NeuroImage 124 (2016) 581–590

Contents lists available at ScienceDirect

NeuroImage

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

Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception

Gavin M. Bidelman [a,b,*], Megan Howell [b]

[a] Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA
[b] School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

ARTICLE INFO

Article history:
Received 6 March 2015
Accepted 9 September 2015
Available online 16 September 2015

Keywords:
Auditory scene analysis
Event-related brain potentials (ERPs)
Speech processing
Hemispheric laterality
Neural compensation
Reverse hierarchy theory (RHT)
Speech-in-noise (SIN) perception

ABSTRACT

Previous studies suggest that at poorer signal-to-noise ratios (SNRs), auditory cortical event-related potentials are weakened, prolonged, and show a shift in the functional lateralization of cerebral processing from left to right hemisphere. Increased right hemisphere involvement during speech-in-noise (SIN) processing may reflect the recruitment of additional brain resources to aid speech recognition or alternatively, the progressive loss of involvement from left linguistic brain areas as speech becomes more impoverished (i.e., nonspeech-like). To better elucidate the brain basis of SIN perception, we recorded neuroelectric activity in normal hearing listeners to speech sounds presented at various SNRs. Behaviorally, listeners obtained superior SIN performance for speech presented to the right compared to the left ear (i.e., right ear advantage). Source analysis of neural data assessed the relative contribution of region-specific neural generators (linguistic and auditory brain areas) to SIN processing. We found that left inferior frontal brain areas (e.g., Broca's areas) partially disengage at poorer SNRs but responses do not right lateralize with increasing noise. In contrast, auditory sources showed more resilience to noise in left compared to right primary auditory cortex but also a progressive shift in dominance from left to right hemisphere at lower SNRs. Region- and ear-specific correlations revealed that listeners' right ear SIN advantage was predicted by source activity emitted from inferior frontal gyrus (but not primary auditory cortex). Our findings demonstrate changes in the functional asymmetry of cortical speech processing during adverse acoustic conditions and suggest that "cocktail party" listening skills depend on the quality of speech representations in the left cerebral hemisphere rather than compensatory recruitment of right hemisphere mechanisms.

1. Introduction

Speech communication rarely occurs in quiet environments as nearly all real-world listening situations (e.g., classrooms, cocktail parties, restaurants) contain some degree of noise interference (Helfer and Wilber, 1990). Extracting relevant information from undesirable auditory scenes is hindered by additional competing sounds to target speech. Indeed, language and literacy skills can be compromised when learning in noisy environments (e.g., Bronzaft, 2002). Additive noise acts as a simultaneous masker, obscuring less intense portions of the speech signal and reducing its signal-to-noise ratio (SNR). Reduced SNR prevents audible access to salient speech cues (e.g., temporal envelope) normally exploited for robust comprehension (Shannon et al., 1995).

Noise exclusion deficits are magnified with age and hearing loss (Harris and Swenson, 1990; Hazrati and Loizou, 2012; Nabelek, 1988). Yet, current hearing aids provide little benefit for speech-in-noise (SIN) understanding despite restoring audiometric thresholds (Chmiel and Jerger, 1996). It is now well accepted that SIN perception cannot be reliably predicted from the audiogram (Killion and Niquette, 2000). Moreover, SIN perception is problematic and highly variable among individuals without substantial hearing impairment (Divenyi and Haupt, 1997; Frisina and Frisina, 1997; Middelweerd et al., 1990) and even normal-hearing young adults (Song et al., 2011, 2012). These findings challenge conventional and longstanding views that speech intelligibility is determined solely by audibility, i.e., peripheral hearing status (Humes and Christopherson, 1991; Plomp, 1986; van Rooij et al., 1989). Rather, hearing sensitivity alone is inadequate to account for SIN perception issues (Humes and Christopherson, 1991; Parbery-Clark et al., 2011). Consequently, a growing body of evidence suggests that central auditory processing plays a critical role in mediating robust perceptual SIN abilities.

Auditory event-related brain potentials (ERPs) offer a precise temporal window to understand how noise affects the neural representation for speech and how central auditory brain mechanisms influence SIN listening skills. Noise-induced changes in the magnitude and timing of the auditory cortical ERPs have been reported by comparing responses to clean relative to noise-degraded speech sounds. The cortical encoding of auditory stimuli amidst noise reflects a complex interaction

* Corresponding author at: School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA. Fax: +1 901 525 1282.
E-mail address: g.bidelman@memphis.edu (G.M. Bidelman).

http://dx.doi.org/10.1016/j.neuroimage.2015.09.020
1053-8119/

between the types of signal/noise, as well as the evoking stimulus paradigm (e.g., sequential vs. oddball paradigm) (Billings et al., 2010). Nevertheless, component waves of the ERPs can be suppressed (i.e., delayed and reduced in amplitude) (Baltzell and Billings, 2014; Billings et al., 2009, 2010) or facilitated (i.e., enhanced in amplitude) (Alain et al., 2009; Parbery-Clark et al., 2011) depending on the type (e.g., white noise, multi-talker babble) and effectiveness of a concurrent noise in masking the target signal. Importantly, behavioral SIN skills are directly related to the magnitude of these noise-related changes in neural activity (Bennett et al., 2012; Bidelman and Dexter, 2015; Billings et al., 2013; Parbery-Clark et al., 2011). Collectively, these studies demonstrate that early cortical neural representations are sensitive to the SNR of the speech signal. More critically, they suggest that noise inhibits the robust encoding of speech acoustics, resulting in the delivery of impoverished neural representation(s) to perceptual mechanisms operating downstream.

Auditory scalp-recorded potentials reflect the engagement of multiple brain networks overlapping in both space and time. As such, it is difficult to ascribe noise-related changes in a particular ERP deflection to a single neural generator. Nevertheless, gross changes in cerebral activation and functional asymmetry (i.e., hemispheric weighting) have been reported during SIN perception. Under normal circumstances, the auditory system shows a prominent leftward lateralization for speech processing, consistent with the well-known functional bias and left hemisphere (LH) specialization for linguistic functions (Tervaniemi and Hugdahl, 2003; Zatorre et al., 1992). In noisy listening conditions however, ERPs show a progressive increase in rightward activation. This reallocation in neural activity has been interpreted as suggesting that right hemisphere (RH) brain areas are recruited to aid degraded speech recognition (e.g., Shtyrov et al., 1998, 1999). However, it is unclear from previous studies if this enhancement in RH activity reflects additional compensatory processing to assist SIN understanding (e.g., Bidelman and Dexter, 2015; Du et al., 2014; Wong et al., 2009) or alternatively, a loss of linguistic function and hence residual engagement of resources that are more specialized to process non-speech sounds (e.g., Bidelman and Dexter, 2015; Zendel et al., 2015).

To further elucidate the neural basis of SIN listening, we recorded neuroelectric activity in young adult listeners while listening to speech sounds presented in ongoing noise at various SNRs. We applied a distributed source analysis to ERP responses to evaluate region-specific source generator differences and lateralization for the neural encoding of acoustically-impoverished speech. Comparing listeners' electrical brain responses to their behavioral performance allowed us to directly assess the degree to which isolated neural substrates (auditory vs. linguistic brain areas) contribute to behavioral SIN abilities. Consistent with previous reports, we hypothesized that noise would both weaken and prolong auditory cortical responses (e.g., Billings et al., 2009, 2010) and modulate functional lateralization (e.g., Shtyrov et al., 1998) dependent on the speech SNR. However, extending previous findings, we predicted that SIN perception would decline concomitant with diminished neural activity in either linguistic or auditory brain areas in the left hemisphere. This finding would support the notion that SIN perception is primarily driven not by an enhancement (i.e., compensation) of neural processing from RH, per se, but rather, a loss in quality of neural representation within (left) linguistic brain regions.

2. Methods

2.1. Participants

Twelve, young adults (mean ± SD age: 24.7 ± 2.7 years) participated in the experiment. All had obtained a similar level of formal education (at least a collegiate-level undergraduate degree) and were monolingual speakers of American English. Musical training is known to amplify the auditory evoked potentials (e.g., Bidelman et al., 2011; Musacchia et al., 2008; Zendel and Alain, 2009) and improve SIN listening skills (Bidelman and Krishnan, 2010; Parbery-Clark et al., 2009; Zendel et al., 2015). Hence, all participants were required to have minimal formal musical training (1.3 ± 1.8 years) and none within the past five years. Air conduction audiograms confirmed normal hearing thresholds (i.e., ≤25 dB HL) at octave frequencies between 250 and 8000 Hz. Subjects also reported no history of hearing or neuropsychiatric disorders. Each gave informed written consent in compliance with a protocol approved by the University of Memphis Institutional Review Board and were reimbursed monetarily for their time.

2.2. Stimuli

Cortical auditory ERPs were elicited by a 300 ms/vCv/speech token/ama/ (cf. Bidelman, 2015; Shannon et al., 1999) (Fig. 1). The stimulus was a natural production recorded by a male speaker. The 50 ms nasal (/m/) was flanked by each vowel phoneme (/a/), both 125 ms in duration. The pitch prosody fell gradually over the duration of the token from an F0 of 120 Hz to 88 Hz. Vowel formant frequencies (F1–F3) were 830, 1200, and 2760 Hz, respectively. The intensity of the token was relatively fixed across its time course. In addition to this no noise "clean" stimulus (SNR = +∞ dB), noise-degraded speech stimuli were created by adding multitalker noise babble (Nilsson et al., 1994) to the clean token at SNRs of +10 and +5 dB. Importantly, SNR was manipulated by changing the level of the masker rather than the level of the signal. This ensured that SNR was inversely correlated with overall sound level (Binder et al., 2004). The babble was presented continuously throughout the (noise) experimental runs (i.e., the noise was not time-locked to the stimulus presentation) and was initiated ~5 s prior to delivery of the target speech stimuli. Continuous noise more closely mimics real-world situations whereby a listener is faced with extracting target signals above a blanket of competing background interference (e.g., cocktail party scenario) (e.g., Alain et al., 2012). The use of babble is also desirable as it has a similar deleterious effect on speech perception as other forms of interference (e.g., white noise) but a larger effect on neural encoding (Kozou et al., 2005).

2.3. Behavioral speech-in-noise task

We measured listeners' speech reception thresholds in noise using the QuickSIN test (Killion et al., 2004). The QuickSIN provides an efficient means to measure noise-degraded speech understanding and provides a standardized behavioral measure of SIN listening skills. In the present study, participants were presented with two lists of six sentences with five keywords per sentence embedded in four-talker

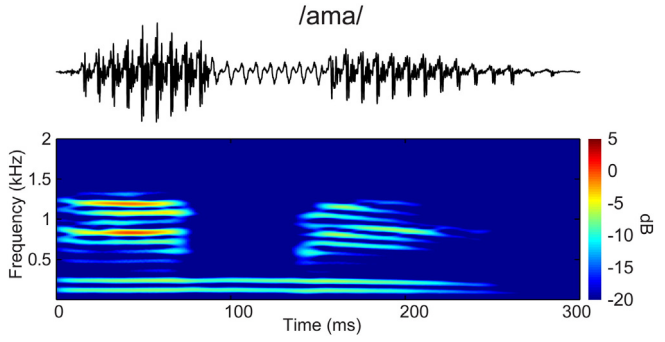

Fig. 1. Speech stimulus used to elicit cortical ERPs. Top, time waveforms of the vCv token/ama/; bottom, spectrogram. Tokens were 300 ms based on natural production of a male speaker. The 50 ms nasal (/m/) was flanked by each vowel phoneme (/a/), both 125 ms in duration. The pitch fell gradually over the duration of the token from an F0 of 120 Hz to 88 Hz. Vowel formant frequencies (F1–F3) were fixed at 830, 1200, and 2760 Hz, respectively. Noise babble was parametrically added to this clean token to achieve SNRs of +10 and +5 dB.

babble noise. Sentences were presented at 70 dB SPL using pre-recorded signal-to-noise ratios (SNRs) which decreased in 5 dB steps from 25 dB (very easy) to 0 dB (very difficult). After each sentence presentation, participants repeated the sentence and were awarded one point for each correctly repeated keyword. "SNR loss" (computed in dB) was determined by subtracting the total number of correctly recalled words from 25.5. The resulting value represents the SNR required to correctly identify 50% of the keywords in the target sentences (Killion et al., 2004). SNR loss was computed from two lists and averaged to improve precision. Left and right ears were tested separately.

2.4. Electrophysiological recordings

2.4.1. Data acquisition and preprocessing

Participants reclined comfortably in an IAC electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts), ignore the sounds they hear (to divert attention to the auditory stimuli), and were allowed to watch a muted subtitled movie to maintain a calm yet wakeful state. Stimulus presentation was controlled by MATLAB® 2013b (The MathWorks, Inc., Natick, MA, USA) routed to a TDT RP2 interface (Tucker-Davis Technologies) and delivered binaurally at an intensity of 81 dB SPL through ER-30 insert earphones (Etymotic Research). The lowpass frequency response of the headphone apparatus was corrected with a dual channel 15 band graphical equalizer (dbx EQ Model 215s; Harman) to achieve a flat frequency response out to 4 kHz. Stimulus intensity was calibrated using a Larson-Davis SPL meter (Model LxT) and measured in a 2-cc coupler (IEC 60126). Left and right ear channels were calibrated separately.

Listeners heard 2000 exemplars of each SNR stimulus presented with fixed, rarefaction polarity. SNR condition order was randomized within and across participants. The interstimulus interval (i.e., stimulus offset-to-onset) was jittered randomly between 480 and 720 ms (20-ms steps, rectangular distribution) to avoid rhythmic entrainment of the EEG (Luck, 2005, p. 168). The choice of these parameters ensured a balance between response recording time and minimizing stimulus-specific refractory (Picton et al., 1978) and component habituation effects (Picton et al., 1977) inherent to cortical auditory ERPs.

Neuroelectric activity was recorded from 64 sintered Ag/AgCl electrodes at standard 10-10 locations around the scalp (Oostenveld and Praamstra, 2001). EEGs were digitized using a sampling rate of 5000 Hz (SynAmps RT amplifiers; Compumedics Neuroscan). This high sampling frequency was necessary to record the extended bandwidth of the brainstem frequency-following response (not reported herein). Responses were then stored to disk for offline analysis. Electrodes placed on the outer canthi of the eyes and the superior and inferior orbit were used to monitor ocular activity. During online acquisition, all electrodes were referenced to an additional sensor placed ~1 cm posterior to Cz. However, data were re-referenced offline to a common average reference. Contact impedances were maintained below 5 kΩ throughout the duration of the experiment. EEGs were then digitally filtered (1.5–20 Hz, zero-phase) for response visualization and quantification.

Subsequent preprocessing was performed in Curry 7 (Compumedics Neuroscan) and custom routines coded in MATLAB. Data visualization and scalp topographies were computed using EEG/ERPLAB (Delorme and Makeig, 2004; Lopez-Calderon and Luck, 2014). Prior to artifact correction, excessively noisy channels were interpolated. Ocular artifacts (saccades and blink artifacts) were then corrected in the continuous EEG using a principal component analysis (PCA) (Wallstrom et al., 2004). The PCA decomposition provided a set of independent components which best explained the topography of the blink/saccadic artifacts. The scalp projection of the first two PCA loadings was subtracted from the continuous EEG traces to nullify ocular contamination in the final ERPs. Cleaned EEGs were then epoched (−200–550 ms), baseline-corrected (i.e., the mean pre-stimulus voltage was subtracted from each epoch time point), and subsequently averaged in the time domain to obtain ERPs for each SNR condition per participant. In total, ERPs reflect the average of 2000 presentations of each stimulus condition. The entire experimental protocol including behavioral and electrophysiological testing took ~2 h to complete.

2.4.2. ERP source analysis

Neuronal sources of evoked potentials must be inferred given the volume-conducted nature of the scalp-recorded EEG and "cross-talk" between adjacent sensor measurements. To more directly assess generator characteristics underlying speech-evoked ERPs and SIN processing we performed a distributed source analysis (e.g., Bidelman and Dexter, 2015). In this study, a distributed source approach is preferable to a simpler paired dipole model as we expected foci to vary as a function of speech SNR (e.g., Shtyrov et al., 1999); changes in location with stimulus SNR would render direct comparison of time-varying source strength spurious, as they would reflect signals generated in different brain areas. A distributed source approach allowed us to examine detailed neural activation within fixed regions of interest (ROIs). Source reconstruction was implemented in the MATLAB package Brainstorm (Tadel et al., 2011). We used a realistic, boundary element model (BEM) volume conductor (Fuchs et al., 1998, 2002) standardized to the MNI brain (Mazziotta et al., 1995). A BEM is less prone to spatial errors than other head models (e.g., concentric spherical conductor) (Fuchs et al., 2002).

The well-established sLORETA inverse solution (Pascual-Marqui, 2002) was used to estimate the distributed neuronal current density underlying the recorded sensor data. This algorithm models the inverse solution as a large number of elementary dipole generators distributed over nodes on a mesh of the cortical surface. When constrained to neocortical layers, the aggregate strength of source activity can be projected spatiotemporally onto the neuroanatomy, akin to functional maps in fMRI. The resultant activation maps represent the transcranial current source density underlying the scalp-recorded potentials as seen from the cortical surface. We used the default settings in Brainstorm's implementation of sLORETA (Tadel et al., 2011). From each sLORETA map, we extracted the time-course of source activity within two predefined ROIs: (1) bilateral primary auditory cortex (A1) (i.e., Heschl's gyrus, HG), (2) bilateral insula situated in inferior frontal gyrus (IFG). These ROIs were chosen to directly contrast source activity in brain regions subserving lower-order auditory processing and higher-order linguistic functions (e.g., Broca's area); they also allowed us to compare current results to those of other recent studies using identical ROIs (e.g., Bidelman and Dexter, 2015; Du et al., 2014). ROI parcellation was based on anatomical segmentations defined in FreeSurfer's Desikan–Killiany Atlas (Desikan et al., 2006) and the OpenMEEG BEM head model (Gramfort et al., 2010) as implemented in Brainstorm (Tadel et al., 2011)—see also, Fig. 1 of Klein and Tourville (2012). Resulting source waveforms reflect the neural activity (current, measured in μAmm) as seen within each anatomical ROI.

Peak amplitude and latency were measured for the prominent deflections of the source waveforms (P1, N1, P2, N2) in specific time windows. Following conventions in previous studies (e.g., Bidelman et al., 2013; Irimajiri et al., 2005), P1 was defined as the maximum positive wave between 50 and 80 ms, N1 the negative-going trough between 85 and 150 ms, P2 as the positive-going peak between 150–250 ms, and N2 as the negativity between 220 and 350 ms. All latency measures were corrected for the acoustic delay of the headphone transducer.

2.4.3. Laterality of auditory cortical source responses

Changes in hemispheric laterality of speech processing with additive noise were assessed by comparing neuroelectric source activity from the left and the right hemispheres. A laterality index (%) was computed as $LI = 100*(R-L)/(R+L)$, where R and L are the N1 response magnitudes (i.e., |amplitude|) from left and right hemispheres, respectively

(cf. King et al., 1999). Responses eliciting symmetric cortical activity produce a value of zero; asymmetric responses yield a positive or negative value for right or left hemisphere dominance, respectively.

2.5. Statistical analysis

Unless otherwise specified, two-way, generalized linear mixed-model ANOVAs were conducted on all dependent variables (GLIMMIX Procedure, SAS® 9.4, SAS Institute, Inc., Cary, NC, USA). Stimulus SNR (3 levels; clean, +10 dB, +5 dB) and hemisphere (2 levels; LH, RH) functioned as the within-subjects factor; subjects served as a random factor. Tukey–Kramer multiple comparisons controlled Type I error inflation. Significance in the laterality of source activity was first assessed using *t*-tests against zero (i.e., null hypothesis of bilateral symmetric response). We then compared shifts in laterality with increasing noise levels using a one-way, mixed-model ANOVA. An a priori significance level was set at $\alpha = 0.05$ for all statistical analyses. Normality and homogeneity of variance assumptions were confirmed prior to statistical inference.

Correlational analyses were used to determine the extent to which cortical responses could predict behavioral SIN performance, as measured via the QuickSIN. Spearman rank correlations were employed as this measure is preferable for small sample sizes and is less prone to artifacts due to outliers. Bootstrap resampling (Efron and Tibshirani, 1993) was used to validate significant brain-behavior correlations and estimate 95% confidence intervals (CIs). Participants were randomly resampled (with replacement) $N = 1000$ times from the original dataset. From each bootstrap surrogate, the correlation between neural and behavioral measures was recomputed. The probability density distribution of these statistics enabled us to confirm reliability of the observed correlations and estimate their variance.

3. Results

3.1. Behavioral speech-in-noise performance

Behavioral SIN scores, as measured via the QuickSIN are shown in Fig. 2. In general, participants achieved 0.5–1 dB better recognition for speech delivered to the right as compared to the left ear [paired samples *t*-test (two-tailed): $t_{11} = 2.66$, $p = 0.022$]. Consistent with the well-known asymmetry for speech perception in normal (i.e., quiet) listening conditions (Hiscock and Kinsbourne, 2011; Kimura, 1961), this finding indicates a similar right-ear advantage for degraded speech recognition.

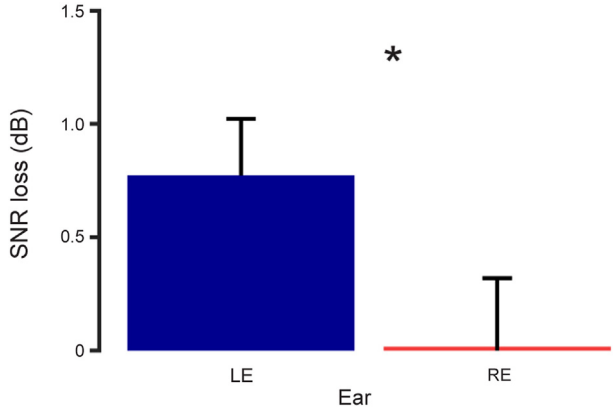

Fig. 2. Behavioral QuickSIN scores for left vs. right ear. Listeners achieved better SIN perception for speech delivered to the RE as compared to the LE consistent with the well-known right-ear advantage in (quiet) speech perception. $^{*}p < 0.05$; errorbars = ±1 s.e.m.

3.2. Cortical ERPs during SIN processing

Scalp topographies and time waveforms of the cortical ERPs (electrode data) as function of noise level are shown in Fig. 3. Visual inspection of scalp maps suggested weaker responses in the positive deflections (e.g., P1, P2) but stronger responses in the negative deflections (e.g., N1, N2) with increasing noise level (i.e., decreasing SNR) (see also Fig. S2). Hemispheric laterality also shifted rightward with increasing noise (Fig. S3). See Supplemental Material for analysis of the scalp data. The volume-conducted nature of sensor-space (i.e., electrode) recordings did not allow us to separate the underlying sources that contribute to these apparent functional changes. Consequently, subsequent analyses were conducted in source space to directly assess the neural generator characteristics underlying SIN processing.

3.3. Source activity during SIN perception

Whole brain source activations for clean and noise-degraded speech processing are shown in Fig. 4. Maps represent the statistical contrast (*t*-value, paired test) of sLORETA neural activations between the clean and +5 dB SNR conditions at 115 ms (~N1) and 230 ms (~N2) after stimulus onset. Maps were masked for multiple comparisons via false-discovery rate (FDR) on the family of contrasts under investigation ($\alpha = 0.05$) (Benjamini and Hochberg, 1995) and projected onto the semi-inflated cortical surface of the MNI brain (Collins et al., 1998). As indicated by the N1, clean speech recruited auditory cortex (HG) bilaterally but also engaged linguistic brain regions in IFG (Insula/Broca's area). This finding confirms a significant contribution of frontal brain areas to the scalp-recorded auditory evoked responses to speech (see also Fig. S3). Contrastively, we found that noise-degraded speech showed stronger (N2) responses in HG of the right hemisphere, consistent with a rightward lateralization observed in the topographies of electrode-level data. This apparent rightward shift in hemispheric laterality of the speech-evoked ERPs could represent a stronger recruitment of RH with noise-degradation (e.g., Fig. S3; Shtyrov et al., 1998; Shtyrov et al., 1999) or a disengagement of linguistic brain regions with the progressive loss of intelligibility (e.g., Fig. 4; Bidelman and Dexter, 2015; Zendel et al., 2015). Analysis of ROI-specific source waveforms extracted from within HG and IFG was used to tease apart these two interpretations.

3.4. Auditory cortical source responses during SIN processing

Fig. 5 illustrates source waveforms extracted from left and right primary auditory cortex (Heschl's gyrus). An ANOVA revealed a simple main effect of SNR on P1 [$F_{2, 55} = 6.02$, $p = 0.0043$] and N2 amplitudes [$F_{2, 55} = 8.05$, $p = 0.0009$] but no hemispheric effects ($ps > > 0.05$). Paralleling sensor-space results (i.e., Fig. S2), source P1 decreased and N2 increased in magnitude for noisier speech. P2 amplitudes were invariant to SNR [$F_{2, 55} = 2.16$, $p = 0.12$] and were similar across hemispheres [$F_{1, 55} = 0.06$, $p = 0.44$]. In contrast to other waves, N1 amplitudes showed a significant SNR × hemisphere interaction [$F_{2, 55} = 4.32$, $p = 0.0181$] (Fig. 5B). Follow-up contrasts revealed prominent hemispheric differences in the +10 dB SNR condition. That is, at even moderate noise levels, source activity elicited by noise-degraded speech was generally larger (i.e., more positive) in the left compared to the right hemisphere.

With regard to latency, P1 occurred earlier in the right compared to the left hemisphere [$F_{1, 55} = 7.11$, $p = 0.01$]. Also consistent with electrode-level responses, N1 and P2 source latencies were modulated by SNR, being prolonged for degraded compared to clean speech [N1: $F_{2, 55} = 6.21$, $p = 0.0037$; P2: $F_{2, 55} = 47.20$, $p < 0.0001$]. Neither a SNR nor a hemisphere effect was observed for N2 latency. Collectively, these findings imply that (i) the neural encoding of speech within auditory cortex is less efficient with increasing noise (i.e., delayed

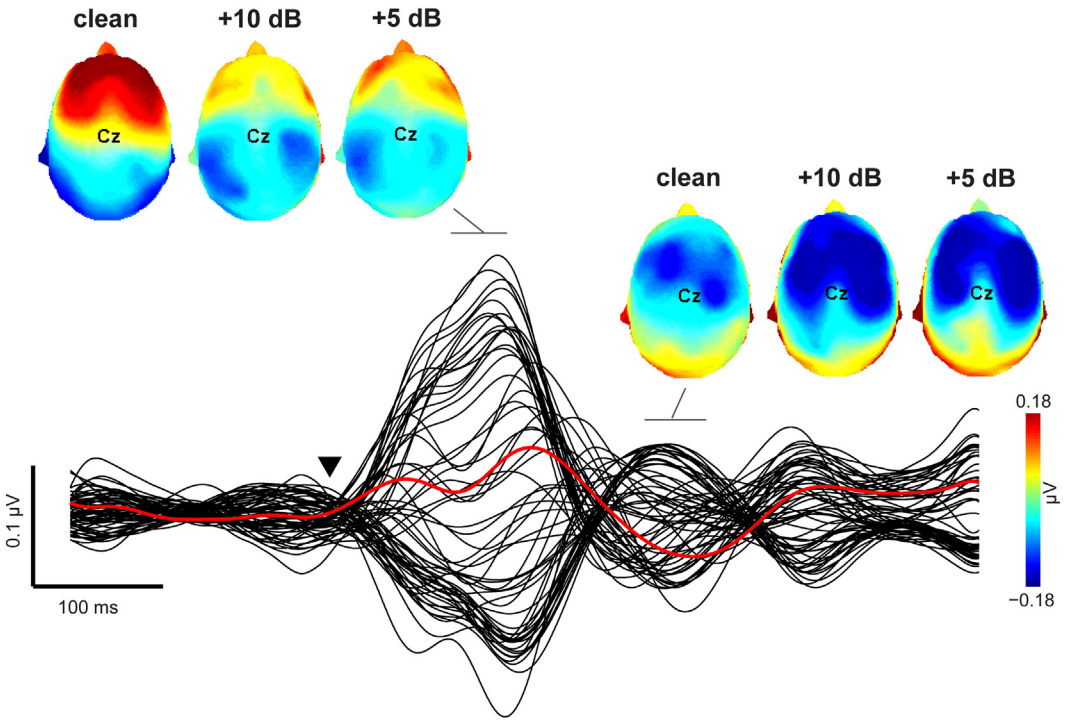

Fig. 3. Scalp topographies and butterfly plot of ERP time waveforms (electrode-level data) as a function of SNR. Cortical responses appeared as a series of obligatory positive and negative deflections (P1-N1-P2-N2 "waves") within the first ~150 ms after the time-locked stimulus. Response waves were modulated in both amplitude and latency with changes in speech SNR. Scalp maps illustrate voltage distributions at the latency of the P2 and N2 waves, respectively. Later and weaker responses with increasing noise suggest a progressive loss of efficiency in the early cortical processing of speech (see also Figs. S1 and S2). Arrow, onset of the time-locking speech token.

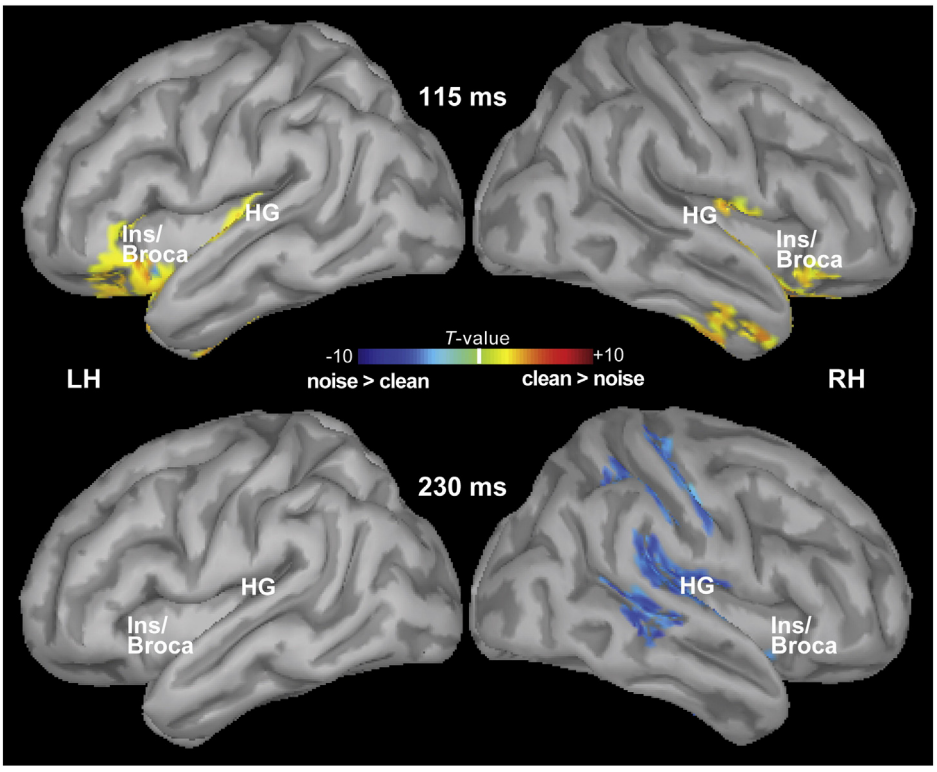

Fig. 4. Cortical source activations for clean and noise-degraded speech processing. Maps show FDR-corrected ($p < 0.05$) activations at 115 ms (~N1) and 230 ms (~N2) after stimulus onset between the clean and +5 dB SNR conditions projected onto the semi-inflated cortical surface of the MNI brain (Collins et al., 1998). Hot colors: brain areas where clean speech > noisy speech; cool colors: noise > clean. Clean speech recruits auditory cortex (HG) bilaterally but also engages linguistic brain regions in IFG regions (Ins./Broca's area). Contrastively, noise-degraded speech shows stronger responses in HG of the RH. HG, Heschl's gyrus; IFG, inferior frontal gyrus; Ins, Insula; LH/RH, left/right hemisphere.

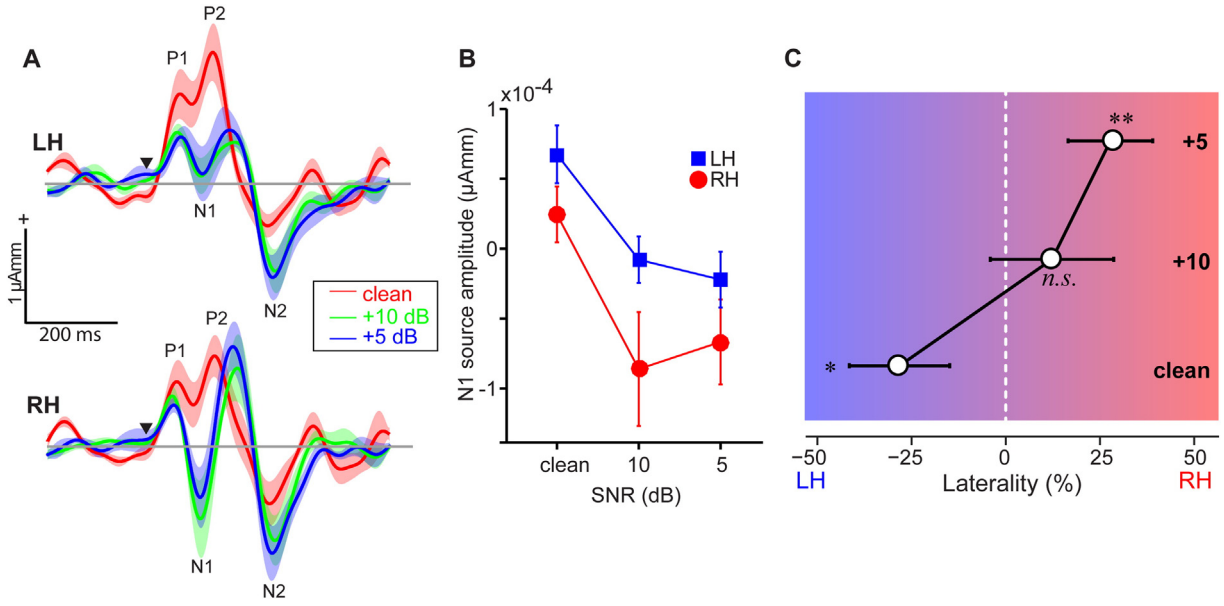

Fig. 5. Auditory cortical source activity to noise-degraded speech. (A) Source waveforms extracted from the left (top) and right (bottom) primary auditory cortex (Heschl's gyrus). Arrow, onset of the time-locking speech token. Consistent with electrode-level data (see SI Material), source P1 became weaker (less positive) and N1 stronger (more negative) with increasing noise. (B) Source N1 showed a hemisphere × SNR interaction; LH responses showed less noise-related change than RH indicating more resilience in LH speech processing. (C) Hemispheric laterality for noise-degraded speech processing. Laterality of the auditory N1 shows a leftward dominance (i.e., LH > RH) for clean speech. Functional responses become more symmetric and then shift rightward for progressively noisier speech. $^{*}p < 0.05$; $^{**}p < 0.01$; shading = ±1 s.e.m.; LH/RH, left/right hemisphere.

responses) and (ii) there is less noise-related change (i.e., more resilience) in response magnitude in left compared to right auditory cortex.

Functional laterality of SIN processing in auditory cortex is shown in Fig. 5C. Note that |LI| ≥ 20% is typically considered "lateralized" (Seghier, 2008). Quantitative analyses confirmed a shift in hemispheric dominance with increasing noise levels. Whereas clean speech was strongly lateralized in left hemisphere [$t_{11} = -2.11$, $p = 0.028$], the introduction of noise produced a shift toward the right hemisphere. Specifically, laterality diminished (became more bilateral) at +10 dB SNR [$t_{11} = 0.75$, $p = 0.76$] and began showing a strong rightward dominance at +5 dB SNR [$t_{11} = 2.45$, $p = 0.01$]. These observations were confirmed by an ANOVA, which revealed a significant effect of SNR on source laterality [$F_{2, 22} = 3.63$, $p = 0.04$]. Follow-up contrasts confirmed a significant change in speech laterality with increasing noise [linear contrast of SNR: $t_{11} = 2.83$, $p = 0.0165$]. Together, these findings indicate that auditory cortical responses are weighted more heavily in the left than the right hemisphere when processing clean, unadulterated speech. Furthermore, the leftward asymmetry in A1 responsiveness shifts progressively rightward during adverse speech listening. Complementary laterality effects were observed for sensor-space recordings (Fig. S3).

3.5. Inferior frontal gyrus source responses during SIN processing

Fig. 6 illustrates source waveforms extracted from left and right IFG. IFG responses were more variable than in the A1 ROI; only the N1 wave was clearly identifiable for subsequent analysis (Fig. 6A). An ANOVA conducted on IFG N1 source amplitudes revealed a simple main effect of hemisphere [$F_{2, 55} = 8.77$, $p = 0.0045$] (Fig. 6B). The sole main effect of hemisphere indicates that the N1 in frontal sources was larger (i.e., more negative) in LH across the board. Laterality of the IFG N1 revealed that clean speech was marginally lateralized in the left hemisphere [$t_{11} = -1.68$, $p = 0.06$] and significantly left lateralized with moderate noise (+10 dB SNR: $t_{11} = -2.27$, $p = 0.02$). Responses became bilateral for the +5 dB SNR condition [$t_{11} = -1.07$, $p = 0.85$].

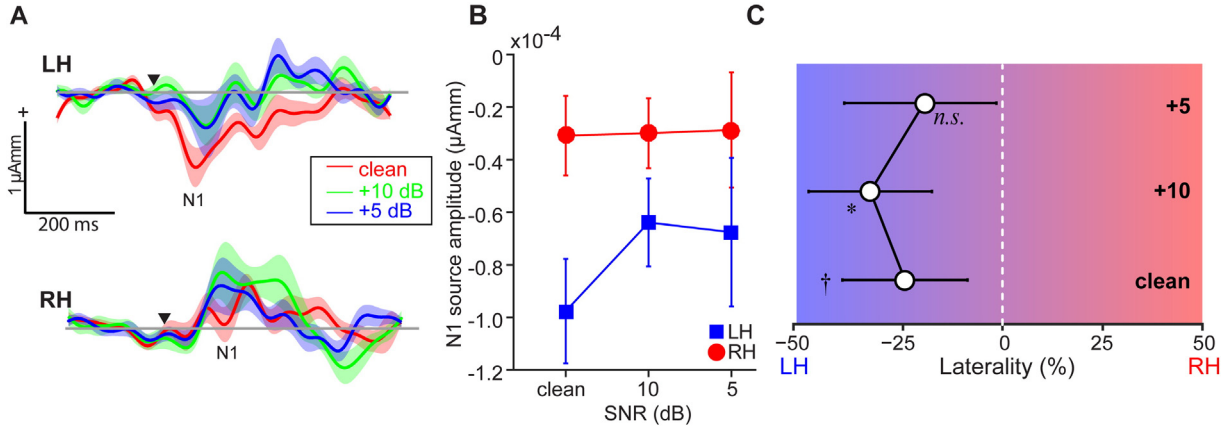

Fig. 6. Inferior frontal gyrus (Ins./Broca's) source activity to noise-degraded speech (otherwise as in Fig. 5). $^{\dagger}p < 0.06$; $^{*}p < 0.05$; shading = ±1 s.e.m.; LH/RH, left/right hemisphere.

3.6. Brain-behavior correspondences

To elucidate the neuroanatomical regions that drive the observed right ear advantage in SIN perception (Fig. 2), we assessed brain-behavior correspondences separately for each ear. Right and left ear QuickSIN scores were regressed against N1 source amplitudes extracted from the two ROIs (i.e., A1 and IFG). For each ROI, responses were collapsed across hemispheres and the two noise conditions (+10 and +5 dB SNR) and regressed separately against each ear's behavioral QuickSIN score. The pooling of hemispheres was motivated by the fact that even monaural stimulus presentation produces bilateral auditory cortical activity (Schonwiesner et al., 2007). N1 amplitudes were used in this analysis given that this was the only source component that showed a significant hemisphere × SNR interaction and was clearly visible across both ROIs.

N1 responses extracted within primary auditory cortex were not correlated with behavior for either ear [LE: $r = 0.12$, $p = 0.70$; RE: −0.15, $p = 0.65$] (data not shown). However, we found a strong association between behavioral scores and IFG (Ins./Broca's) source activity for right [$r = 0.63$, $p = 0.02$] but not left ear SIN listening [$r = 0.23$, $p = 0.47$] (Fig. 7). Despite the smaller sample size of the dataset, bootstrap and permutation resampling tests confirmed a significant difference between right and left ear correlations ($p < 0.0001$; Fig. 7 insets). These findings are consistent with the notions that SIN processing is driven by (i) higher order linguistic brain areas beyond sensory cortex (Bidelman and Dexter, 2015) and (ii) auditory information relayed from the right ear (Bidelman and Bhagat, 2015).

4. Discussion

4.1. Neural correlates of SIN perception

Consistent with the pattern of noise-induced changes described in previous reports (e.g., Billings et al., 2009; Parbery-Clark et al., 2011), we found that nearly all of the cortical ERPs (i.e., P1, N1, P2) showed modulations in latency and/or amplitude with changes in speech SNR at the sensor-level (i.e., electrode data; Fig. S2). However, previous dipole source analysis suggests that the mean "center of gravity" of speech evoked ERP (i.e., auditory responses) changes dramatically with SNR (e.g., Shtyrov et al., 1999). Indeed, our data revealed significant contributions of frontal brain areas during speech processing (e.g., Figs. 4 and S3) and thus, an overlap in the contributions from "auditory" and "linguistic" neural generators. This blurring of sources in scalp potentials precludes firm interpretations between auditory ERPs (sensor data) and SIN perception reported in previous studies.

Our analysis of auditory ERPs in source space allowed us to tease apart the contributions of various neural generators engaged during noise degraded speech processing. Our findings show activity restricted to primary auditory cortex undergoes distinct changes in both its amplitude and latency characteristics with noise. Nearly all defections of the auditory ERPs became progressively weaker and delayed in more challenging listening conditions (i.e., poorer SNRs). Our results thus extend previous ERP studies (Billings et al., 2009, 2010) by demonstrating these noise-related modulations result primarily from degraded neural representations within A1.

As noted by Billings et al. (2009), the fact that the majority of ERPs are affected by SNR–despite their distinct neural generators (Picton et al., 1999)–suggest that the observed changes in cortical speech processing might be driven by more peripheral auditory structures (e.g., subcortical processing). Indeed, previous studies examining auditory brainstem responses show decreased amplitude and increased latencies of subcortical responses when processing noise-degraded compared to clean speech signals (Anderson et al., 2011; Bidelman and Krishnan, 2010; Parbery-Clark et al., 2009). Moreover, the magnitude of noise-related changes in brainstem processing predicts listeners' success in SIN perception including recognition (Bidelman and Bhagat, 2015; Parbery-Clark et al., 2009) and discrimination tasks (Bidelman and Krishnan, 2010). In the current study, we found parallel degradations in auditory cortical source activity with additive noise. It is conceivable that this type of degraded signal analysis in early auditory cortex is at least partially inherited or influenced by structures much lower in the auditory pathway including the brainstem (Bidelman and Alain, 2015; Bidelman and Krishnan, 2010; Song et al., 2011) and/or cochlea (Bidelman and Bhagat, 2015).

In the current study, we found an increase in N1 negativity with low-level noise. Stronger N1 responses with decreasing SNR may reflect a partial facilitatory enhancement of speech with increasing noise levels, i.e., a physiological release from masking (Androulidakis and Jones, 2006). Consistent with our observations, low-intensity background noise has been shown in some studies to magnify cortical responses elicited by non-speech auditory stimuli (Alain et al., 2009, 2012, 2014). N1 response facilitation could emerge via efferent-induced modulations (Alain et al., 2009; Bidelman and Bhagat, 2015; Winslow et al., 1987). Indeed, medial olivocochlear (MOC) efferent neurons from the

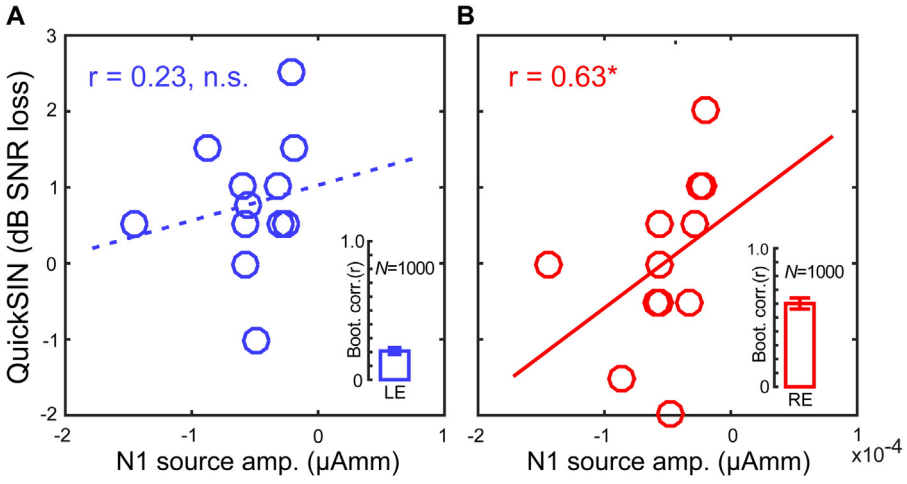

Fig. 7. Brain-behavior correlations underlying SIN perception. N1 amplitudes extracted in Ins./Broca's were pooled across hemispheres and the two noise conditions. Solid lines denote significant correlations; dotted lines, insignificant relationships. IFG (Ins./Broca's) source N1 amplitudes predict behavioral QuickSIN performance for SIN perception for right (B) but not left ear (A) listening. Larger N1 responses to speech in noise are associated with better SIN perception. Insets show bootstrapped estimates ($N = 1000$) of brain-behavior correlations. errorbars = 95% CIs, $^*p < 0.05$.

lower brainstem are thought to improve hearing in adverse listening conditions by playing an "antimasking" role in cochlear processing to improve signal detection and sensitivity in noise (Guinan, 2006; Scharf et al., 1997). The larger ERP amplitudes we observed with even minimal noise (e.g., +10 dB SNR) may reflect enhanced neural representations, vis-à-vis the efferent system, output from these lower level structures (i.e., cochlea or caudal brainstem nuclei) which are further tuned or at least maintained in cortical responses upstream. Indeed, as indexed by otoacoustic emissions, we have recently shown that SIN listening skills are well predicted by MOC efferent control (i.e., "antimasking") to the right cochlea (Bidelman and Bhagat, 2015) suggesting that this low-level system plays an important role in degraded speech perception.

Alternatively, noise may elongate the temporal integration window for auditory information (e.g., Alain et al., 2012) which could also account for more robust ERP responses in the noise conditions. Changes in the ERPs across SNRs could also be related to other stimulus factors. For instance, changes in inter-trial phase coherence with increasing noise levels (Ponjavic-Conte et al., 2013) would affect the degree of neural synchrony and hence amplitude of speech-evoked responses. These alternate interpretations could also account for the observed changes in ERP waves for noisier speech signals.

4.2. Brain-behavior relations underlying SIN perception

Previous studies examining auditory ERPs at a sensor level have shown a relationship between the N1 component (latency) and superior SIN perception (Billings et al., 2013; Parbery-Clark et al., 2009). However as noted earlier, the diffuse, volume-conducted nature of scalp-recorded potentials makes it impossible to ascribe a single neural generator to a particular ERP deflection. Indeed, the N1 is composed of multiple sub-components (Näätänen and Picton, 1987) including sources from the frontal lobes (Picton et al., 1999). Consequently, electrode-level recordings obscure potential relations between pure auditory neural activity and SIN behavior as they contain other, non-lemniscal and non-auditory brain activity. Frontal contributions are also evident in the current data. At both the electrode- and source levels, we observed significant contribution of frontal brain regions to the auditory ERPs (e.g., Figs. 4 and S3).

The current study replicates but further extends these previous studies by isolating region-specific cortical activity and revealing the nature of these brain-behavior relations. Notably, we found that (i) inferior frontal brain areas (but not primary auditory regions) predicted behavioral SIN skills and (ii) IFG (Ins./Broca) activation was associated with QuickSIN scores in the right but not left ear. These findings support the notion that the quality of neural representation(s) and speech information routed to the left hemisphere is critical for degraded speech perception and may account for the right ear dominance in SIN listening observed in this and previous studies (Bidelman and Bhagat, 2015). Our results suggest that the correspondence between N1 and SIN performance noted in previous work (e.g., Billings et al., 2013; Parbery-Clark et al., 2009) might not be driven entirely by auditory neural representations, per se. Instead, our data lead us to infer that the degree to which linguistic brain areas continue to be engaged, dictate how well listeners are able to extract speech in noisy listening environments.

4.3. Auditory vs. frontal engagement during SIN

The contrastive pattern of responses between auditory and frontal brain regions suggests that spatially distinct areas of cortex coding speech might be differentially vulnerable to noise. The differential recruitment of frontal vs. more temporal generators as a function of noise closely aligns with the notion of the reverse hierarchy theory (RHT) (Ahissar and Hochstein, 2004). Originally applied to the visual domain, RHT is a theoretical framework positing that with increasing task demand (e.g., poorer SNR), the brain performs a progressive backward search from higher-order information to lower-level inputs in search of representations with more optimal signal-to-noise ratio. Under the RHT, poorer neural representations that arise with noise in frontal, linguistic brain regions would force a backward search toward the sensory input (i.e., auditory cortex) in order to arrive at more favorable speech representations. While successful speech analysis may ultimately be determined by left IFG engagement (e.g., Fig. 7), the stronger resilience of left auditory cortex to noise (Fig. 5B) implies that the brain may rely more heavily on early sensory representations in LH to compensate for impoverished speech. This interpretation is supported by work showing that poorer perceivers of nonnative speech contrasts have attenuated auditory mismatch responses in frontal cortex relative to good perceivers (Bidelman and Dexter, 2015; Diaz et al., 2008) and the increased engagement of superior temporal brain regions in monolingual speakers during degraded listening conditions (Bidelman and Dexter, 2015).

4.4. Hemispheric laterality during SIN: recruitment of right- or loss of left-hemisphere involvement?

In addition to intracerebral frontal–temporal changes within each hemisphere, we found that noise degraded speech produced stronger responses in RH compared to LH (Figs. 4 and 5C). This finding is broadly consistent with previous reports demonstrating global, interhemispheric changes in speech activation during SIN listening and prima facie, that RH is progressively recruited to aid degraded speech perception in noisy listening environments (e.g., Okamoto et al., 2007; Shtyrov et al., 1998, 1999). However, stronger RH activity could indicate recruitment of compensatory processing to boost SIN understanding (e.g., Du et al., 2014; Shtyrov et al., 1999; Wong et al., 2009) or alternatively, the loss of linguistic function and hence residual engagement of RH resources that are more specialized to process non-speech sounds (e.g., Rinne et al., 1999). That is, a rightward shift may reflect a progressive diminishment of LH contribution, a consequence of its higher susceptibility to fluctuations in the noise envelope (Hiraumi et al., 2008) and/or the fact that stimuli become progressively more non-speech like at lower SNRs (cf. Rinne et al., 1999).

In this regard, separating the contributions of auditory and frontal ROIs helps tease apart these two interpretations. Laterality analysis of primary auditory source activity indicated a progressive shift in neural activity from left to right hemisphere with increasing noise. Our findings thus clarify previous studies that have implied increased RH recruitment for degraded speech signals (e.g., Shtyrov et al., 1998, 1999) by demonstrating this modal shift is due to changes in processing within left/right primary auditory regions. Indeed, we found that the normal leftward asymmetry in frontal sources was merely lost at poorer SNRs (becoming bilateral symmetric), whereas auditory cortical sources continued toward right lateralization with increasing noise. The asymmetry of auditory-specific neural responses is also evident in the differential change in source N1, which showed less reduction with noise in the left compared to the right auditory cortex (Fig. 5B). Collectively, our results suggest that during degraded speech processing, LH frontal (i.e., linguistic) brain areas disengage with increasing noise, relying more heavily on left auditory cortex to analyze impoverished speech signals.

Alternatively, we cannot rule out the interpretation that increased loudness of the babble masker might be more taxing in terms of neural processing, e.g., decreasing the auditory system's ability to spectrally glimpse speech (Cooke, 2006). Spectral glimpsing might draw on right-hemispheric auditory resources more than on left-hemispheric auditory resources (Scott et al., 2009). In this interpretation, the observed hemispheric shift would not only depend on the quality of the neural representation of target speech in the left hemisphere, but also on the way the competing background noise itself is processed. Regardless of the underlying mechanism, our data demonstrate changes in the functional asymmetry of cortical speech processing during adverse

acoustic conditions and suggest that "cocktail party" listening skills are largely dependent on the quality of speech representations in the left cerebral hemisphere rather than compensatory recruitment of additional right hemisphere resources.

Acknowledgments

This work was supported by grants from the American Hearing Research Foundation (AHRF) and American Academy of Audiology Foundation (AAAF) awarded to G.M.B.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.09.020.

References

Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. 8, 457–464.
Alain, C., Quan, J., McDonald, K., Van Roon, P., 2009. Noise-induced increase in human auditory evoked neuromagnetic fields. Eur. J. Neurosci. 30, 132–142.
Alain, C., McDonald, K., Van Roon, P., 2012. Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. Hear. Res. 283, 126–135.
Alain, C., Roye, A., Salloum, C., 2014. Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. Front. Syst. Neurosci. 8, 8.
Anderson, S., Parbery-Clark, A., Yi, H.G., Kraus, N., 2011. A neural basis of speech-in-noise perception in older adults. Ear Hear. 32, 750–757.
Androulidakis, A.G., Jones, S.J., 2006. Detection of signals in modulated and unmodulated noise observed using auditory evoked potentials. Clin. Neurophysiol. 117, 1783–1793.
Baltzell, L.S., Billings, C.J., 2014. Sensitivity of offset and onset cortical auditory evoked potentials to signals in noise. Clin. Neurophysiol. 125, 370–380.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300.
Bennett, K.O., Billings, C.J., Molis, M.R., Leek, M.R., 2012. Neural encoding and perception of speech signals in informational masking. Ear Hear. 32, 231–238.
Bidelman, G.M., 2015. Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. Hear. Res. 323, 68–80.
Bidelman, G.M., Alain, C., 2015. Hierarchical neurocomputations underlying concurrent sound segregation: connecting periphery to percept. Neuropsychologia 68, 38–50.
Bidelman, G.M., Bhagat, S.P., 2015. Right ear advantage drives the link between olivocochlear efferent "antimasking" and speech-in-noise listening benefits. Neuroreport 26, 483–487.
Bidelman, G.M., Dexter, L., 2015. Bilinguals at the "cocktail party": dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-noise recognition deficits. Brain Lang. 143, 32–41.
Bidelman, G.M., Krishnan, A., 2010. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Res. 1355, 112–125.
Bidelman, G.M., Gandour, J.T., Krishnan, A., 2011. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J. Cogn. Neurosci. 23, 425–434.
Bidelman, G.M., Moreno, S., Alain, C., 2013. Tracing the emergence of categorical speech perception in the human auditory system. NeuroImage 79, 201–212.
Billings, C.J., Tremblay, K.L., Stecker, G.C., Tolin, W.M., 2009. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. Hear. Res. 254, 15–24.
Billings, C.J., Bennett, K.O., Molis, M.R., Leek, M.R., 2010. Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. Ear Hear. 32, 53–60.
Billings, C.J., McMillan, G.P., Penman, T.M., Gille, S.M., 2013. Predicting perception in noise using cortical auditory evoked potentials. J. Assoc. Res. Otolaryngol. 14, 891–903.
Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A., Ward, B.D., 2004. Neural correlates of sensory and decision processes in auditory object identification. Nat. Neurosci. 7, 295–301.
Bronzaft, A.L., 2002. Noise Pollution: A Hazard to Physical and Mental Well-Being. John Wiley & Sons, Inc., N.Y.
Chmiel, R., Jerger, J., 1996. Hearing aid use, central auditory disorder, and hearing handicap in elderly persons. J. Am. Acad. Audiol. 7, 190–202.
Collins, D.L., Zijdenbos, A.P., Kollokian, V., et al., 1998. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imaging 17, 463–468.
Cooke, M., 2006. A glimpsing model of speech perception in noise. J. Acoust. Soc. Am. 119, 1562–1573.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134, 9–21.
Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980.
Diaz, B., Baus, C., Escera, C., Costa, A., Sebastian-Galles, N., 2008. Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. Proc. Natl. Acad. Sci. U. S. A. 105, 16083–16088.
Divenyi, P.L., Haupt, K.M., 1997. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation. Ear Hear. 18, 189–201.
Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc. Natl. Acad. Sci. U. S. A. 111, 1–6.
Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.
Frisina, D.R., Frisina, R.D., 1997. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. Hear. Res. 106, 95–104.
Fuchs, M., Drenckhahn, R., Wischmann, H.-A., Wagner, M., 1998. An improved boundary element method for realistic volume-conductor modeling. IEEE Trans. Biomed. Eng. 45, 980–997.
Fuchs, M., Kastner, J., Wagner, M., Hawes, S., Ebersole, J., 2002. A standardized boundary element method volume conductor model. Clin. Neurophysiol. 113, 702–712.
Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. Biomed. Eng. Online 45.
Guinan Jr., J.J., 2006. Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear Hear. 27, 589–607.
Harris, R.W., Swenson, D., 1990. Effects of reverberation and noise on speech recognition by adults with various amounts of sensorineural hearing impairment. Audiology 29, 314–321.
Hazrati, O., Loizou, P.C., 2012. The combined effects of reverberation and noise on speech intelligibility by cochlear implant listeners. Int. J. Audiol. 51, 437–443.
Helfer, K., Wilber, L., 1990. Hearing loss, aging, and speech perception in reverberation and in noise. J. Speech Hear. Res. 33, 149–155.
Hiraumi, H., Nagamine, T., Morita, T., Naito, Y., Fukuyama, H., Ito, J., 2008. Effect of amplitude modulation of background noise on auditory-evoked magnetic fields. Brain Res. 1239, 191–197.
Hiscock, M., Kinsbourne, M., 2011. Attention and the right-ear advantage: what is the connection? Brain Cogn. 76, 263–275.
Humes, L.E., Christopherson, L., 1991. Speech identification difficulties of hearing-impaired elderly persons: the contributions of auditory processing deficits. J. Speech Hear. Res. 34, 686–693.
Irimajiri, R., Golob, E.J., Starr, A., 2005. Auditory brain-stem, middle- and long-latency evoked potentials in mild cognitive impairment. Clin. Neurophysiol. 116, 1918–1929.
Killion, M., Niquette, P., 2000. What can the pure-tone audiogram tell us about a patient's SNR loss? Hear. J. 53, 46–53.
Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., Banerjee, S., 2004. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 116, 2395–2405.
Kimura, D., 1961. Cerebral dominance and the perception of verbal stimuli. Can. J. Psychol. 15, 166–170.
King, C., Nicol, T., McGee, T., Kraus, N., 1999. Thalamic asymmetry is related to acoustic signal complexity. Neurosci. Lett. 267, 89–92.
Klein, A., Tourville, J., 2012. 101 labeled brain images and a consistent human cortical labeling protocol. Front. Neurosci. 6.
Kozou, H., Kujala, T., Shtyrov, Y., Toppila, E., Starck, J., Alku, P., Naatanen, R., 2005. The effect of different noise types on the speech and non-speech elicited mismatch negativity. Hear. Res. 199, 31–39.
Lopez-Calderon, J., Luck, S.J., 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front. Hum. Neurosci. 8.
Luck, S., 2005. An Introduction to the Event-Related Potential Technique. MIT Press, Cambridge, MA, USA.
Mazziotta, J.C., Toga, A.W., Evans, A., Lancaster, J.L., Fox, P.T., 1995. A probabilistic atlas of the human brain: theory and rational for its development. NeuroImage 2, 89–101.
Middelweerd, M.J., Festen, J.M., Plomp, R., 1990. Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram. Audiology 29, 1–7.
Musacchia, G., Strait, D., Kraus, N., 2008. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear. Res. 241, 34–42.
Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375–425.
Nabelek, A.K., 1988. Identification of vowels in quiet, noise, and reverberation: relationships with age and hearing loss. J. Acoust. Soc. Am. 84, 476–484.
Nilsson, M., Soli, S.D., Sullivan, J.A., 1994. Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am. 95, 1085–1099.
Okamoto, H., Stracke, H., Ross, B., Kakigi, R., Pantev, C., 2007. Left hemispheric dominance during auditory processing in a noisy environment. BMC Biol. 5, 52.
Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112, 713–719.
Parbery-Clark, A., Skoe, E., Kraus, N., 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. J. Neurosci. 29, 14100–14107.
Parbery-Clark, A., Marmel, F., Bair, J., Kraus, N., 2011. What subcortical–cortical relationships tell us about processing speech in noise. Eur. J. Neurosci. 33, 549–557.
Pascual-Marqui, R.D., 2002. Standardized low resolution brain electromagnetic tomography (sLORETA). Methods Find. Exp. Clin. Pharmacol. 24D, 5–12.
Picton, T.W., Woods, D.L., Baribaeu-Braun, J., Healy, T.M.G., 1977. Evoked potential audiometry. J. Otolaryngol. 6, 90–119.

Picton, T.W., Woods, D.L., Proulx, G.B., 1978. Human auditory sustained potentials. II. Stimulus relationships. Electroencephalography and Clinical Neurophysiology 45 pp. 198–210.
Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neuro Otol. 4, 64–79.
Plomp, R., 1986. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. J. Speech Hear. Res. 29, 146–154.
Ponjavic-Conte, K.D., Hambrook, D.A., Pavlovic, S., Tata, M.S., 2013. Dynamics of distraction: competition among auditory streams modulates gain and disrupts inter-trial phase coherence in the human electroencephalogram. PLoS One 8, e53953.
Rinne, T., Alho, K., Alku, P., Holi, M., Sinkkonen, J., Virtanen, J., Bertrand, O., Naatanen, R., 1999. Analysis of speech sounds is left-hemisphere predominant at 100–150 ms after sound onset. Neuroreport 10, 1113–1117.
Scharf, B., Magnan, J., Chays, A., 1997. On the role of the olivocochlear bundle in hearing: 16 case studies. Hear. Res. 103, 101–122.
Schonwiesner, M., Krumbholz, K., Rubsamen, R., Fink, G.R., von Cramon, D.Y., 2007. Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. Cereb. Cortex 17, 492–499.
Scott, S.K., Rosen, S., Beaman, C.P., Davis, J.P., Wise, R.J., 2009. The neural processing of masked speech: evidence for different mechanisms in the left and right temporal lobes. J. Acoust. Soc. Am. 125, 1737–1743.
Seghier, M.L., 2008. Laterality index in functional MRI: methodological issues. Magn. Reson. Imaging 26, 594–601.
Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270, 303–304.
Shannon, R.V., Jensvold, A., Padilla, M., Robert, M.E., Wang, X., 1999. Consonant recordings for speech testing. J. Acoust. Soc. Am. 106, L71–L74.
Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R.J., Naatanen, R., 1998. Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: magnetic mismatch negativity study. Neurosci. Lett. 251, 141–144.
Shtyrov, Y., Kujala, T., Ilmoniemi, R.J., Naatanen, R., 1999. Noise affects speech signal processing differently in the cerebral hemispheres. Neuroreport 10, 2189–2192.
Song, J.H., Skoe, E., Banai, K., Kraus, N., 2011. Perception of speech in noise: neural correlates. J. Cogn. Neurosci. 23, 2268–2279.
Song, J.H., Skoe, E., Banai, K., Kraus, N., 2012. Training to improve hearing speech in noise: biological mechanisms. Cereb. Cortex 22, 1180–1190.
Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 1–13.
Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. Brain research. Brain Res. Rev. 43, 231–246.
van Rooij, J.C., Plomp, R., Orlebeke, J.F., 1989. Auditive and cognitive factors in speech perception by elderly listeners. I: development of test battery. J. Acoust. Soc. Am. 86, 1294–1309.
Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F., Fox, N.A., 2004. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. Int. J. Psychophysiol. 53, 105–119.
Winslow, R.L., Sachs, M.B., 1987. Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve response to tones in noise. J. Neurophysiol. 57, 1002–1021.
Wong, P.C.M., Jin, J.X., Gunasekera, G.M., Abel, R., Lee, E.R., Dhar, S., 2009. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia 47, 693–703.
Zatorre, R., Evans, A., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256, 846–849.
Zendel, B.R., Alain, C., 2009. Concurrent sound segregation is enhanced in musicians. J. Cogn. Neurosci. 21, 1488–1498.
Zendel, B.R., Tremblay, C.D., Belleville, S., Peretz, I., 2015. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. J. Cogn. Neurosci. 27, 1044–1059.

Supplemental Methods: Electrode-level responses

ERP electrode response analysis

For the purpose of data reduction, we collapsed the sensor data into five electrode clusters covering frontocentral areas of the scalp. The mean response of four adjacent electrodes within each cluster defined the five "super electrodes" across the head surface that included a left frontal (AF3, F5, F3, F1), right frontal (AF4, F2, F4, F6), left temporal (FC5, FC3, C5, C3), central (FCz, C1, Cz, C2), and right temporal (FC4, FC6, C4, C6) cluster. These clusters were selected as the auditory ERPs are maximal at the vertex and carry a prominent frontocentral scalp distribution (Picton et al., 1999; Woods, 1995). Within each electrode cluster, peak amplitude and latency were measured for the prominent deflections of the cortical response (P1, N1, P2, N2) as in the source space analysis. Analysis of the scalp data was restricted to the central electrode cluster near Cz where ERPs showed the most stereotyped morphology.

ERP laterality

We assessed laterality of the scalp-recorded waveforms at frontral electrode sites given the expected (dis-)engagement of linguistic brain regions with changes in speech SNR (Bidelman and Dexter, 2015). We extracted electrode recordings from two specific channels (F5 and FC6) in the left and right hemisphere (e.g., Herwig et al., 2003, see also http://www.brainm.com/software/pubs/dg/BA_10-20_ROI_Talairach/nearesteeg.htm). These channels were selected given their proximal location to Brodmann area 44/45 (Broca's area) and its right hemisphere homologue based on the Talairach brain atlas (Talairach and Tournoux, 1988). We then compared left and right hemisphere responses using amplitude and latency measures as described above. A laterality index (%) was computed as LI = 100*(R-L)/(R+L), where R and L are the response measures from left and right hemispheres, respectively (cf. King et al., 1999). Responses eliciting symmetric cortical activity produce a value of zero; asymmetric responses yield a positive or negative value for right or left hemisphere dominance, respectively. P2-N2 amplitudes were used as the dependent measure for all laterality tests as this was the most prominent signature of the electrode-level data at frontal electrode sites (see Fig. S1).

Supplemental Results

ERPs (sensor-space) waveforms are shown in Fig. S1 for the various electrode clustes. Quantitative analysis of electrode-level recordings revealed that the amplitudes of the earlier waves including the P1 [$F_{2, 22} = 4.97$, $p = 0.0165$] and N1 [$F_{2, 22} = 10.59$, $p = 0.0006$] were modulated by speech SNR. Whereas P1 decreased in magnitude with increasing noise, N1 increased in magnitude, becoming more negative at lower SNRs (Fig. S2A). Neither the P2 [$F_{2, 22} = 2.38$, $p = 0.1163$] nor N2 [$F_{2, 22} = 1.98$, $p = 0.16$] components showed appreciable amplitude changes with SNR.

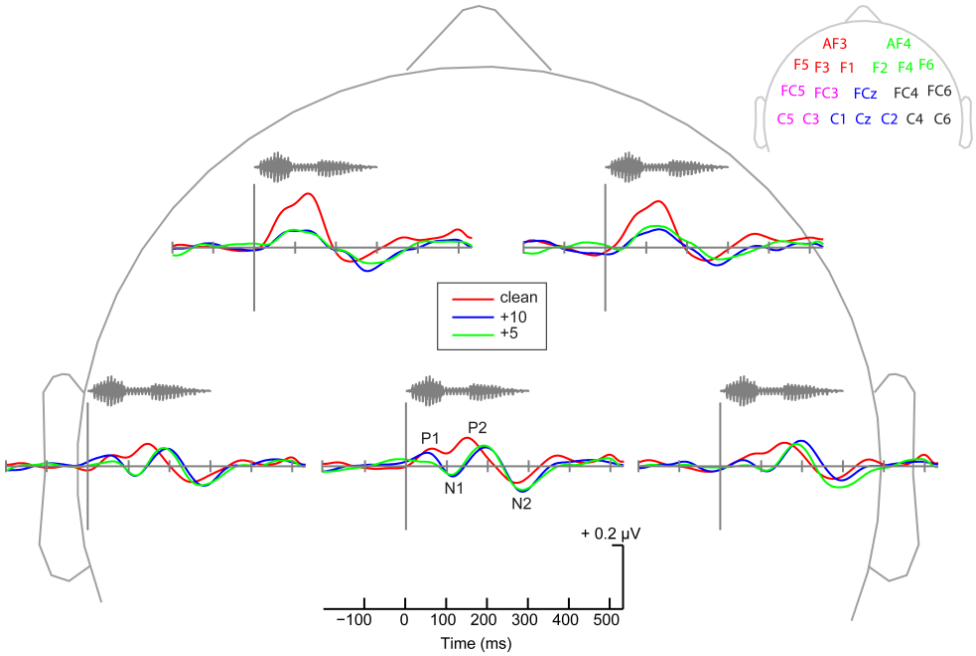

Figure S1: Distribution of the speech-evoked ERPs as a function of SNR in five electrode regions of interest around the scalp. Grey insets, time waveforms of the speech stimulus. Top right, electrode cluster definition. Prominent latency and amplitude modulations are observed at the vertex (central cluster), particularly in the N1 and later deflections.

ERP latency was similarly modulated by noise level (Fig. S2B). With the exception of N2, all component waves showed noise-related changes in latency, occurring later for poorer compared to more favorable SNRs [P1: $F_{2, 22} = 4.69$, $p = 0.02$; N1: $F_{2, 22} = 5.59$, $p = 0.011$; P2: $F_{2, 22} = 14.67$, $p < 0.001$; N2: $F_{2, 22} = 0.04$, $p = 0.96$].

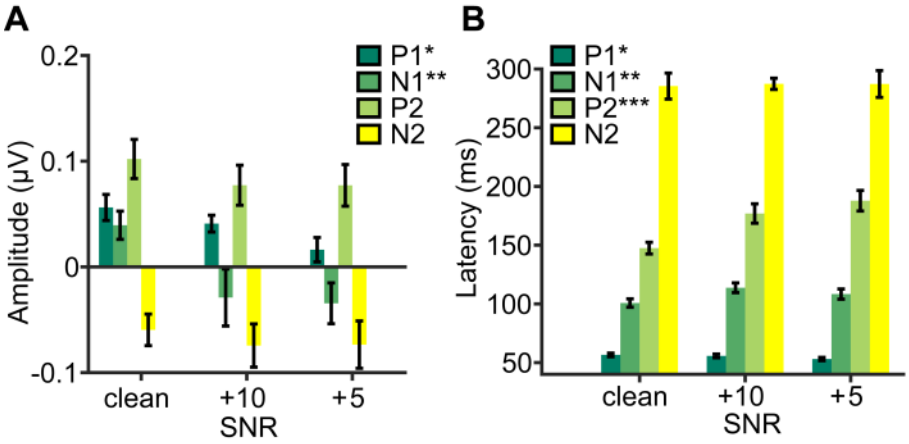

Figure S2: Amplitude (A) and latency (B) profiles for the speech-evoked ERPs with increasing noise. Prominent amplitude decrements and prolonged latencies are seen in the variance component waves at poorer speech SNRs. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Hemispheric laterality of scalp-recorded ERPs during SIN processing

A comparison of speech-ERPs recorded in right vs. left hemispheres is shown in Figure S3A-C as a function of speech SNR. Traces reflect neuroelectric activity recorded from electrodes proximal to Broca's area and its right hemisphere homologue (Herwig et al., 2003; Talairach and Tournoux, 1988). For clean speech, more robust responses are observed in the left hemisphere compared to the right (i.e., LH > RH). However, consist with scalp topographies (i.e., Fig. 3), noise-degraded speech produced a rightward asymmetry (i.e., RH > LH) in the distribution of the scalp potentials. These findings indicate a modal shift in hemispheric dominance in frontal brain regions with increasing noise level (i.e., decreasing SNR).

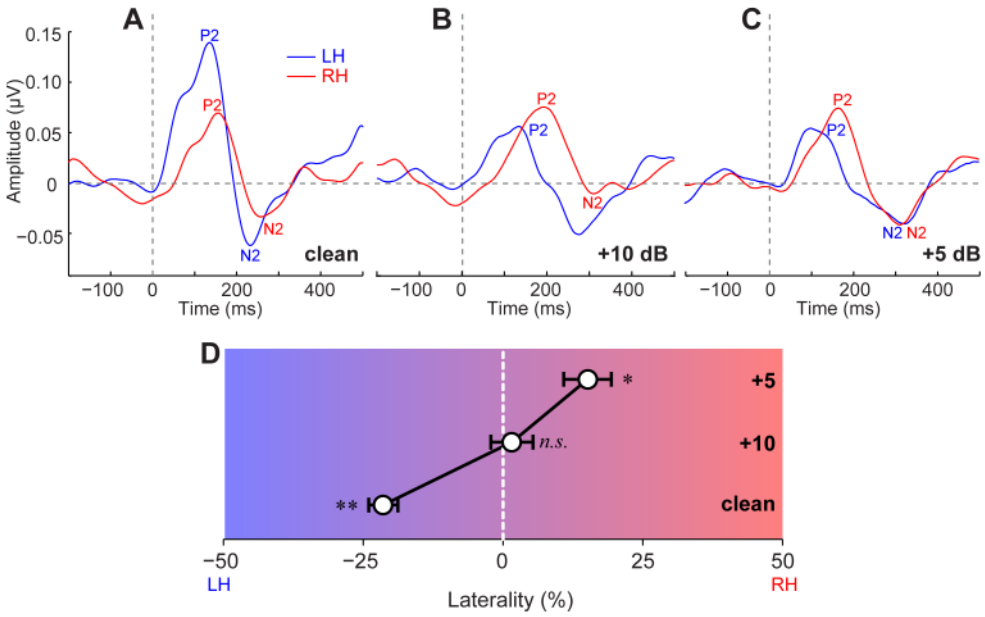

Figure S3: Hemispheric laterality for noise-degraded speech processing. (A-C) Hemispheric comparison of cortical ERPs recorded proximal to Broca's area (electrode F5) and its right hemisphere homologue (electrode FC6) for (A) clean and degraded speech at (B) +10 dB SNR and (C) +5 dB SNR. (D) Laterality of the cortical ERPs (P2-N2 magnitudes). A leftward laterality (i.e., LH > RH) is observed for clean speech. The functional asymmetry in auditory processing becomes more symmetric and then shifts rightward for progressively noisier speech.*$p < 0.05$, **$p < 0.01$

Functional laterality of the cortical ERPs during SIN processing is shown in Figure S3D. Note that $|LI| \geq 0.2$ is typically considered "lateralized" (Seghier, 2008). Quantitative analyses confirmed a shift in hemispheric dominance with increasing levels of noise. Whereas clean speech was strongly lateralized in the left hemisphere [t_{11} = -3.51, p= 0.0024], the introduction of noise produced a shift toward the right hemisphere; laterality diminished (became more bilateral) at +10 dB SNR [t_{11} = 0.21, p= 0.58] and began showing a rightward dominance at +5 dB SNR [t_{11} = 1.69, p= 0.05]. This was supported

by an ANOVA, which revealed a significant effect of SNR on ERP laterality [$F_{2, 22} = 12.36$, $p = 0.0003$]. Follow-up contrasts confirmed significant change in speech laterality with increasing noise [linear contrast of SNR: $t_{11} = 3.71$, $p = 0.0034$]. Together, these findings indicate that auditory cortical responses (at the scalp) are weighted more heavily in the left than right hemisphere when processing clean, unadulterated speech. Furthermore, the leftward response asymmetry becomes more bilateral and shifts progressively rightward during adverse speech listening.

References

Bidelman, G.M., Dexter, L., 2015. Bilinguals at the "cocktail party": Dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-noise recognition deficits. Brain and Language 143, 32-41.

Herwig, U., Satrapi, P., Schonfeldt-Lecuona, C., 2003. Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. Brain Topography 16, 95-99.

King, C., Nicol, T., McGee, T., Kraus, N., 1999. Thalamic asymmetry is related to acoustic signal complexity. Neuroscience Letters 267, 89–92.

Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiology and Neuro-Otology 4, 64-79.

Seghier, M.L., 2008. Laterality index in functional MRI: Methodological issues. Magnetic Resonance Imaging 26, 594-601.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system: an approach to cerebral imaging. Thieme Medical Publishers, New York.

Woods, D.L., 1995. The component structure of the N1 wave of the human auditory evoked potential. Electroencephalography and Clinical Neurophysiology 44, 102-109.