

Perceptual Demand Modulates Activation of Human Auditory Cortex in Response to Task-irrelevant Sounds

Merav Sabri, Colin Humphries, Matthew Verber, Jain Mangalathu, Anjali Desai, Jeffrey R. Binder, and Einat Liebenthal

Abstract

■ In the visual modality, perceptual demand on a goal-directed task has been shown to modulate the extent to which irrelevant information can be disregarded at a sensory-perceptual stage of processing. In the auditory modality, the effect of perceptual demand on neural representations of task-irrelevant sounds is unclear. We compared simultaneous ERPs and fMRI responses associated with task-irrelevant sounds across parametrically modulated perceptual task demands in a dichotic-listening paradigm. Participants performed a signal detection task in one ear

(Attend ear) while ignoring task-irrelevant syllable sounds in the other ear (Ignore ear). Results revealed modulation of syllable processing by auditory perceptual demand in an ROI in middle left superior temporal gyrus and in negative ERP activity 130–230 msec post stimulus onset. Increasing the perceptual demand in the Attend ear was associated with a reduced neural response in both fMRI and ERP to task-irrelevant sounds. These findings are in support of a selection model whereby ongoing perceptual demands modulate task-irrelevant sound processing in auditory cortex. ■

INTRODUCTION

The stage in the information processing stream at which task-irrelevant information can be disregarded has been the topic of longstanding debate in cognitive science between theorists advocating early selection (Treisman, 1969; Broadbent, 1958) and those advocating late selection (Duncan & Humphreys, 1992; Duncan, 1980; Norman, 1968; Deutsch & Deutsch, 1963). In early selection models, attention shuts down or attenuates processing of irrelevant information at an early sensory-perceptual stage of processing. In late selection models, attention acts only after incoming relevant and irrelevant information has been fully processed. The load model of attention combines aspects of both views and holds that the level of perceptual demand (load) required for processing task-relevant stimuli determines the extent to which irrelevant information can be disregarded (Lavie, 1995, 2005; Lavie, Hirst, de Fockert, & Viding, 2004; Lavie & Tsal, 1994). While relying on the idea of limited attentional resources (Duncan, Martens, & Ward, 1997), the load model predicts that high perceptual load depletes attentional resources resulting in reduced perception of distractors (early selection view), whereas under low perceptual load, unused resources are directed automatically toward processing of irrelevant distractors (late selection view). There is considerable behavioral evidence in support of the load model, at least in the visual

modality (Lavie, 1995, 2005, 2010; Lavie et al., 2004; Lavie & Tsal, 1994; cf. Benoni & Tsal, 2010). Importantly, neuroimaging studies have shown that perceptual demand modulates neural activity associated with irrelevant visual distractors (e.g., faces, moving dots, letters, flickering checkerboards) in the direction predicted by the model, namely smaller neural response in sensory-perceptual networks for distractors under high perceptual load and larger responses when perceptual load is low (Schwartz et al., 2005; Yi, Woodman, Widders, Marois, & Chun, 2004; Berman & Colby, 2002; O'Connor, Fukui, Pinski, & Kastner, 2002; Vuilleumier, Armony, Driver, & Dolan, 2001; Rees, Frith, & Lavie, 1997). Nevertheless, these findings can be explained with at least one alternative theory according to which the control of attention is based on an inhibition mechanism (as opposed to limited resources) that becomes stronger as attention activity for relevant stimuli is increased with task demands (LaBerge, 1995, 2002).

In the auditory modality, the extent to which task-irrelevant information is processed has been studied widely with behavioral measures starting with Cherry's classic dichotic listening experiments (Koch, Lawo, Fels, & Vorlander, 2011; Dark, Johnston, Myles-Worsley, & Farah, 1985; Johnston & Heinz, 1979; Moray, 1959; Cherry, 1953). The effect of perceptual demand in an auditory central task on sensory-perceptual processing of irrelevant sounds has not been studied systematically with neuroimaging. There is some evidence for greater processing of task-irrelevant sound features in auditory cortex when the demands of

an auditory task is higher (Sabri, Liebenthal, Waldron, Medler, & Binder, 2006), inconsistent with findings in the visual modality. One reason could be the lack of spatial separation between relevant and irrelevant information in this study. Facilitatory effects of high perceptual demand were observed in a visual Stroop task, where the target word and distractor color were contained within a single stimulus (Chen, 2003). In such paradigms, the greater attention channeled to task targets under high demand is also directed to the irrelevant information contained in them (Lavie, 2005). A recent dichotic listening study, whereby relevant and irrelevant information were presented to opposite ears, observed greater activity for the latter in auditory cortex as task demands decreased (Rinne, 2010). However, this effect was weak and did not reach significance, possibly due to relatively low statistical power ($n = 9$).

Here we investigated the extent to which sensory-perceptual processing of task-irrelevant sounds is modulated by the perceptual demand of a primary auditory task, in a dichotic listening paradigm, using simultaneous recordings of ERPs and fMRI. In the primary task (detection of tone in noise), signal-to-noise ratio (SNR) was modulated parametrically to create four perceptual load levels, while keeping the noise level constant. Task-relevant and -irrelevant information was spatially separated using dichotic presentation. To examine the effects of perceptual demand on task-irrelevant information, neural responses to ignored syllables were compared between the lowest and highest loads in the ERP and in localizer-defined speech-sensitive area in auditory cortex. To determine if the load manipulation was related linearly to the BOLD signal in auditory cortex and to the ERP elicited by the syllables, contrasts weighted by the four load levels were employed. Our findings corroborate and extend those in the visual modality, demonstrating reduced activity for

task-irrelevant sounds in sensory-perceptual auditory ROI under high compared with low perceptual demand. These findings clarify the mechanism by which the brain manages the processing of multiple sources of auditory information and provide support for a model involving selection at a sensory-perceptual processing stage as modulated by perceptual demand.

METHODS

Participants

Participants were 24 healthy adults (10 men, mean age = 24 years, $SD = 3$) with no history of neurological or hearing impairments and normal or corrected-to-normal visual acuity. The participants were native English speakers, and all were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Data from eight participants were excluded from ERP analysis (six due to noisy EEG and two due to equipment failure). Data from one participant were excluded from fMRI analysis (due to excessive motion artifact). Informed consent was obtained from each participant before the experiment, in accordance with the Medical College of Wisconsin Institutional Review Board.

Task Design and Procedure

The study employed an event-related design with individual trials blocked by condition and a dichotic listening paradigm. There were 10 simultaneous ERP/fMRI dichotic-listening runs, each divided into eight blocks of 51 sec/block. Each block was composed of seventeen 1.2-sec trials (Figure 1). Image acquisition (1.8 sec) followed immediately each trial. In the Attend ear, stimulation consisted of a white noise burst (Noise; 1.2 sec) with a 50-msec, 800-Hz

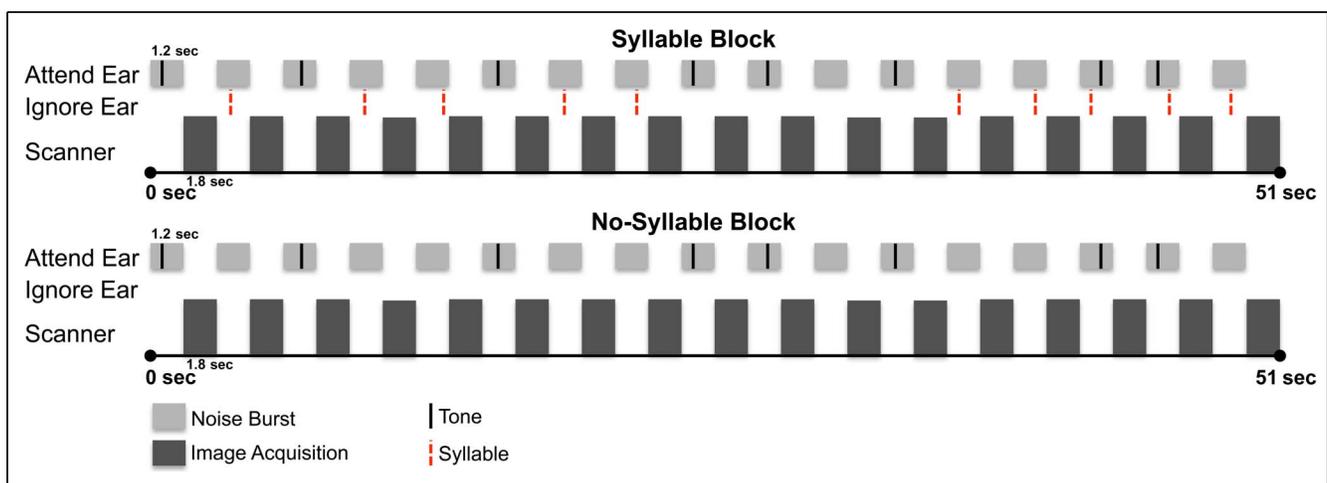


Figure 1. A schematic illustration of stimulus presentation in Syllable (top) and No-Syllable (bottom) blocks. Light gray bars represent the 1.2-sec noise bursts. Dark gray bars represent 1.8-sec of image acquisition. Black tick marks represent eight signal tones in noise bursts in the Attend stream. The additional nine noise bursts did not include a tone. Dashed red tick marks represent task-irrelevant syllables in the Ignore stream in Syllable blocks. The Load was fixed within each block. Attend and Ignore ear designation was fixed within a run.

signal tone (Tone; $p = .47$) embedded in eight of the trials. The Tone was presented at a random time ranging from 200 to 1000 msec after the beginning of the trial. The SNR between the Tone and Noise was modulated parametrically to create four perceptual demand/load conditions ranging from low to high (Load 1, Load 2, Load 3, Load 4). The Noise was presented at a fixed intensity (112 dB) with the amplitude of the Tone varying to produce the desired SNR (88, 89, 90, 91 dB). The SNR (Load) was fixed within each 51-sec block.

In the Ignore ear, half of the blocks included syllables. In Syllable blocks, 10 different task-irrelevant syllables (/ba/, /da/, /bi/, /di/, /bu/, /du/, /be/, /de/, /bo/, /do/), each 180 msec in duration, were presented to the Ignore ear at a random time ranging from 200 to 1000 msec after the beginning of the trial. The tones (in the Attend ear) and syllables (in the Ignore ear) were presented such that they did not overlap temporally. Within a single 1.2-sec trial, either a syllable (in the Ignore ear) or a Tone (in the Attend ear) was presented, except for two trials, which included both (ISI \geq 500 msec) in random order. In No-Syllable control blocks, speech sounds were not presented. Trials were randomized within each block. Eight blocks (4 Syllable, 4 No-Syllable) were delivered randomly within each run. The presentation order of the four load conditions was randomized with equal probability. Each block was followed by a 12-sec rest period. The ISI of the syllables was jittered exponentially between 3 and 15 sec. In the entire experiment, there were 100 ignored syllable events per load condition.

During the experiment, participants performed a signal detection task in the Attend ear and were instructed to ignore the irrelevant speech sounds presented to the other ear. Attend and Ignore ear designation was fixed within a run. Participants were instructed to press Button 1 upon detection of a tone and Button 2 when they did not hear a tone. They were told that approximately half of the noise bursts in a block included a tone and that some of them would be harder to detect. The ear of delivery for the signal detection task was equiprobable and randomized between the runs. A cross-hair was presented in the middle of the screen to assist in minimizing eye movement.

An event-related localizer run, designed to identify areas sensitive to speech stimuli, followed the 10 dichotic-listening runs. In the localizer run, participants discriminated between randomly presented 180-msec binaural tones and syllables by pressing Buttons 1 and 2, respectively. The syllables were identical to those used in the dichotic-listening runs. Tones were 10 logarithmically spaced sinewaves ranging from 200 to 4000 Hz. Stimulation consisted of randomly presented 40 syllable and 40 tone events, occurring during the 1.2 sec between image acquisitions. ISI was jittered exponentially between 3 and 9 sec (mean = \sim 5 sec).

The syllables were recorded from a male native English speaker and normalized according to loudness. Sounds

were delivered through MRI-compatible STAX SR-003 electrostatic ear inserts (STAX, Saitama Prefecture, Japan). The visual fixation stimulus was projected through an Epson LCD video projector onto an angled mirror located just above the eyes. Stimulus delivery was controlled by a personal computer running Presentation software (Neuro-behavioral Systems, Inc., Albany, CA).

fMRI Acquisition and Analysis

Images were acquired on a 3T GE Excite scanner (GE Medical Systems, Milwaukee, WI). Functional data consisted of T2*-weighted, gradient-echo, echo-planar images (echo time = 20 msec, flip angle = 77° , acquisition time = 1.8 sec, delay = 1.2 sec), obtained using clustered acquisition at 3-sec intervals. Sound stimulation (Noise alone, Noise and Tone, or Syllable) was presented during the 1.2-sec period between image acquisitions to avoid perceptual masking by the acoustic noise of the scanner. Functional images were composed of 35 axially oriented 3 mm slices with a 0.5-mm interslice gap covering the whole brain, with field of view = 192 mm and 64×64 matrix, resulting in $3.0 \times 3.0 \times 3.5$ voxel dimensions. A total of 1720 images were acquired across the 10 dichotic-listening runs (172 per run). A total of 168 images were acquired in the localizer run. High-resolution anatomical images of the entire brain were obtained using a 3-D spoiled gradient-echo sequence (SPGR) as a set of 130 contiguous axial slices with $0.938 \times 0.938 \times 1.0$ mm voxel dimensions.

Image analysis was conducted using the AFNI software package (Cox, 1996). Within-subject analysis consisted of spatial registration to minimize motion artifacts (Cox & Jesmanowicz, 1999) and coregistration of functional and anatomy images (Saad et al., 2009). In the dichotic-listening runs, analyses focused on task-irrelevant syllables. Voxel-wise multiple linear regression was applied to individual time series, with reference functions separately representing the occurrence of a syllable, a tone in the Syllable and No-Syllable blocks, or syllable and tone in the four load conditions. Another regressor was added to code Noise alone trials. The shape and magnitude of the hemodynamic response (HRF) were estimated using the program 3dDeconvolve. Coefficient maps were generated for Syllables in each load condition representing the lags of the HRF. The individual coefficient maps were projected into standard stereotaxic space (Talairach & Tournoux, 1988) by linear resampling and then smoothed with a Gaussian kernel of 6 mm FWHM.

ROI Analysis

The localizer run was analyzed in a similar fashion. The reference functions in the multiple regression represented the occurrence of a syllable or a tone. A general linear test

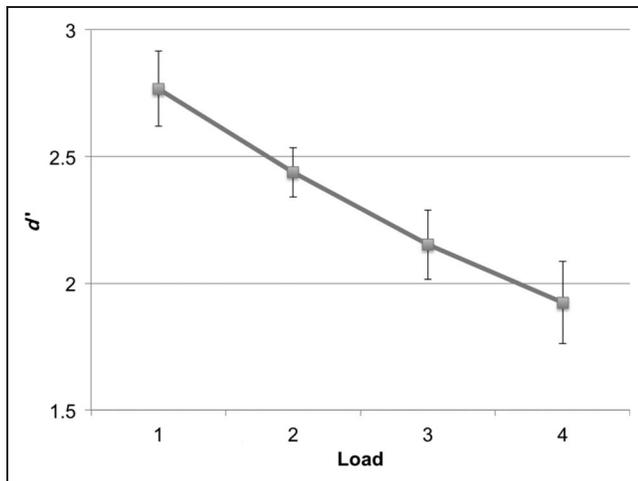


Figure 2. Behavioral performance (d') in the primary signal detection task, in each load condition, in Syllable blocks. Error bars indicate *SEM*.

between syllables and tones was conducted at the response peak to obtain regions sensitive specifically to speech sounds. Group maps were created using a random-effects analysis. The group maps were thresholded at a voxel-wise $p < .01$ and corrected for multiple comparisons by removing clusters smaller than 1008 μl , resulting in a corrected map-wise two-tailed $\alpha = .05$. This cluster threshold was determined through Monte Carlo simulations that provide the chance probability of spatially contiguous voxels exceeding the voxel-wise p threshold.

An ROI analysis was carried out within speech-sensitive area in auditory cortex. An ROI in middle left superior temporal gyrus (STG) was identified based on the localizer. The average BOLD signal in the identified ROI was extracted for the task-irrelevant syllables in each load condition at the peak height of the HRF, for each participant, and subjected to a paired t test between the two extreme loads. In addition, a test for linear trend of the loads ($1 > 2 > 3 > 4$) was performed on the mean signals using a repeated-measures ANOVA with a weighted contrast vector.

ERP Acquisition and Analysis

Sixty-four-channel EEG activity was acquired using the Maglink system (Neuroscan, Inc.) in a continuous mode and the Quik-Cap electrode positioning system (Neuroscan, Inc.). Activity was recorded at full bandwidth and digitally sampled at 500 Hz per channel. Electrode sites conformed to the International 10–20 System with CPz serving as the reference. Vertical eye movements and electrocardiogram were each monitored with bipolar recordings. Interelectrode resistance was kept below 5 k Ω .

EEG analysis was conducted using the Scan 4.3 software package (Compumedics Neuroscan), focusing on task-irrelevant syllables. Initial within-subject analysis consisted of bandpass filtering at 0.1–30 Hz, ballistocardiogram artifact removal, creating epochs of -100 to $+450$ msec from each sound onset, baseline-correction

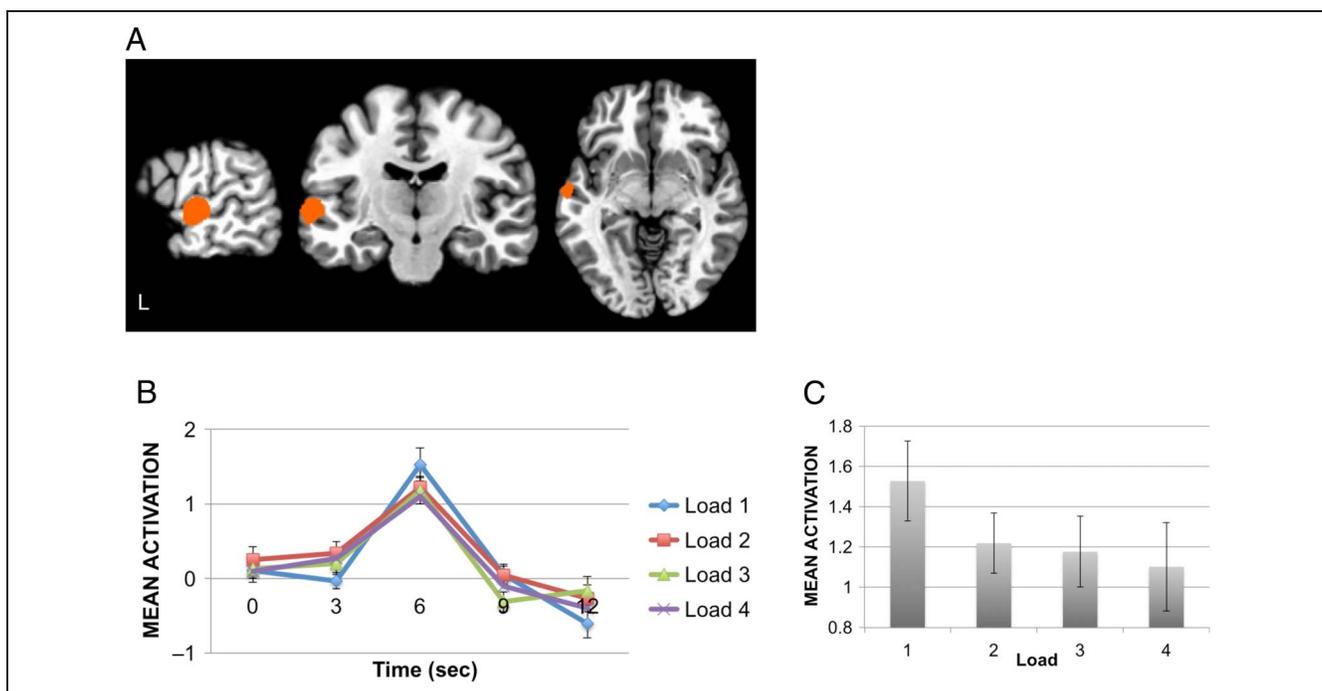


Figure 3. (A) The syllable-sensitive ROI as identified in the contrast Syllables $>$ Tones in the localizer run. (B) fMRI activation in the syllable-sensitive ROI by irrelevant syllables as a function of perceptual load. (C) fMRI activation in the syllable-sensitive ROI as a function of perceptual load at the HRF peak. Error bars indicate *SEM*.

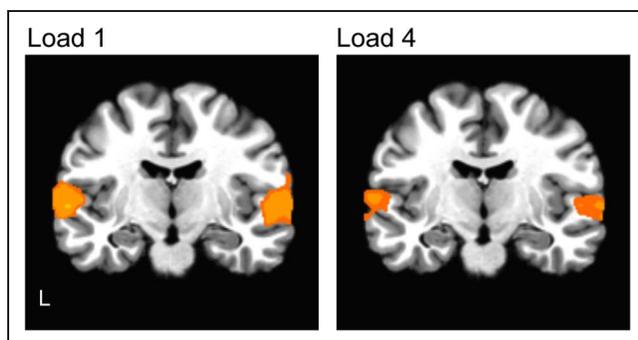


Figure 4. Whole-brain analyses: Statistical parametric maps of irrelevant syllables activation ($p < .05$, corrected) in Load 1 and Load 4 against baseline.

of each epoch by removing the mean voltage value of the whole sweep, and rejection of epochs with voltage values exceeding $\pm 150 \mu\text{V}$. The remaining epochs were then averaged according to each load condition. Each waveform was baseline corrected by subtracting the mean voltage of the prestimulus period from each point in the post stimulus interval. Grand-averaged waveforms were computed for syllable events in the four load conditions. The resulting waveforms were digitally rereferenced to the mastoids. Group level analyses were performed using

MATLAB (MathWorks, Inc., Natick, MA) and STATISTICA (StatSoft, Inc., Tulsa, OK). Mean amplitudes were extracted for each participant and averaged across 16 frontal electrodes (F7, F8, AF7, AF8, F5, F6, F3, F4, AF3, AF4, FP1, FP2, F1, F2, Fz, FPz) in the 130–230 msec time window in each condition and subjected to a paired t test between the two extreme loads and a repeated-measures ANOVA with a weighted contrast vector to test for linear trend of the loads.

RESULTS

Behavioral Performance

The d' ($z[\text{hit}] - z[\text{false alarm}]$) measure of perceptual sensitivity was calculated for each load. Signal detection performance in Syllable blocks varied by load [$F(3, 69) = 20.264, p < .001$], with a linear decrease in d' as perceptual load increased [$F(1, 23) = 63.98, p < .001$] (Figure 2). An ANOVA with Load (Loads 1, 2, 3, 4) and Block Type (Syllable, No-Syllable) as repeated-measures revealed main effect of Load [$F(3, 69) = 26.374, p < .001$]. The effect of Block Type and the Interaction was not significant [$F(1, 23) = .158, p = .69; F(3, 69) = .615, p = .61$], confirming no predictive relationship between relevant and irrelevant sound delivery.

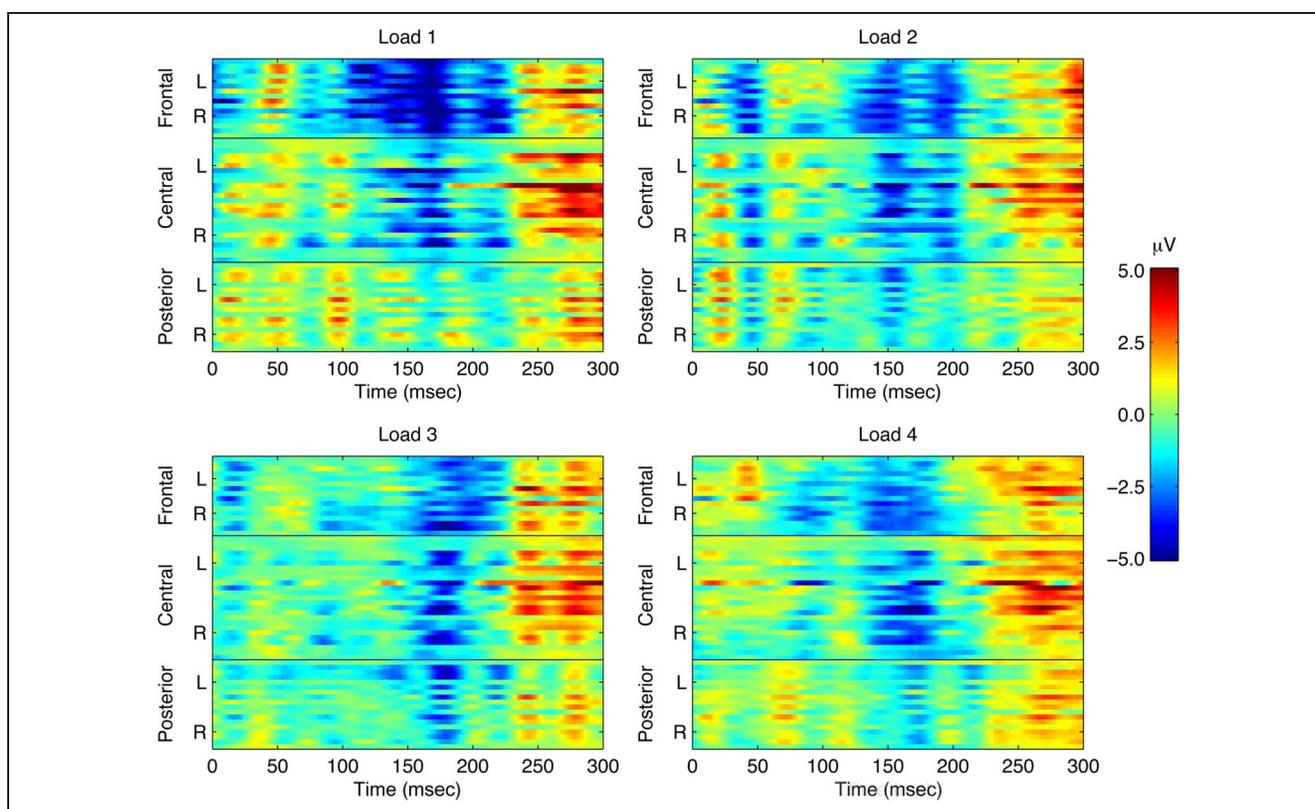


Figure 5. Spatio-temporal maps from 60 electrodes: Grand-averaged ERPs of irrelevant syllables at each electrode as a function of load. The y axis represents the frontal, central, and posterior electrodes. Each group of electrodes (frontal, central, posterior) is arranged top to bottom according to their lateral position from left (L) to right (R) with the midline electrode in the middle. The color scale represents the amplitude in μV .

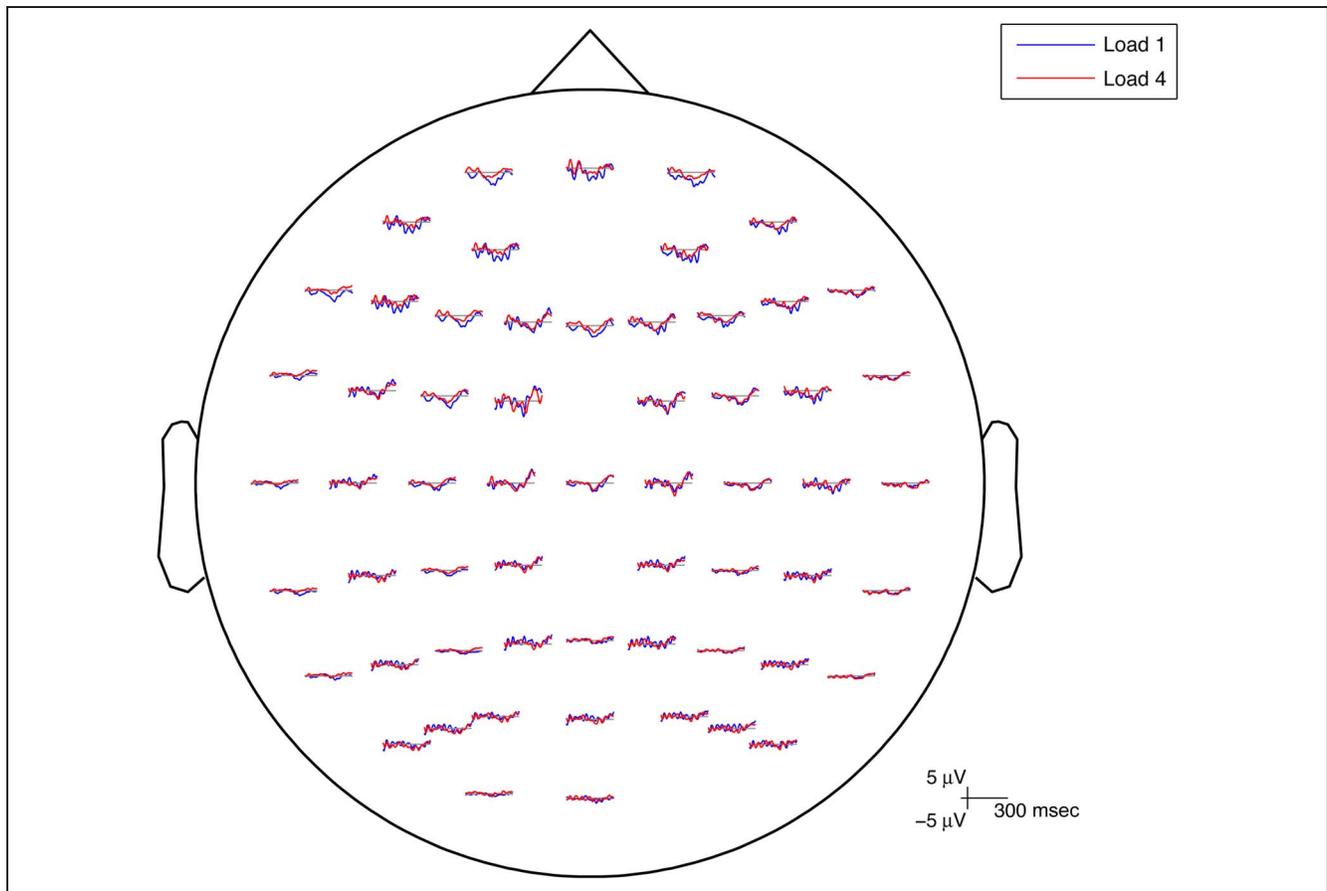


Figure 6. Group-averaged ERP waveforms superimposed for irrelevant syllables in Load 1 and Load 4. Electrode sites conformed to the International 10–20 System.

fMRI

Localizer: Syllables > Tones

The focus of the current study is on the effect of perceptual demand on sensory-perceptual processing of task-irrelevant speech sounds. To identify neural regions specifically related to speech processing, we contrasted Syllable and Tone activation in the localizer run. The contrast Syllables–Tones is presented in Figure 3A. Greater significant activation for syllables over tones was observed in one cluster ($x = -59, y = -13, z = -2$; threshold $z > 2.57$, cluster-corrected $\alpha = .05$, 1008 μl) that included the anterior-middle portion of the left STG and STS and the anterior-lateral portion of Heschl’s gyrus (HG). No other statistically significant activation clusters were observed. There were no significant areas of activation for tones over syllables.

Load Effects on Irrelevant Syllable Processing in Speech-sensitive Auditory Region Defined in the Localizer

The Syllable over Tone cluster identified in the localizer run was used as an ROI in the dichotic-listening runs. The average BOLD signal in the left STG ROI, as a function of perceptual load, is depicted in Figure 3B. Mean activation

at the HRF peak (6 sec; Figure 3C) was significantly different between the low load (Load 1) where the activation was strongest and the high load (Load 4) where the activation was lowest [$t(22) = 2.2233, p = .036$]. As the level of load increased, the BOLD signal for irrelevant syllables decreased, as indicated by a linear trend [$F(1, 22) = 4.49, p = .04$].

Whole-Brain Analyses

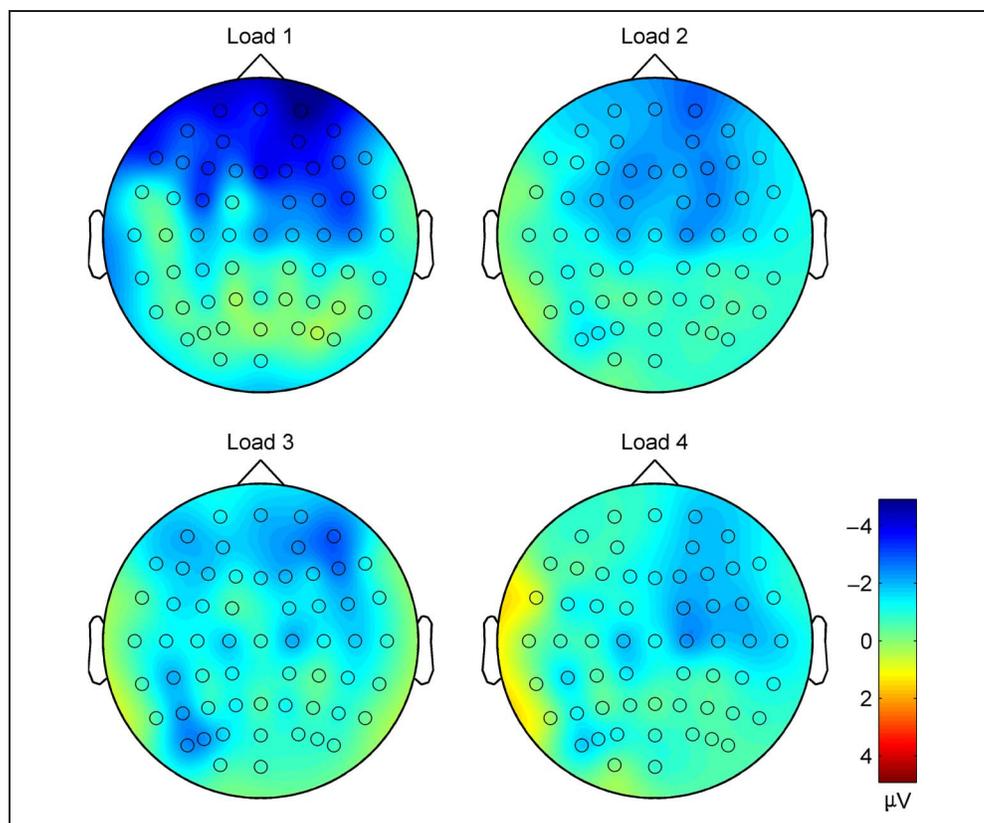
A whole-brain analysis was performed to examine whether there were differential load activations beyond the defined ROI. There were no significant differences in activations across the load conditions at a corrected whole-brain threshold (threshold $z > 1.96$, cluster correction $\alpha = .05$, 5040 μl). The extent of activation in auditory cortex for Load 1 and Load 4 against baseline is depicted in Figure 4 (threshold $z > 3.29$, cluster correction $\alpha = .05$, 347 μl).

Event-related Potentials

Load Effects on Irrelevant Syllable Processing

A fronto-central negativity was observed in the N1 time window in response to irrelevant syllables in all load

Figure 7. Mean scalp distribution for irrelevant syllables in the 130–230 msec time window, in the four load conditions. The color scale represents the amplitude in μV .



conditions (Figures 5 and 6). Differences across load conditions for irrelevant syllables were observed approximately 130–230 msec after stimulus onset, predominantly in frontal electrodes (Figure 7). This effect was quantified by computing the mean amplitude in this time range on frontal electrodes, for each perceptual load and for each participant (Figure 8). The mean negativity was significantly higher in amplitude in the low load (Load 1) compared with the high load (Load 4) condition [$t(15) = 3.49, p =$

.003]. The test for linear trend of the loads was also significant [$F(1, 15) = 10.77, p = .0005$].

DISCUSSION

The extent of processing of task-irrelevant syllable sounds was assessed using fMRI and ERP measures of brain activity. Modulation of syllable processing by auditory perceptual demands was observed in an ROI encompassing primarily the middle portion of the left STG, and in a negative ERP with onset at 130 msec, the N1 component. High perceptual load, as determined in a psychophysical auditory task, was associated with a reduced neural response in the fMRI and ERP for task-irrelevant syllables, whereas a low load level produced the greatest responses. A linear trend was observed in the fMRI and ERP data, demonstrating increased neural response for task-irrelevant syllables as task demands decreased.

The N1 component is a potential elicited in response to auditory stimulation and associated with sensory processing (Näätänen & Picton, 1987). The amplitude of N1 increases with attention (Woldorff & Hillyard, 1991; Näätänen, 1990; Sams, Aulanko, Aaltonen, & Näätänen, 1990; Hansen, Dickstein, Berka, & Hillyard, 1983; Hillyard, Hink, Schwent, & Picton, 1973), suggesting that this component is susceptible to top-down influences. In the current study, the largest negativity in the N1 time

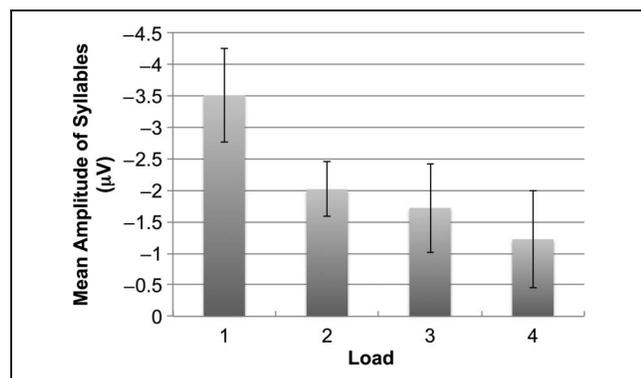


Figure 8. Mean amplitude on frontal electrodes (“frontal group” in Figure 5) for irrelevant syllables at 130–230 msec poststimulus, as a function of perceptual load. Error bars indicate SEM.

range was observed under the lowest perceptual load, in line with the fMRI results, suggesting greater sensory processing of task-irrelevant complex sounds in that condition. The sources of the N1 were estimated previously to include parts of HG and anterior-middle and posterior STG/planum temporale depending on sound characteristics and dipole estimate methods (Ahveninen et al., 2011; Jääskeläinen et al., 2004; Picton et al., 1999; Fujiwara, Nagamine, Imai, Tanaka, & Shibasaki, 1998; Scherg, Vasjar, & Picton, 1989; Scherg & Von Cramon, 1986). Portions of the left anterior-middle STG/anterior HG region were encompassed in the fMRI ROI. It is likely that the effects of perceptual demands on N1 are reflected to some extent in the differential BOLD signal observed in this ROI.

A converging body of evidence from neuroimaging studies suggests that the middle STG/STS, specifically in the left hemisphere, plays a prominent role in phonemic perception and prelexical processing (DeWitt & Rauschecker, 2012; Leaver & Rauschecker, 2010; Liebenthal et al., 2010; Specht, Osnes, & Hugdahl, 2009; Liebenthal, Binder, Spitzer, Possing, & Medler, 2005; Davis & Johnsrude, 2003; Specht & Reul, 2003). Attention to phonetic material enhances activation in this region (Ahveninen et al., 2011; Woods, Herron, Kang, Cate, & Yund, 2011; Woods & Alain, 2009). The posterior part of STG, planum temporale (not included here as an ROI), has been implicated in processing of spectrally and temporally complex sounds, independent of phonetic contents (Specht & Reul, 2003; Jancke, Wustenberg, Scheich, & Kaplan Layer, 2002; Binder, Frost, Hammeke, Rao, & Cox, 1996). The linear trend observed between load level and BOLD signal in the middle STG region suggests that processing of unattended sounds is reduced as demands increase (i.e., successful selection). At lower loads, irrelevant sounds are processed regardless of task instructions possibly due to availability of attentional resources or capacity (Lavie, 1995, 2005, 2010; Duncan et al., 1997; Lavie & Tsal, 1994) or reduced inhibition (LaBerge, 1995, 2002).

Our findings are consistent with imaging studies of perceptual demand manipulations in attention paradigms in the visual modality (Schwartz et al., 2005; Yi et al., 2004; Berman & Colby, 2002; O'Connor et al., 2002; Vuilleumier et al., 2001; Rees et al., 1997). The study was not designed to resolve the controversy between competing theories in accounting for the effects of perceptual demand, namely limited resources versus inhibition. According to the load theory, in conditions of high perceptual load, the processing of irrelevant information is gated at an early sensory-perceptual stage due to limited perceptual resources, in line with early selection accounts (Treisman, 1969; Broadbent, 1958). In conditions of low perceptual load, however, available resources are thought to automatically “spill over” toward processing of irrelevant information until available resources are exhausted (Lavie et al., 2004; Lavie, 1995; Lavie & Tsal, 1994), requiring an additional late selection mechanism (Lavie, 2005, 2010; Yi et al., 2004). According to the inhibition account, inhibition weakens as attention

for targets decreases due to low perceptual demands, resulting in greater processing of task-irrelevant information.

The pattern of results reported here might be specific to the type of task employed (perceptual). It has been demonstrated in the visual modality that cognitive control demand (e.g., working memory task) modulates processing of task-irrelevant stimuli in the opposite direction than perceptual demand (Lavie, 2005; Lavie & De Fockert, 2005; Yi et al., 2004; de Fockert, Rees, Frith, & Lavie, 2001). High load on cognitive control was associated with failure of selection. Future studies are needed to investigate the effects of cognitive control demand on processing task-irrelevant information in the auditory modality.

Acknowledgments

We thank Suzanne Pendl for assistance with data collection, Doug Ward for statistical assistance, and two anonymous reviewers for their valuable comments and suggestions. This work was supported by the National Institute on Deafness and Other Communication Disorders (R03 DC008399; R01 DC006287) and the Clinical and Translational Science Award (CTSA) program of the National Center for Research Resources (UL1RR031973).

Reprint requests should be sent to Dr. Merav Sabri, Department of Neurology, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, or via e-mail: msabri@mcw.edu.

REFERENCES

- Ahveninen, J., Hämäläinen, M., Jääskeläinen, I. P., Ahlfors, S. P., Huang, S., Lin, F.-H., et al. (2011). Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proceedings of the National Academy of Sciences*, *108*, 4182–4187.
- Benoni, H., & Tsal, Y. (2010). Where have we gone wrong? Perceptual load does not affect selective attention. *Vision Research*, *50*, 1292–1298.
- Berman, R. A., & Colby, C. L. (2002). Auditory and visual attention modulate motion processing in area MT+. *Cognitive Brain Research*, *14*, 64–74.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*, 1239–1247.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.
- Chen, Z. (2003). Attentional focus, processing load, and Stroop interference. *Perception and Psychophysics*, *65*, 888–900.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Cox, R. W., & Jesmanowicz, A. (1999). Real-time 3D image registration of functional MRI. *Magnetic Resonance in Medicine*, *42*, 1014–1018.
- Dark, V. J., Johnston, W. A., Myles-Worsley, M., & Farah, M. J. (1985). Levels of selection and capacity limits. *Journal of Experimental Psychology: General*, *114*, 472–497.
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*, 3423–3431.

- de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*, 1803–1806.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*, 80–90.
- DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, E505–E514.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*, 272–300.
- Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 578–588.
- Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, *387*, 808–810.
- Fujiwara, N., Nagamine, T., Imai, M., Tanaka, T., & Shibasaki, H. (1998). Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. *Cognitive Brain Research*, *7*, 99–109.
- Hansen, J. C., Dickstein, P. W., Berka, C., & Hillyard, S. A. (1983). Event-related potentials during selective attention to speech sounds. *Biological Psychology*, *16*, 211–224.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 6809–6814.
- Jancke, L., Wustenberg, T., Scheich, H., & Kaplan Layer, J. (2002). Phonetic perception and the temporal cortex. *Neuroimage*, *15*, 733–746.
- Johnston, W. A., & Heinz, S. P. (1979). Depth of nontarget processing in an attention task. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 168–175.
- Koch, I., Lawo, V., Fels, J., & Vorlander, M. (2011). Switching in the cocktail party: Exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1140–1147.
- LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- LaBerge, D. (2002). Attentional control: Brief and prolonged. *Psychological Research*, *66*, 220–233.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 451–468.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75–82.
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*, 143–148.
- Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*, 669–674.
- Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339–354.
- Lavie, N., & Tsai, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics*, *56*, 183–197.
- Leaver, A. M., & Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: Effects of acoustic features and auditory object category. *The Journal of Neuroscience*, *30*, 7604–7612.
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex*, *15*, 1621–1631.
- Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., & Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. *Cerebral Cortex*, *20*, 2958–2970.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *The Quarterly Journal of Experimental Psychology*, *11*, 56–60.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*, 201–288.
- Näätänen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sounds: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, *75*, 522–536.
- O'Connor, D. H., Fukui, M. M., Pinsk, M. A., & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience*, *5*, 1203–1209.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., et al. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology and Neuro-Otology*, *4*, 64–79.
- Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, *278*, 1616–1619.
- Rinne, T. (2010). Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *The Open Neuroimaging Journal*, *4*, 187–193.
- Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., & Cox, R. W. (2009). A new method for improving functional-structural MRI alignment using local Pearson correlation. *Neuroimage*, *44*, 839–848.
- Sabri, M., Liebenthal, E., Waldron, E. J., Medler, D. A., & Binder, J. R. (2006). Attentional modulation in the detection of irrelevant deviance: A simultaneous ERP/fMRI study. *Journal of Cognitive Neuroscience*, *18*, 689–700.
- Sams, M., Aulanko, R., Aaltonen, O., & Näätänen, R. (1990). Event-related potentials to infrequent changes in synthesized phonetic stimuli. *Journal of Cognitive Neuroscience*, *2*, 344–357.
- Scherg, M., Vasjar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience*, *1*, 336–355.
- Scherg, M., & Von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*, *65*, 344–360.
- Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & Driver, J. (2005). Attentional load and sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex*, *15*, 770–786.
- Specht, K., Osnes, B., & Hugdahl, K. (2009). Detection of differential speech-specific processes in the temporal lobe using fMRI and a dynamic “sound morphing” technique. *Human Brain Mapping*, *30*, 3436–3444.
- Specht, K., & Reul, J. (2003). Functional segregation of the temporal lobes into highly differentiated subsystems for

- auditory perception: An auditory rapid event-related fMRI-task. *Neuroimage*, *20*, 1944–1954.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*, 282–299.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron*, *30*, 829–841.
- Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*, 170–191.
- Woods, D. L., & Alain, C. (2009). Functional imaging of human auditory cortex. *Current Opinion in Otolaryngology & Head & Neck Surgery*, *17*, 407–411.
- Woods, D. L., Herron, T., Kang, X., Cate, A. D., & Yund, E. W. (2011). Phonological processing in human auditory cortical fields. *Frontiers in Human Neuroscience*, *5*, 1–15.
- Yi, D. J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*, 992–996.