Semantic versus perceptual interactions in neural processing of speech-in-noise

CrossMark

Narly Golestani [a,b,*], Alexis Hervais-Adelman [b], Jonas Obleser [c], Sophie K. Scott [a]

[a] Institute of Cognitive Neuroscience, University College London, 17 Queen Square, WC1N 3AR London, UK
[b] University Medical School, 1, Rue Michel-Servet, CH-1211 Geneva, Switzerland
[c] Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A, 04103 Leipzig, Germany

ARTICLE INFO

ABSTRACT

Native listeners make use of higher-level, context-driven semantic and linguistic information during the perception of speech-in-noise. In a recent behavioral study, using a new paradigm that isolated the semantic level of speech by using words, we showed that this native-language benefit is at least partly driven by semantic context (Golestani et al., 2009). Here, we used the same paradigm in a functional magnetic resonance imaging (fMRI) experiment to study the neural bases of speech intelligibility, as well as to study the neural bases of this semantic context effect in the native language. A forced-choice recognition task on the first of two auditorily presented semantically related or unrelated words was employed, where the first, 'target' word was embedded in different noise levels. Results showed that activation in components of the brain language network, including Broca's area and the left posterior superior temporal sulcus, as well as brain regions known to be functionally related to attention and task difficulty, was modulated by stimulus intelligibility. In line with several previous studies examining the role of linguistic context in the intelligibility of degraded speech at the sentence level, we found that activation in the angular gyrus of the left inferior parietal cortex was modulated by the presence of semantic context, and further, that this modulation depended on the intelligibility of the speech stimuli. Our findings help to further elucidate neural mechanisms underlying the interaction of context-driven and signal-driven factors during the perception of degraded speech, and this specifically at the semantic level.

© 2013 Elsevier Inc. All rights reserved.

Introduction

In studying the neural implementation of spoken language processing, it is important to consider the complexity of linguistic processes. For example, one can ask how higher-order, semantic versus lower-order, perceptual processes interact during the processing of noisy speech — a phenomenon that is ubiquitous in our daily lives, and how the brain supports the interaction of these complementary cognitive and perceptual dimensions (Mattys et al., 2009). It is known that in one's native language, speech comprehension is often successful even when hearing noisy, or degraded speech (Nabelek and Donahue, 1984; Takata and Nabelek, 1990; van Wijngaarden et al., 2002). Further, using the Speech Perception in Noise (SPIN) paradigm (Bilger et al., 1984; Kalikow et al., 1977), in which the predictability of the final word in sentences is manipulated, it has been shown that this native language advantage can arise from the use of higher-level linguistic, *contextual information* (Florentine, 1985a; Mayo et al., 1997). These original studies further showed that lower SNRs (or higher noise levels) are associated with a greater context benefit (Mayo et al., 1997). We are thus capable of making use of

linguistic context to compensate for poor signal quality. Linguistic context includes semantic and syntactic information, as well as pragmatic and prosodic information. Recently, using a new paradigm that isolates the semantic level of speech (i.e., at the word rather than at the sentence level), we showed that this native language intelligibility advantage for the perception of degraded speech is at least in part specifically driven by *semantic* context (Golestani et al., 2009). We used an auditory version of the retroactive word priming paradigm (Bernstein et al., 1989) in which we presented pairs of semantically related or unrelated words, the first of which was embedded in different levels of noise. We found that performance was better on related compared to unrelated trials during native language speech processing, and further, we showed that this benefit of context increases as SNR decreases. These results support the idea that semantic context, *specifically*, contributes to the native-language advantage for speech-in-noise processing.

Here, we used the same paradigm in a functional magnetic resonance imaging (fMRI) experiment to study the neural bases of a) speech intelligibility at the word level, and of b) semantic context effects during degraded native language speech processing. Previous functional imaging studies have examined the neural bases of speech intelligibility at the sentence level during native language processing by comparing activation arising during the processing of distorted to clear speech. In many of these studies speech is distorted by noise-vocoding (Shannon

* Corresponding author at: Brain and Language Lab, Department of Clinical Neuroscience, University of Geneva, 1 Rue Michel-Servet, CH-1211 Geneva, Switzerland.
E-mail address: Narly.Golestani@Unige.ch (N. Golestani).

et al., 1995), or by parametrically modulating speech intelligibility by manipulating the number of channels used to synthesize noise-vocoded speech (Davis and Johnsrude, 2003; Shannon et al., 1995). These studies have shown that the left inferior frontal gyrus (IFG), as well as regions along the left anterior to posterior temporal lobe are more activated when speech is degraded (challenging) but intelligible compared to when hearing unintelligible sounds matched for complexity (Davis and Johnsrude, 2003; Narain et al., 2003; Obleser and Kotz, 2010; Obleser et al., 2007a; Scott et al., 2000, 2006). Two studies explicitly also examined activations during processing of clear speech, and found the left IFG to be more strongly recruited if sentences were degraded but intelligible compared to when they were unintelligible but also compared to when they were clearly audible, suggesting that this region is recruited most strongly when successful but effortful speech processing is required (Davis and Johnsrude, 2003; Davis et al., 2011). Even in the absence of sentence-level context, it has been shown that the LIFG responds significantly more to potentially-intelligible noise-vocoded words than to clear words (Hervais-Adelman et al., 2012). The left IFG's known role in phonetic, semantic and syntactic processing and also in verbal working memory makes it a likely candidate in a compensatory process whereby such linguistic information (phonetic, semantic, and syntactic) as well as verbal working memory are used to assist comprehension of degraded speech.

Left temporal cortex activations in the above studies range from the anterior to posterior superior temporal gyrus (STG), superior temporal sulcus (STS), and middle temporal gyrus (MTG), with the posterior activations extending to the classic temporo-parietal Wernicke's area (Davis et al., 2011; Narain et al., 2003; Okada et al., 2010; Scott et al., 2000). A subset of the studies on speech intelligibility have also found the left anterior hippocampus to be sensitive to speech intelligibility, and this finding has been related to the role of this brain region in verbal memory and in processing meaning in verbal stimuli (Davis and Johnsrude, 2003; Davis et al., 2011). Together, this literature suggests that left frontal regions work in concert with left anterior to posterior temporal regions involved in higher-level binding of meaning in connected speech to assist the effortful comprehension of degraded speech (Davis and Johnsrude, 2003; Davis et al., 2011; Obleser et al., 2007a; Okada et al., 2010).

Several studies have been performed to explicitly assess the neural processes involved when utilizing 'top-down' linguistic information, or linguistic context to assist in comprehending degraded speech. One such study by Obleser et al. (2007a) used noise-vocoded sentences that were subjected to different levels of degradation (using 2, 8, or 32 filter bands), and that were either linguistically predictable or not (Obleser et al., 2007a). Consistent with the above studies on intelligibility, it was shown that regions including bilateral temporal lobes as well as the left IFG were overall more engaged when sentences from the above-referenced SPIN set (Kalikow et al., 1977) were predictable compared to when they were not, and when they were degraded but intelligible. Behaviorally, it was shown that sentence predictability (i.e. linguistic context) assisted speech comprehension most when sentences were degraded at an intermediate (at 8-band noise-vocoding) level. Based on this behavioral finding, neural activation at this level of degradation was assessed. Results revealed that regions including the left angular gyrus, the left anterior temporal lobe and the left ventral IFG (Brodmann's area 47) were involved only if speech comprehension was likely to succeed despite adverse listening conditions, and under circumstances that allowed the use of higher-order linguistic information. In a related study, short, minimal sentences were used in which the 'cloze' probability differed: the pairs of sentences all ended in the same keyword, but the preceding context was either predictable or not. The sentences were once again degraded using noise vocoding. In this study also, the left angular gyrus was shown to be involved when speech comprehension was successful, but this time either because of increased signal quality or because of the presence of linguistic context (Obleser and Kotz, 2010). The above

studies involved sentence-level stimuli, and thus do not allow us to specify the level of speech (semantic, syntactic, prosodic) at which angular gyrus involvement is important in using context to extract meaning from degraded speech. A recent M/EEG study of noise-vocoded speech (Sohoglu et al., 2012) showed that activity in the LIFG and left STS was enhanced when listeners received congruent prior information (a transcription of the forthcoming degraded word) compared to either incongruent or neutral (a letter string) prior information. This paradigm exploited identity priming rather than semantic priming, and hints that the angular gyrus is specifically involved when semantic information may be used to help degraded speech perception.

The purpose of the present study was to isolate semantic influences on top-down word comprehension during the processing of speech-in-noise, and to provide a better understanding of the neural processes underlying such semantic context effects. We first predicted better behavioral performance at higher signal-to-noise ratios (SNRs) — that is, better performance at lower compared to higher noise levels. In line with our behavioral study, we also predicted that we would find a semantic context effect (i.e.; benefit); that is, we predicted better behavioral performance during related compared to unrelated trials. Lastly, we predicted that this behavioral benefit of semantic relatedness would be most apparent at lower compared to higher SNRs. Examining brain regions whose activation is modulated by SNR will provide further insights into the brain networks involved in the effort of handling the challenge posed by acoustically degraded words. We expected that activation in brain regions including the left IFG and the left superior/middle temporal cortex would be modulated by speech intelligibility, and that activation in regions including the left angular gyrus would be further modulated by the presence of semantic context. Thus, we expected greater involvement of higher-level components of the language network, such as the angular gyrus, during more context-driven processing (i.e.; during semantically related trials). Conversely, we predicted greater involvement of lower-level components of the language network, such as the auditory cortex, during more stimulus-driven processing (i.e.; during semantically unrelated trials).

Materials and methods

Participants

Nine right-handed native French speakers (3 men) participated in the study. None had previously taken part in an experiment using this paradigm. Participants had a homogeneous language background; all had learned a second language in school (English, German, or Spanish) from the ages of 10 to 18, and a third language (English, German, or Spanish) from the ages of 12 to 18. None spoke a second or third language proficiently, and none had been regularly exposed to a language other than French before the age of 11, apart from sporadic and variable exposure to English by watching TV and being exposed to other popular media (music etc.).

Participants were recruited in the larger framework of a bilingual study, and all had been living in an English-speaking environment (London, UK) for six months to 3 years prior to scanning. In the current study, however, we only analyze and report responses to native-language French stimuli. All participants gave written informed consent to participate in the study, which was approved by the regional ethical committee.

Stimulus materials

A forced-choice visual recognition task on the first of two auditorily presented semantically related or unrelated words was employed, where the first, 'target' word was embedded in different noise levels (see below for task details). Details of stimulus parameters and synthesis have previously been reported (Golestani et al., 2009). While both

English and French stimuli were presented, only the French trials are analyzed. Also, in the behavioral study (Golestani et al., 2009), we had additionally used an SNR of −4 dB, but we excluded this from the present study since in the previous study there was no main effect of context (semantic relatedness) in the native language at this SNR. Excluding this SNR allowed us to include more trials per condition in the present fMRI study.

French semantically related and unrelated word pairs were selected from the database by Ferrand and Alario (1998). Two sets of stimuli were generated (i.e. two sets of 520 word pairs) such that in list 1, a specific target was followed by a related prime, and that in list 2, it was followed by an unrelated prime (see Table 1 for examples of stimulus items). The lexical frequency of related and unrelated primes was matched across lists because more frequent (or common) words are more likely to be recognized than less common ones (Bradlow and Pisoni, 1999). The number of syllables of related and unrelated primes was also matched because pilot testing revealed that longer words are more likely to be recognized than shorter ones, perhaps because in longer utterances there is usually more phonetic information that survives the noise. Word frequency information was taken from the 'Lexique 3' database (New et al., 2004) (www.lexique.org). Half of the participants were presented with one list (e.g. list 1 in Table 1) and the other half with the other list (e.g. list 2 in Table 1). This was done to ensure that results were not due to stimulus-specific effects but rather to the manipulations of interest.

The visual foils used in the recognition phase of the task were semantically matched to the target word (i.e. the degraded one, to be recognized). Visual foils were also matched with targets with respect to number of syllables; this was done in order to ensure that participants did not use this information to recognize the target from the foil.

The stimuli were digitally recorded by a multilingual female speaker in an anechoic chamber using a sampling rate of 44.1 kHz with 16 bit quantization. The microphone was positioned 30 cm from the speaker's mouth, at 15° to the mid-sagittal line. The final set of stimuli was created off-line by editing the words at zero crossing before and after each word. Recordings were normalized with respect to root-mean-squared amplitude and had an average duration of 1.1 s.

An equal proportion of targets from each list was embedded at the following SNRs: −7 dB (highest level of noise), −6 dB, and −5 dB, and no noise. We used speech-shaped noise which approximated the average long term spectrum of the speech of adult males and females. SNR was varied by manipulating the ratio of a mixture of the noise and the stimulus. The RMS amplitude for the final stimuli was fixed, and the RMS amplitude of the noise and the word was scaled relative to one-another to produce the desired SNRs, and the resulting combined sound was scaled to the desired RMS. Thus, the RMS amplitude of the stimuli was equal over all SNR conditions. Signal processing was carried out in Matlab (The Mathworks, Natick, MA.)

Experimental and scanning procedure

Before scanning, all participants performed approximately ten practice trials with stimuli that were not used in the fMRI experiment outside the scanner (at different SNRs), and the accuracy of their responses was verified in order to ensure that they understood the task at hand. Participants were told that they would have 1.5 s and not longer to respond on every trial, and that they should always try to respond within that time window. Imaging was performed on a 3.0-Tesla Philips (Best, The Netherlands) Intera scanner equipped with a six element SENSE head coil for radiofrequency signal detection, fitted with a B0-dependent auditory stimulus delivery system (MR-Confon, Magdeburg, Germany). Two series of 162 gradient-echo images of blood-oxygenation-level-dependent (BOLD) signal were acquired (TR = 5 s, TA = 2 s, 2.5 × 2.5 × 3.25 mm voxels). Auditory words were presented during the 3 second silent period between scans. High-resolution T1 weighted anatomical scans (0.94 mm × 0.94 mm × 1.2 mm voxels) were also obtained.

The following procedure was implemented for each participant during the two scanning blocks in French. There was a one minute period of rest between blocks. E-prime (Psychology Software Tools, Pittsburgh, PA) was used to present visual and auditory stimuli, and to collect responses. For each trial, participants heard a pair of words, half of which were semantically related and half of which were semantically unrelated. These stimuli were presented at a comfortable listening level, that was unchanged from participant to participant. The first, 'target' word was degraded by being presented at different SNRs (−7, −6, −5 dB, and no-noise), whereas the second, 'prime' word was always clearly audible. Subsequently, participants saw two visually presented words, one being the target and the other being a semantically related foil. They were required to decide which of the two visually presented words corresponded to the target by making a button press response (left button for word on left side of screen, and right button for word on right side of screen).

The following conditions were included: semantic context (related and unrelated), SNR (−7, −6, −5 dB, and no noise), and silent off-trials, resulting in a total of 9 conditions. There were 30 related and 30 unrelated word pairs at each of the four SNRs (i.e. 30 stimuli per condition), resulting in a total of 120 related and 120 unrelated word pairs. Each participant therefore performed a total of 240 trials. The SNR was blocked into mini-blocks of 5 trials each, and 'relatedness' (context) was mixed within mini-blocks in order to ensure that participants would not adopt different response strategies across relatedness conditions.

Data analysis

Data were preprocessed and analyzed using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK). Preprocessing included the following steps: 1) rigid realignment of the first EPI volumes of each session to one-another, followed by realignment of all EPI volumes in a session to the first image of that session, 2) "unwarping" of images to take account of susceptibility-by-movement artifacts, 3) coregistration of structural volumes to the first EPI image of the series, 4) normalization of the structural scan to a standard structural template (SPM5's T1 canonical template), 5) normalization of the EPI volumes by application of the parameters estimated in the normalization step, and 6) spatial smoothing using a Gaussian kernel of 8 mm^3.

For each participant a first level analysis was carried out using a general linear model (GLM) in which every scan was coded for condition. Null events were left unmodeled and formed an implicit baseline, as suggested by Josephs and Henson (1999). Each run was modeled separately and the effect of block was coded as a separate regressor. Each event was modeled using the canonical hemodynamic response function in SPM5. Six parameters were appended in each run to code for the effects of movement (x, y and z translations and x, y and z rotations derived from the rigid realignment step of the pre-processing). A high-pass filter (cutoff 128 s) and correction for first-order serial autocorrelation were applied.

The analysis of group data was achieved by entering the parameter estimates for each of the 8 conditions from each participant into a single random effect model. Using the contrast vectors described by Henson

Table 1
Examples of stimulus and test word pairs.

List 1 (related prime)				List 2 (unrelated prime)			
Target	Prime	Target	Foil	Target	Prime	Target	Foil
coq[a]	*poule*[a]	coq	œuf [a]	*coq*	*radio*[a]	coq	œuf

Note: Words in italics represent auditorily presented words (stimulus words), and words not in italics represent visually presented words (test words).
[a] English translations for French words: coq = rooster, poule = hen, œuf = egg, radio = radio.

and Penny (2005), an analysis of variance (ANOVA) with two factors (context: 2 levels; SNR: 4 levels) was carried out on the imaging data to test for the main effects of context and SNR, and any interaction between the two. In order to test for regions that show a response that is modulated by SNR, we carried out a correlational analysis in which we sought a linear effect of SNR. However, because the SNR = ∞ in the clear condition, SNR values were transformed into a measure of the proportion of signal and noise in the mixture.[1] The transformed values were then used as a contrast vector to find any regions showing a response that was reliably positively or negatively linearly correlated with SNR.

Unless otherwise stated, all results reported are significant at a family-wise error (FWE) corrected level of $p < 0.05$. Coordinates of peak activations are in the MNI (Montreal Neurological Institute) space.

Results

Behavioral results

Fig. 1 presents the reaction times obtained during task performance in the scanner. A 2-way (SNR by context) repeated-measures analysis of variance was performed on the reaction times, excluding the no noise condition. The no noise condition was excluded in order to be able to examine the effect of SNR only in conditions where noise was present, so as to ensure that the effect of SNR is not mainly driven by all conditions containing noise compared to the no noise one. As predicted, results revealed a main effect of SNR ($F_{(2, 16)} = 3.76$, $p < 0.05$), with faster reaction times at higher SNRs. The main effect of SNR was explained by a strong linear trend ($F_{(1, 8)} = 15.93$, $p < 0.01$), demonstrating that higher SNRs were associated with proportionately faster performance. There was also a predicted main effect of context ($F_{(1, 8)} = 5.77$, $p < 0.05$), with faster reaction times on semantically related compared to unrelated trials. Last, there was a weak trend towards an SNR by context interaction ($F_{(2, 16)} = 2.32$, $p = 0.13$). In Fig. 1 it can be seen that this is due to faster reaction times at SNR = −7 on related compared to unrelated trials. Planned comparisons confirmed the prediction of a benefit of context at the lowest SNRs ($F_{(1, 16)} = 6.62$, $p < 0.05$). There was no benefit of semantic context at any of the other SNRs (at SNR = −6: $F_{(1, 16)} = 0.00$, $p > 0.05$, at SNR = −5: $F_{(1, 16)} = 0.00$, $p > 0.05$, at SNR = no noise: $F_{(1, 24)} = 0.94$, $p > 0.05$). Given the benefit of context at this and not at other SNRs, these planned comparisons suggest a differential effect of context at different SNRs. Specifically, they confirm the predicted benefit of semantic relatedness at the lowest SNR.

The accuracy measures obtained during scanning were not as sensitive as reaction times to the context effects. Accuracy values ranged from 77 to 86% correct across the six different conditions containing noise (i.e. excluding the no noise conditions), with an overall mean of 82% correct and a standard deviation of 2.8%. These results demonstrate that although the task was a difficult one, performance was above chance (on a binomial distribution with $p = 0.5$, equivalent to random responding, the probability of scoring 77% correct by chance over 30 trials per condition = 0.002), and a relatively good proportion of items were likely comprehended by participants despite the items having been embedded in noise. For the no noise conditions, accuracy values ranged from 94 to 97% correct, with an overall mean of 96% and a standard deviation of 1.3%. A 2-way (SNR by context) repeated-measures analysis of variance performed on the accuracy measures showed the predicted main effect of SNR ($F_{(3,24)} = 20.5$, $p < 0.001$), with better performance at higher compared to lower SNRs. The SNR effect was explained by strong linear ($F_{(1,8)} = 27.61$, $p < 0.001$) and quadratic ($F_{(1,8)} = 73.34$, $p < 0.001$) trends, demonstrating that higher SNRs were associated with proportionately better performance. Unlike the

[1] Proportion of signal in stimulus: No noise = 1, SNR −5 = 0.3162, SNR −6 = 0.2512, and SNR −7 = 0.1995.
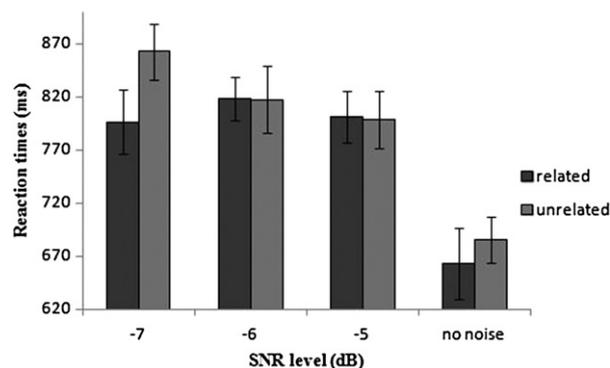

Fig. 1. Mean reaction times of behavioral performance across the group. Error bars indicate standard error of the mean (SEM).

results of the previous behavioral study (Golestani et al., 2009), the main effect of context was not significant ($F_{(1,8)} = 0.17$, $p > 0.05$), and there was no significant SNR by context interaction ($F_{(3,24)} = 0.88$, $p > 0.05$).

Imaging results

A positive correlation between brain activation and SNR (or alternatively, a negative correlation between activation and noise, as shown on the red scale in Fig. 2) was observed in a network of regions including portions of the left angular gyrus, and the posterior cingulate gyrus and adjacent precuneus. At a lower statistical threshold of $p < 0.001$ (uncorrected for multiple comparisons) the right angular gyrus, the medial orbitofrontal gyrus and the anterior cingulate cortex also appeared, revealing a highly symmetrical network. As can be seen from the contrast estimate graphs showing the pattern of signal change across conditions in Fig. 2, the correlation with SNR in this network is driven by progressively greater relative deactivation when the stimuli are presented at lower SNRs (i.e. in greater amounts of noise).

A negative correlation between activation and SNR (or alternatively, a positive correlation between activation and noise, as shown on the blue scale in Fig. 2) was found in the anterior insula/frontal operculum bilaterally. At a lower statistical threshold of $p < 0.001$ (uncorrected for multiple comparisons), the left IFG (*pars triangularis* and *pars opercularis*), the lower bank of the left posterior superior temporal sulcus (pSTS) extending into the MTG, and the right IFG (*pars triangularis*) also appear. It can been seen from the activation plots in Fig. 2 that these regions are relatively more engaged when the stimuli are more masked by noise. In other words, activation in this speech-related network is greater when speech processing is made more challenging by noise. These results are summarized in Table 2.

The above results were uncovered by seeking a linear effect of SNR in our data. The main effect of SNR (i.e. not looking for a linear relationship specifically) within a 2-way ANOVA on the imaging results shows activation in exactly the same regions as reported above, at a threshold of $p < 0.001$ uncorrected.

In parallel with the semantic context effect in the behavioral results, the imaging results showed the predicted effect of context in the angular gyrus of the left inferior parietal cortex, and at the lower threshold, in its right hemisphere homologue. These regions were more active during the related compared to the unrelated conditions. There was also an effect of context in the left and right mid-STS, with the left hemisphere result appearing at a more lenient threshold. In these regions, signal change was greater in unrelated compared to related trials. Table 3 and Fig. 3 summarize these results.

In line with the behavioral data indicating a differential impact of context at different SNRs, the imaging results revealed a significant SNR by context interaction in the angular gyrus of the left inferior
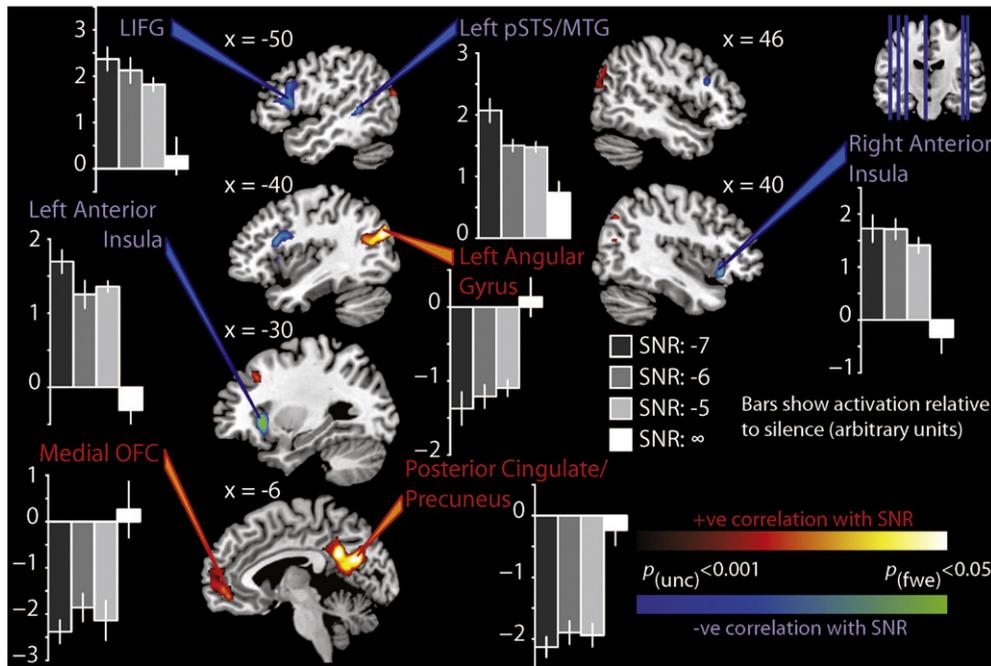
Fig. 2. Correlation with SNR. Positive correlations are displayed in red and reveal regions where activity is increased with increasing SNR. Negative correlations are displayed in blue and reveal regions whose activity increases with decreasing SNR. Bar plots represent signal change across conditions at locations indicated in Table 2. Note that for display purposes, images are presented at a threshold of p < 0.001, uncorrected. Error bars indicate SEM, corrected for within-subject comparisons (Loftus and Masson, 1994).

parietal cortex only, at the MNI coordinates X = −42, Y = −62, Z = 36. Table 4 and Fig. 4 summarize these results. Here, contrast estimate graphs show the pattern of signal change for the interaction effect: there is a relatively greater recruitment of this region in the related versus the unrelated conditions at higher compared to lower SNRs. Higher SNRs therefore appear to result in a greater context effect in this region. Note that this is opposite to the pattern in the behavioral findings, where the effect of context is greater at lower SNRs.

Discussion

Here, we used a semantic retroactive priming paradigm with degraded words to examine the modulation of brain activation by semantic context during effortful speech processing. Analysis of the behavioral results obtained during scanning revealed that, as expected, performance was faster when words were presented at higher compared to

lower SNRs. The imaging results showed that activation in components of the brain language network including Broca's area, extending to the left anterior insula/frontal operculum medially, the left pSTS/MTG, as well as brain regions known to be functionally related to attention and task difficulty (components of the default-mode network, see below), was modulated by stimulus intelligibility.

The behavioral data also revealed an overall benefit of semantic context on performance, and they further showed that at the lowest SNR (−7 dB), performance was faster on semantically related compared to unrelated trials. These findings replicate a previously reported behavioral benefit of semantic context during the processing of native language speech in noise, shown there also to be particularly beneficial at relatively lower compared to higher SNRs (Golestani et al., 2009). The imaging results revealed a differential response pattern as a function of semantic context. Specifically, there was an effect of context in the left angular gyrus and in the mid-STS bilaterally. We also found that the strength of the neural context effect was modulated by SNR in the left angular gyrus. We will now discuss the possible significance of these respective findings.

Signal-to-noise ratio (SNR) effects

Regions negatively correlated with SNR (or positively with noise)
A number of previous functional imaging studies have examined the neural bases of speech intelligibility at the sentence level (Davis

Table 2
Weighted effect of SNR.

Structure	MNI coordinates			Z	Direction of correlation with noise
Left inferior frontal gyrus (IFG)	−46	20	16	4.99[a]	Positive
Right IFG	50	22	24	4.17[a]	Positive
Left anterior insula/frontal operculum (FO)	−30	22	−4	5.31	Positive
Right anterior insula/FO	34	24	−6	5.03	Positive
Left angular gyrus	−40	−78	28	4.89	Negative
Right angular gyrus	44	−70	38	4.23[a]	Negative
Posterior cingulate gyrus	−2	−50	28	5.19	Negative
	6	−50	18	4.84	Negative
Precuneus	−2	−68	26	5.09	Negative
Anterior cingulate/medial orbitofrontal gyrus	8	56	−2	4.36[a]	Negative
Left posterior STS/MTG	−52	−46	8	4.48[a]	Positive

Note: Results are presented at a family-wise error (FWE) corrected level of p < 0.05.
  [a] Indicates a result appearing at a lower voxel-wise threshold of 0.001, uncorrected.

Table 3
Main effect of context.

Structure	MNI coordinates			Z
Left angular gyrus	−42	−64	36	4.93
Right angular gyrus	56	−54	40	4.00[a]
Right mid-STS	62	−10	−10	5.22
	58	−16	−8	4.91
Left mid-STS	−52	−12	−10	4.21[a]

Note: Results are presented at a family-wise error (FWE) corrected level of p < 0.05.
  [a] Indicates a result appearing at a lower voxel-wise threshold of p < 0.001, uncorrected.
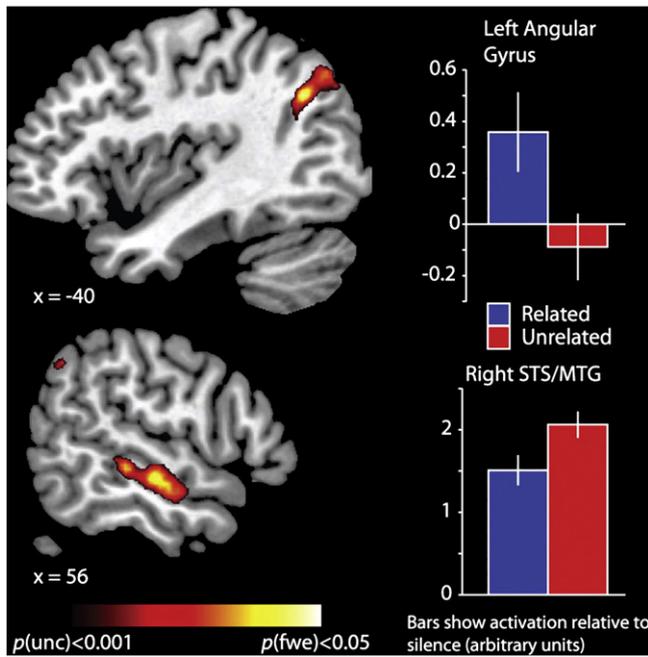
Fig. 3. Main effect of semantic context. Bar plots represent signal change across conditions at locations indicated in Table 3. Error bars indicate SEM, corrected for within-subject comparisons.

and Johnsrude, 2003; Davis et al., 2011; Giraud et al., 2004; Narain et al., 2003; Obleser et al., 2007a; Okada et al., 2010; Scott et al., 2000, 2006). Consistent with these studies, our results show that activation in a fronto-temporal network including the left IFG and the medially adjacent anterior insula, as well as left posterior STS/MTG is modulated by speech intelligibility.

We find greater engagement of the left IFG at lower SNRs (i.e. when there is more noise) compared to when the target words are heard more clearly, suggesting that more effortful processing is required at lower SNRs. Further, when the words are presented in no noise, left IFG involvement is minimal (see Fig. 1). This is consistent with the results of two studies by Davis and Johnsrude (2003) and Davis et al. (2011) using sentences that were distorted to different extents, ranging from being clear, to being degraded but intelligible, to being degraded but unintelligible. In both studies, inverted 'U'-shaped signal change patterns were shown in the left IFG as a function of the degree of degradation: activation was found to be strongest when sentences were degraded but intelligible compared to the two other conditions (Davis and Johnsrude, 2003; Davis et al., 2011). These findings suggest that greatest involvement of this region is observed when speech can be comprehended, but effortfully (Peelle et al., 2010). Similarly, in an elegant study by Giraud et al. (2004), neural responses to natural sentences, to broad-band speech envelope sentences (intelligible but effortfully after training) and to narrow-band speech envelope sentences (unintelligible, even after training) were compared. It was found that after training, Broca's area responded to natural speech and to broadband speech envelope noises, but that the response to natural speech dropped rapidly whereas the response to broadband speech

envelope noises remained sustained throughout stimulus presentation (Giraud et al., 2004). This finding together with ours and with those of Davis and Johnsrude (2003) and Davis et al. (2011) suggests that the left IFG is recruited during the processing of degraded speech, very probably to support comprehension. Obleser et al. (2007a) showed an increase in left IFG involvement with increasing intelligibility, but in this study responses to clear speech were not evaluated (Obleser et al., 2007a).

The left IFG has been shown to play a role in linguistic (phonetic, semantic, syntactic, and prosodic) processing, in verbal working memory, and in semantic and contextual integration (Baumgaertner et al., 2002; Braver et al., 1997; Buckner et al., 1995; Dapretto and Bookheimer, 1999; Gold and Buckner, 2002; Gough et al., 2005; Keller et al., 2001; Paulesu et al., 1993; Rodd et al., 2005). These are processes which can be expected to be more strongly recruited during sentence compared to word processing. Nonetheless, we still observe involvement of the LIFG during the low SNR condition, demonstrating the role of this region during effortful speech processing even when there are lesser demands on higher-level linguistic processing during our word-level task. This finding is consistent with the results of Hervais-Adelman et al. (2012) who demonstrated involvement of the LIFG in response to listening to potentially-comprehensible degraded words, as compared to clear words and to incomprehensible degraded words. Left IFG involvement during degraded word processing is also consistent with the results of another study, where words were degraded by being interrupted by white noise bursts centered around the fricatives and affricatives. It was found that the left IFG is involved in the 'repair' process, whereby participants had the illusion of hearing natural words when in fact they were hearing degraded words (Shahin et al., 2009). Also work on processing sentences containing semantic ambiguity has shown the involvement of regions including the left IFG during the disambiguation of high-ambiguity sentences, in support of the role of this region in selecting contextually appropriate word meanings during processing that is more demanding (i.e. that requires disambiguation) (Rodd et al., 2005). A recent study has shown that response bias can modulate the speech motor system during speech processing (Venezia et al., 2012). Our findings in the left IFG are unlikely to be driven by response bias, however, because the present 2-alternative forced-choice task was unlikely to become systematically biased (i.e., participants had to choose one of two words on each trial, and the pairs of words differed from trial to trial).

We find evidence for greater activation in the left posterior STS/MTG with increasing degradation levels. The direction of this SNR effect is the same as that in the left IFG. Previous studies using sentences have shown greater involvement of regions along the temporal lobe, spanning from the anterior STS to the middle/lateral temporal gyrus (near Heschl's gyrus) to the posterior STG/STS/MTG during the processing of degraded but intelligible speech (Davis and Johnsrude, 2003; Davis et al., 2011; Narain et al., 2003; Scott et al., 2000, 2006). In one study where neural responses to degraded speech were explicitly compared to those of clear speech, and where the response profile in the posterior STG is shown, it appears as though similar to the left IFG, the response in this region is maximal when speech is degraded but intelligible, and that the response diminishes somewhat during listening to clear speech (Davis et al., 2011). This is in line with our finding that the response in this region is diminished when speech is effortlessly perceived. In the present study we do not find evidence for modulation by SNR in the left anterior STG/STS. This may due to methodological issues related to fMRI in the anterior temporal lobe, which is known to be more susceptible to signal loss due to magnetic susceptibility artifacts (Devlin et al., 2000).

The left posterior MTG has previously been implicated in semantic processing (Chao et al., 1999; Moore and Price, 1999; Mummery et al., 1998; Perani et al., 1999), including modality independent semantic processing (e.g.; in response to pictures and words) (Vandenberghe et al., 1996). More recently, it has been proposed that this region is

Table 4
SNR × context interaction.

Structure	MNI coordinates			Z
Left angular gyrus	−42	−62	36	5.13
(2nd cluster)	−48	−60	22	4.14[a]

Note: Results are presented at a family-wise error (FWE) corrected level of p < 0.05.
[a] Indicates a result appearing at a lower voxel-wise threshold of p < 0.001, uncorrected for multiple comparisons.
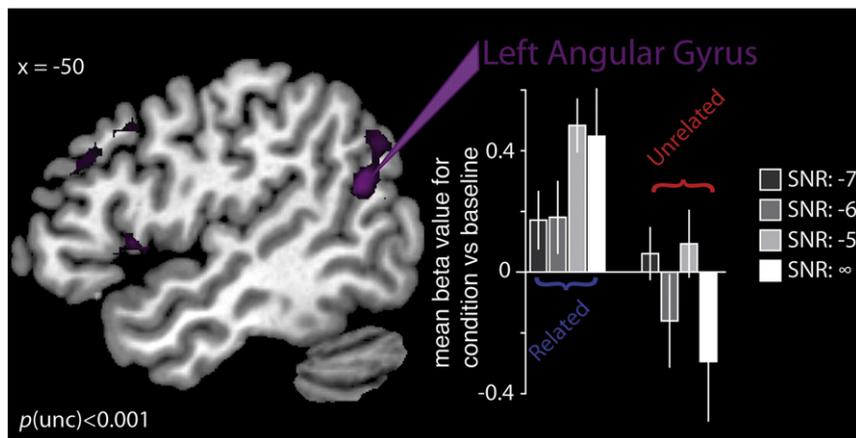
Fig. 4. SNR by context interaction. Bar plots represent signal change across conditions at locations indicated in Table 4. Error bars indicate SEM, corrected for within-subject comparisons.

important for linking sound to meaning, or that it plays a role as a 'lexical interface' (Hickok and Poeppel, 2004, 2007; Turken and Dronkers, 2011). Our results, which show stronger recruitment of this region during degraded word processing, suggest that there are greater demands on semantic processing and on the lexical interface during more effortful processing.

Speech in noise processing can be expected to require greater attention during the process of speech recognition (Darwin, 2008), and a recent review has addressed the different factors that are involved during speech processing under adverse conditions (Mattys et al., 2012). During lower compared to higher SNR conditions, we do not find greater involvement of brain regions known to be associated with attention, task difficulty or executive resources per se, such as subcortical regions, the posterior parietal cortex, the anterior cingulate cortex, or the dorsolateral prefrontal cortex (Hervais-Adelman et al., 2011; Paus et al., 1998; Petersen and Posner, 2012; Posner and Petersen, 1990). It is important to highlight however the contribution of top-down factors related to more effortful processing under conditions where speech is relatively more noisy, effects which may partly be sustained by regions such as the left IFG (see above).

Regions positively correlated with SNR (or negatively with noise)

We also find relative *suppression* of the BOLD signal during perception of speech in noise compared to no noise in brain regions including the inferior parietal cortices bilaterally, the anterior and posterior cingulate gyri, the precuneus, and the medial orbitofrontal gyrus. These regions are components of the 'resting baseline', or default network, which has been shown to be active during the resting state (i.e.; when no task is being performed) (Cavanna, 2007; Fransson and Marrelec, 2008; Margulies et al., 2009; Raichle et al., 2001). Studies have further shown that these very regions are relatively suppressed during the performance of a variety of tasks, reflecting blood flow decreases in brain regions and modalities that are not relevant to task performance (Gusnard et al., 2001; Haxby et al., 1994; Raichle et al., 2001; Seghier et al., 2010; Shulman et al., 1997). Such deactivations have previously been explained by the 'limited capacity' model of cognitive processing whereby during the performance of a particular task, attentional mechanisms select among competing mental processes (Handy, 2000; Posner and Presti, 1987; Shulman et al., 1997). Here, we find that the default network regions are relatively more suppressed during task performance when speech is presented in noise compared to no noise. We suggest that this is because during the former, more difficult condition, greater neural resources are allocated to other brain regions, ones that are relevant to successful task performance. Another study has also shown deactivation in a network very similar to ours during performance of a lexical

decision task on auditorily presented words preceded by subliminal primes (Kouider et al., 2010).

Semantic context effects

In parallel with the behavioral benefit of semantic context on task performance during scanning, we found that the left angular gyrus was overall more active during semantically related compared to unrelated conditions. Thus, we find that when there is a semantic relationship between the words, this region is recruited more strongly. Further, we find that this effect of semantic context in the left angular gyrus is modulated by SNR: the effect of context is stronger when stimuli are clear or only somewhat degraded compared to when they are presented at lower SNRs. Note that this pattern is different from what we find in the behavioral results, where the effect of context is greater at lower SNRs. This angular gyrus result suggests that under circumstances during which adequate resources can be devoted to successful word processing (i.e.; when words are clear or only somewhat degraded), greater recruitment of this region underlies processing of words within a semantic context (i.e.; when they are followed by semantically related words).

The role of the left angular gyrus in making use of linguistic context when processing degraded speech at the sentence level has previously been shown (Obleser and Kotz, 2010; Obleser et al., 2007a). In two studies by Obleser and colleagues, fMRI was used to examine the effects of higher-level, linguistic information (i.e.; predictability of the final word) during the perception of noise-vocoded sentences (Obleser and Kotz, 2010; Obleser et al., 2007a). In the Obleser et al. (2007a) study, it was found that activation in the left angular gyrus was modulated by the presence of linguistic context (predictability) but only at intermediate degradation levels. Specifically, this region was more activated when context was present compared to when it was not, at an intermediate level of degradation (Obleser et al., 2007a). In the more recent study using short, minimal sentences, it was found that this same region was best described with a "neurometric" function reflecting listeners' predictability benefit during comprehension; that is, its involvement was not simply tied to a specific degradation level (Obleser and Kotz, 2010). We find that the left angular gyrus is more strongly recruited when the degraded target word is followed by a semantically related prime word compared to when it is not, but mainly at high SNRs (−5 dB and during the no noise condition). Together our findings and the ones of preceding studies suggest that the left angular gyrus plays a role in utilizing linguistic context in order to assist the perception of intelligible speech, and of degraded speech. Such semantic processing and integration (linking meaning across words) to assist comprehension of

degraded speech can fail if the signal quality is too poor (when there is too much noise), and thus less angular gyrus involvement can be observed (Obleser and Kotz, 2010).

In the two above-described studies by Obleser and Kotz (2010) and Obleser et al. (2007a), sentence-level stimuli were used. As such, higher-level information other than semantics such as syntax and prosody also contained in the sentences likely contributed to the observed context effects. Our results, obtained with word-level stimuli, extend the above and show that the left angular gyrus plays a role in making use specifically of *semantic* information contained at the word level (Raettig and Kotz, 2008). A related, non-exclusive explanation for our effects is the possibility that we are observing retroactive semantic priming effects (i.e. when word pairs are related but not when they are unrelated), but mainly when the stimuli are perceived with relative ease.

Consistent with these interpretations, the left angular gyrus has previously been implicated in semantic processing (Binder et al., 1997, 2003; Chee et al., 1999; Price et al., 1999; Pugh et al., 1996; Seghier et al., 2010; Sharp et al., 2004), and has been shown to be part of a left perisylvian network that is activated during lexical decisions on auditorily presented words (Kouider et al., 2010). Consistent with the possible role of the left angular gyrus in semantic priming is the finding that in children and adolescents, there is a greater activation of the left inferior parietal lobule during semantic relatedness judgments when word pairs are strongly semantically related compared to when they are not semantically related (Chou et al., 2006). It has been suggested that this region is involved in feature integration and in semantic categorization, allowing semantic relationships between words to be determined (Grossman et al., 2003).

We also find that the left and right mid-STS were differentially activated during related compared to unrelated trials. Unlike the above results, these regions were more active in the unrelated compared to the related conditions. The left and right STS are known to play an important role in speech processing. Using fMRI, it has been shown that regions centered around the STS bilaterally respond more to speech stimuli than to tones, but also that these regions respond to words, pseudowords, and reversed speech, suggesting that these regions respond to lower-level, acoustic aspects of speech processing (Binder et al., 2000). Further, also using fMRI, it has been shown that dorsal temporal areas closer to primary auditory cortex in the STG respond bilaterally and equally to phonemic and nonphonemic sounds matched for complexity (Liebenthal et al., 2005). When speech sounds are perceived phonemically however, activation is observed in the left middle and anterior STS more than when sounds are not perceived phonemically, suggesting that these regions play a role in phonetic perception (Liebenthal et al., 2005; Obleser et al., 2007b). Our results thus suggest the existence of a complimentary aspect of the expected neural context effect. We find that lower-level, auditory regions responsible for acoustic/phonetic processing are more involved when semantic context is lacking and hence a more in-depth analysis at these lower processing levels is needed, compared to when a semantic context is present. In other words, we find that low-level auditory regions are relatively more involved during conditions that favor more perceptual, stimulus-driven processing.

General discussion and conclusions

It is of interest that some of our results regarding the effects of noise and of semantic context are convergent with those of previous studies even though in many of these, speech was degraded by noise-vocoding (Davis and Johnsrude, 2003; Davis et al., 2011; Hervais-Adelman et al., 2012; Obleser et al., 2007a; Okada et al., 2010; Scott et al., 2000, 2006) and not by embedding stimuli in noise (Davis et al., 2011; Giraud et al., 2004; Zekveld et al., 2012) as we have done. Also, most of these previous studies have used sentence-level stimuli whereas we have used word pairs. The convergence of our results

with those of previous studies suggests the implication of similar networks underlying speech intelligibility (LIFG and left STS) and semantic context (left angular gyrus) for noise-vocoding versus speech in noise, and for sentence-level versus word-level stimuli.

Our results have important implications for understanding the neural bases of language and communication that are generalizable to real life situations, where speech and communication often take place in noisy external and internal environments, as well as for cochlear implant users who hear speech in a manner that is distorted relative to normal-hearing individuals. Our results also speak to the more general question of how sensory (lower-level) and cognitive (higher-level) processes and resources interact in the brain: during the processing of speech in noise, the presence of semantic context engages a higher-level component of the language network, namely the left angular gyrus relatively more. Moreover, the lack of such semantic context engages lower-level, auditory brain regions, namely the bilateral STS, relatively more. Such a complementary interplay of perceptual and cognitive resources during linguistic processing likely extends to domains other than speech, and to modalities other than audition. For example, it may apply to the processing of written language, where poor input (e.g.; misspelled words) can be compensated for by expectancies (Rodd, 2004). Similar mechanisms likely operate in vision (Bar, 2004) and in memory (Jenkins and Ranganath, 2010; Montaldi and Mayes, 2010). More generally, there may be individual differences in the relative use of stimulus-driven versus context-driven processing, and changes throughout the lifespan might result in a relative shift from the former to the latter.

The exact contributions of different brain regions in using context to comprehend degraded speech can be further elucidated by seeking convergent evidence regarding the respective roles of these regions in the various required subfunctions. For example, a very recent study has shown the roles of the left anterior temporal lobe and of the left angular gyrus in basic linguistic combinatorials during both listening and reading (Bemis and Pylkkänen, 2012), demonstrating the modality independent contributions of these regions in the type of integration that is required when utilizing linguistic context to comprehend degraded speech. Our findings have extended this and other previous work in regard to the role of the left angular gyrus in showing that its involvement in combining meaning, or in using context to assist speech and language processing under demanding circumstances, can be demonstrated even at the level of word processing.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5, 617–629.
Baumgaertner, A., Weiller, C., Buchel, C., 2002. Event-related fMRI reveals cortical sites involved in contextual sentence integration. Neuroimage 16, 736–745.
Bemis, D.K., Pylkkänen, L., 2012. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhs170 (Epub ahead of print).
Bernstein, I.H., Bissonnette, V., Vyas, A., Barclay, P., 1989. Semantic priming: subliminal perception or context? Percept. Psychophys. 45, 153–161.
Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., Rzeczkowski, C., 1984. Standardization of a test of speech perception in noise. J. Speech Hear. Res. 27, 32–48.

Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., Prieto, T., 1997. Human brain language areas identified by functional magnetic resonance imaging. J. Neurosci. 17, 353–362.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cereb. Cortex 10, 512–528.

Binder, J.R., McKiernan, K.A., Parsons, M.E., Westbury, C.F., Possing, E.T., Kaufman, J.N., Buchanan, L., 2003. Neural correlates of lexical access during visual word recognition. J. Cogn. Neurosci. 15, 372–393.

Bradlow, A.R., Pisoni, D.B., 1999. Recognition of spoken words by native and non-native listeners: talker-, listener-, and item-related factors. J. Acoust. Soc. Am. 106, 2074–2085.

Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., Noll, D.C., 1997. A parametric study of prefrontal cortex involvement in human working memory. Neuroimage 5, 49–62.

Buckner, R.L., Raichle, M.E., Petersen, S.E., 1995. Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. J. Neurophysiol. 74, 2163–2173.

Cavanna, A.E., 2007. The precuneus and consciousness. CNS Spectr. 12, 545–552.

Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat. Neurosci. 2, 913–919.

Chee, M.W., O'Craven, K.M., Bergida, R., Rosen, B.R., Savoy, R.L., 1999. Auditory and visual word processing studied with fMRI. Hum. Brain Mapp. 7, 15–28.

Chou, T.L., Booth, J.R., Bitan, T., Burman, D.D., Bigio, J.D., Cone, N.E., Lu, D., Cao, F., 2006. Developmental and skill effects on the neural correlates of semantic processing to visually presented words. Hum. Brain Mapp. 27, 915–924.

Dapretto, M., Bookheimer, S.Y., 1999. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 24, 427–432.

Darwin, C.J., 2008. Listening to speech in the presence of other sounds. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 1011–1021.

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

Davis, M.H., Ford, M.A., Kherif, F., Johnsrude, I.S., 2011. Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. J. Cogn. Neurosci. 23, 3914–3932.

Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Wilson, J., Moss, H.E., Matthews, P.M., Tyler, L.K., 2000. Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. Neuroimage 11, 589–600.

Ferrand, L., Alario, X., 1998. Normes d'associations verbales pour 366 noms d'objets concrets. Annee Psychol. 98, 689–739.

Florentine, M., 1985. Speech perception in noise by fluent, non-native listeners. Proceedings of the Acoustical Society of Japan, p. H-85-16.

Fransson, P., Marrelec, G., 2008. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: evidence from a partial correlation network analysis. Neuroimage 42, 1178–1184.

Giraud, A.L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M.O., 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.

Gold, B.T., Buckner, R.L., 2002. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron 35, 803–812.

Golestani, N., Rosen, S., Scott, S.K., 2009. Native-language benefit for understanding speech-in-noise: the contribution of semantics. Biling.-Lang. Cogn. 12, 385–392.

Gough, P.M., Nobre, A.C., Devlin, J.T., 2005. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J. Neurosci. 25, 8010–8016.

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Moore, P., Gee, J., Detre, J., Alsop, D., 2003. Neural basis for verb processing in Alzheimer's disease: an fMRI study. Neuropsychology 17, 658–674.

Gusnard, D.A., Raichle, M.E., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. Nat. Rev. Neurosci. 2, 685–694.

Handy, T.C., 2000. Capacity theory as a model of cortical behavior. J. Cogn. Neurosci. 12, 1066–1069.

Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., Grady, C.L., 1994. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J. Neurosci. 14, 6336–6353.

Henson, R.N., Penny, W.D., 2005. ANOVAs and SPM. Institute of Cognitive Neuroscience: Wellcome Department of Imaging Neuroscience, London.

Hervais-Adelman, A.G., Moser-Mercer, B., Golestani, N., 2011. Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. Front. Psychol. 2, 234.

Hervais-Adelman, A.G., Carlyon, R.P., Johnsrude, I.S., Davis, M.H., 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. Lang. Cognit. Process. 27, 1145–1166.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99.

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

Jenkins, L.J., Ranganath, C., 2010. Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. J. Neurosci. 30, 15558–15565.

Josephs, O., Henson, R.N., 1999. Event-related functional magnetic resonance imaging: modelling, inference and optimization. Philos. Trans. R. Soc. Lond. B Biol. Sci. 354, 1215–1228.

Kalikow, D.N., Stevens, K.N., Elliott, L.L., 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J. Acoust. Soc. Am. 61, 1337–1351.

Keller, T.A., Carpenter, P.A., Just, M.A., 2001. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. Cereb. Cortex 11, 223–237.

Kouider, S., de Gardelle, V., Dehaene, S., Dupoux, E., Pallier, C., 2010. Cerebral bases of subliminal speech priming. Neuroimage 49, 922–929.

Liebenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T., Medler, D.A., 2005. Neural substrates of phonemic perception. Cereb. Cortex 15, 1621–1631.

Loftus, G.R., Masson, M.E.J., 1994. Using confidence-intervals in within-subject designs. Psychon. Bull. Rev. 1, 476–490.

Margulies, D.S., Vincent, J.L., Kelly, C., Lohmann, G., Uddin, L.Q., Biswal, B.B., Villringer, A., Castellanos, F.X., Milham, M.P., Petrides, M., 2009. Precuneus shares intrinsic functional architecture in humans and monkeys. Proc. Natl. Acad. Sci. U. S. A. 106, 20069–20074.

Mattys, S.L., Brooks, J., Cooke, M., 2009. Recognizing speech under a processing load: dissociating energetic from informational factors. Cogn. Psychol. 59, 203–243.

Mattys, S.L., Davis, M., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: a review. Lang. Cognit. Process. 27, 953–978.

Mayo, L.H., Florentine, M., Buus, S., 1997. Age of second-language acquisition and perception of speech in noise. J. Speech Lang. Hear. Res. 40, 686–693.

Montaldi, D., Mayes, A.R., 2010. The role of recollection and familiarity in the functional differentiation of the medial temporal lobes. Hippocampus 20, 1291–1314.

Moore, C.J., Price, C.J., 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. Brain 122, 943–962.

Mummery, C.J., Patterson, K., Hodges, J.R., Price, C.J., 1998. Functional neuroanatomy of the semantic system: divisible by what? J. Cogn. Neurosci. 10, 766–777.

Nabelek, A.K., Donahue, A.M., 1984. Perception of consonants in reverberation by native and non-native listeners. J. Acoust. Soc. Am. 75, 632–634.

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.

New, B., Pallier, C., Brysbaert, M., Ferrand, L., 2004. Lexique 2: a new French lexical database. Behav. Res. Methods Instrum. Comput. 36, 516–524.

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

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

Obleser, J., Zimmermann, J., Van Meter, J., Rauschecker, J.P., 2007b. Multiple stages of auditory speech perception reflected in event-related FMRI. Cereb. Cortex 17, 2251–2257.

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.

Paulesu, E., Frith, C.D., Frackowiak, R.S., 1993. The neural correlates of the verbal component of working memory. Nature 362, 342–345.

Paus, T., Koski, L., Caramanos, Z., Westbury, C., 1998. Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. Neuroreport 9, R37–R47.

Peelle, J.E., Johnsrude, I.S., Davis, M.H., 2010. Hierarchical processing for speech in human auditory cortex and beyond. Front. Hum. Neurosci. 4.

Perani, D., Schnur, T., Tettamanti, C., Gorno-Tempini, M., Cappa, S.F., Fazio, F., 1999. Word and picture matching: a PET study of semantic category effects. Neuropsychologia 37, 293–306.

Petersen, S.E., Posner, M.I., 2012. The attention system of the human brain: 20 years after. Annu. Rev. Neurosci. 35, 73–89.

Posner, M.I., Petersen, S.E., 1990. The attention system of the human brain. Annu. Rev. Neurosci. 13, 25–42.

Posner, M.I., Presti, D.E., 1987. Selective attention and cognitive control. Trends Neurosci. 10, 12–17.

Price, C.J., Mummery, C.J., Moore, C.J., Frakowiak, R.S., Friston, K.J., 1999. Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients. J. Cogn. Neurosci. 11, 371–382.

Pugh, K.R., Shaywitz, B.A., Shaywitz, S.E., Constable, R.T., Skudlarski, P., Fulbright, R.K., Bronen, R.A., Shankweiler, D.P., Katz, L., Fletcher, J.M., Gore, J.C., 1996. Cerebral organization of component processes in reading. Brain 119 (Pt 4), 1221–1238.

Raettig, T., Kotz, S.A., 2008. Auditory processing of different types of pseudo-words: an event-related fMRI study. Neuroimage 39, 1420–1428.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676–682.

Rodd, J.M., 2004. When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. Psychon. Bull. Rev. 11, 434–439.

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.

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 (Pt 12), 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.

Seghier, M.L., Fagan, E., Price, C.J., 2010. Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. J. Neurosci. 30, 16809–16817.

Shahin, A.J., Bishop, C.W., Miller, L.M., 2009. Neural mechanisms for illusory filling-in of degraded speech. Neuroimage 44, 1133–1143.

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

Sharp, D.J., Scott, S.K., Wise, R.J., 2004. Monitoring and the controlled processing of meaning: distinct prefrontal systems. Cereb. Cortex 14, 1–10.

Shulman, G.L., Corbetta, M., Buckner, R.L., Raichle, M.E., Fiez, J.A., Miezin, F.M., Petersen, S.E., 1997. Top-down modulation of early sensory cortex. Cereb. Cortex 7, 193–206.

Sohoglu, E., Peelle, J.E., Carlyon, R.P., Davis, M.H., 2012. Predictive top-down integration of prior knowledge during speech perception. J. Neurosci. 32, 8443–8453.

Takata, Y., Nabelek, A.K., 1990. English consonant recognition in noise and in reverberation by Japanese and American listeners. J. Acoust. Soc. Am. 88, 663–666.

Turken, A.U., Dronkers, N.F., 2011. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. Front. Syst. Neurosci. 5, 1.

van Wijngaarden, S.J., Steeneken, H.J., Houtgast, T., 2002. Quantifying the intelligibility of speech in noise for non-native talkers. J. Acoust. Soc. Am. 112, 3004–3013.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., Frackowiak, R.S.J., 1996. Functional anatomy of a common semantic system for words and pictures. Nature 383, 254–256.

Venezia, J.H., Saberi, K., Chubb, C., Hickok, G., 2012. Response bias modulates the speech motor system during syllable discrimination. Front. Psychol. 3, 157.

Zekveld, A.A., Rudner, M., Johnsrude, I.S., Heslenfeld, D.J., Rönnberg, J., 2012. Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. Brain Lang. 122 (2), 103–113.