Acoustic Detail Guides Attention Allocation in a Selective Listening Task

Malte Wöstmann[1,2], Erich Schröger[3], and Jonas Obleser[1]

Abstract

■ The flexible allocation of attention enables us to perceive and behave successfully despite irrelevant distractors. How do acoustic challenges influence this allocation of attention, and to what extent is this ability preserved in normally aging listeners? Younger and healthy older participants performed a masked auditory number comparison while the EEG was recorded. To vary selective attention demands, we manipulated perceptual separability of spoken digits from a masking talker by varying acoustic detail (temporal fine structure). Listening conditions were adjusted individually to equalize stimulus audibility as well as the overall level of performance across participants. Accuracy increased, and response times decreased with more acoustic detail. The decrease in response times with more acoustic detail was stronger in the group of older participants. The onset of the distracting speech masker triggered a prominent contingent negative variation (CNV) in the EEG. Notably, CNV magnitude decreased parametrically with increasing acoustic detail in both age groups. Within identical levels of acoustic detail, larger CNV magnitude was associated with improved accuracy. Across age groups, neuropsychological markers further linked early CNV magnitude directly to individual attentional capacity. Results demonstrate for the first time that, in a demanding listening task, instantaneous acoustic conditions guide the allocation of attention. Second, such basic neural mechanisms of preparatory attention allocation seem preserved in healthy aging, despite impending sensory decline. ■

INTRODUCTION

Listening to one talker despite distracting speakers ("cocktail party problem"; Cherry, 1953) requires selective attention, that is, preferential processing of a specific signal at the expense of distractor signals (Kerlin, Shahin, & Miller, 2010). The demand on selective auditory attention is particularly high if listening conditions are compromised because of hearing loss (Tun, McCoy, & Wingfield, 2009) or signal degradation (Wild et al., 2012). It is unknown how and to what extent listeners of different ages retain the ability to flexibly allocate attention to changing stimulus acoustics. Here, the EEG was recorded to trace neural signatures of selective attention deployment, while younger (20–30 years) and older (60–70 years) healthy listeners performed an effortful selective listening task, in which varying degrees of acoustic degradation implicitly signaled task difficulty.

Fluctuations in cortical excitability have been proposed to regulate auditory selective attention (Lakatos et al., 2013; Schroeder & Lakatos, 2009), by lowering sensory thresholds for relevant stimuli. Cortical excitability is enhanced by the depolarization of pyramidal neurons, causing slow cortical potentials of negative amplitude in the EEG (He & Raichle, 2009). One well-studied slow potential is the contingent negative variation (CNV; Walter, Cooper, Aldridge, McCallum, & Winter, 1964), which occurs after a warning signal during the anticipation of an imperative stimulus (e.g., Chennu et al., 2013; Zanto et al., 2011). The CNV magnitude is lowered when participants' selective attention to task-relevant stimuli is impaired by distractors (Travis & Tecce, 1998; Tecce & Scheff, 1969; McCallum & Walter, 1968). In turn, larger CNV magnitudes at stimulus onset improve detectability of visual (O'Connell et al., 2009) and auditory (Rockstroh, Muller, Wagner, Cohen, & Elbert, 1993) targets. These findings suggest that CNV magnitude correlates with selective attention, possibly through an enhancement of excitability in task-relevant cortical neural networks (Raichle, 2011). It is thus a timely endeavor to exploit the CNV for a refined understanding of selective auditory attention in younger and older listeners.

To study the CNV in a well-controlled, nonetheless ecologically valid selective listening situation, participants performed an auditory number comparison task (Moyer & Landauer, 1967) masked by a distracting talker. To vary the effort of selective attention (Shinn-Cunningham & Best, 2008), perceptual separability of digits and masker was altered by parametrically degrading temporal fine structure (TFS; Moore, 2008)—an acoustic feature found highly relevant for listening against fluctuating maskers (Hopkins & Moore, 2009, 2010). Critically, the onset of the masker served as a warning stimulus in the present

---

[1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2]International Max Planck Research School on Neuroscience of Communication, Leipzig, Germany, [3]University of Leipzig, Germany

design, because the degree of acoustic degradation in the masker implicitly signaled task difficulty and allowed a graded allocation of attention to compensate for unfavorable acoustic conditions. Thus, the dependent neural measure in this study was the CNV evoked by the onset of the speech masker.

In this attention-demanding selective listening task, we expected improved performance with more preserved acoustic detail. Decreased CNV magnitude with more acoustic detail would indicate that participants adaptively allocate less attention as the signal quality improves. To further tighten the link between the CNV and mechanisms of auditory attention, we anticipated, first, absent or reduced CNV modulation in a control experiment when acoustic detail would not cue task difficulty and, second, a correlation between CNV magnitude and a behavioral marker of individual attentional capacity. Through careful adjustments of stimulus intensities to participants' individual requirements, we were able to investigate the neural mechanisms of auditory attention allocation independent of age differences in signal audibility or overall performance level. We asked whether healthy aging would affect the flexible allocation of attention to changing acoustic conditions.

METHODS

Participants

Twenty younger (age range = 20–30 years, mean age = 25.7 years; nine women) and twenty older (age range = 60–70 years, mean age = 64 years; 11 women) healthy, right-handed German native speakers participated in the main experiment. Data of 38 participants were included in the final analysis (see below). Participants gave informed consent and were financially compensated for participation. Procedures were approved by the local ethics committee of the University of Leipzig Medical Faculty.

Speech Materials

German-spoken digits from 21 to 99 (excluding multiples of 10) were recorded from a trained female speaker (sampling rate = 44.1 kHz). All digits contained four syllables and had an average length of 1.125 sec (SD = 0.056 sec). The distracting masker was extracted from a German audiobook (Oscar Wilde, "Der junge König") spoken by a female talker (sampling rate = 44.1 kHz). To increase the energetic overlap of masker and spoken digits, silent periods longer than 70 msec were removed automatically from the masker (using a customised MATLAB script R2013a; MathWorks, Inc., Natick, MA). The resulting audio file had a length of 29′52″, from which we extracted 1000 random snippets with a length of 6 sec.

For each stimulus, two spoken target digits (referred to as S1 and S2) and one masker snippet (referred to as masker) were selected randomly. Intensities of digits and masker were modified to realize different target-to-masker ratios (TMRs, which were individually titrated; see below). For this purpose, root-mean-squared (RMS) masker intensity was fixed at −30 dB full scale (dBFS), whereas digit intensity was further reduced (using the AttenuateSound function from the psychoacoustics toolbox for MATLAB). For example, for a TMR of −15 dBFS and given the masker intensity of −30 dBFS, intensities of S1 and S2 were set to RMS −45 dBFS. Finally, digit and masker signals were combined.

To modify the amount of acoustic detail (TFS), the combined signal (composed of masker and digits) was divided in frequency space into 16 overlapping channels (using a gammatone filterbank implemented in the auditory toolbox for MATLAB; Slaney, 1993). Channel center frequencies increased exponentially from 0.08 to 10 kHz. TFS was preserved in all channels below and including six TFS preservation cutoffs (0, 0.11, 0.21, 0.4, 0.76, and 1.45 kHz) and degraded above (Figure 1A). Thus, TFS was always degraded in channels above 1.45 kHz but was systematically degraded across conditions in channels at and below 1.45 kHz. We did not preserve TFS above 1.45 kHz, as we observed the largest performance increase up to this frequency in a behavioral pretest (n = 12). All channels below and including the TFS preservation cutoff were left unchanged (i.e., "intact"; Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). In channels above a given TFS preservation cutoff, the speech envelope (slow amplitude modulations < 16 Hz) was extracted using the Hilbert transform (Smith, Delgutte, & Oxenham, 2002). The envelope was used to modulate a sinusoidal tone with random starting phase at the channel center frequency. The resulting signal was filtered again with the initial filters to remove out-of-channel frequency components (Lunner, Hietkamp, Andersen, Hopkins, & Moore, 2012). The RMS amplitude of the signal in each channel was equalized to this channel's RMS after initial filtering. Finally, intact and modified channels were combined, yielding six different TFS preservation levels. Note that a TFS preservation of 0 kHz meant that TFS was entirely degraded in all 16 channels (Figure 1A, top), whereas a TFS preservation of 1.45 kHz meant that TFS was preserved in channels below and including 1.45 kHz and was degraded in all channels above (Figure 1A, bottom).

In essence, our manipulation substantially degraded the fast spectrotemporal fluctuations in higher frequencies, while leaving the slow temporal envelope fluctuations largely intact (Shamma & Lorenzi, 2013). Lower levels of TFS preservation made the signal sound tinny and artificial, rendering perceptual segregation of masker and digits perceptually more demanding. Importantly, speech with degraded TFS in all channels ("vocoded speech") is intelligible if presented in quiet, provided that the number of channels is sufficiently high (Obleser, Eisner, & Kotz, 2008; Obleser, Wise, Alex Dresner, & Scott, 2007; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995).
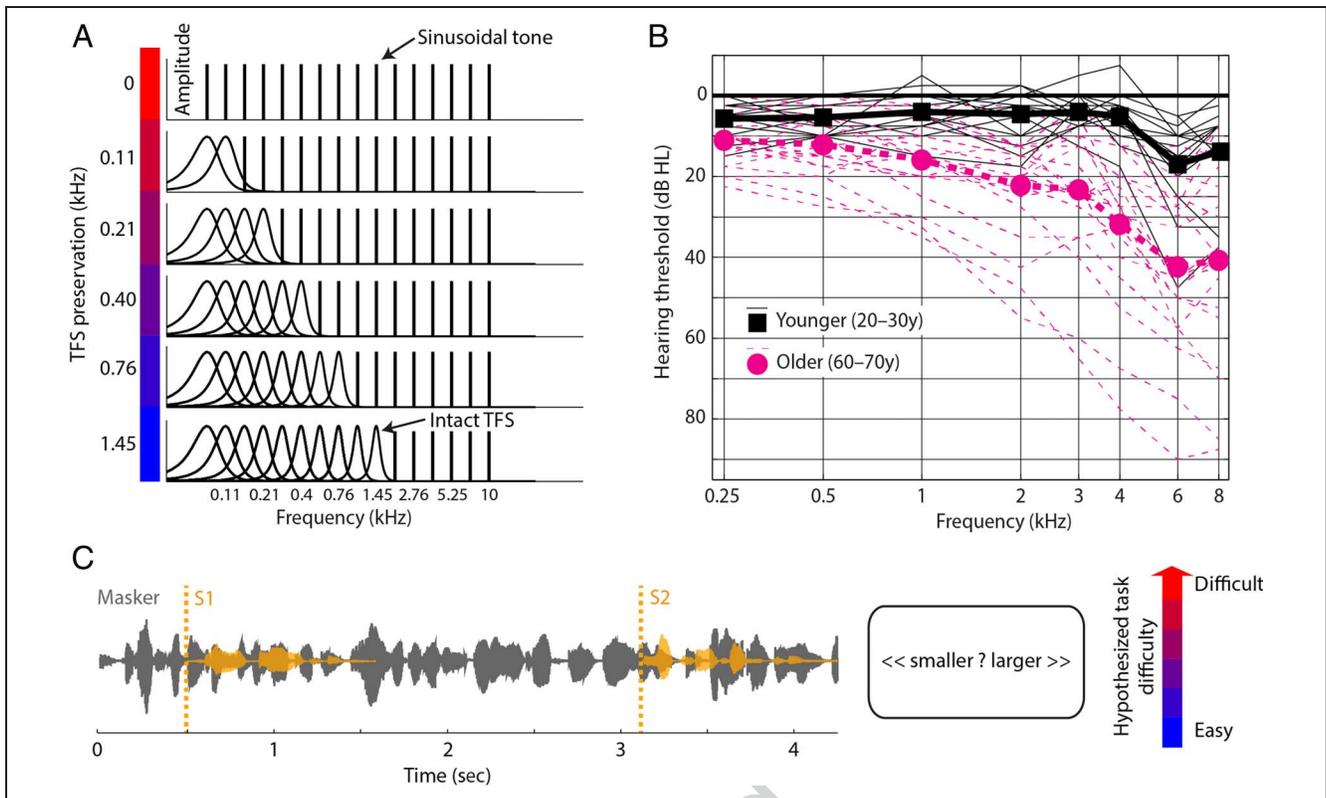
Figure 1. Manipulation of acoustic detail and hearing thresholds. (A) Schematic illustration of TFS manipulation. TFS was preserved in channels at lower frequencies and replaced by sinusoidal tones at higher frequencies. The TFS preservation level (highest channel with intact TFS) varied over six levels (0, 0.11, 0.21, 0.4, 0.76, and 1.45 kHz). Amplitudes in all channels are equalized for illustration purpose only. (B) Pure-tone air-conduction audiometric thresholds averaged over both ears for 18 younger (black) and 20 older (magenta) participants. Thick lines show average thresholds for the two age groups. (C) Trial design of the auditory number comparison task. Two spoken digits (S1 and S2) were presented against a distracting speech masker. Task difficulty was hypothesized to increase with lower levels of TFS preservation. y = years; HL = hearing level.

Hearing Acuity

To assess an objective measure of hearing acuity, participants' pure-tone air-conduction audiometric thresholds (at frequencies of 0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz) were assessed by a trained audiologist separately for both ears in steps of 5-dB hearing level using a clinical audiometer (according to the procedures described in British Society for Audiology, 2011). Participants did not show inter-aural asymmetries (≥20-dB difference between both ears at more than two frequencies). Hearing thresholds of younger and older participants are shown in Figure 1B. Notably, none of the participants were using a hearing aid nor were any of them subjectively aware of significant hearing impairments.

Individual Adjustments of Materials

One of the main rationales of this study was to investigate the effect of acoustic signal or age on attention allocation while controlling for potentially confounding between-subject differences in signal audibility or overall task performance level. Before the actual experiment, we thus adjusted stimulus intensities to individual requirements to assure a comparable level of task performance across

(younger and older) participants on stimulus materials under the most severe degradation (TFS preservation of 0 kHz). In the EEG experiment, we then systematically enhanced the degree of preserved TFS in stimulus materials. We explored in how far younger and older participants' behavioral responses and neural markers of attention allocation were sensitive to these changes in the degree of TFS preservation.

First, to equate audibility of materials despite considerable interindividual differences in hearing thresholds (Figure 1B), overall stimulus intensity was adapted to hearing abilities. To this end, a frequency-specific amplification based on hearing thresholds from 0.25 to 6 kHz was applied to all materials using the CAMEQ procedure (Moore, Alcantara, & Glasberg, 1998). In essence, this procedure raises signal intensity at frequencies that showed elevated hearing thresholds.

Second, because performance levels in auditory tasks cannot be matched between age groups by controlling for pure-tone audiometric thresholds alone (see Pichora-Fuller, Schneider, & Daneman, 1995), we individually adjusted the TMR (Schneider, Daneman, Murphy, & See, 2000). To this end, we varied the TMR systematically while participants performed the auditory number comparison task on materials without preserved TFS (0 kHz) in an

adaptive tracking procedure (two-down one-up procedure; targeting approximately 71% accuracy; Levitt, 1971). Testing started at a favorable TMR of +10 dB. This made it rather easy for all participants to perform the number comparison task initially. After two successive correct trials, TMR was decreased (two-down), reducing intelligibility of digits. After one incorrect trial, TMR was increased (one-up). Younger participants performed three sessions of adaptive tracking; and older participants, four sessions of adaptive tracking. The individual TMR used in the actual experiment was estimated from the average results of all tracking sessions.

Processing Speed

Processing speed was assessed with a standard visual test for attentional capacities (d2-R; Brickenkamp, Schmidt-Atzert, & Liepmann, 2010). Participants had to mark target letters in 12 lists containing targets and highly similar nontargets. They were instructed to perform the task "as fast and accurately as possible" and were given 20 sec to work on each list, after which they were prompted to switch immediately to the subsequent list. As a test score, we calculated the sum of processed targets on all lists ("BZO" score, possibly ranging between 0 and 308) with high scores indicating high processing speed (Bates & Lemay, 2004).

Working Memory

Working memory capacity was assessed with the auditory backward digit span (BSpan) test (subtest of the WAIS-Revised; Wechsler, 1984). On each trial, participants were presented a list of spoken digits between 1 and 9. Digits were spoken by a female voice at a rate of approximately one digit/sec and presented at ~75 dB SPL. Participants' task was to repeat the digits in reverse order. The test had seven levels with list lengths increasing from two to eight digits. Each level was composed of two items. Participants' responses were marked as correct only if all digits were repeated in the correct order. Testing stopped when participants performed incorrectly on both items for a particular list length. The individual BSpan score was calculated as the sum of correctly completed items, possibly ranging between 0 and 14.

Experimental Procedure

Participants were instructed to perform the number comparison as fast and accurately as possible. Each trial started with the presentation of the two response options ("kleiner," smaller; "größer," larger) on a computer screen. Auditory stimulation with the manipulated speech masker started after 1.5 sec. Spoken digits (S1 and S2) were placed 0.5 and 3.125 sec, respectively, after masker onset resulting in an average delay interval of 1.5 sec between S1

offset and S2 onset. All audio files ended simultaneously with S2 offset and had a length of ~4.25 sec (Figure 1C). Participants indicated via button press on a response box whether the second digit was smaller (left button pressed with left thumb) or larger (right button pressed with right thumb) than the first. Next, they rated their confidence in this response on a 3-point scale (1 = *unconfident*, 3 = *confident*). The next trial started self-paced with an additional button press. Behavioral data were recorded by Presentation software (Neurobehavioral Systems).

Each participant performed 300 trials, 50 for each TFS preservation level. For each trial, it was determined randomly whether the second digit was in fact smaller or larger than the first. The experiment was divided in five blocks. Each block contained 10 trials for each TFS preservation level in random order, meaning that the level of TFS preservation changed from trial to trial. Blocks were separated by short breaks. The experiment lasted approximately 70 min.

Behavioral Data Analysis

Data from two younger participants were excluded from all analyses because of technical problems during data acquisition and below-chance performance (38% correct), respectively.

To analyze differences in the individual adjustments of materials between age groups, the effect of age group on individually titrated TMR was analyzed with an independent-samples t test. The relationship between working memory and TMR was analyzed using a Pearson's correlation (Figure 2B).

To quantify participants' performance in the auditory number comparison task, accuracy on each trial (correct vs. incorrect) was weighted by confidence ratings to get a more fine-grained measure of task performance (Kitayama, 1991). As a result, correct responses were transformed to 100% weighted accuracy in case of high confidence ratings, to 80% in case of medium confidence, and to 60% in case of low confidence. Similarly, incorrect responses yielded 40% weighted accuracy for low confidence ratings, 20% for medium confidence, and 0% for high confidence ratings. In the remainder of this article, we use, for simplicity, the term "accuracy" to refer to accuracy weighted by confidence ratings. As a second measure of task performance, we analyzed participants' response times in the number comparison task. In detail, response times corresponded to the time interval between the onset of the second digit and participants' button press to indicate whether the second digit was smaller or larger than the first.

For statistical analyses, we calculated linear coefficients characterizing the linear change (slope) of accuracy and response times over the six levels of TFS preservation for each participant (predictor values: −2.5, −1.5, −0.5, 0.5, 1.5, 2.5; using the polyfit function in MATLAB). To test for significant effects of TFS preservation on performance
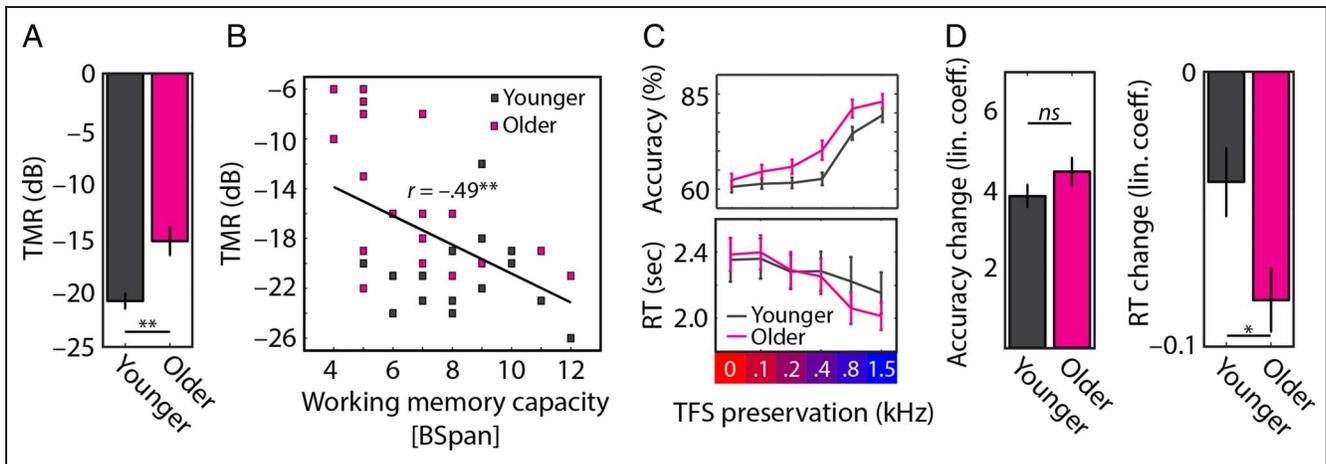
Figure 2. Stimulus adjustments and task performance. (A) Average TMR used for the individual stimulus adjustments for younger and older participants. (B) Scatterplot of TMR as a function of working memory capacity (auditory BSpan score) for younger (black) and older (magenta) participants. Note that only 33 of the 38 data points are visible as some points overlap. (C) Accuracy increased, and response times decreased in the auditory number comparison with higher levels of TFS preservation (both $ps < .001$). Accuracy was weighted by confidence ratings. (D) Bars show linear coefficients, which quantify the change in accuracy (left) and response times (right) with each level of TFS preservation. The speedup of response times with higher levels of TFS preservation was significantly stronger in older participants. *$p < .05$; **$p < .01$. Error bars show ±1 SE. lin. coeff. = linear coefficient.

measures, the distribution of linear coefficients was tested against zero (using a one-sample t test). To test for effects of age group, we compared younger and older participants' linear coefficients, overall (condition-independent) accuracy measures, and overall response times (using independent-samples t tests).

EEG Recording and Analyses

EEG was recorded at a 500-Hz sampling rate with a DC 135-Hz pass band (TMS international, Enschede, The Netherlands). Twenty-eight electrodes (Ag/Ag–Cl) were placed at the following positions (Easycap, Herrsching, Germany): Fpz, Fp1, Fp2, Fz, F3, F4, F7, F8, FC3, FC4, FT7, FT8, Cz, C3, C4, T7, T8, CP5, CP6, Pz, P3, P4, P7, P8, O1, O2, left mastoid (A1), and right mastoid (A2). The reference electrode was placed at the tip of the nose; and the ground electrode, at the sternum. The EOG was recorded from vertical and horizontal bipolar montages. All electrode resistances were kept below 5 kΩ.

Offline data were analyzed using MATLAB and the field-trip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Epochs were extracted from the continuous signal around masker onset (−2 to 6.5 sec). Epochs were low-pass filtered at 100 Hz and baseline corrected by subtracting the mean amplitude in the time interval of −0.1 to 0 sec. An independent component analysis was performed on the epoched data. Components corresponding to eye blinks, saccadic eye movements, muscle activity, electrode drifts, and heartbeat were identified and rejected by inspection of the components' topographies, frequency spectra, and time courses. Remaining artifact-contaminated trials were deleted after visual inspection of EEG waveforms at all electrodes. On average, $7 ± 1\%$ (SE) of trials were rejected from further analyses. Before statistical analyses, data
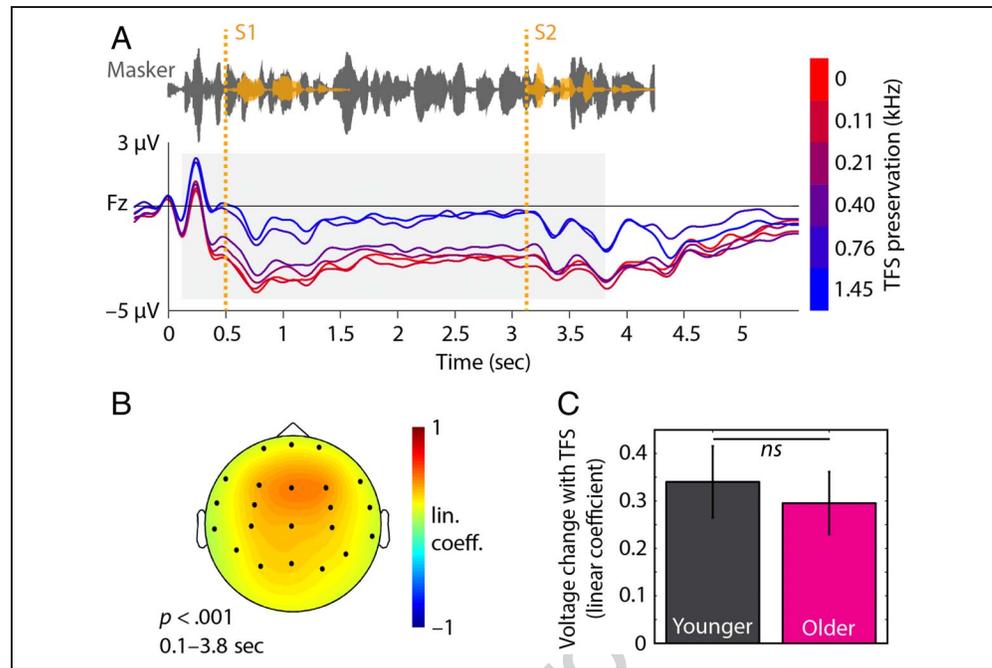
were further low-pass filtered at 20 Hz (fourth-order Butterworth filter, zero phase shift).

To calculate the ERP, the time-locked average over all artifact-free trials (irrespective of whether the number comparison was performed correctly or incorrectly) was computed separately for the six TFS preservation levels for each participant. To detect significant effects of TFS preservation on ERP amplitude, a two-level statistical analysis was applied (cf. Wilsch, Henry, Herrmann, Maess, & Obleser, 2014; Obleser, Wostmann, Hellbernd, Wilsch, & Maess, 2012). On the first (individual) level, EEG recordings from all trials at 28 scalp electrodes and between 0 and 4.25 sec (relative to masker onset) were submitted to a parametric regression t test for independent samples (implemented in the *ft_timelockstatistics* function in fieldtrip). For this regression, we used linearly spaced zero-centered predictor values (−2.5, −1.5, −0.5, 0.5, 1.5, 2.5) to model the monotonic change of ERP amplitude over six levels of TFS preservation. For each participant, we obtained an electrode–time matrix of linear coefficients characterizing the linear change (slope) of ERP amplitude with increasing TFS preservation.

On the second (group) level, individual matrices of linear coefficients were tested for significant differences from zero using a cluster-based permutation dependent samples t test (Maris & Oostenveld, 2007). First, this test clustered t values of adjacent points in electrode–time space with a $p < .05$, considering a minimum of three neighboring electrodes as a cluster. Next, the summed t value of each cluster was computed and compared against the distribution of 1000 iteratively and randomly drawn clusters from permuted-labels data. The cluster p value resulted from the proportion of Monte Carlo iterations in which the summed t statistic of the observed cluster was exceeded. As we performed this analysis as

Figure 3. CNV changes with acoustic detail. (A) Schematic trial procedure and grand-averaged ERPs at electrode Fz, averaged over all participants for six levels of TFS preservation. ERP waveforms are low-pass filtered at 5 Hz for illustration purpose only. The gray area highlights the period of the significant increase in ERP amplitude with lower levels of TFS preservation, as revealed by the cluster test (see text). (B) Topography shows average linear coefficients (quantifying the change in ERP amplitude with each level of TFS preservation) for significant electrode–time points in the cluster. Positive linear coefficients indicate that ERP amplitude became larger (i.e., less negative) with higher levels of TFS preservation. Black circles indicate electrodes belonging to the significant cluster (22 of the 28 scalp electrodes). (C) Average linear coefficients in the significant cluster did not differ between age groups. *ns*, not significant. Error bars show ±1 between-subject *SE*.

a two-sided test (for clusters exhibiting positive and negative effects), clusters with $p < .025$ were considered significant. Linear coefficients significantly larger than zero would indicate that ERP amplitude became more positive with higher levels of TFS preservation. The analysis revealed one extensive significant cluster (Figure 3).

To test whether the effect of TFS preservation on ERP amplitude in the significant cluster differed between age groups, individual linear coefficients were averaged over electrodes and time points of the significant cluster and submitted to an independent-samples *t* test with the between-subject factor age group (Figure 3C). To directly compare the two age groups in their exhibited ERP amplitude change with higher levels of TFS preservation during the entire trial (not only in the significant cluster), younger and older participants' individual matrices of linear coefficients were submitted to another cluster-based permutation independent samples *t* test (between-subject factor: age group).

To test whether CNV magnitude in individual trials was related to accuracy in the number comparison task, we performed a median split on single-trial CNV magnitude in the significant cluster. We calculated the mean accuracy for trials with a small and large CNV magnitude for each participant and level of TFS preservation (Figure 4). For statistical analysis, a repeated-measures ANOVA (within-subject factors: TFS preservation and CNV magnitude; between-subject factor: age group) was applied to these data.

We analyzed whether the magnitude of the CNV would correlate with neuropsychological markers of individual attentional capacity. To this end, we focused on the early CNV (0.1–0.5 sec) before S1 onset, which was independent of processing task-relevant digits but thought to reflect the preparatory allocation of attention for the ensuing number comparison task. We correlated overall early CNV magnitude (i.e., averaged over all electrodes of the significant cluster and over all conditions) with d2-R
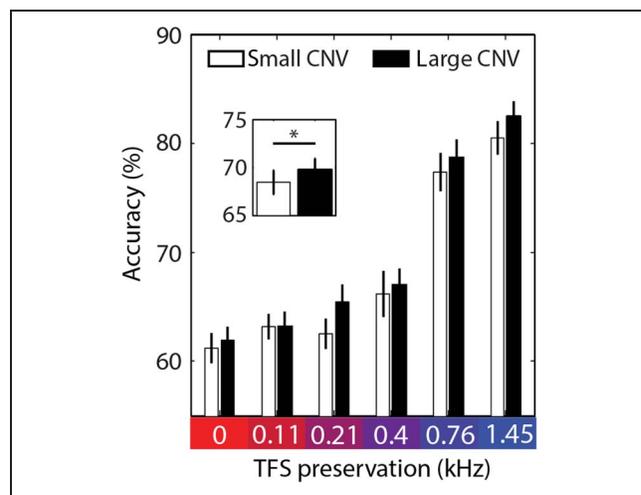

Figure 4. Larger CNV amplitude was associated with better task performance. Accuracy in the number comparison task across levels of TFS preservation was higher in trials with large (black bars) compared with small (white bars) CNV magnitude. The inset highlights this main effect of CNV magnitude on accuracy. Accuracy was weighted by confidence ratings. *$p < .05$.

Figure 5. Early CNV in the main and control experiments (*n* = 6). (A) Average ERPs of six participants relative to masker onset (0 sec) for six levels of TFS preservation in the main experiment, where the acoustic detail in the masker cued task difficulty. (B) ERPs in the control experiment, where acoustic detail in the masker was uninformative about task difficulty. ERP waveforms are low-pass filtered at 5 Hz for illustration purpose only. Gray-shaded areas indicate the time interval in which the acoustic stimulation in the main and control experiments was identical. This time interval of the early CNV (0.1–0.5 sec) was used for statistical analyses.
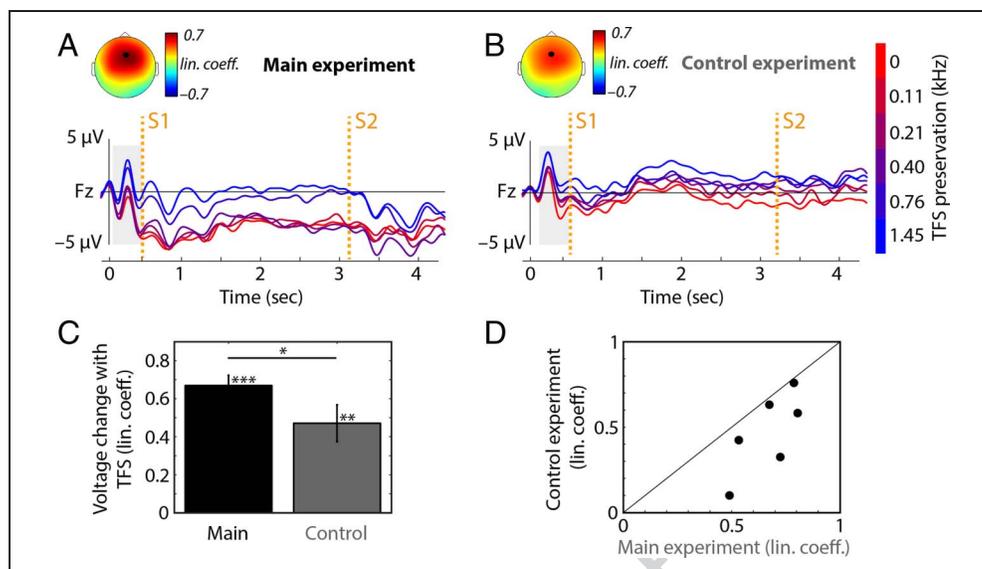
(C) Average linear coefficients capturing the change of early CNV magnitude with higher levels of TFS preservation in the main (black) and control (grey) experiments. Error bars show ±1 *SE*. (D) Scatterplot of linear coefficients from individual participants in the main and control experiments. Points below the diagonal show that the effect of TFS preservation on the early CNV were stronger in the main experiment compared with the control experiment. **p < .01; ***p < .001.

scores for processing speed. To control for a possible confound of entering two different groups of participants (younger and older) in one correlation analysis, we also controlled for the effect of age group in a partial correlation (Figure 6). Effects of age group on overall early CNV magnitude and d2-R scores were analyzed with independent-samples *t* tests.

Control Experiment

In a control experiment, we slightly altered the acoustic processing scheme to obtain masker signals identical to the main experiment, but to preserve the TFS of the spoken target digits. Masker and target digits were submitted to the TFS manipulation (Figure 1A) separately, such that acoustic detail (TFS) was only manipulated in the speech masker (over the same six levels as before) but was always preserved up to 1.45 kHz (i.e., maximally intact) in spoken digits.

We hypothesized that task difficulty would be unaffected by these varying masker signals because the task-relevant digits were always maximally intact. Thus, changing acoustic detail in the masker was expected to be no longer an indicative cue on task difficulty in the control experiment. All other experimental and analysis procedures, however, were identical to the main experiment. Importantly, the acoustic stimulation before S1 onset was physically identical in the main experiment and in the control experiment. Therefore, we restricted the analysis of ERP data to the time interval of the early CNV before S1 (0.1–0.5 sec). We reinvited six (three younger, three older) participants 8–12 months after participating

in the main experiment. All six had shown a prominent CNV effect in the main experiment (Figure 5A).

For statistical analysis, we computed average linear coefficients for the monotonic change in CNV amplitude with higher levels of TFS preservation before S1 onset at electrode Fz in the main and control experiments for each participant. This allowed us to quantify precisely the effect of acoustic detail on CNV in the individual, which allows for compelling within-subject comparisons despite the comparably low number of participants reinvited for the control experiment. Finally, distributions of linear coefficients from the main and control experiments were tested against zero (using one-sample *t* tests) and compared between the main and control experiments (using a paired *t* test).

RESULTS

Individual Adjustments of Speech Materials

Figure 2A shows younger and older participants' average TMR resulting from the individual adjustments of speech materials. As expected, average TMR was significantly lower for younger compared with older participants ($t(36) = 3.60, p = .001$), showing that younger participants were able to perform the number comparison task under more compromised acoustic conditions.

Figure 2B shows individual TMRs as a function of working memory capacity measured with the BSpan test. The correlation was significant ($r = -.49, p = .002$; controlling for age group: $p = .018$), indicating that participants with a smaller working memory capacity required a higher

TMR in the auditory number comparison task. When the correlation was computed separately for the two age groups, it reached significance only for older ($r = -.52, p = .018$) but not for younger ($r = -.03, p = .903$) participants, showing that the relationship between TMR and working memory capacity was mainly driven by the group of older participants. Generally, younger participants performed significantly better in the working memory test compared with older participants ($t(36) = 2.19, p = .035$).

Performance Profits from Acoustic Detail

Figure 2C shows response times and accuracy in the number comparison task for younger (black) and older (magenta) participants. Across age groups, participants showed significantly increasing accuracy ($t(37) = 17.81, p < .001$) and decreasing response times ($t(37) = -6.95, p < .001$) as more acoustic detail (TFS) was preserved. The TFS-induced improvement in accuracy did not differ significantly between age groups (Figure 2D; $t(36) = 1.35, p = .186$). Contrary, response times decreased significantly stronger with more TFS in older compared with younger participants ($t(36) = 2.53, p = .016$). Although Figure 2C indicates an overall higher accuracy for older participants, this main effect only approached significance ($t(36) = 1.95, p = .059$). When we analyzed age effects on performance measures separately for unweighted accuracy values and confidence ratings, we found that better performance in older adults was driven by higher overall unweighted accuracy ($t(36) = 2.47, p = .018$) rather than higher confidence ratings ($t(36) = 1.27, p = .211$). Overall response times did not differ significantly between age groups ($t(36) = 0.28, p = .783$).

CNV Magnitude Is Modulated by Acoustic Detail

Figure 3A shows the grand-averaged ERP for six levels of acoustic detail (i.e., TFS preservation). The onset of the speech masker triggered a sustained negative voltage deflection (CNV), which was smaller in magnitude for higher levels of TFS preservation. Notably, this CNV magnitude difference was sustained over the entire trial duration and declined after the offset of the acoustic stimulation.

Statistical analysis revealed one significant electrode–time cluster capturing the effect of decreasing CNV magnitude with more acoustic detail in speech materials ($p < .001$; Figure 3B). The cluster was composed a large number of mainly frontocentral electrodes and was significant from ~0.1 up to ~3.8 sec after masker onset (Figure 3A, gray-shaded area). This cluster exhibited a positive effect, indicating that CNV magnitude decreased (i.e., it became more positive in amplitude) with higher levels of TFS preservation. Linear coefficients in Figure 3B and C quantify the change in CNV magnitude (in μv) as TFS preservation was enhanced by one level. The effect of TFS preservation on CNV magnitude did not differ significantly

between younger and older participants ($t(36) = 0.47, p = .639$; Figure 3C).

One additional positive cluster approached significance ($p = .036$, with $\alpha = 0.025$ for two-sided testing). This cluster showed a topography similar to the significant cluster (Figure 3B) and appeared in the end of the trial after the significant cluster (3.85–4.25 sec). This cluster was not considered in further analyses. No significant clusters exhibiting a significant effect of age group on ERP amplitude change with higher levels of TFS preservation were found ($p > .1$ for all clusters).

CNV Magnitude Predicts Task Performance

Figure 4 shows participants' accuracy in the number comparison task separately for trials exhibiting a small or large CNV magnitude at electrodes and time points of the significant cluster. Across all six levels of TFS preservation, average accuracy was higher in those trials that showed a large CNV compared with trials with a small CNV. Statistical analysis revealed a significant main effect of CNV magnitude on accuracy ($F(1, 36) = 6.67, p = .014$). This main effect was also significant when we analyzed the impact of CNV amplitude on unweighted accuracy measures ($F(1, 36) = 7.89, p = .008$) and confidence ratings ($F(1, 36) = 8.24, p = .007$) separately. There were no significant two-way or three-way interactions between age group, TFS preservation, and CNV magnitude (all $ps > .05$).

Early CNV Dynamics and Cued Task Difficulty

An important finding in this study was that the significant cluster capturing the CNV effect became significantly well before the onset of the first digit (S1; Figure 3A, gray-shaded area). A critical question was whether this early CNV (0.1–0.5 sec) was a marker of cued task difficulty or of the acoustic detail in speech materials. In a control experiment, we thus tested to what degree the early CNV was modulated when acoustic detail was manipulated but cued task difficulty was held constant. To this end, acoustic detail varied only in the masker but was fixed in the target digits. Thus, varying acoustic detail in the masker should not cue task difficulty as task-relevant digits were always maximally intact. For the six participants tested in the control experiment, accuracy did not change with the degree of TFS preservation in the masker ($t(5) = -0.34, p = .75$; average accuracy = 54%, average unweighted accuracy = 61%) indicating constant task difficulty across conditions.

For the analysis of the early CNV, it was critical that the acoustic stimulation before S1 was identical in the main and control experiments. Thus, any difference in early CNV modulation between the main and control experiments could not be because of differences in the acoustic stimulation. Figure 5A and B show average CNVs ($n = 6$) for the main and control experiments, respectively. In the

main experiment, early CNV (0.1–0.5 sec) magnitude at electrode Fz decreased (i.e., amplitude became more positive) when more TFS was preserved in the speech materials ($t(5) = 12.49$, $p < .001$). Crucially, even in the control experiment, where task demands were constant over conditions, early CNV amplitude decreased with more preserved TFS ($t(5) = 4.85$, $p = .005$). This finding suggests that the early CNV is sensitive to varying degrees of preserved TFS in the masker even if varying acoustics do not cue task difficulty. Most important for this study however, the early CNV modulation in the main experiment, where preserved TFS cued task difficulty, was significantly stronger compared with that in the control experiment ($t(5) = 2.92$, $p = .033$; Figure 5C). In summary, the early CNV is sensitive to acoustic manipulations as such, but it is even modulated stronger if these acoustic manipulations implicitly cue task difficulty.

Figure 5D shows mean linear coefficients, quantifying the change in early CNV amplitude at electrode Fz with higher levels of TFS preservation, for each of the six participants in the main experiment contrasted with the control experiment. The fact that all points fall below the diagonal demonstrates that all six participants showed a stronger CNV modulation in the main experiment compared with the control experiment, indicating the high consistency of this effect across participants.

Early CNV Magnitude and Individual Attentional Capacity

Finally, we reasoned that the magnitude of the early CNV reflecting participants' attentional preparation for the ensuing number comparison task should be directly related to individual attentional capacity. Figure 6 shows overall (condition-independent) early CNV magnitude
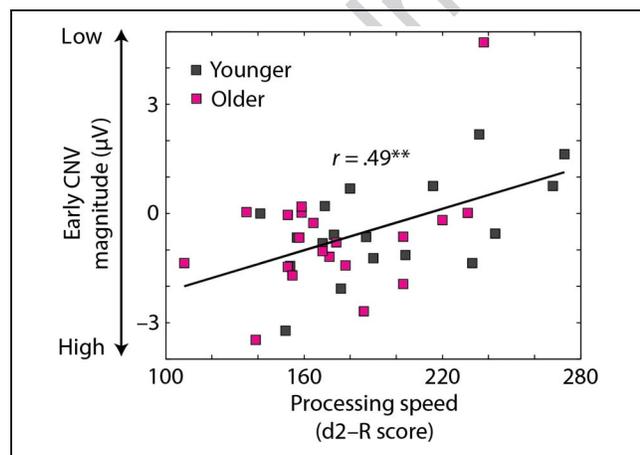

Figure 6. Processing speed predicts early CNV magnitude. Scatterplot of overall (i.e., condition-independent) early CNV magnitude (0.1–0.5 sec; averaged over all electrodes of the significant cluster) as a function of individual d2-R scores for processing speed for younger (black) and older (magenta) participants. As the CNV is a negative scalp potential, smaller μv values on the *y* axis indicate higher CNV magnitude. **$p < .01$.

(0.1–0.5 sec) in the main experiment as a function of d2-R scores for processing speed, an established neuro-psychological marker for attentional capacity. The correlation was significant ($r = .49$, $p = .002$; controlling for age group: $p = .002$), indicating that participants with higher processing speed showed smaller (i.e., more positive) early CNV magnitudes. As is discernible from the scatterplot in Figure 6, younger and older participants overlapped largely in both measures of early CNV magnitude and d2-R scores. Statistical analyses revealed no significant difference of early CNV magnitude between age groups ($t(36) = 0.58$, $p = .568$) but a tendency for higher d2-R scores in younger participants ($t(36) = 1.92$, $p = .063$).

DISCUSSION

How flexibly can changing acoustics trigger the allocation of attention in a selective listening situation, and how is this attention allocation process affected by healthy aging? Here, we tested the hypothesis that variations in the instantaneous acoustic conditions would signal task difficulty and implicitly cue the allocation of attention in younger (20–30 years) and older (60–70 years) participants. EEG recordings of the CNV served as an index of auditory selective attention.

Acoustic Detail Guides the Allocation of Attention

The most important finding was a strong dependence of CNV magnitude on preserved acoustic detail (TFS) in speech materials. This is a new observation extending previous knowledge on the CNV as a marker of attention allocation: It demonstrates, first, that CNV magnitude is directly and parametrically dependent on the TFS of the acoustic signal; second, however, this dependency is modulated by the task relevance of this acoustic cue itself (see control experiment and in-depth discussion below).

As acoustic detail was parametrically preserved from the low frequencies, participants' task performance improved (Figure 2C), and CNV magnitude decreased (Figure 3). These findings suggest that, when the perceptual segregation of digits and masker became less effortful because of more preserved TFS (Hopkins & Moore, 2009, 2010; Hopkins, Moore, & Stone, 2008; Moore, 2008), the task was less attention demanding as reflected in smaller CNV magnitude (Chennu et al., 2013; Zanto et al., 2011; Travis & Tecce, 1998; Tecce, Savignano-Bowman, & Meinbresse, 1976; Wilkinson & Ashby, 1974; Tecce & Scheff, 1969; McCallum & Walter, 1968). On a neuronal level, enhanced CNV magnitude in conditions with less acoustic detail could reflect a lowering of perceptual thresholds through an enhanced cortical excitability in task-relevant cortical networks (Raichle, 2011; He & Raichle, 2009; O'Connell et al., 2009; Rockstroh et al., 1993). In line with this interpretation, combined EEG–fMRI studies revealed a positive relationship between

BOLD activity and CNV magnitude (Scheibe, Ullsperger, Sommer, & Heekeren, 2010; Hinterberger et al., 2005; Nagai et al., 2004), suggesting an enhanced information flow between thalamus and cortex during the CNV period. Our finding of improved task performance in trials with a large CNV magnitude (Figure 4) further supports the view that a larger CNV indicates increased selective attention, which, in turn, leads to improved processing of auditory targets embedded in a speech masker.

Figure 3A shows that the significant modulation of the ERP started as early as 0.1 sec after masker onset, covering the time range of early auditory evoked potentials (N1 and P2; Picton & Hillyard, 1974). Statistical analysis revealed only a single electrode–time cluster exhibiting a significant effect of acoustic detail covering almost the entire trial (foreperiod, target encoding, and retention), as it is typical for slow cortical potentials like the CNV. This finding suggested that the CNV was superimposed on early ERP components, and we did not analyze these early evoked potentials in isolation. Instead, we focused largely on the early CNV, emerging right after the onset of the speech masker but before the onset of the first digit (S1). Critically, the early CNV was independent of processing task-relevant digits but thought to solely reflect participants' preparation for the number comparison task. In trials with minimal preserved acoustic detail, the speech masker before S1 onset implicitly cued a high task difficulty. Listeners could take advantage of this implicit cue and allocate more selective attention to overcome the unfavorable acoustic conditions. We presumed that the early CNV modulation (0.1–0.5 sec) reflected participants' graded allocation of auditory attention as the speech masker implicitly signaled task difficulty.

However, this interpretation implies that the early CNV modulation as a function of acoustic detail should be significantly reduced if acoustic detail in the masker does not cue task difficulty. To test this hypothesis, we conducted a control experiment (Figure 5) in which acoustic detail of the masker did not cue task difficulty. In the control experiment, performance did not improve with more acoustic detail showing that task difficulty was unaffected by acoustic detail. Most importantly, the early CNV effect was significantly stronger when acoustic detail cued task difficulty (main experiment), compared with a setting where acoustic detail was uninformative about task difficulty (control experiment). The fact that this pattern of results was consistent over all participants tested in the control experiment (Figure 5D) justifies the relatively small sample of six participants in the control experiment. In general, this finding corroborates our conjecture that the early CNV is an indicator of preparatory selective attention allocation triggered by expected task difficulty.

In the control experiment where acoustic detail varied but did not cue task difficulty, the early CNV effect was decreased but not entirely absent. It is thus conceivable, in line with previous research, that degraded acoustic conditions automatically increase the allocation of atten-tion (Obleser et al., 2012; Obleser & Weisz, 2012) even if the degradation applies only to task-irrelevant materials (Winkler, Teder-Salejarvi, Horvath, Näätänen, & Sussman, 2003). Note that, in everyday listening situations, acoustic degradations resulting from reverberations, background noise, or phone lines apply to all transmitted signals (target and masking signals). Therefore, an automatic increase in the allocation of auditory selective attention in adverse acoustic conditions is an effective mechanism to compensate for compromised acoustic conditions.

One important point in our study is to consider whether the observed negative voltage deflection (Figure 3) can indeed be considered a CNV. In most classical CNV paradigms, a warning stimulus triggers a negative-going CNV that peaks at the expected time point of a later-occurring target stimulus. In our study, however, the warning stimulus (masker onset) was followed by two consecutive target stimuli (S1 and S2). The early occurrence of the first target stimulus 0.5 sec after masker onset is a possible reason why our negative voltage deflection did not considerably increase in magnitude after S1 onset. Besides, although the CNV, in its narrow sense, varies with changing "attention to" or "anticipation of" a target stimulus, our negative voltage deflection was also sensitive to changes of acoustic detail alone (control experiment, Figure 5). As described above, we consider it likely that more adverse listening conditions automatically enhanced the allocation of attention, reflected in a stronger negative voltage deflection. Finally, our negative voltage deflection shows a number of properties of typical CNVs because it (a) shows up as a sustained negative voltage deflection strongest over frontocentral electrode sites, (b) is associated with improved task performance if its higher in magnitude (Figure 4), and (c) could be directly linked to markers of selective attention (Figure 6). Thus, despite the fact that our negative voltage deflection differs slightly from the classical CNV in the narrow sense, we still consider it appropriate to be referred to as a CNV.

Early CNV Magnitude Reflects Individual Attentional Capacities

Evidence for a close relation between individual cognitive capacities and the magnitude of slow cortical potentials (see also Vogel, McCollough, & Machizawa, 2005) was given by the significant correlation of overall (condition-independent) early CNV magnitude and the d2-R score for processing speed (Figure 6; Brickenkamp et al., 2010). In the d2-R test, visual target items compete with highly similar distractors for limited processing resources (Bates & Lemay, 2004; Desimone & Duncan, 1995). Better participants succeed at selectively attending to targets while ignoring distractors. They can thus process more target items and achieve higher d2-R scores. Here, participants with good selective attention abilities showed smaller (i.e., more positive) overall early CNV magnitudes. Generally, this finding adds weight to the

interpretation of the early CNV as a direct electrophysiological index of preparatory selective attention allocation. In particular, this result suggests that the effort of selective attention in a demanding listening task was lower for participants with higher selective attention abilities. In conclusion, the strong link between attentional capacities and CNV magnitude emphasizes the importance of taking into account individual cognitive capabilities for the investigation and treatment of subject-specific listening abilities in acoustically demanding situations.

Age Affects Required Acoustic Conditions and Response Times

In contrast to prior studies, which found age differences both in CNV dynamics (Zanto et al., 2011; Loveless & Sanford, 1974) and in the accuracy of detecting changes in TFS (Hopkins & Moore, 2011; Grose & Mamo, 2010), we found age effects rather in the individual adjustments of speech materials required before experimental testing and in response times. First, for several older participants, hearing acuity was reduced (especially at higher frequencies) compared with younger participants (Figure 1B). As overall stimulus intensities were adjusted to individual hearing thresholds (CAMEQ procedure; Moore et al., 1998), these older participants were listening to overall more amplified materials during the experiment. Second, older participants required, on average, a significantly higher TMR to reach a similar performance level as younger participants (Figure 2A). This result confirms prior research showing that older listeners usually require higher signal-to-noise ratios to hear individual words in noise than do younger listeners (Pichora-Fuller, 2003; Schneider et al., 2000; Murphy, McDowd, & Wilcox, 1999; Pichora-Fuller et al., 1995). The need for less attention-demanding listening conditions in older participants might speak for a decline in attentional control, causing difficulties in attending relevant and ignoring irrelevant sound sources (Passow et al., 2014). Third, the speedup of response times with higher levels of TFS preservation was stronger in older compared with younger participants (Figure 2C and D). Thus, older participants show an enhanced sensitivity to changes in spectral detail (see also Schvartz, Chatterjee, & Gordon-Salant, 2008), implying that older listeners' task performance is particularly dependent on stimulus-inherent features in the acoustic materials. However, as we did not find concomitant differences in CNV dynamics between age groups, it is an open issue for future studies to relate this difference in behavior to neural changes in the elderly.

The finding that older participants performed poorer in the auditory working memory test (BSpan) compared with younger participants confirms the general trajectory of decline in memory functioning with age (Fisk & Warr, 1996; Salthouse & Kersten, 1993). More important, however, individual working memory capacity significantly predicted the relative intensity of spoken digits (TMR) determined in the individual adjustments of stimulus materials (Figure 2B). Participants with a smaller working memory capacity required more favorable acoustic conditions (higher TMR) to perform the number comparison task. Research has shown that limited resources of the working memory system must be allocated to processing and temporary maintenance and manipulation of speech information (Lunner, Rudner, & Ronnberg, 2009; McCoy et al., 2005). We presume that participants with fewer memory resources required more favorable encoding conditions to free resources needed for the retention and numerical comparison of digits. In general, this finding demonstrates the tight link between sensory and higher cognitive abilities (Li & Lindenberger, 2002). In summary, aging in and by itself is not critically affecting the ability to allocate attention in a task-adaptive manner, as long as listening conditions are adjusted to individual sensory acuity and working memory capacity.

Conclusions

Dynamics of the early CNV reveal that the instantaneous acoustic conditions in a selective listening task cue the adaptive allocation of auditory selective attention (Fritz, Elhilali, David, & Shamma, 2007) in younger and older listeners. This preparatory allocation of attention for an ensuing task is shown to be partly automatic (driven by characteristics of the signal), but it depends to large extents on the expected task difficulty conveyed by the signal itself (Figure 5). The effort of selective attention allocation during the task depended on listeners' individual selective attention abilities (Figure 6). Listeners' age is not critically affecting these processes, as long as listening conditions are adjusted to individual sensory acuity and working memory capacity, suggesting that basic mechanisms of preparatory attention allocation are preserved in healthy aging.

REFERENCES

Bates, M. E., & Lemay, E. P., Jr. (2004). The d2 test of attention: Construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society, 10,* 392–400.

Brickenkamp, R., Schmidt-Atzert, L., & Liepmann, D. (2010). *d2-revision aufmerksamkeits-und konzentrationstest.* Göttingen, Germany: Hogrefe Verlag.

British Society for Audiology. (2011). *Recommended procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking.* UK: Berkshire.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibanez, A., et al. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience, 33,* 11194–11205.

Cherry, C. E. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Fisk, J. E., & Warr, P. (1996). Age and working memory: The role of perceptual speed, the central executive, and the phonological loop. *Psychology and Aging, 11,* 316–323.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology, 17,* 437–455.

Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and Hearing, 31,* 755–760.

He, B. J., & Raichle, M. E. (2009). The fMRI signal, slow cortical potential and consciousness. *Trends in Cognitive Sciences, 13,* 302–309.

Hinterberger, T., Veit, R., Wilhelm, B., Weiskopf, N., Vatine, J. J., & Birbaumer, N. (2005). Neuronal mechanisms underlying control of a brain–computer interface. *European Journal of Neuroscience, 21,* 3169–3181.

Hopkins, K., & Moore, B. C. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *Journal of the Acoustical Society of America, 125,* 442–446.

Hopkins, K., & Moore, B. C. (2010). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *Journal of the Acoustical Society of America, 127,* 1595–1608.

Hopkins, K., & Moore, B. C. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *Journal of the Acoustical Society of America, 130,* 334–349.

Hopkins, K., Moore, B. C., & Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *Journal of the Acoustical Society of America, 123,* 1140–1153.

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party." *Journal of Neuroscience, 30,* 620–628.

Kitayama, S. (1991). Impairment of perception by positive and negative affect. *Cognition & Emotion, 5,* 255–274.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron, 77,* 750–761.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49(Suppl. 2),* 467+.

Li, K. Z., & Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience & Biobehavioral Reviews, 26,* 777–783.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 18866–18869.

Loveless, N. E., & Sanford, A. J. (1974). Effects of age on the contingent negative variation and preparatory set in a reaction-time task. *Journal of Gerontology, 29,* 52–63.

Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K., & Moore, B. C. (2012). Effect of speech material on the benefit of temporal fine structure information in speech for young normal-hearing and older hearing-impaired participants. *Ear and Hearing, 33,* 377–388.

Lunner, T., Rudner, M., & Ronnberg, J. (2009). Cognition and hearing aids. *Scandinavian Journal of Psychology, 50,* 395–403.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190.

McCallum, W. C., & Walter, W. G. (1968). The effects of attention and distraction on the contingent negative variation in normal and neurotic subjects. *Electroencephalography and Clinical Neurophysiology, 25,* 319–329.

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology: A, Human Experimental Psychology, 58,* 22–33.

Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology, 9,* 399–406.

Moore, B. C., Alcantara, J. I., & Glasberg, B. R. (1998). Development and evaluation of a procedure for fitting multi-channel compression hearing aids. *British Journal of Audiology, 32,* 177–195.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215,* 1519–1520.

Murphy, D. R., McDowd, J. M., & Wilcox, K. A. (1999). Inhibition and aging: Similarities between younger and older adults as revealed by the processing of unattended auditory information. *Psychology and Aging, 14,* 44–59.

Nagai, Y., Critchley, H. D., Featherstone, E., Fenwick, P. B., Trimble, M. R., & Dolan, R. J. (2004). Brain activity relating to the contingent negative variation: An fMRI investigation. *NeuroImage, 21,* 1232–1241.

Obleser, J., Eisner, F., & Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *Journal of Neuroscience, 28,* 8116–8123.

Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex, 22,* 2466–2477.

Obleser, J., Wise, R. J., Alex Dresner, M., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience, 27,* 2283–2289.

Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience, 32,* 12376–12383.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience, 29,* 8604–8611.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011,* 156869.

Passow, S., Westerhausen, R., Hugdahl, K., Wartenburger, I., Heekeren, H. R., Lindenberger, U., et al. (2014). Electrophysiological correlates of adult age differences in attentional control of auditory processing. *Cerebral Cortex, 24,* 249–260.

Pichora-Fuller, M. K. (2003). Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension. *International Journal of Audiology, 42(Suppl. 1),* S59–S67.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America, 97,* 593–608.

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials: II. Effects of attention. *Electroencephalography and Clinical Neurophysiology, 36,* 191–199.

Raichle, M. E. (2011). The restless brain. *Brain Connectivity, 1,* 3–12.

Rockstroh, B., Muller, M., Wagner, M., Cohen, R., & Elbert, T. (1993). "Probing" the nature of the CNV. *Electroencephalography and Clinical Neurophysiology, 87,* 235–241.

Salthouse, T. A., & Kersten, A. W. (1993). Decomposing adult age differences in symbol arithmetic. *Memory & Cognition, 21,* 699–710.

Scheibe, C., Ullsperger, M., Sommer, W., & Heekeren, H. R. (2010). Effects of parametrical and trial-to-trial variation in prior probability processing revealed by simultaneous electroencephalogram/functional magnetic resonance imaging. *Journal of Neuroscience, 30,* 16709–16717.

Schneider, B. A., Daneman, M., Murphy, D. R., & See, S. K. (2000). Listening to discourse in distracting settings: The effects of aging. *Psychology and Aging, 15,* 110–125.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences, 32,* 9–18.

Schvartz, K. C., Chatterjee, M., & Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *Journal of the Acoustical Society of America, 124,* 3972–3988.

Shamma, S., & Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *Journal of the Acoustical Society of America, 133,* 2818–2833.

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

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification, 12,* 283–299.

Slaney, M. (1993). *An efficient implementation of the Patterson-Holdsworth Auditory Filter Bank.* Cupertino, CA.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature, 416,* 87–90.

Tecce, J. J., Savignano-Bowman, J., & Meinbresse, D. (1976). Contingent negative variation and the distraction—Arousal hypothesis. *Electroencephalography and Clinical Neurophysiology, 41,* 277–286.

Tecce, J. J., & Scheff, N. M. (1969). Attention reduction and suppressed direct-current potentials in the human brain. *Science, 164,* 331–333.

Travis, F., & Tecce, J. J. (1998). Effects of distracting stimuli on CNV amplitude and reaction time. *International Journal of Psychophysiology, 31,* 45–50.

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging, 24,* 761–766.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438,* 500–503.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature, 203,* 380–384.

Wechsler, D. (1984). *Manual for the Wechsler Adult Intelligence Scale-Revised.* San Antonio, TX: Psychological Corp.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience, 32,* 14010–14021.

Wilkinson, R. T., & Ashby, S. M. (1974). Selective attention, contingent negative variation and the evoked potential. *Biological Psychology, 1,* 167–179.

Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., & Obleser, J. (2014). Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cerebral Cortex.*

Winkler, I., Teder-Salejarvi, W. A., Horvath, J., Näätänen, R., & Sussman, E. (2003). Human auditory cortex tracks task-irrelevant sound sources. *NeuroReport, 14,* 2053–2056.

Zanto, T. P., Pan, P., Liu, H., Bollinger, J., Nobre, A. C., & Gazzaley, A. (2011). Age-related changes in orienting attention in time. *Journal of Neuroscience, 31,* 12461–12470.