
L MFG (BA 46)	-30, 28, 36	4.15
L posterior STG	-46, -36, 18	4.09
R IFG (BA 44)	52, 6, 14	4.00

Note: Specifications refer to peak voxels. SMG, supramarginal gyrus; MFG, middle frontal gyrus; L, left; R, right.

detail in bilateral anterior STS to be pronounced for low-cloze sentences, whereas such an increase is almost absent for high-cloze probability sentences. In middle to posterior left STG, a strong intelligibility modulation in high- and low-cloze sentences is observed.

A whole-brain interaction test confirmed that the intelligibility effect in the anterior left STS was entirely driven by the low-cloze sentences (main peak in left anterior STS) (Montreal Neurological Institute [MNI] coordinates: -54, -8, -14; $z = 4.5$; Fig. 2, Table 1).

In response to low-cloze probability sentences, bilateral (middle to posterior) STS and left IFG (BA 44, pars opercularis) exhibited stronger activations, expressed statistically as a main effect of cloze probability (Fig. 3). In contrast, left occipital clusters (cuneus, BA 18) and the posterior cingulate were more responsive to high-cloze than to low-cloze sentences.

In the left IFG, a cloze probability differentiation gradually appeared as sentences became increasingly intelligible: the better a sentence's signal quality the more the low-cloze verb-object combinations drove the activation of the left IFG.

This interaction was also shown in a whole-volume interaction test ("intelligibility-dependent increase in low-cloze but a decrease in high-cloze sentences"): the left IFG (BA 44) was activated strongest in this test (MNI coordinates: -60, 14, 14, $z = 4.88$; Table 1). At 16-band signal quality, the gap between low cloze and high cloze was most evident (Fig. 3).

In the behavioral data obtained from the short postscanning test, the interaction between signal quality and cloze probability also attained significance ($F_{2,30} = 4.6$, $\epsilon = 0.79$, $P = 0.028$). In line with previous reports on noise-vocoded speech (Shannon et al. 1995, 2004), participants had obviously gained important perceptual experience during the 35-min in-scanner listening component: In the posttest, a trend ($P < 0.1$) toward a difference of high- versus low-cloze sentences surfaced even

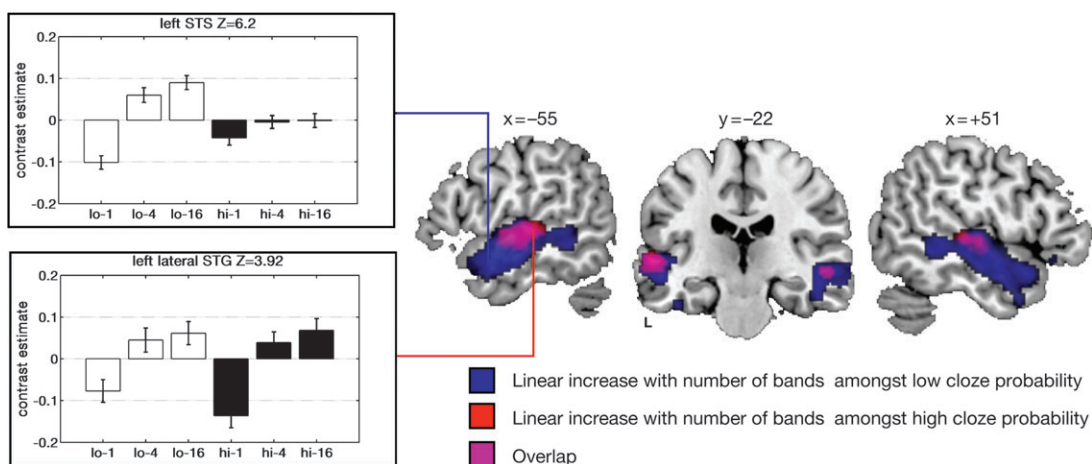


Figure 2. Effects of signal degradation on temporal lobe activation separately for low- and high-cloze probability sentences, thresholded at $P < 0.001$ (cluster extent $k > 87$ voxels; identical with Table 1), plotted onto left and right sagittal and coronal slices of a T_1 -weighted template brain image. The monotonic increase of activation is strongest in bilateral anterolateral STG/STS, extending also in posterior STS regions. It appears almost exclusively driven by low-cloze sentences (blue), whereas intelligibility modulation among high-cloze sentences (red, overlap in purple) yields activation confined to mid-STG and STS regions. This is reflected in the contrast estimate bar graphs from peak voxels in the left anterior STS (top left panel) and the mid- to posterior STG (left bottom panel; white bars—low cloze, black bars—high cloze).

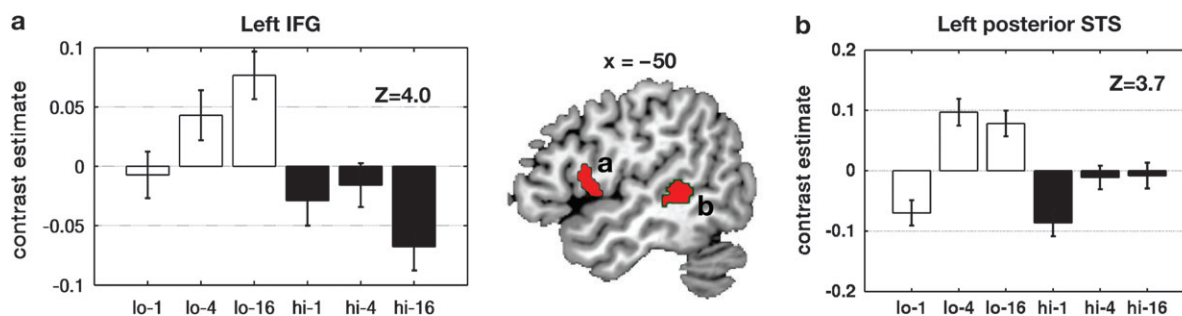


Figure 3. Regions exhibiting an effect of low- > high-cloze probability, thresholded at $P < 0.001$ (cluster extent $k > 87$ voxels; identical with Table 1), plotted onto left and right sagittal (top middle panels) and axial slices (bottom middle panels) of a T_1 -weighted template brain image. Left IFG (most likely located in BA 44; site a) and bilateral posterior STS (site b) exhibit enhanced activation under low-cloze conditions. Notably, however, patterns of activation across conditions differ, and left IFG (a, bottom left panel) shows the most language-specific pattern: With better signal quality, the expected signature of semantic computation (low > high cloze) becomes evident.

at 1-band speech. Also, the significant behavioral gain from high-cloze probability at 4-band speech was comparably weak for the short and comparably uniform sentences (ca., 13% difference in performance). In line with this, no brain area showed a significant increase when directly comparing 4-band high- and low-cloze sentences.

Based on the manipulation of semantic expectancy in Obleser et al. (2007), we assumed that activation in the left angular gyrus (BA 39) and adjacent IPC areas should be linked to conditions where an increase in speech comprehension takes place because of improved signal quality (see main effect of intelligibility, Table 1), especially so with facilitating semantic expectancy being present (i.e., in the high-cloze conditions): Participants were significantly better at comprehending 4-band high-cloze speech than 4-band low-cloze speech. The relative difference to the (ceiling level) 16-band speech comprehension was thus 2 times greater for low-cloze speech. We therefore formulated a statistical parametric mapping (SPM) contrast that directly tested for an increase in 4- to 16-band speech in low cloze (36%, or a factor of 1.65) compared with the somewhat smaller increase for 4- to 16-band speech in high cloze (18%, or a factor of 1.26; cf., Fig. 1). Notably, a single brain region was identified, this being the left BA 39 (Fig. 4).

Discussion

In the current study, we set out to further disentangle the interaction of speech signal processing in areas of the temporal cortex from semantic expectancy processing in areas of the inferior frontal and inferior parietal cortices.

Intelligibility and Expectancy

The first main finding is an important addition to the well-known intelligibility effect in the superior temporal cortex (STS/STG). The anterior and posterior parts of bilateral STS responded more strongly to more intelligible speech, which is in line with a whole series of imaging studies using different intelligibility modulations (Scott et al. 2000; see also, e.g., Davis and Johnsruide 2003; Liebenthal et al. 2005; Zekveld et al. 2006). However, this modulation was exclusively accounted for by the subset of low-cloze sentences. In contrast, when sentences were semantically highly predictable, the intelligibility modulation was restricted to middle regions of STG/STS. This suggests that the activation extent in the superior temporal cortex is restricted by high semantic-phonological expectancy constraints that become available over the course of a sentence.

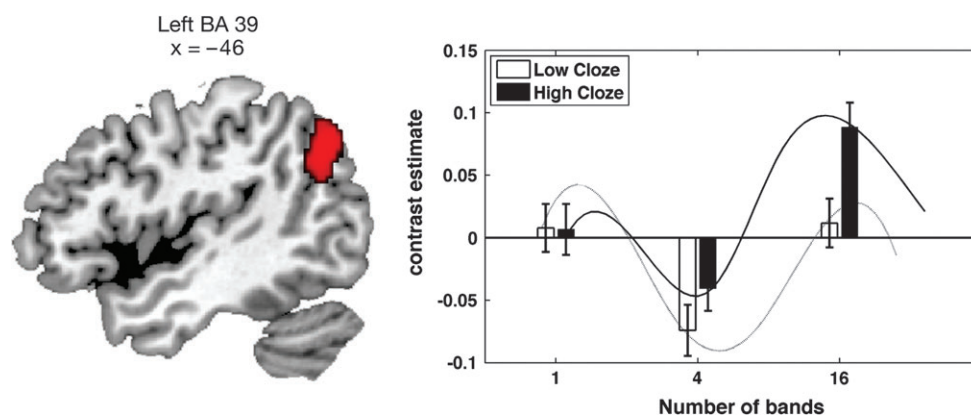


Figure 4. Activation pattern of the left IPC (angular gyrus, BA 39). Contrast estimates (right panel) for all conditions in left BA 39 (thresholded at $P < 0.001$ and a cluster extent $k > 87$ voxels; identical with Table 1) are shown. The spline interpolations (based on the current data as well as 2-, 8-, and 32-band speech as reported in Oleser et al. 2007) demonstrate the inverted *U*-shaped response behavior of the inferior parietal gyrus with changing speech degradation levels. This is more pronounced for sentences that allow for a high semantic expectancy (high cloze; black bars and solid line).

A second finding relates to the role of the IFG in speech intelligibility. The left IFG responded distinctly to low-cloze sentences, indicating an extra processing effort (analogous to the mechanisms discussed in the N400 literature; for review, see Kutas and Federmeier 2000), but this only occurred as sentences became increasingly intelligible.

From a psycholinguistic perspective, in low-cloze sentences, the left IFG appears to drive the semantic computation necessary to integrate the sentential components in order to derive meaning (Kan and Thompson-Schill 2004)—a function that can become obsolete when semantic expectancy is highly constrained. This hinges on the repeatedly implied function of the left IFG, namely, BA 44 and 45, in semantic and contextual integration (Baumgaertner et al. 2002; Rodd et al. 2005; Maess et al. 2006; Zempleni et al. 2007).

Interestingly, in comparison to previous studies, we find this result in varying degrees in response to the parametrically degraded speech signals. It thus appears that only if the speech signal is sufficiently intelligible can the left IFG actively participate in sentential integration (she weighs the flour). Thus, these processes of sentential integration depend on the successful outcome of signal decoding processes.

However, a very different picture emerges for high-cloze probability sentences (she sifts the flour). Here, sentential integration is effortless due to semantic expectancy. The concomitant brain activation is solely centered in bilateral mid-STG and STS, and neither the IFG nor the anterior STS or the posterior STS/MTG are involved (It is understood that qualifications about certain brain structures being “involved” or “activated,” in an fMRI study, are only valid within the bounds of a given significance threshold.). Figure 2 clearly demonstrates this for the condition of sufficient signal quality (16 band), where low- and high-cloze probability sentences are equally well comprehended (Fig. 1*b*), whereas the anterior STS shows a strong preference for low- over high-cloze probability ($z = 4$ at $-56, -6, -14$; Fig. 2).

If such a subtle expectancy manipulation is sufficient to downregulate the involvement of anterior and posterior STS as well as the IFG, their position in a “default” sentence comprehension network becomes relative. The question that arises is how the speech system exploits the semantic constraints in handling degraded speech and thus attains comprehension through only relatively restricted brain areas (bilateral mid-STG/STS).

A likely explanation is rooted in the fact that common “semantic” associations are naturally accompanied by common phonological pairings. Therefore, the brain can as well exploit and build up a strong “phonological” expectancy when confronted with high semantic expectancy (in high-cloze sentences). Given the compromised signal quality throughout the experiment, it is especially likely that the auditory comprehension system focuses on exploiting such phonological and segmental expectancies. As a consequence, the comprehension of high-cloze probability sentences might be already resolved at this level of information processing, and the conversion of acoustic input to a phonological code is likely to happen within the boundaries of mid-STG and STS (Jacquemot et al. 2003; Lieberthal et al. 2005; Oleser and Eisner 2009).

In sum, our data help to specify the interplay of the anterior and the posterior STS in speech comprehension. Our data show that neither one is inevitably strongly involved when sentences are heard. Rather, both regions are part of a network that processes and integrates sentence information, which is entirely in line with previous studies (e.g., Ferstl et al. 2002; Spitsyna et al. 2006). However, when high semantic and high phonological expectancies are formed, phonological predictions can become the driving factor in analyzing the incoming speech signal, thus reducing and constraining activation to mid-STG/STS.

With regard to hemispheric differences, it is of note that the left and right STS do not show quantitative or qualitative differences in their response patterns, which corroborates the overriding finding in the literature on bilateral STS activation in fMRI studies of speech perception (e.g., Binder et al. 2000; Davis and Johnsrude 2003; Oleser et al. 2007, 2008). In the IFG, the left-hemispheric predominance fits very well with a language-specific role of the left IFG in processing these degraded stimuli. Finally, the reported activation in the IPC is clearly left lateralized, as previously observed (Oleser et al. 2007; Raettig and Kotz 2008). This is suggestive of a comparably high-level and language-specific role the inferior parietal cortex fulfills in the comprehension of degraded speech, as outlined in the following section.

Role of the Inferior Parietal Cortex

There are several studies that—in different frameworks—have tied the interaction of phonological and semantic processes in

language perception and production to the inferior parietal lobe (IPC, subsuming supramarginal as well as angular gyrus; for discussion, see Dronkers et al. 2004; Mechelli et al. 2004; Golestani and Pallier 2007; Lee et al. 2007; Raettig and Kotz 2008). In the current data, the left IPC showed strongest activation differences between conditions showing the most pronounced difference in cloze-probability- and intelligibility-dependent comprehension performance (Figs 1 and 4). This supports the hypothesis that the left IPC is a key brain structure in facilitating sentence integration under compromised speech signal quality.

Interestingly, another auditory fMRI study comparing derivational pseudowords to opaque pseudowords (eluphant vs. thratofant; Raettig and Kotz 2008) yielded the same peak cluster in IPC. What these experiments have in common is that meaning is extracted successfully from a compromised or ambiguous signal. The left IPC appears instrumental in this process: It is a suitable and plausible brain structure to connect long-term memory and high-order concepts (“association cortex of association cortices,” N. Geschwind) to items that are either held in verbal short-term memory (supramarginal gyrus; Paulesu et al. 1993; Lee et al. 2007; for review, see Jacquemot and Scott 2006) or reflect acoustically ambiguous, difficult-to-process sensory input (posterior STG; Griffiths and Warren 2002; Obleser and Eisner 2009). Direct evidence for this also comes from a very recent functional connectivity study on magnetoencephalographic data, where inferior parietal activations were found to Granger-cause (Granger 1969) activation in left posterior STG in processing lexically ambiguous phonemes (Gow et al. 2008).

With respect to a more fine-grained functional distinction of BA 39 (supramarginal) and 40 (angular gyrus), we see an overlay of activations that reflect the largest comprehension increase in accompanying behavioral data in the current study as well as one previously published (Obleser et al. 2007). Both activations have their peak in the angular gyrus (BA 39) rather than supramarginal gyrus (BA 40; e.g., Lee et al. 2007). However, a functional distinction on the basis of group fMRI data remains difficult. Further studies utilizing transient disruption via transcranial magnetic stimulation or possibly data from patients with circumscribed lesions in either substructure of the IPC will be needed (along the lines of Dronkers et al. 2004) to cross-validate the differential contributions to speech comprehension made by the angular gyrus.

The activation pattern exhibited by the left BA 39 across all conditions (Fig. 4) concurs with our interpretation of the left IPC as a postsensory interface structure that taps long-term semantic knowledge/memory. When such access is not possible, however (as in very degraded speech), forming an abstract speech representation might fail (Obleser and Eisner 2009) and “weaker” IPC involvement is observed.

Differences to Previous Studies

The most salient difference between the previous brain imaging study by Obleser et al. (2007) combining noise-vocoded speech and variables of semantic context and the current one lies in the shift of the locus from 8-band speech to 4-band speech (Fig. 1*a*, which we took into account and adapted our in-scanner conditions accordingly, Fig. 1*b*). Given the behavioral data, this was a warranted step, and the shift of contextual facilitation to

a more degraded, generally less intelligible, signal level was most likely rooted in the very predictable minimal sentence context we used in order to focus on effects of cloze probability.

Our data also call for further efforts to disentangle the contributions of functional and anatomical subdivisions of the inferior frontal and inferior parietal cortices to speech comprehension. Both the left IFG, with BA 44 and 45, and the left IPC, with the supramarginal gyrus (BA 40) and the angular gyrus (39), often appear in the language-related literature, but most inferences as to their specific contributions are based on explicit language tasks involving verbal working memory (e.g., Jacquemot and Scott 2006; Buchsbaum and D’Esposito 2008) or explicit categorization and response selection tasks (e.g., Kan and Thompson-Schill 2004).

The most reliable data for functional distinctions of BA 44 versus BA 45 or BA 44 versus the frontal operculum in language-related functions stem from explicit tasks involving active decision making, speech production (Heim et al. 2005, 2007), or grammar learning (Friederici et al. 2006). There is no simple way to relate these findings to their possible role in treating degraded speech input; BA 44 activation, although clearly more dorsal (inferior–superior coordinate of $z = 30$), was previously reported in a study for effortful auditory search processes (that were carefully disentangled from speech comprehension itself, also using a form of noise analogues to speech; Giraud et al. 2004). However, no contextual or semantic factors were varied—factors which showed the most interesting effect in BA 44 in the current study. Probability atlas data render BA 44 the most likely locus of the found IFG cluster (with a 60–65% likelihood for left BA 44; Eickhoff et al. 2005).

Conclusions

Using well-controlled sentence material in a parametric set of acoustic degradation, this study furthers our understanding of the speech comprehension process by providing 3 important findings. First, the left IFG and the bilateral STS are not equivocally involved in processing sentences, but their level of activation appears to be adjusted to the semantic and phonological demands of the sentences processed, with fewer neural resources being allocated to highly predictable sentences. Second, a certain amount of signal intelligibility is required before such processing differences can surface, and the differences increase in left IFG with increasing intelligibility. Finally, the left IPC might serve as a mediating structure that facilitates speech comprehension when signal intelligibility is compromised; yet, expectancies formed over the course of a sentence facilitate comprehension, most likely through top-down activation of semantic concepts (angular gyrus) and by providing a short-term verbal memory store for lexical candidate items (supramarginal gyrus).

These findings demonstrate that speech intelligibility and semantic processing have a much more intimate relationship in neural terms than previously accounted for. They offer important contributions to our current understanding of the functional neuroanatomy of speech comprehension.

Funding

The German Science Foundation (DFG) to S.A.K.; postdoctoral elite grant of the Landesstiftung Baden-Württemberg gGmbH to J.O.

Notes

J.O. and S.A.K. are employed by the Max Planck Society Germany. Ulrike Barth, Mandy Naumann, Annett Wiedemann, Domenica Wilfling, and Simone Wipper were helpful in acquiring the behavioral and functional brain imaging data. Rosie Wallis proofread the manuscript. We also thank our anonymous reviewers who helped to substantially improve the manuscript. *Conflict of Interest:* None declared.

Address correspondence to Jonas Obleser, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, Leipzig 04103, Germany. Email: jonas@obleser.de.

References

- Baumgaertner A, Weiller C, Buchel C. 2002. Event-related fMRI reveals cortical sites involved in contextual sentence integration. *Neuroimage*. 16:736–745.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex*. 10:512–528.
- Buchsbaum BR, D'Esposito M. 2008. The search for the phonological store: from loop to convolution. *J Cogn Neurosci*. 20:762–778.
- Connolly JF, Phillips NA, Forbes KA. 1995. The effects of phonological and semantic features of sentence-ending words on visual event-related brain potentials. *Electroencephalogr Clin Neurophysiol*. 94: 276–287.
- Davis MH, Coleman MR, Absalom AR, Rodd JM, Johnsrude IS, Matta BF, Owen AM, Menon DK. 2007. Dissociating speech perception and comprehension at reduced levels of awareness. *Proc Natl Acad Sci USA*. 104:16032–16037.
- Davis MH, Johnsrude IS. 2003. Hierarchical processing in spoken language comprehension. *J Neurosci*. 23:3423–3431.
- Davis MH, Johnsrude IS. 2007. Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear Res*. 229:132–147.
- Dronkers NF, Wilkins DP, Van Valin RD, Jr, Redfern BB, Jaeger JJ. 2004. Lesion analysis of the brain areas involved in language comprehension. *Cognition*. 92:145–177.
- Edmister WB, Talavage TM, Ledden PJ, Weisskoff RM. 1999. Improved auditory cortex imaging using clustered volume acquisitions. *Hum Brain Mapp*. 7:89–97.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*. 25:1325–1335.
- Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. 2007. Multiple effects of sentential constraint on word processing. *Brain Res*. 1146:75–84.
- Ferstl EC, Guthke T, von Cramon DY. 2002. Text comprehension after brain injury: left prefrontal lesions affect inference processes. *Neuropsychology*. 16:292–308.
- Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A. 2006. The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc Natl Acad Sci USA*. 103:2458–2463.
- Giraud AL, Kell C, Thierfelder C, Sterzer P, Russ MO, Preibisch C, Kleinschmidt A. 2004. Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cereb Cortex*. 14:247–255.
- Golestani N, Pallier C. 2007. Anatomical correlates of foreign speech sound production. *Cereb Cortex*. 17:929–934.
- Gow DW, Jr, Segawa JA, Ahlfors SP, Lin FH. 2008. Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates. *Neuroimage*. 43:614–623.
- Granger CWJ. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 37:424–438.
- Griffiths TD, Warren JD. 2002. The planum temporale as a computational hub. *Trends Neurosci*. 25:348–353.
- Gunter TC, Friederici AD, Schriefers H. 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *J Cogn Neurosci*. 12:556–568.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. “Sparse” temporal sampling in auditory fMRI. *Hum Brain Mapp*. 7:213–223.
- Hannemann R, Obleser J, Eulitz C. 2007. Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Res*. 1153:134–143.
- Heim S, Alter K, Ischebeck AK, Amunts K, Eickhoff SB, Mohlberg H, Zilles K, von Cramon DY, Friederici AD. 2005. The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords. *Brain Res Cogn Brain Res*. 25:982–993.
- Heim S, Eickhoff SB, Ischebeck AK, Friederici AD, Stephan KE, Amunts K. 2007. Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM. *Hum Brain Mapp*. doi: 10.1002/hbm.20512.
- Hoen M, Pachot-Clouard M, Segebarth C, Dominey PF. 2006. When Broca experiences the Janus syndrome: an ER-fMRI study comparing sentence comprehension and cognitive sequence processing. *Cortex*. 42:605–623.
- Jacquemot C, Pallier C, LeBihan D, Dehaene S, Dupoux E. 2003. Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J Neurosci*. 23:9541–9546.
- Jacquemot C, Scott SK. 2006. What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci*. 10:480–486.
- Kalikow DN, Stevens KN, Elliott LL. 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am*. 61:1337–1351.
- Kan IP, Thompson-Schill SL. 2004. Selection from perceptual and conceptual representations. *Cogn Affect Behav Neurosci*. 4:466–482.
- Kotz SA, Cappa SF, von Cramon DY, Friederici AD. 2002. Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study. *Neuroimage*. 17:1761–1772.
- Kutas M, Federmeier KD. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn Sci*. 4:463–470.
- Kutas M, Hillyard SA. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature*. 307:161–163.
- Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Pitcher K, Crinion J, Mechelli A, Frackowiak RS, et al. 2007. Anatomical traces of vocabulary acquisition in the adolescent brain. *J Neurosci*. 27:1184–1189.
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. 2005. Neural substrates of phonemic perception. *Cereb Cortex*. 15:1621–1631.
- Maess B, Herrmann CS, Hahne A, Nakamura A, Friederici AD. 2006. Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Res*. 1096:163–172.
- Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, Price CJ. 2004. Neurolinguistics: structural plasticity in the bilingual brain. *Nature*. 431:757.
- Obleser J, Eisner F. 2009. Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci*. 13:14–19.
- Obleser J, Eisner F, Kotz SA. 2008. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J Neurosci*. 28:8116–8123.
- Obleser J, Wise RJ, Dresner MA, Scott SK. 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. *J Neurosci*. 27:2283–2289.
- Paulesu E, Frith CD, Frackowiak RS. 1993. The neural correlates of the verbal component of working memory. *Nature*. 362:342–345.
- Raettig T, Kotz SA. 2008. Auditory processing of different types of pseudowords: an event-related fMRI study. *Neuroimage*. 39:1420–1428.
- Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb Cortex*. 15:1261–1269.
- Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*. 123: 2400–2406.
- Scott SK, Rosen S, Lang H, Wise RJ. 2006. Neural correlates of intelligibility in speech investigated with noise vocoded speech—a positron emission tomography study. *J Acoust Soc Am*. 120:1075–1083.

- Shannon RV, Fu QJ, Galvin J, 3rd. 2004. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol.* 552(Suppl):50-54.
- Shannon RV, Zeng FG, Kamath V, Wyganski J, Ekelid M. 1995. Speech recognition with primarily temporal cues. *Science.* 270:303-304.
- Shinn-Cunningham BG, Wang D. 2008. Influences of auditory object formation on phonemic restoration. *J Acoust Soc Am.* 123:295-301.
- Sivonen P, Maess B, Lattner S, Friederici AD. 2006. Phonemic restoration in a sentence context: evidence from early and late ERP effects. *Brain Res.* 1121:177-189.
- Slotnick SD, Moo LR, Segal JB, Hart J, Jr. 2003. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Res Cogn Brain Res.* 17:75-82.
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJ. 2006. Converging language streams in the human temporal lobe. *J Neurosci.* 26:7328-7336.
- Taylor WL. 1953. Cloze procedure: a new tool for measuring readability. *Journal Q.* 30:414-433.
- van Atteveldt NM, Formisano E, Goebel R, Blomert L. 2007. Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage.* 36:1345-1360.
- van Linden S, Vroomen J. 2007. Recalibration of phonetic categories by lipread speech versus lexical information. *J Exp Psychol Hum Percept Perform.* 33:1483-1494.
- Warren JD, Scott SK, Price CJ, Griffiths TD. 2006. Human brain mechanisms for the early analysis of voices. *Neuroimage.* 31:1389-1397.
- Zekveld AA, Heslenfeld DJ, Festen JM, Schoonhoven R. 2006. Top-down and bottom-up processes in speech comprehension. *Neuroimage.* 32:1826-1836.
- Zempleni MZ, Renken R, Hoeks JC, Hoogduin JM, Stowe LA. 2007. Semantic ambiguity processing in sentence context: evidence from event-related fMRI. *Neuroimage.* 34:1270-1279.