

Published in final edited form as:

J Cogn Neurosci. 2013 September ; 25(9): 1553–1562. doi:10.1162/jocn_a_00416.

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

Department of Neurology, Medical College of Wisconsin, 8701 Watertown Plank Rd, Milwaukee, WI, USA

Abstract

In the visual modality, perceptual demand on a goal-directed task have 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 **event-related potentials** (ERP) and functional magnetic resonance imaging (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 a region of interest in middle left superior temporal gyrus and in negative ERP activity 130–230 ms 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**.

Keywords

Attention; Auditory; Perceptual demand/load; fMRI; ERP

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 (Broadbent, 1958; Treisman, 1969) and those advocating late selection (Deutsch and Deutsch, 1963; Duncan, 1980; Duncan and Humphreys, 1992; Norman, 1968). 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 et al., 2004; Lavie and Tsai, 1994). While relying on the idea of limited attentional resources (Duncan et al., 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 and Tsal, 1994) (cf. Benoni and 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 (Berman and Colby, 2002; O'Connor et al., 2002; Rees et al., 1997; Schwartz et al., 2005; Vuilleumier et al., 2001; Yi et al., 2004). 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 (Cherry, 1953; Dark et al., 1985; Johnston and Heinz, 1979; Koch et al., 2011; Moray, 1959). 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 et al., 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 event-related potentials (ERPs) and functional imaging (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 blood oxygen level dependent (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 region of interest under high compared to 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.

Material and Methods

Subjects

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 8 subjects were excluded from ERP analysis (6 due to noisy EEG, 2 due to equipment failure). Data from 1 subject were excluded from fMRI analysis (due to excessive motion artifact). Informed consent was obtained from each participant prior to 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 ten simultaneous ERP/fMRI dichotic-listening runs, each divided into eight blocks of 51-s/block. Each block was composed of seventeen 1.2-s trials (Figure 1). Image acquisition (1.8 s) followed immediately each trial. In the Attend ear, stimulation consisted of a white noise burst (Noise; 1.2 s) with a 50-ms, 800-Hz signal tone (Tone; $p=0.47$) embedded in eight of the trials. The Tone was presented at a random time ranging from 200 to 1000 ms 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-s 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 ms in duration, were presented to the Ignore ear at a random time ranging from 200 to 1000 ms 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-s trial, either a syllable (in the Ignore ear) or a Tone (in the Attend ear) was presented, except for two trials, which included both ($ISI \geq 500$ ms) 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-s rest period. The inter-stimulus interval of the syllables was jittered exponentially between 3 and 15 s. 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-ms 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 ten 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 s between image acquisitions. ISI was jittered exponentially between 3 and 9 s (mean = ~5 s).

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 (Neurobehavioral 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 ms, flip angle = 77°, acquisition time = 1.8 s, delay = 1.2 s), obtained using clustered acquisition at 3-s intervals. Sound stimulation (Noise alone, Noise and Tone, or Syllable) was presented during the 1.2 s 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 FOV = 192 mm and 64 × 64 matrix, resulting in 3.0 × 3.0 × 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 × 0.938 × 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 and Jesmanowicz, 1999) and co-registration 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 and Tournoux, 1988) by linear re-sampling, and then smoothed with a Gaussian kernel of 6 mm FWHM.

Region of interest (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 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 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 subject, 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. Inter-electrode resistance was kept below 5k 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 ms from each sound onset, baseline-correction 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 pre-stimulus period from each point in the post stimulus interval. Grand-average waveforms were computed for syllable events in the four load conditions. The resulting waveforms were digitally re-referenced to the mastoids. Group level analyses were performed using MATLAB (MathWorks, Inc.) and STATISTICA (StatSoft, Inc.). Mean amplitudes were extracted for each subject 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 ms 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 (Load 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 were not significant [$F(1, 23)=.158$, $p=.69$; $F(3, 69)=.615$, $p=.61$], confirming no predictive relationship between relevant and irrelevant sound delivery.

Functional Magnetic Resonance Imaging

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 superior temporal gyrus and sulcus (STG, 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 s; 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 conditions (Figures 5, 6). Differences across load conditions for irrelevant syllables were observed approximately 130–230 ms 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 subject (Figure 8). The mean negativity was significantly higher in amplitude in the low load (Load 1) compared to 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 a region of interest encompassing primarily the middle portion of the left STG, and in a negative event-related potential with onset at 130 ms, 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 and Picton, 1987). The amplitude of N1 increases with attention (Hansen et al., 1983; Hillyard et al., 1973; Näätänen, 1990; Sams et al., 1990; Woldorff and Hillyard, 1991), suggesting that this component is susceptible to top-down influences. In the current study, the largest negativity in the N1 time-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; Fujiwara et al., 1998; Jääskeläinen et al., 2004; Picton et al., 1999; Scherg et al., 1989; Scherg and 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 (Davis and Johnsrude, 2003; DeWitt and Rauschecker, 2012; Leaver and Rauschecker, 2010; Liebenthal et al., 2005; Liebenthal et al., 2010; Specht et al., 2009; Specht and Reul, 2003). Attention to phonetic material enhances activation in this region (Ahveninen et al., 2011; Woods and Alain, 2009; Woods et al., 2011). 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 (Binder et al., 1996; Jancke et al., 2002; Specht and Reul, 2003). 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 (Duncan et al., 1997; Lavie, 1995, 2005, 2010; Lavie and 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 (Berman and Colby, 2002; O'Connor et al., 2002; Rees et al., 1997; Schwartz et al., 2005; Vuilleumier et al., 2001; Yi et al., 2004). 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 (Broadbent, 1958; Treisman, 1969). 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, 1995; Lavie et al., 2004; Lavie and 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 (de Fockert et al., 2001; Lavie, 2005; Lavie and De Fockert, 2005; Yi et al., 2004). 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].

References

- Ahveninen J, Hämäläinen M, Jääskeläinen IP, Ahlfors SP, Huang S, Lin FH, Raji T, Sams M, Vasios CE, Belliveau JW. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proceedings of the National Academy of Sciences*. 2011; 108:4182–4187.

- Benoni H, Tsal Y. Where have we gone wrong? Perceptual load does not affect selective attention. *Vision Research*. 2010; 50:1292–1298. [PubMed: 20430048]
- Berman RA, Colby CL. Auditory and visual attention modulate motion processing in area MT+ *Brain Research Cognitive Brain Research*. 2002; 14:64–74. [PubMed: 12063130]
- Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW. Function of the left planum temporale in auditory and linguistic processing. *Brain*. 1996; 119:1239–1247. [PubMed: 8813286]
- Broadbent, DE. *Perception and Communication*. London: Pergamon Press; 1958.
- Chen Z. Attentional focus, processing load, and Stroop interference. *Perception and Psychophysics*. 2003; 65:888–900. [PubMed: 14528898]
- Cherry EC. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*. 1953; 25:975–979.
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 1996; 29:162–173. [PubMed: 8812068]
- Cox RW, Jesmanowicz A. Real-time 3D image registration of functional MRI. *Magnetic Resonance in Medicine*. 1999; 42:1014–1018. [PubMed: 10571921]
- Dark VJ, Johnston WA, Myles-Worsley M, Farah MJ. Levels of selection and capacity limits. *Journal of experimental psychology General*. 1985; 114:472–497. [PubMed: 2934499]
- Davis MH, Johnsrude IS. Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*. 2003; 23:3423–3431. [PubMed: 12716950]
- de Fockert JW, Rees G, Frith CD, Lavie N. The role of working memory in visual selective attention. *Science*. 2001; 291:1803–1806. [PubMed: 11230699]
- Deutsch JA, Deutsch D. Attention: Some theoretical considerations. *Psychological Review*. 1963; 70:80–90. [PubMed: 14027390]
- DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:E505–514. [PubMed: 22308358]
- Duncan J. The locus of interference in the perception of simultaneous stimuli. *Psychological Review*. 1980; 87:272–300. [PubMed: 7384344]
- Duncan J, Humphreys G. Beyond the search surface: visual search and attentional engagement. *Journal of experimental psychology Human perception and performance*. 1992; 18:578–588. [PubMed: 1593236]
- Duncan J, Martens S, Ward R. Restricted attentional capacity within but not between sensory modalities. *Nature*. 1997; 387:808–810. [PubMed: 9194561]
- Fujiwara N, Nagamine T, Imai M, Tanaka T, Shibasaki H. Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. *Cognitive Brain Research*. 1998; 7:99–109. [PubMed: 9774711]
- Hansen JC, Dickstein PW, Berka C, Hillyard SA. Event-related potentials during selective attention to speech sounds. *Biological Psychology*. 1983; 16:211–224. [PubMed: 6615954]
- Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science*. 1973; 182:177–180. [PubMed: 4730062]
- Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levänen S, Lin FH, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW. Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:6809–6814. [PubMed: 15096618]
- Jancke L, Wustenberg T, Scheich H, Kaplan Layer J. Phonetic perception and the temporal cortex. *Neuroimage*. 2002; 15:733–746. [PubMed: 11906217]
- Johnston WA, Heinz SP. Depth of nontarget processing in an attention task. *Journal of experimental psychology Human perception and performance*. 1979; 5:168–175. [PubMed: 528926]
- Koch I, Lawo V, Fels J, Vorlander M. Switching in the cocktail party: exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*. 2011; 37:1140–1147. [PubMed: 21553997]
- LaBerge, D. *Attentional Processing: The Brain's Art of Mindfulness*. Harvard University Press; Cambridge, MA: 1995.

- LaBerge D. Attentional control: brief and prolonged. *Psychological Research*. 2002; 66:220–233. [PubMed: 12466921]
- Lavie N. Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*. 1995; 21:451–468. [PubMed: 7790827]
- Lavie N. Distracted and confused?: selective attention under load. *Trends Cogn Sci*. 2005; 9:75–82. [PubMed: 15668100]
- Lavie N. Attention, Distraction, and Cognitive Control Under Load. *Current Directions in Psychological Science*. 2010; 19:143–148.
- Lavie N, De Fockert J. The role of working memory in attentional capture. *Psychon Bull Rev*. 2005; 12:669–674. [PubMed: 16447380]
- Lavie N, Hirst A, de Fockert JW, Viding E. Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*. 2004; 133:339–354. [PubMed: 15355143]
- Lavie N, Tsai Y. Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics*. 1994; 56:183–197. [PubMed: 7971119]
- Leaver AM, Rauschecker JP. Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *The Journal of Neuroscience*. 2010; 30:7604–7612. [PubMed: 20519535]
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. Neural substrates of phonemic perception. *Cerebral Cortex*. 2005; 15:1621–1631. [PubMed: 15703256]
- Liebenthal E, Desai R, Ellingson MM, Ramachandran B, Desai A, Binder JR. Specialization along the Left Superior Temporal Sulcus for Auditory Categorization. *Cerebral Cortex*. 2010; 20:2958–2970. [PubMed: 20382643]
- Moray N. Attention in dichotic listening: Affective cues and the influence of instructions. *The Quarterly Journal of Experimental Psychology*. 1959; 11:56–60.
- Näätänen R. 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*. 1990; 13:201–288.
- Näätänen R, Picton TW. The N1 wave of the human electric and magnetic response to sounds: A review and a analysis of the component structure. *Psychophysiology*. 1987; 24:375–425. [PubMed: 3615753]
- Norman DA. Toward a theory of memory and attention. *Psychological Review*. 1968; 75:522–536.
- O'Connor DH, Fukui MM, Pinsk MA, Kastner S. Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience*. 2002; 5:1203–1209.
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Picton TW, Alain C, Woods DL, John MS, Scherg M, Valdes-Sosa P, Bosch-Bayard J, Trujillo NJ. Intracerebral sources of human auditory-evoked potentials. *Audiology and Neuro-Otology*. 1999; 4:64–79. [PubMed: 9892757]
- Rees G, Frith CD, Lavie N. Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*. 1997; 278:1616–1619. [PubMed: 9374459]
- Rinne T. Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *The open neuroimaging journal*. 2010; 4:187–193. [PubMed: 21760872]
- Saad ZS, Glen DR, Chen G, Beauchamp MS, Desai R, Cox RW. A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage*. 2009; 44:839–848. [PubMed: 18976717]
- Sabri M, Liebenthal E, Waldron EJ, Medler DA, Binder JR. Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. *Journal of Cognitive Neuroscience*. 2006; 18:689–700. [PubMed: 16768370]
- Sams M, Aulanko R, Aaltonen O, Näätänen R. Event-related potentials to infrequent changes in synthesized phonetic stimuli. *J Cognitive Neuroscience*. 1990; 2:344–357.
- Scherg M, Vasjar J, Picton TW. A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience*. 1989; 1:336–355. [PubMed: 23971985]

- Scherg M, Von Cramon D. Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*. 1986; 65:344–360. [PubMed: 2427326]
- Schwartz S, Vuilleumier P, Hutton C, Maravita A, Dolan RJ, Driver J. 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*. 2005; 15:770–786. [PubMed: 15459076]
- Specht K, Osnes B, Hugdahl K. Detection of differential speech-specific processes in the temporal lobe using fMRI and a dynamic “sound morphing” technique. *Human Brain Mapping*. 2009; 30:3436–3444. [PubMed: 19347876]
- Specht K, Reul J. Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task. *Neuroimage*. 2003; 20:1944–1954. [PubMed: 14683700]
- Talairach, J.; Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme Medical; New York: 1988.
- Treisman AM. Strategies and models of selective attention. *Psychological Review*. 1969; 76:282–299. [PubMed: 4893203]
- Vuilleumier P, Armony JL, Driver J, Dolan RJ. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*. 2001; 30:829–841. [PubMed: 11430815]
- Woldorff MG, Hillyard SA. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*. 1991; 79:170–191. [PubMed: 1714809]
- Woods DL, Alain C. Functional imaging of human auditory cortex. *Current Opinion in Otolaryngology & Head & Neck Surgery*. 2009; 17:407–411. [PubMed: 19633556]
- Woods DL, Herron T, Kang X, Cate AD, Yund EW. Phonological Processing In Human Auditory Cortical Fields. *Frontiers in human neuroscience*. 2011; 5:1–15. [PubMed: 21283556]
- Yi DJ, Woodman GF, Widders D, Marois R, Chun MM. Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nature Neuroscience*. 2004; 7:992–996.

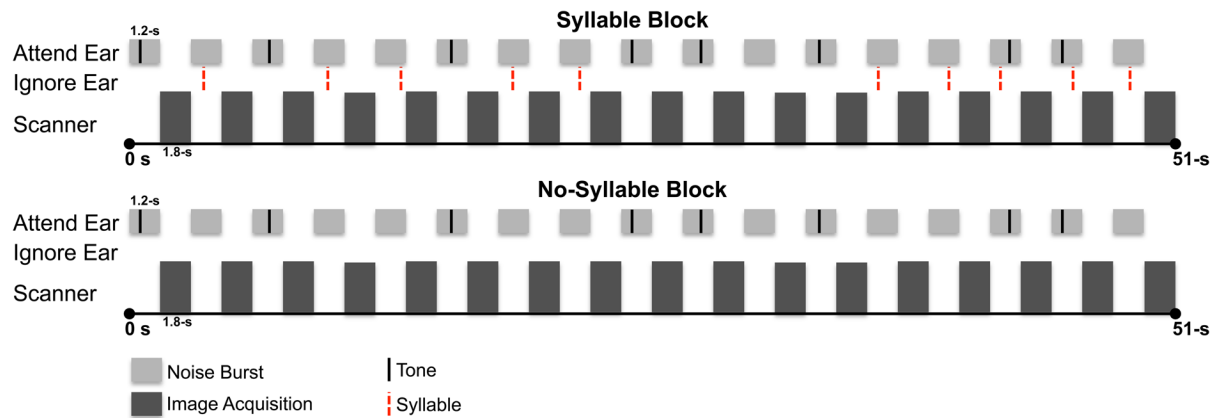


Figure 1.

A schematic illustration of stimulus presentation in Syllable (Top) and No-Syllable (Bottom) blocks. Light gray bars represent the 1.2-s noise bursts. Dark gray bars represent 1.8-s of image acquisition. Black tic marks represent 8 signal tones in noise bursts in the Attend stream. The additional 9 noise bursts did not include a tone. Dashed red tic 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.

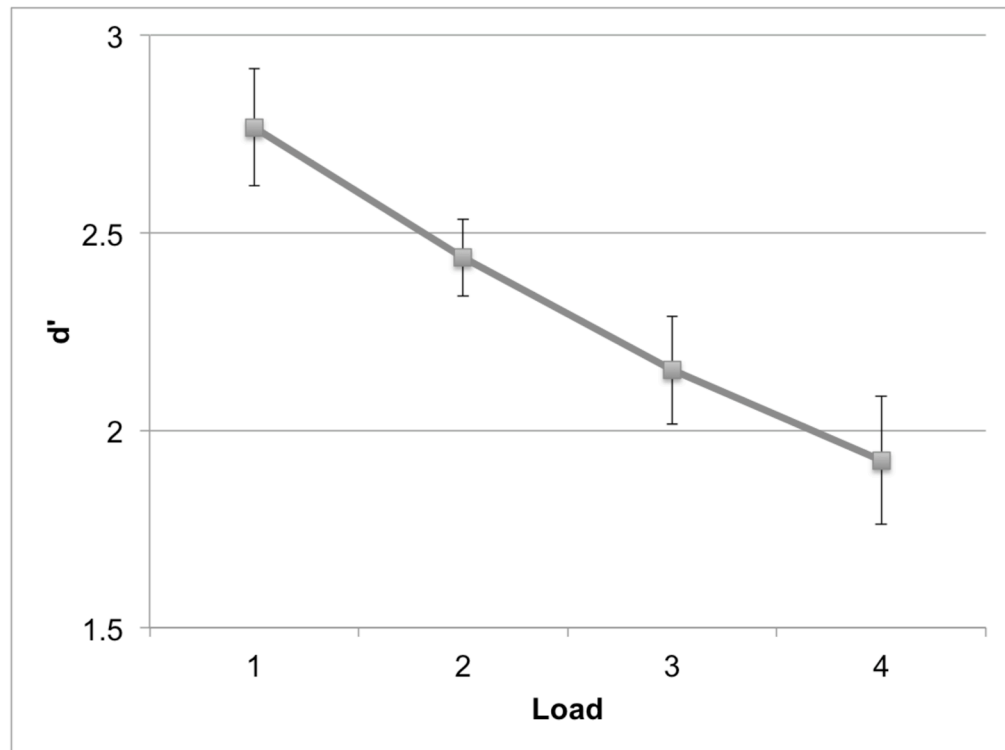


Figure 2. Behavioral performance (d') in the primary signal-detection task, in each load condition, in Syllable blocks. Error-bars indicate standard error (SEM).

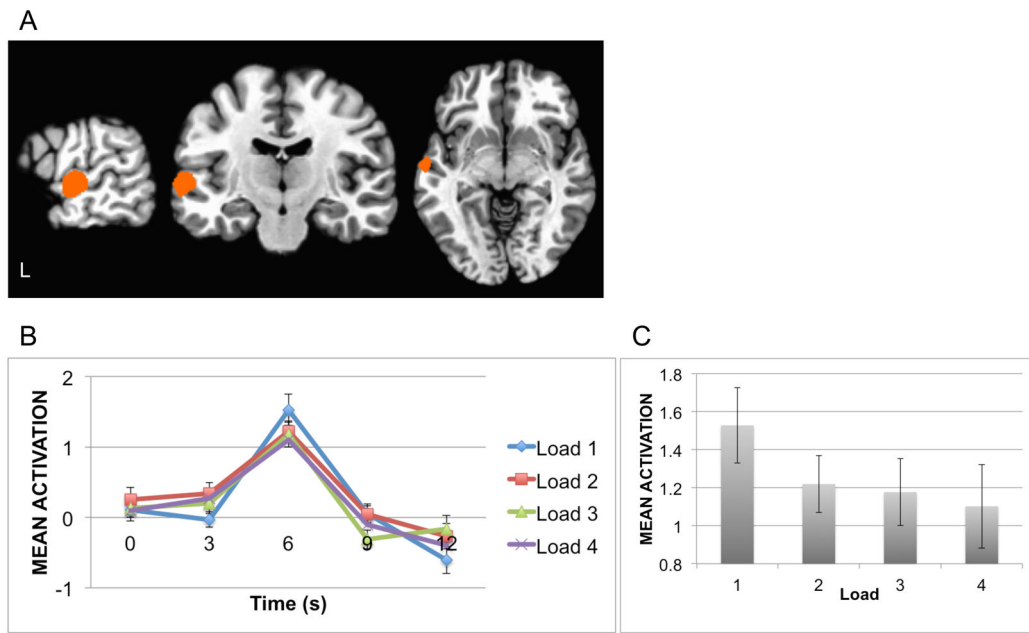


Figure 3.

(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 standard error (SEM).

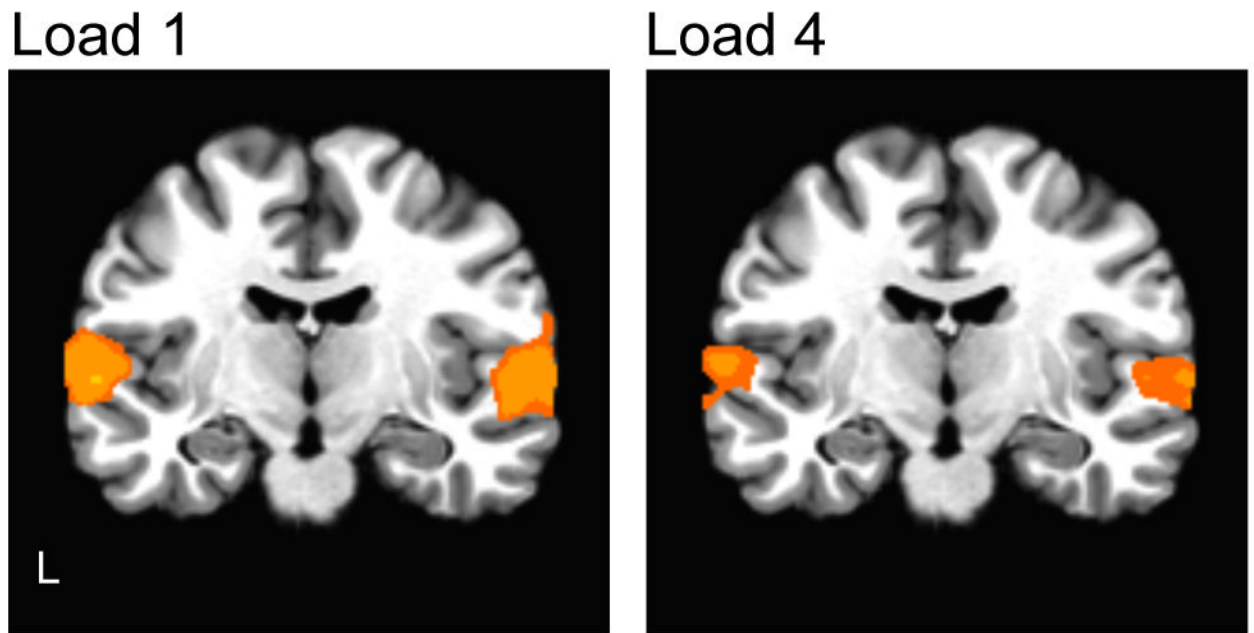


Figure 4. Whole-brain analyses: Statistical parametric maps of irrelevant syllables activation ($p < .05$, corrected) in Load 1 and Load 4 against baseline.

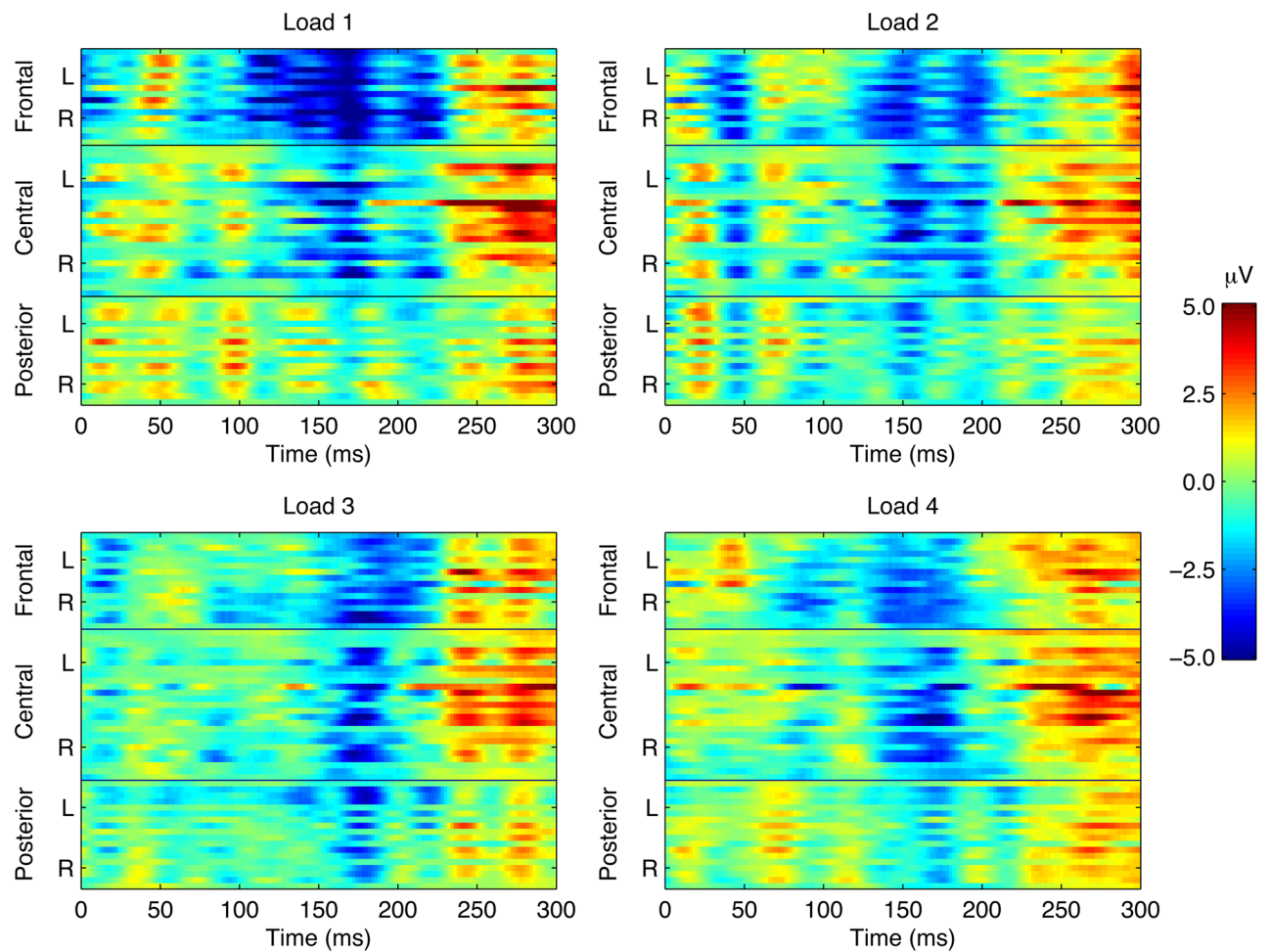


Figure 5.

Spatio-temporal maps from 60 electrodes: Grand average 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 .

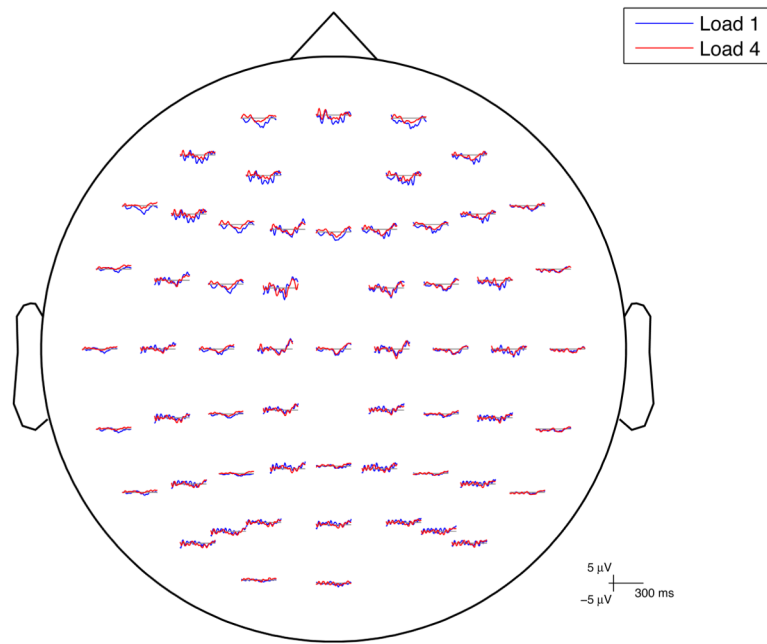


Figure 6. Group average ERP waveforms superimposed for irrelevant syllables in Load 1 and Load 4. Electrode sites conformed to the International 10–20 System.

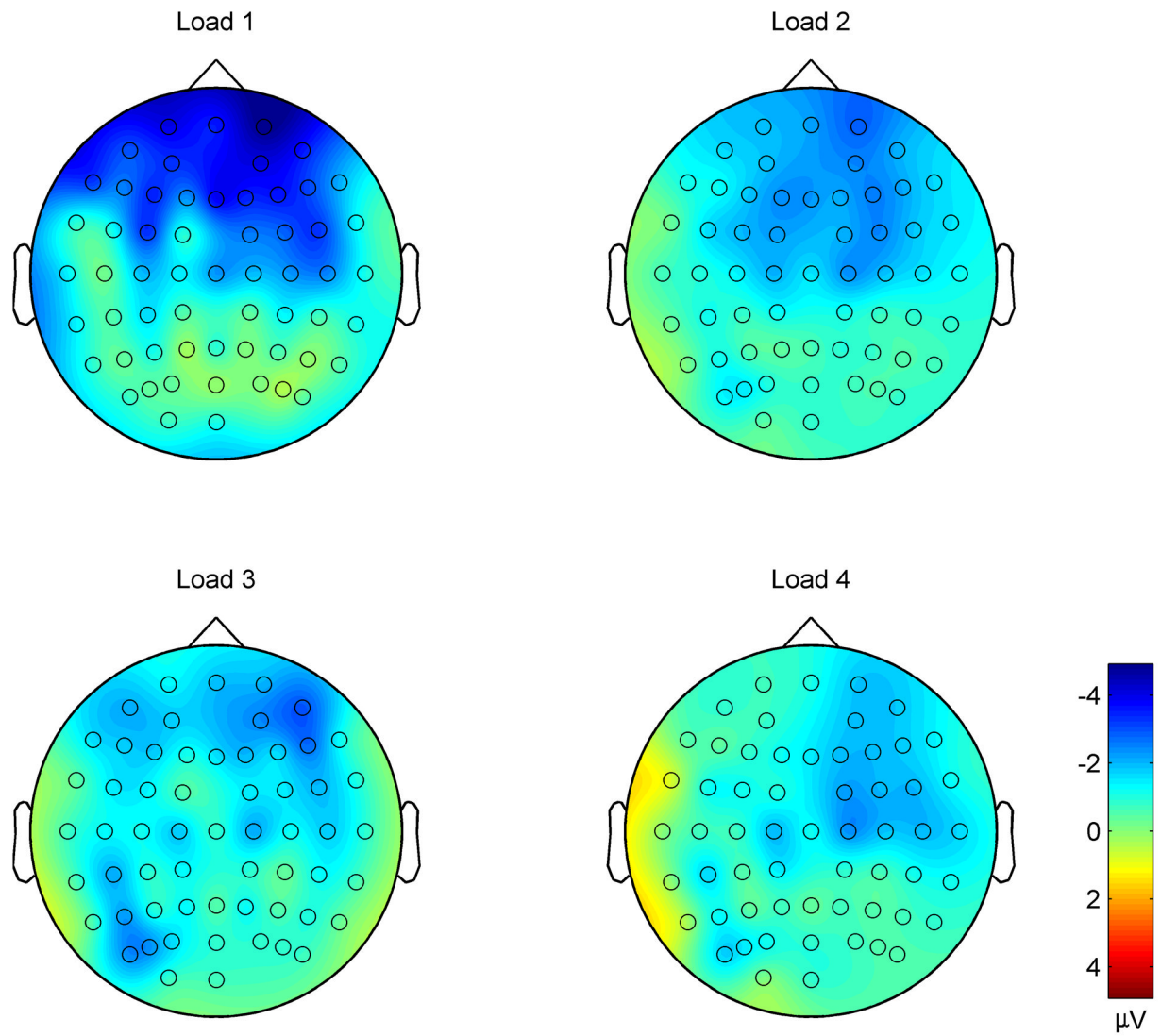


Figure 7. Mean scalp distribution for irrelevant syllables in the 130–230 ms time window, in the four load conditions. The color scale represents the amplitude in μV .

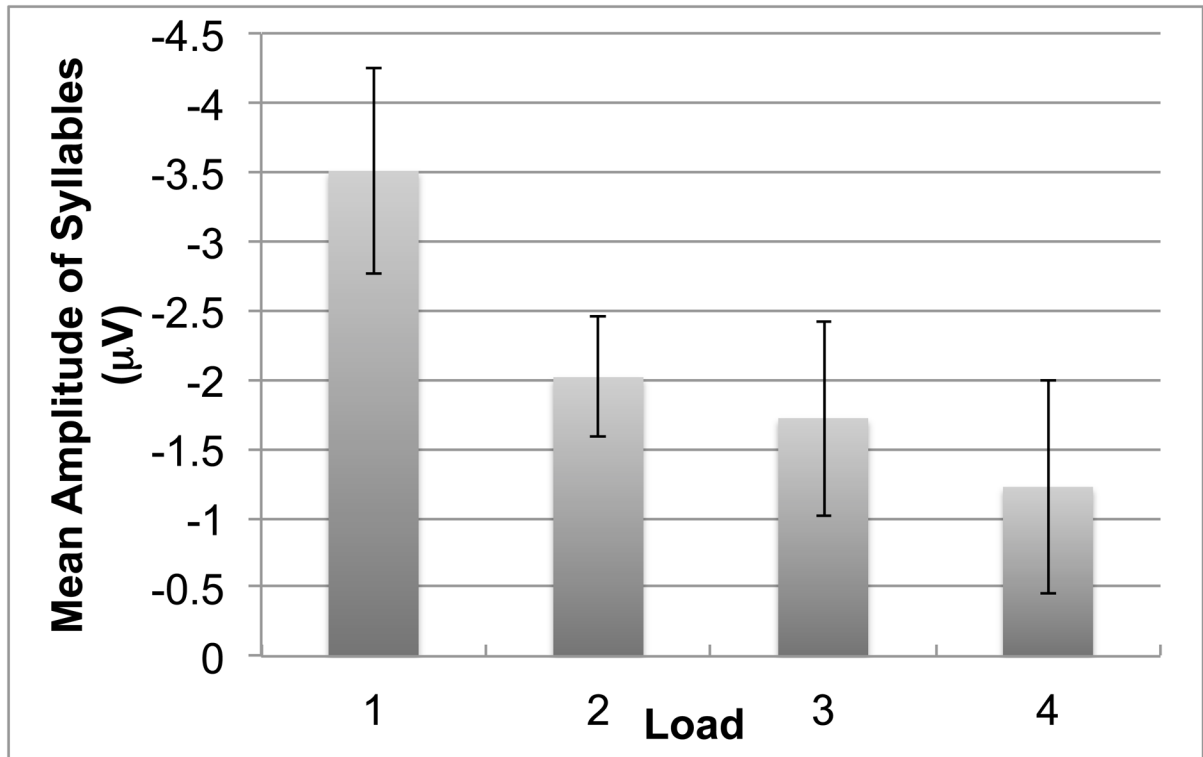


Figure 8. Mean amplitude on frontal electrodes (“frontal group” in Figure 5) for irrelevant syllables at 130–230 ms post-stimulus, as a function of perceptual load. Error-bars indicate standard error (SEM).