

Adverse Listening Conditions and Memory Load Drive a Common Alpha Oscillatory Network

Jonas Obleser,¹ Malte Wöstmann,¹ Nele Hellbernd,¹ Anna Wilsch,¹ and Burkhard Maess²

¹Max Planck Research Group “Auditory Cognition” and ²Research and Development Unit “Magnetoencephalography,” Max Planck Institute for Human Cognitive and Brain Sciences, D-04103 Leipzig, Germany

How does acoustic degradation affect the neural mechanisms of working memory? Enhanced alpha oscillations (8–13 Hz) during retention of items in working memory are often interpreted to reflect increased demands on storage and inhibition. We hypothesized that auditory signal degradation poses an additional challenge to human listeners partly because it draws on the same neural mechanisms. In an adapted Sternberg paradigm, auditory memory load and acoustic degradation were parametrically varied and the magnetoencephalographic response was analyzed in the time–frequency domain. Notably, during the stimulus-free delay interval, alpha power monotonically increased at central–parietal sensors as functions of memory load (higher alpha power with more memory load) and of acoustic degradation (also higher alpha power with more severe acoustic degradation). This alpha effect was superadditive when highest load was combined with most severe degradation. Moreover, alpha oscillatory dynamics during stimulus-free delay were predictive of response times to the probe item. Source localization of alpha power during stimulus-free delay indicated that alpha generators in right parietal, cingulate, supramarginal, and superior temporal cortex were sensitive to combined memory load and acoustic degradation. In summary, both challenges of memory load and acoustic degradation increase activity in a common alpha-frequency network. The results set the stage for future studies on how chronic or acute degradations of sensory input affect mechanisms of executive control.

Introduction

Adverse listening situations are challenging. Acoustical adversities range in severity from phone lines and noisy environments to age-related hearing loss and cochlear implants. However, the neural consequences of simultaneous adverse listening conditions (i.e., an acoustic degradation of the speech input) and cognitive effort (i.e., processing and memorizing this acoustic input) are unresolved. From behavioral work, degraded acoustic signals are known to draw on a listener’s cognitive resources (Rabbitt, 1968; Pichora-Fuller and Singh, 2006). Specifically, working memory (Pichora-Fuller et al., 1995; Rudner et al., 2011) and selective attention capacities (Shinn-Cunningham and Best, 2008; Kerlin et al., 2010) are affected by acoustic degradation.

How does the neural system respond to degraded acoustic speech input? Acoustic degradation elicits perceptual uncertainty (Burkholder et al., 2005), which is likely to perpetuate into memory storage phases and interfere with resources required for memory load (Rabbitt, 1968; Wingfield et al., 2005; Piquado et al., 2010). As speech unfolds, auditory encoding and memory storage are intertwined processes hard to separate. Therefore, an adapted auditory Sternberg

task (Sternberg, 1966) was used to study the neural traces of acoustic degradation and memory load in a silent delay period without auditory input. We combined three levels of memory load (2, 4, or 6 digits to be retained in memory) with three levels of acoustic degradation (digits were spectrally reduced to 16, 8, or only 4 bands; Drulman, 1995; Shannon et al., 1995) in a magnetoencephalography (MEG) study.

Our main hypothesis is concerned with the stimulation-free delay period. We expect acoustic degradation to increase perceptual uncertainty and thus to allocate more cognitive resources during the silent delay period, for storage of encoded (but potentially inaccurate) items and/or inhibition of interfering information. Importantly, we expect that exacerbations of acoustic degradation will exert a similar neural effect as increases in the number of presented items. We expect these effects in a common alpha (8–13 Hz) oscillatory network. Enhanced alpha oscillations have become a well-documented neural substrate of increased cognitive effort, in line with a functional, inhibitory role of alpha in controlling or gating local circuits of neural activity (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Foxe and Snyder, 2011; Weisz et al., 2011). Accordingly, alpha enhancement as a function of memory load has been demonstrated unanimously in the Sternberg task, in which items have to be briefly retained in short-term memory, and higher memory load is known to elicit longer response times (Jensen et al., 2002; Leiberg et al., 2006; Nenert et al., 2012; Sander et al., 2012). Thus, the load effects on alpha power during delay and on response time form a reliable starting point in the present study; however, our study claims that acoustic degradation also draws on this alpha oscillatory network, effectively increasing storage and inhibition demands. In accordance with the “functional inhibition” hypothesis, we expect

Received Sept. 26, 2011; revised July 9, 2012; accepted July 13, 2012.

Author contributions: J.O. designed research; J.O., M.W., N.H., A.W., and B.M. performed research; J.O., M.W., N.H., A.W., and B.M. analyzed data; J.O., A.W., and B.M. wrote the paper.

Research was supported by the Max Planck Society. The authors are grateful to Maren Grigutsch, Christian Obermeier, and Nathan Weisz, who all provided technical support and helpful comments at various stages of this project. Yvonne Wolff helped acquire the data.

The authors declare no conflict of interest.

Correspondence should be addressed to Jonas Obleser, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany. E-mail: obleser@cbs.mpg.de.

DOI:10.1523/JNEUROSCI.4908-11.2012

Copyright © 2012 the authors 0270-6474/12/3212376-08\$15.00/0

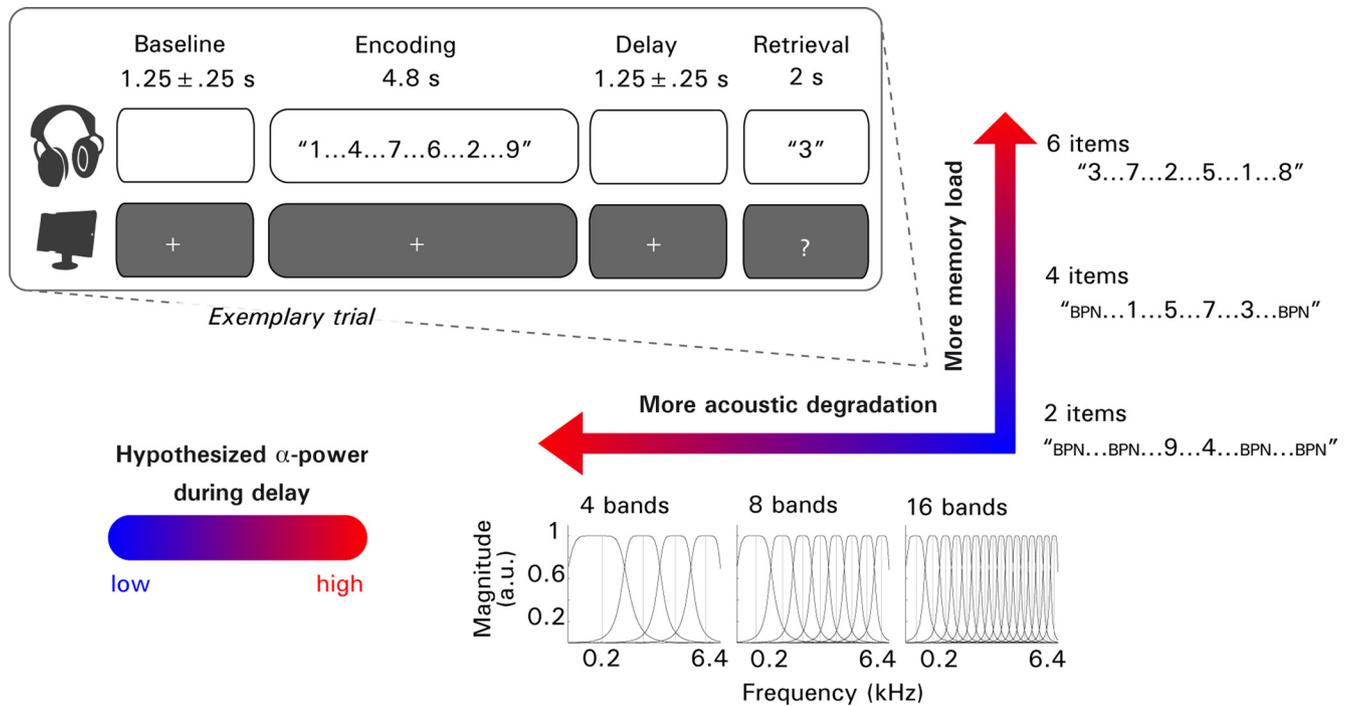


Figure 1. Combining memory load and acoustic degradation in a memory paradigm. A schematic trial of the adapted Sternberg task is shown in the top left. Memory load, operationalized as number of items to be retained [2, 4, or 6 items, flanked by bandpassed noise (BPN) to ensure presentation of 6 auditory events]; speech degradation of the presented items in a given trial, operationalized as the number of bands used for spectral degradation (16, 8, or only 4 bands). The bottom right indicates the vocoding filter frequencies and bandwidths. The color coding indicates the hypothesized parametric change of alpha power during stimulus-free delay.

that alpha activity during stimulus-free delay is predictive of the expected response time differences. Source localization of alpha changes during the delay phase will allow us to infer the neural generators of this alpha-tuned network.

Materials and Methods

Participants. Twenty healthy right-handed participants (10 females), who had no previous experience with noise-vocoded speech and who reported no known hearing deficit, took part in this experiment (age range, 20–32 years; mean age, 25.9 years). Data of 18 participants could be included in the final analyses (see below). All procedures were approved by the local ethics committee (University of Leipzig) and were in line with the Helsinki Declaration of Ethical Principles for Medical Research Involving Human Subjects.

Study design and stimuli. In an adapted auditory version of the Sternberg paradigm (Sternberg, 1966), we used a 3 × 3 design of the orthogonal factors memory load (2, 4, or 6 items in the to-be-memorized set of items of a trial) and acoustic degradation (16, 8, or only 4 bands in noise vocoding of the items; see below).

Stimuli used for short-term retention in each trial were edited from acoustic recordings of spoken digits. To this end, a trained female speaker recorded the German digit words from “null” (“zero”) to “neun” (“nine”). Recordings were performed in a soundproof chamber and digitized at 44.1 kHz. Offline editing included cutting at zero crossings and root-mean-square amplitude normalization; the single-digit audio file durations were not changed but reflected the naturally spoken digit length (588 ± 63 ms, mean ± SD).

Final audio files were additionally submitted to a noise-vocoding algorithm in MATLAB (MathWorks) to create audio versions of variable degradation for each digit. Noise vocoding is an effective technique to manipulate the spectral detail while preserving the temporal envelope of the speech signal (Drullman, 1995; Shannon et al., 1995) and render it more or less intelligible in a graded and controlled way, depending on the number of bands used. Less bands yield a less intelligible speech signal. The technique has been used widely in behavioral and brain imaging studies previously (Scott et al., 2000, 2006; Faulkner et al., 2001; Davis

and Johnsrude, 2003; Obleser et al., 2007). In vocoding, the bands were equally spaced using the Greenwood formula (as implemented by Rosen et al., 1999); the filter center frequencies were linearly spaced on the log-frequency axis. Figure 1 depicts the filter center frequencies and bandwidths for 4-band, 8-band, and 16-band speech. The pass band for filtering into channels/bands and envelope extraction was set to 0.07–9 kHz; the low-pass filter cutoff for the temporal envelope extraction was set at 256 Hz. Based on previous research as well as pilot digit recognition tests with naive participants, 4-band, 8-band, and 16-band versions were used in the final MEG study; of these conditions, 4-band speech was assumed to be most effortful to the perceptual–cognitive system, whereas correct identification of the digit (from this small set of 10 digits) from such degraded audio was still possible (see behavioral performance below). This is important because we did not aim at manipulating speech intelligibility per se but first and foremost aimed at manipulating the perceptual uncertainty and the concomitant effort evoked by degraded speech (Pisoni, 2000; Burkholder et al., 2005; Pichora-Fuller and Singh, 2006).

Procedure. Figure 1 presents a schematic trial timeline. Each trial contained four intervals of interest. The baseline interval started the trial (fixation cross for 1000 ms plus a random interval of 0–500 ms). During the encoding interval, a fixation cross and the 2, 4, or 6 auditory digits were presented. In the 2- and 4-digit sets, digits were flanked by items of loudness-matched 1 kHz, 1/3 octave bandwidth bandpassed noise with a duration reflecting the average digit length (i.e., 588 ms) to ensure presentation of six sounds. All sound files were played with a sound-onset asynchrony of 800 ms, resulting in a total encoding duration of 4800 ms.

This was followed by the delay interval during which the items had to be retained in memory (the fixation cross stayed on screen; thus, the delay phase started implicitly after the last of the six sounds). This delay interval lasted 1000 ms at minimum, plus some varying extra time (the identical random length of 0–500 ms chosen for the baseline period in a given trial).

Finally, the retrieval interval followed. Participants were presented simultaneously with a question mark and an additional digit (in the same acoustic degradation level as the digits of this trial during encoding) and had to decide with a button press (left or right index finger) whether it

appeared during encoding (response window of 2000 ms). The “yes/no”–left/right hand assignment was counterbalanced across participants, and yes answers were correct in 50% of all trials. Immediate feedback (“correct,” “wrong,” “too slow”) was presented on-screen after the button press. A trial was followed by an encouragement to blink and to proceed self-paced with a button press. Subjects could further pause at their own discretion between blocks. Total duration was ~50 min/participant.

For each cell of the 3 (memory load) \times 3 (acoustic degradation) design (Fig. 1B) over the course of five runs, we acquired on average 25 trials (~225 trials in total) in randomized order. With respect to the comparably low number of trials per experimental cell, please note that all analyses were run as parametric analyses (Obleser et al., 2008). Also, as outlined in detail below, we used proper first-level (i.e., subject-specific) statistics rather than raw power change estimates. This effectively standardizes all estimates for the variance across trials and thus avoids some of the problems potentially arising from the comparably low number of trials.

Data recording and analyses. Participants lay supine in an electromagnetically shielded room (Vacuumschmelze). Magnetoencephalographic signals were recorded using a 306-channel MEG (Vectorview; Elekta Neuromag Oy). The electrooculogram was also recorded using two bipolar (horizontal and vertical) electrode pairs.

The magnetic fields were recorded at a sampling rate of 1000 Hz and were online filtered with a bandpass of 0.03–330 Hz. During acquisition, the position of the participant’s head was registered by five head-position indicator coils. The signal space separation method (SSS; Taulu et al., 2004) was applied to suppress external noise. The default settings for signal–interference separation were used (i.e., an SSS basis with $L_{in} = 8$ and $L_{out} = 3$). Also, the SSS method was used to correct for differences in head position between runs, that is, data from all experimental runs were recomputed to assume the same head position as the beginning of the first run. Magnetometer data were only used for interference suppression and head-position correction. All additional data analyses were performed on the planar gradiometer recordings only.

Offline, the data were analyzed using MATLAB and the FieldTrip toolbox (Oostenveld et al., 2011). The continuous signal was low-pass filtered at 200 Hz and epoched. There were trials in nine conditions according to variations in the factors memory load (2, 4, or 6 digits to retain) and degradation (16, 8, or only 4 bands of spectral resolution of the audio recordings of the digits). For time–frequency analysis, epochs of -1 to $+2$ s were extracted from the signal for baseline, delay, and retrieval intervals and epochs of -1 to $+6$ s for the encoding interval. These long epochs were extracted to circumvent windowing problems in the time–frequency analysis; the intervals analyzed statistically were shorter (see below). A z -score-based algorithm, available in the FieldTrip toolbox, automatically rejected epochs contaminated by eye movements using a cutoff value of 4. Remaining artifact-contaminated epochs were rejected by visual inspection of all channels. Two participants’ datasets were discarded as a result of technical problems while recording the MEG. Thus, for further data analyses, we used the data of 18 subjects.

On average, >80% of all trials per participant could be retained for additional analyses. Importantly, no significant differences in rejection rate per condition occurred (repeated-measures ANOVA with factors memory load and acoustic degradation, all $F \leq 1$). Thus, it is safe to conclude that no substantial signal/noise ratio confounds were present in additional analyses.

To obtain time–frequency representations of the data, trial data were convoluted with a family of Morlet wavelets (7 cycles width), and the power spectra for all intervals were estimated from 3 to 30 Hz in 1 Hz steps and from -1 to $+2$ s around the baseline, delay, and retrieval onsets, in steps of 20 ms, and -1 to $+6$ s around the encoding onset. Time- and frequency-specific power estimates were computed for each participant separately for all nine conditions. The time intervals forwarded to statistical analysis were as follows: (1) 0–5 s after encoding onset; (2) 0.4–1 s after delay onset, skipping the beginning and the variable end of the delay period; and (3) 0–1 s after probe onset. All these estimates were baseline corrected (relative change to the mean power estimate of 0.4–1 s after baseline onset; Fig. 1) to attain measures of relative power change per trial for the respective time interval.

Finally, data from the 204 gradiometers (i.e., 102 pairs of gradiometer channels) were combined in each participant (using the *ft_combinepla-*

nar function) in all nine conditions, and time–frequency-specific power estimates at 102 locations were available for statistical analysis.

Statistical analysis. We pursued a two-level statistical analysis. At the first or individual level, time–frequency representations from all trials were submitted to a parametric regression test for independent samples (as implemented in *ft_freqstatistics*). By setting the contrast coefficients accordingly, we obtained time–frequency–sensor matrices for each subject that contained t values for either a memory load effect ($6 > 4 > 2$ items) or an acoustic degradation effect ($4 > 8 > 16$ bands). Attaining first-level statistics rather than using only average power change compared to baseline has the advantage that the effects that are subsequently entered into group statistics are effectively standardized for across-trial variance. Last, matrices of t values were transformed to z -values, because degrees of freedom slightly varied across subjects as a result of slight differences in number of trials.

At the second or group level, we submitted the individual z -maps to a massed cluster-based permutation t test (testing for significant differences from zero, two-sided, dependent samples, with 1000 iterations; as outlined by Maris and Oostenveld, 2007). In essence, this procedure checks for time–frequency–sensor clusters that show parametric effects of either a decrease or an increase covarying with the manipulated stimulus dimension. The permutation tests for the encoding, delay, and retrieval interval included all time–frequency bins across all frequencies (3–30 Hz) and across all 102 sensors.

Note that this procedure protects at the cluster level against an inflated false-positive rate otherwise arising from multiple comparisons. In short, the approach first checks for time–frequency bins that show a significant permuted t statistic (i.e., the “cluster entry criterion,” $p < 0.05$) and then searches for time–frequency–sensor clusters of bins that behave similarly, considering a minimum of three neighboring sensors as a cluster. The resulting test statistic assesses significance by comparing the observed cluster-level statistic (summed t values per cluster) to the distribution of all randomized cluster-level statistics, with the final p value resulting from the proportion of Monte Carlo iterations in which the cluster-level statistic was exceeded (for details, see Maris and Oostenveld, 2007).

In all intervals of interest (Figs. 1A, 2), two different massed t tests were set up: one testing for linear effects of memory load and another testing for linear effects of acoustic degradation. The time–frequency cluster tests did not readily allow testing for a potential interaction of the two factors. Therefore, we tested for a potential interaction in the stimulus–delay period by extracting mean alpha power per condition and subject from those sensors belonging to both the significant memory and acoustic degradation cluster. On these data, we calculated a 3 (memory load levels) \times 3 (acoustic degradation levels) repeated-measures ANOVA. A Greenhouse–Geisser corrected p value is reported for this interaction test.

Behavioral responses were analyzed in a 3 (memory load levels) \times 3 (acoustic degradation levels) repeated-measures ANOVA. Here also, Greenhouse–Geisser-corrected p values are reported.

Source localization. MEG more than electroencephalography suffers from individual variations of head-to-sensor arrangement. In contrast, source space analysis overcomes some of this variability because each subjects’ reconstructed brain can be spatially normalized using the MRI and coregistered to the same reference space. Therefore, to overcome this and other inherent limitations when interpreting sensor-space topographies (Lopes da Silva, 2004), alpha activity during delay was projected to source space using an adaptive spatial filter in the frequency domain [dynamic imaging of coherent sources (DICS); Gross et al., 2001]. The DICS technique is based on the cross-spectral density (CSD) matrix, which was obtained in every trial and condition by applying a multitaper fast Fourier transform estimate of the time windows and frequencies of interest (here yielding CSD estimates centered at 10 Hz with a 2 Hz spectral smoothing on each side and a window length of 500 ms centered at 750 ms after delay onset).

For source localization, we chose a procedure used in various previous studies localizing oscillatory power (Medendorp et al., 2007; Haegens et al., 2010; Hipp et al., 2011; Obleser and Weisz, 2012). T1-weighted MR images were aligned with individual head shapes as acquired during MEG. A realistically shaped single-sphere model was used as volume

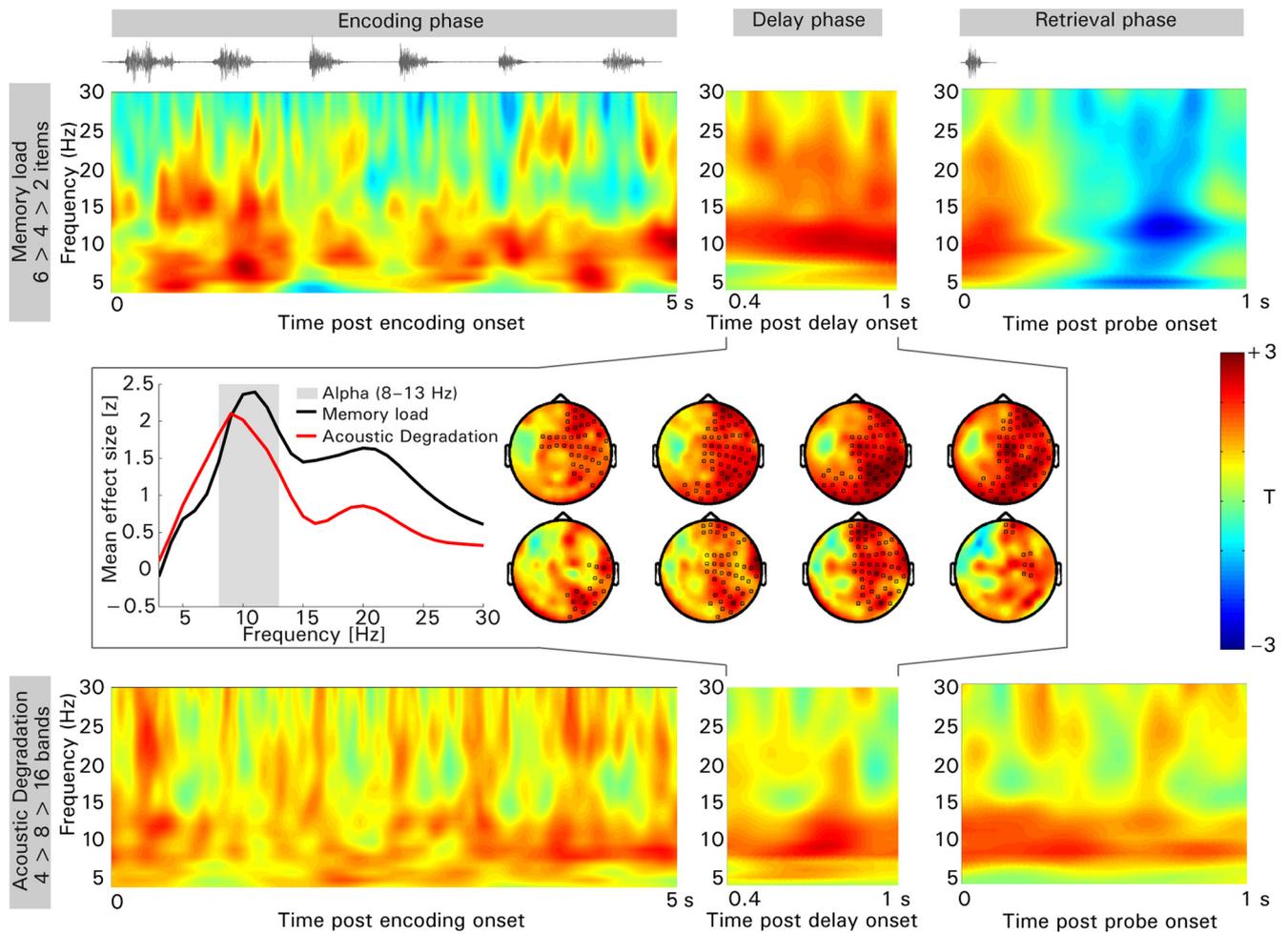


Figure 2. Time–frequency response during encoding, delay, and retrieval phases. All panels show *t* values ($N = 18$) for parametric effects of memory load (top row) and degradation (bottom row), separately for the encoding phase (left column), delay phase (middle column), and the retrieval phase (right column). Waveforms at the top schematically indicate the presented sound. Top and bottom rows show time–frequency plots of *t* values, averaged across sensors that belong to the respective significant cluster. Positive *t* values (i.e., more alpha activity for higher load or more severe degradation, respectively) are shown in hot colors and negative values in cold colors. Line plot, For the delay period (middle panels), the distribution of memory load (black) and degradation (red) effect sizes over frequencies confirms an alpha peak for both effects. The sensors contributing to the significant memory load and degradation clusters are indicated by dots in the series of sensor topographies over time (left middle).

conductor (Nolte, 2003), and the lead field was calculated at each point in a grid with a 1 cm resolution. It is of note that, rather than constructing individual grids, we constructed a template grid in the MNI space template brain (as used in SPM8) and then warped these grids into individual, native space using individuals’ inverse homogenous transformation matrices (as derived when spatially normalizing individual MR images). This allows to do statistics on a common grid.

Using the CSD data, a spatial filter was constructed for each grid point, and the spatial distribution of power was estimated for each condition in each subject. A common filter was constructed from all baseline and delay segments (i.e., based on the CSD matrices of the combined conditions). Subject- and condition-specific solutions reflected relative alpha power change in the delay period for each grid point.

On these grid-point-wise source power changes, we performed source-level cluster statistics (following the same logic as outlined for sensor-level statistics above). To infer tentatively on the brain structures generating the scalp topographies (Fig. 2), we tested for grid points with significant overall alpha increase compared with baseline (i.e., collapsing across conditions). For illustration purposes, significant clusters at the grid level were interpolated to a 3D standard MR template (in MNI space) and plotted (see Fig. 4, top row; bottom row shows the same data being plotted onto brain areas as outlined in the automatic anatomical labeling atlas; Tzourio-Mazoyer et al., 2002).

Results

Time–frequency changes during stimulus-free delay

The main hypothesis of this study concerned the stimulus-free delay period while items were retained in memory (“delay”; Fig. 2, middle; see also Figs. 3, 4). We expected alpha power to increase as a function of number of retained items but also of acoustic degradation of these items and also tested for any interaction of these two manipulations.

Higher memory load elicited an increase in alpha power at right parietal–central sites that was linearly dependent on number of items to retain: one significant positive cluster was found for the analysis of the effect of memory load ($p = 0.007$). However, such a significant cluster was also found for the effect of acoustic degradation ($p = 0.02$). Thus, alpha power during delay was also a function of acoustic degradation during the encoding of the presented items. These two clusters overlapped in topography (Fig. 2), with the cluster for memory load extending more frontally. A significant interaction of these two effects was also present ($F_{(2,34)} = 7.02$, Greenhouse–Geisser $\epsilon = 0.65$, $p = 0.001$). Inspection of the data implies that this interaction is best de-

scribed as an “expansive” function for alpha power (Fig. 3*A,B*). When plotting alpha power as a function of both number of items and degree of acoustic degradation, alpha power appears higher than expected from the two main effects at most severe degradation (4 bands) and at the highest load (6 items). At the same time, alpha power appears to “undershoot” when combining least degradation (16 bands) with least load (2 items only).

The right middle row of Figure 2 shows the frequency specificity of the alpha-band effect. When plotting the average *z*-values across subjects (from the first-level statistics) as line graphs, the effects of degradation and memory load both peak in the alpha-band range (peak difference in hertz; *n.s.*). Both effects also show a subpeak in the beta-band (15–25 Hz) range, but these were not borne out by the cluster-based permutation test as significant by themselves. Thus, during stimulus-free delay, memory load and acoustic degradation are not separable based on exact alpha frequency.

Time–frequency changes during encoding and during retrieval

Although the main hypothesis of this study was concerned with the stimulus-free delay period, we also analyzed the experimental encoding and retrieval phases on the sensor level. In all three post-baseline periods of the experiment (encoding, delay, and retrieval), we tested for significant time–frequency–sensor clusters (Fig. 2).

During encoding, this yielded alpha-frequency clusters, that is, alpha power was found to increase already during encoding depending on the two manipulations. One significant positive cluster was found for the memory load effect ($p = 0.018$). Also, one cluster was found for the acoustic degradation effect ($p = 0.009$). Both clusters showed sensor distributions similar to the one shown for the ensuing delay phase in the middle row of Figure 2.

For brain activity at and after the probe (i.e., during retrieval; Fig. 2, right column), the cluster-based permutation test for effects of memory load revealed one early (0–200 ms) positive cluster ($p = 0.029$) reflecting the same pattern seen during delay. More notable, one late (~600–850 ms) negative cluster ($p = 0.008$) was also observed. That is, alpha power was suppressed as a function of the number of items to be “released from memory”—the inverse effect of what was seen during encoding and delay. In sharp contrast to this, acoustic degradation yielded one extended, positive cluster ($p = 0.034$), indicating that more severe degradation of the probe triggered renewed alpha enhancement (Obleser and Weisz, 2012).

Behavioral results and correlation with delay-phase alpha

As expected in a Sternberg paradigm, listeners performed very well in all memory load conditions, and, as intended, the chosen acoustic degradation levels did not hinder their performance. Average \pm SD performance across conditions was $95 \pm 3\%$ cor-

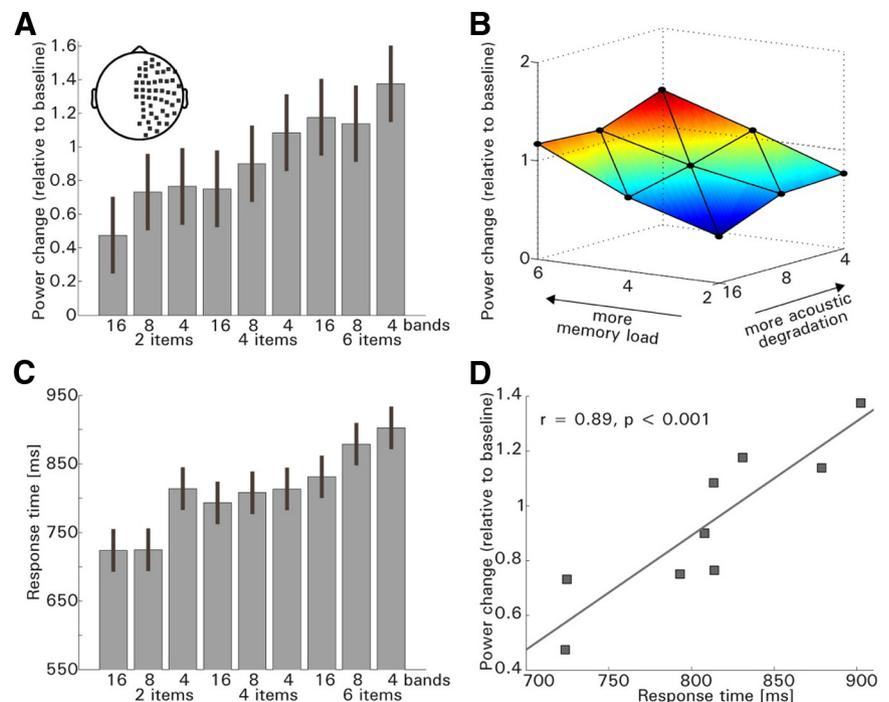


Figure 3. Alpha enhancement during delay and relation to response time. **A**, Mean alpha power change during delay (8–13 Hz, 0.4–1 s; the inset sensor topography shows the used sensors; those sensors were part of both the memory load cluster and the degradation cluster). Means ± 1 within-subject standard error are displayed. Note the cumulative effects of memory load (memory load, left to right) and degradation (less number of bands, also left to right). **B**, Subadditivity and superadditivity of alpha power. Alpha power (data identical to **A**) replotted as a two-dimensional function along the two manipulated dimensions (load, more items; degradation, less number of bands). Note the alpha overshoot (red) for most severe degradation/highest number of items and alpha undershoot (blue) for least degradation/least number of items. **C**, Mean response time per condition. Means ± 1 within-subject standard error are displayed. **D**, Scatter plot of mean response time (*x*-axis; data identical to **C**) versus mean alpha power during delay (*y*-axis; data identical to **A**) with linear regression line.

rect, with average condition performances ranging from 91 to 100%. No effects of memory load or acoustic degradation attained significance.

Figure 3*C* shows the response time data. Response times to the probe confirmed the known Sternberg effect (longer response times to more items, main effect of memory load, $F_{(2,34)} = 22.0$, $\epsilon = 0.79$, $p < 0.0001$) but also showed a main effect of acoustic degradation (longer response times to stronger degradation; $F_{(2,34)} = 18.4$, $\epsilon = 0.99$, $p < 0.0001$). Given the overall very accurate performance, reaction time patterns did not change whether or not only correct trials were analyzed, and the few erroneous trials are included here.

Notably, the mean response times per condition were well predicted by the respective mean alpha power change during the preceding delay phase (Pearson’s $r = 0.889$, $p < 0.001$; Fig. 3*D*).

Source localization

For alpha power changes during stimulus-free delay, we used an adaptive spatial filter in the frequency domain (DICS; Gross et al., 2001) to localize the most likely sources of the alpha power increase. This was done across conditions (delay > baseline) to assess the most likely neuroanatomical brain source for the observed alpha increase during delay. Results are shown in Figure 4.

When testing for voxels showing alpha power increase during delay compared with baseline in a one-sample *t* test, a significant, spatially extensive cluster was found ($p < 0.04$). The cluster ranged from right superior parietal cortex down to right posterior supramarginal gyrus (SMG), extending ventrally as far as right posterior superior temporal gyrus and medially into precu-

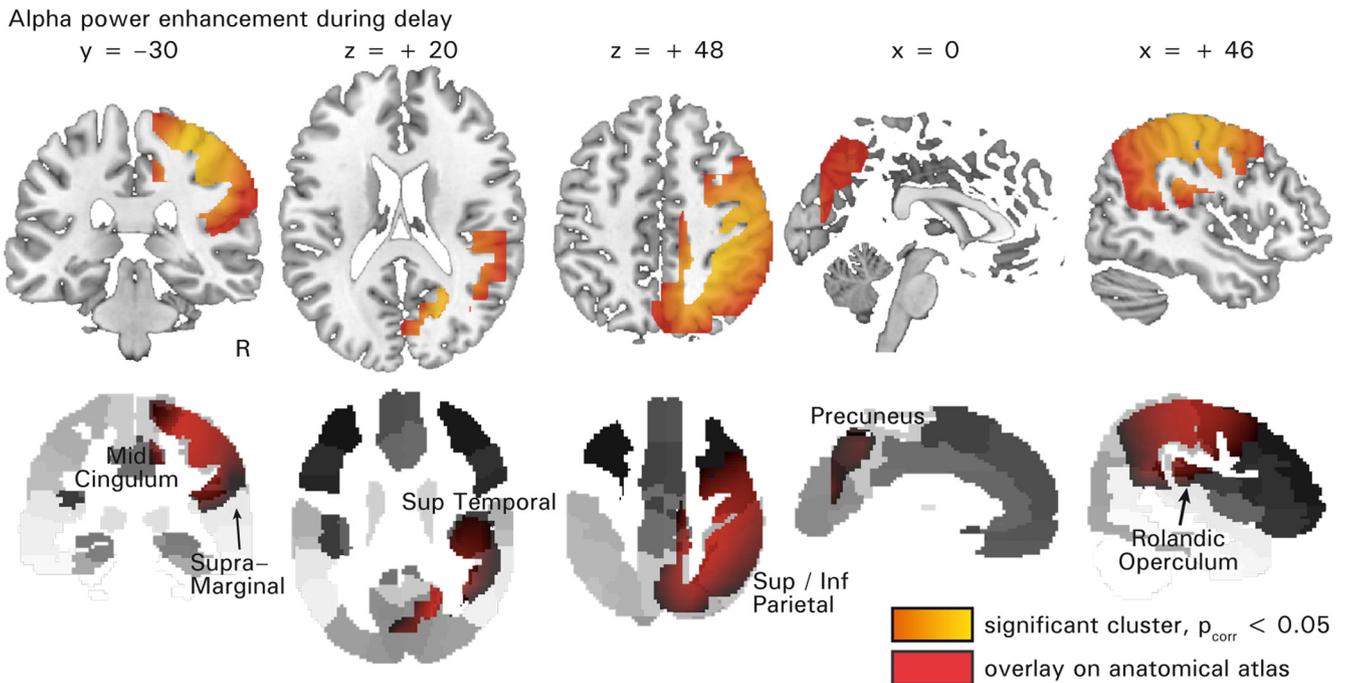


Figure 4. Source localization of alpha enhancement during delay. Top row illustrates the source-level cluster ($p < 0.05$) exhibiting an overall alpha power enhancement compared with baseline, with t values from the significant cluster displayed on a standard MR template. Bottom row shows the data plotted on the automatic anatomical labeling atlas. Sup, superior; Inf, inferior.

neus (Fig. 4). This corroborates well the overall pattern seen in the gradiometer topographies (Fig. 2, middle row).

Beamformer methods are known to encounter difficulties in resolving correlated sources, for example, from simultaneously active bilateral auditory cortex (Dalal et al., 2006). However, this potential problem is unlikely to have occluded bilateral auditory activity here, because the gradiometer sensor maps (Fig. 2) did not hint at bilateral auditory temporal sources as generators of the alpha effects during the delay period. This was then confirmed and neuroanatomically specified in the DICS source localization (Fig. 4).

Discussion

How does acoustic degradation affect the neural mechanisms of auditory working memory? Here, we tested the hypothesis that a neural “executive control” (i.e., storage and/or inhibition) system, expressed as enhanced alpha power during working memory retention, would be additionally affected by acoustic degradation of the to-be-memorized speech material.

The most important finding was a significant right temporo-parietal alpha enhancement during auditory memory retention. This enhancement in ~ 10 Hz power was not only parametrically driven by the memory load (more number of items; cf. Jensen et al., 2002) but also by adversity of the acoustic signal (more severe degradation; Figs. 2, 3). In accordance with the prolonged reaction times (Fig. 3C), the preceding alpha power enhancements during the delay phase reflect the varying cognitive demands. The new finding is that, at this comparably late and stimulation-free stage of memory retention, acoustic degradation does affect this process considerably.

The alpha enhancement attributable to more severe acoustic degradation was found to be additive to the alpha enhancement attributable to more stored items: acoustic degradation (a perceptual challenge) and memory load (a capacity challenge) both draw on a neural system subserving “functional inhibition”

(Klimesch et al., 2007; Jensen and Mazaheri, 2010). Condition-specific alpha power changes illustrate this (Fig. 3A,B): the average alpha power while retaining 2 items presented in severely degraded quality (4-band speech) approximately equals the average alpha power while retaining 4 (rather than only 2) items in least degraded quality (16-band speech).

Importantly, the alpha-change pattern also showed a significant interaction, best characterized as an “expansive” (i.e., non-linear) response function (Miller and Troyer, 2002; Duong and Freeman, 2008; Fig. 3A,B). Summations of memory load and acoustic degradation led to subadditive and superadditive responses of the alpha network, respectively. This is evident when plotting alpha power during stimulus-free delay as either a monotonic function of both manipulations (Fig. 3A) or as a two-dimensional function of both manipulations separately (Fig. 3B). An expansive response function can serve as an internal, sensitivity-enhancing mechanism, and particularly the super-additive alpha response to combinations of many items and severe degradation is commensurate with this interpretation.

Unlike fMRI, magnetoencephalographic source localization in humans offers temporal and spectral specificity (here by localizing only alpha-band activity and only during the delay period). As suggested by the gradiometer topographies (Fig. 2), alpha power in the delay phase emerged mainly from posterior parietal cortex, in line with the often reported posterior alpha network (Foxe et al., 1998; Laufs et al., 2006). As illustrated in Figure 4, the significant cluster extends into the SMG and the temporo-parietal junction (TPJ), both being hotspots in verbal working memory (Paulesu et al., 1993; Jacquemot et al., 2003; Jacquemot and Scott, 2006; Obleser and Eisner, 2009). Obleser and Eisner (2009) have argued that the SMG should operate on a post-categorical “abstract” code, in line with its important role in the “phonological store” concept (Jacquemot and Scott, 2006; Buchsbaum and D’Esposito, 2008). The present data concur with

these suggestions and emphasize how intertwined speech perception and verbal working memory are, particularly when coping with degraded input (Eisner et al., 2010).

The source localization also suggests contributions from precuneus and posterior cingulate cortex. Involvement of this area ties in well with the “functional inhibition” interpretation of alpha activity and is suited to reconcile alpha-oscillation-based theories with more general models on executive control. Inhibitory, “top-down” control over task-irrelevant processes and over potentially erroneous behavior is likely to critically depend on the alpha-frequency domain (Klimesch et al., 2007; Jensen and Mazaheri, 2010), within as well as across brain areas. Concomitantly, recent functional imaging studies have pointed out that suppression of BOLD activity (often reported to be anticorrelated with alpha power, Laufs et al., 2003; Sadaghiani et al., 2010) in cingulate areas (i.e., the “default” network, as well as in the TPJ) is pivotal to successful working memory performance (Anticevic et al., 2010, 2011; Reas et al., 2011). In line with these studies and an inhibitory function of alpha, we find alpha power in temporoparietal and posterior cingulate areas to be affected by our number of items and acoustic degradation manipulations.

Notably, the presented data do not indicate a strong involvement of domain-specific, auditory areas in this alpha network, the only exception being the extension of the significant cluster into the right superior temporal gyrus (BA 41; Fig. 4). The present data rather imply that most neural correlates of increased effort do not manifest in auditory regions (at least not during delay, i.e., when no sound input is present). This is important for the auditory neuroscience of degraded hearing and aging, and functional MRI studies of listening to degraded speech have pointed in this direction before (Sharp et al., 2004; Harris et al., 2009). The current study transcends these important imaging studies in dissecting the influence of (modality-specific) acoustic degradation from the influence of (modality-unspecific) load and in providing a real-time and frequency-specific measure of executive control.

The present results have implications for the neural bases of degraded hearing and aging alike. In degraded hearing and cochlear implants, sensory degradation is without doubt accompanied by increased effort at multiple cognitive processing stages (Wingfield et al., 2005; Pichora-Fuller and Singh, 2006). These additional challenges might draw on neural resources that will in turn be lacking for other cognitive tasks (Heinrich and Schneider, 2011; Obleser and Weisz, 2012). Our study pinpoints this “double taxation” arising from sensory degradation and additional cognitive demands to a cingulate–parietal alpha oscillation network. Future studies can build on this and compare measures of the alpha oscillatory network in chronically hearing-impaired subjects.

The present results also provide evidence that processes, such as storage or inhibition reflected here in enhanced alpha power during working memory retention, are affected by both signal-dependent and capacity-dependent challenges. Recall, however, that the overall difficulty of the Sternberg task was not approaching participants’ capacity limits, even at the worst level of acoustic degradation; participants performed above 90% in all conditions. Neither were the alpha-network dynamics reaching saturation, as expressed by the expansive response behavior (Fig. 3A, B). A necessary next step to further validate our hypothesis (a shared cognitive resource affected by degradation and memory load alike) would be to explore cognitive capacities closer to their near maximum and test the influence of acoustic degradation accordingly. Only behavioral evidence pointing in this direction is available

(Burkholder et al., 2005); the authors tested listeners’ actual capacity limits (digit span) in vocoded speech and predicted this limit from the accuracy to identify the degraded single items. Our setup, instead, allows measuring the neural consequences of degraded (and potentially inaccurate) information.

Finally, recall that alpha power was modulated already during the 5-s-long presentation (i.e., the encoding) of the stimuli, by memory load as well as by acoustic degradation (Fig. 2). This is in line with the hypothesis of increased effort in neurally encoding degraded stimuli. Arguably the most important difference between visuospatial and auditory tasks is the dependence on the unfolding of time (Shamma, 2001): more so than in vision, the time required for encoding six digits (or, more realistically, an ongoing conversation) will trigger a neural cascade of encoding and retention in memory.

The most striking difference between the memory load and acoustic degradation manipulations occurred after the “probe” digit but before the participants’ response. Although the load-dependent alpha showed a “rebound” or suppression of alpha after the trial-final probe digit (a “release from memory”), the degradation-dependent alpha exerted another sustained alpha power increase in response to the degraded probe (again, parametric with degradation severity; Fig. 2, compare right top with right bottom).

In conclusion, alpha oscillations during a stimulus-free delay period pose a composite measure of cognitive effort. High alpha power is thought to reflect the enhanced need for “functional inhibition” (Jensen and Mazaheri, 2010). We show for the first time that this mechanism of enhanced alpha power is not only modulated by changing domain-general requirements such as the number of stored items: challenges arising from mild to severe sensory degradation do affect this system, too, and both manipulations cause an enhancement of oscillatory power in the same time–frequency range, in a nonlinear manner (Fig. 3). Also, these alpha responses appear to be generated by a primarily overlapping set of neuroanatomical structures. These findings have implications for our understanding of alpha-driven executive control networks as well as for our concepts on how elderly and hearing-impaired listeners can encounter their sensory difficulties.

References

- Anticevic A, Repovs G, Shulman GL, Barch DM (2010) When less is more: TPJ and default network deactivation during encoding predicts working memory performance. *Neuroimage* 49:2638–2648.
- Anticevic A, Repovs G, Barch DM (2011) Working memory encoding and maintenance deficits in schizophrenia: neural evidence for activation and deactivation abnormalities. *Schizophr Bull*. Online advance publication. Retrieved July 31, 2012. doi:10.1093/schbul/sbr107.
- Buchsbaum BR, D’Esposito M (2008) The search for the phonological store: from loop to convolution. *J Cogn Neurosci* 20:762–778.
- Burkholder RA, Pisoni DB, Svirsky MA (2005) Effects of a cochlear implant simulation on immediate memory in normal-hearing adults. *Int J Audiol* 44:551–558.
- Dalal SS, Sekihara K, Nagarajan SS (2006) Modified beamformers for coherent source region suppression. *IEEE Trans Biomed Eng* 53:1357–1363.
- Davis MH, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. *J Neurosci* 23:3423–3431.
- Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am* 97:585–592.
- Duong T, Freeman RD (2008) Contrast sensitivity is enhanced by expansive nonlinear processing in the lateral geniculate nucleus. *J Neurophysiol* 99:367–372.
- Eisner F, McGettigan C, Faulkner A, Rosen S, Scott SK (2010) Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *J Neurosci* 30:7179–7186.
- Faulkner A, Rosen S, Wilkinson L (2001) Effects of the number of channels and

- speech-to-noise ratio on rate of connected discourse tracking through a simulated cochlear implant speech processor. *Ear Hear* 22:431–438.
- Foxe JJ, Snyder AC (2011) The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front Psychol* 2:154.
- Foxe JJ, Simpson GV, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9:3929–3933.
- Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R (2001) Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* 98:694–699.
- Haegens S, Osipova D, Oostenveld R, Jensen O (2010) Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Hum Brain Mapp* 31:26–35.
- Harris KC, Dubno JR, Keren NI, Ahlstrom JB, Eckert MA (2009) Speech recognition in younger and older adults: a dependency on low-level auditory cortex. *J Neurosci* 29:6078–6087.
- Heinrich A, Schneider BA (2011) Elucidating the effects of ageing on remembering perceptually distorted word pairs. *Q J Exp Psychol (Hove)* 64:186–205.
- Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69:387–396.
- Jacquemot C, Scott SK (2006) What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci* 10:480–486.
- Jacquemot C, Pallier C, LeBihan D, Dehaene S, Dupoux E (2003) Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J Neurosci* 23:9541–9546.
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186.
- Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex* 12:877–882.
- Kerlin JR, Shahin AJ, Miller LM (2010) Attentional gain control of ongoing cortical speech representations in a “cocktail party.” *J Neurosci* 30:620–628.
- Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev* 53:63–88.
- Laufs H, Kleinschmidt A, Beyerle A, Eger E, Salek-Haddadi A, Preibisch C, Krakow K (2003) EEG-correlated fMRI of human alpha activity. *Neuroimage* 19:1463–1476.
- Laufs H, Holt JL, Elfont R, Krams M, Paul JS, Krakow K, Kleinschmidt A (2006) Where the BOLD signal goes when alpha EEG leaves. *Neuroimage* 31:1408–1418.
- Leiberg S, Lutzenberger W, Kaiser J (2006) Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain Res* 1120:131–140.
- Lopes da Silva F (2004) Functional localization of brain sources using EEG and/or MEG data: volume conductor and source models. *Magn Reson Imaging* 22:1533–1538.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Medendorp WP, Kramer GF, Jensen O, Oostenveld R, Schoffelen JM, Fries P (2007) Oscillatory activity in human parietal and occipital cortex shows hemispheric lateralization and memory effects in a delayed double-step saccade task. *Cereb Cortex* 17:2364–2374.
- Miller KD, Troyer TW (2002) Neural noise can explain expansive, power-law nonlinearities in neural response functions. *J Neurophysiol* 87:653–659.
- Nenert R, Viswanathan S, Dubuc DM, Visscher KM (2012) Modulations of ongoing alpha oscillations predict successful short-term visual memory encoding. *Front Hum Neurosci* 6:127.
- Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol* 48:3637–3652.
- Obleser J, Eisner F (2009) Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci* 13:14–19.
- Obleser J, Weisz N (2012) Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex*. Advance online publication. Retrieved July 31, 2012. doi:10.1093/cercor/bhr325.
- Obleser J, Wise RJ, Alex Dresner M, Scott SK (2007) Functional integration across brain regions improves speech perception under adverse listening conditions. *J Neurosci* 27:2283–2289.
- Obleser J, Eisner F, Kotz SA (2008) Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J Neurosci* 28:8116–8123.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Paulesu E, Frith CD, Frackowiak RS (1993) The neural correlates of the verbal component of working memory. *Nature* 362:342–345.
- Pichora-Fuller MK, Singh G (2006) Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif* 10:29–59.
- Pichora-Fuller MK, Schneider BA, Daneman M (1995) How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 97:593–608.
- Piquado T, Cousins KA, Wingfield A, Miller P (2010) Effects of degraded sensory input on memory for speech: behavioral data and a test of biologically constrained computational models. *Brain Res* 1365:48–65.
- Pisoni DB (2000) Cognitive factors and cochlear implants: some thoughts on perception, learning, and memory in speech perception. *Ear Hear* 21:70–78.
- Rabbitt PM (1968) Channel-capacity, intelligibility and immediate memory. *Q J Exp Psychol* 20:241–248.
- Reas ET, Gimbel SI, Hales JB, Brewer JB (2011) Search-related suppression of hippocampus and default network activity during associative memory retrieval. *Front Hum Neurosci* 5:112.
- Rosen S, Faulkner A, Wilkinson L (1999) Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am* 106:3629–3636.
- Rudner M, Rönnerberg J, Lunner T (2011) Working memory supports listening in noise for persons with hearing impairment. *J Am Acad Audiol* 22:156–167.
- Sadaghiani S, Scheeringa R, Lehongre K, Morillon B, Giraud AL, Kleinschmidt A (2010) Intrinsic connectivity networks, alpha oscillations, and tonic alertness: a simultaneous electroencephalography/functional magnetic resonance imaging study. *J Neurosci* 30:10243–10250.
- Sander MC, Werkle-Bergner M, Lindenberger U (2012) Amplitude modulations and inter-trial phase stability of alpha-oscillations differentially reflect working memory constraints across the lifespan. *Neuroimage* 59:646–654.
- Scott SK, Blank CC, Rosen S, Wise RJ (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123:2400–2406.
- Scott SK, Rosen S, Lang H, Wise RJ (2006) Neural correlates of intelligibility in speech investigated with noise vocoded speech—a positron emission tomography study. *J Acoust Soc Am* 120:1075–1083.
- Shamma S (2001) On the role of space and time in auditory processing. *Trends Cogn Sci* 5:340–348.
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304.
- Sharp DJ, Scott SK, Wise RJ (2004) Monitoring and the controlled processing of meaning: distinct prefrontal systems. *Cereb Cortex* 14:1–10.
- Shinn-Cunningham BG, Best V (2008) Selective attention in normal and impaired hearing. *Trends Amplif* 12:283–299.
- Sternberg S (1966) High-speed scanning in human memory. *Science* 153:652–654.
- Taulu S, Kajola M, Simola J (2004) Suppression of interference and artifacts by the signal space separation method. *Brain Topogr* 16:269–275.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Weisz N, Hartmann T, Müller N, Lorenz I, Obleser J (2011) Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol* 2:73.
- Wingfield A, Tun PA, McCoy SL (2005) Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr Dir Psychol Sci* 14:144–148.