NeuroImage 196 (2019) 261–268

Contents lists available at ScienceDirect

NeuroImage

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

Hearing-impaired listeners show increased audiovisual benefit when listening to speech in noise

Sebastian Puschmann [a,b,c,*,1], Mareike Daeglau [d,e,1], Maren Stropahl [d], Bojana Mirkovic [c,d], Stephanie Rosemann [b,c], Christiane M. Thiel [b,c,f], Stefan Debener [c,d,f]

[a] Montreal Neurological Institute, McGill University, Montreal, Quebec, H3A 2B4, Canada
[b] Biological Psychology Lab, Department of Psychology, School of Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany
[c] Cluster of Excellence Hearing4All, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany
[d] Neuropsychology Lab, Department of Psychology, School of Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany
[e] Neurocognition and Functional Neurorehabilitation Group, Department of Psychology, School of Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany
[f] Research Center Neurosensory Science, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany

ARTICLE INFO

Keywords:
Audiovisual speech
Presbycusis
Cross-modal plasticity
Multisensory processing
EEG

ABSTRACT

Recent studies provide evidence for changes in audiovisual perception as well as for adaptive cross-modal auditory cortex plasticity in older individuals with high-frequency hearing impairments (*presbycusis*). We here investigated whether these changes facilitate the use of visual information, leading to an increased audiovisual benefit of hearing-impaired individuals when listening to speech in noise. We used a naturalistic design in which older participants with a varying degree of high-frequency hearing loss attended to running auditory or audiovisual speech in noise and detected rare target words. Passages containing only visual speech served as a control condition. Simultaneously acquired scalp electroencephalography (EEG) data were used to study cortical speech tracking. Target word detection accuracy was significantly increased in the audiovisual as compared to the auditory listening condition. The degree of this audiovisual enhancement was positively related to individual high-frequency hearing loss and subjectively reported listening effort in challenging daily life situations, which served as a subjective marker of hearing problems. On the neural level, the early cortical tracking of the speech envelope was enhanced in the audiovisual condition. Similar to the behavioral findings, individual differences in the magnitude of the enhancement were positively associated with listening effort ratings. Our results therefore suggest that hearing-impaired older individuals make increased use of congruent visual information to compensate for the degraded auditory input.

1. Introduction

A lasting deprivation from auditory sensory input results in a recruitment of auditory sensory brain regions by the other sensory modalities (Butler and Lomber, 2013; Kral, 2007; Sandmann et al., 2012). Although cross-modal plasticity has been mostly studied in populations with severe and early-onset hearing impairment, recent work in animals provides evidence that similar effects can also occur following moderate and partial hearing loss at adult age (Meredith et al., 2012; Schormans et al., 2017). Complementing these animal data, studies in older human individuals with high-frequency hearing impairments (*presbycusis*) found hearing-loss related shifts of visual and somatosensory responses towards auditory cortex (Campbell and Sharma, 2013; Cardon and Sharma, 2018), changed responses to audiovisual stimulation (Musacchia et al., 2009; Stropahl and Debener, 2017), and increased functional connectivity between auditory and visual sensory brain regions (Puschmann and Thiel, 2017). Behavioral data suggest that these neural changes go along with increased cross-modal distractibility (Guerreiro and Van Gerven, 2017; Puschmann et al., 2014) and a higher susceptibility to the McGurk illusion (Rosemann and Thiel, 2018; Stropahl and Debener, 2017), providing evidence for changed audiovisual processing in hearing-impaired individuals. The perceptual consequences of these

Abbreviations: MoCA, Montreal Cognitive Assessment.

* Corresponding author. Montreal Neurological Institute, Cognitive Neuroscience Unit, 3801 Rue University, Montreal, QC, H3A 2B4, Canada.
E-mail address: sebastian.puschmann@mail.mcgill.ca (S. Puschmann).

[1] These authors contributed equally to the project.

https://doi.org/10.1016/j.neuroimage.2019.04.017
Received 22 September 2018; Received in revised form 2 April 2019; Accepted 4 April 2019
Available online 9 April 2019
1053-8119/

effects for daily-life listening are however not understood at present.

It is assumed that older hearing-impaired listeners make extended use of visual input to facilitate speech perception in adverse listening situations (Hallam and Corney, 2014; Winneke and Phillips, 2011). Previous work has demonstrated that congruent visual input, such as viewing a speaker's face, significantly benefits speech-in-noise perception, both in normal-hearing and hearing-impaired individuals (Ross et al., 2006; Sumby and Pollack, 1954; Tye-Murray et al., 2007a). On the neural level, visual stimulation has been reported to speed up the cortical processing of speech (Arnal et al., 2009; van Wasserhove et al., 2005; Winneke and Phillips, 2011), to reduce the variability in firing patterns (Kayser et al., 2010), and to increase the neural entrainment to the speech envelope (Crosse et al., 2015), thus leading to a more robust representation of the speech signal in auditory sensory brain regions. Congruent visual input can restore the early cortical tracking of speech under severe levels of background noise and may facilitate auditory stream segregation (Atilgan et al., 2018; Crosse et al., 2016b; Zion Golumbic et al., 2013).

Cross-modal plasticity in deaf individuals has been related to superior perceptual performance in the visual and somatosensory modality, thus compensating for the deprived auditory input (Hauthal et al., 2013; Shiell et al., 2014; Stropahl et al., 2015). Building upon this view, cross-modal plasticity and changed audiovisual processing in older individuals with high-frequency hearing impairment may potentially facilitate the use of visual information, thus leading to an enhanced audiovisual benefit. In line with this, Altieri and Hudock (2014) indeed reported a positive relationship between the degree of hearing loss and the audiovisual enhancement in the perception of degraded speech. Other studies however found no such effect (Rosemann and Thiel, 2018; Tye-Murray et al., 2007a). We here reinvestigated this hypothesis using a naturalistic listening paradigm in which older participants with varying degrees of hearing loss attended to a continuous speech stream in background noise and detected rare target words. To account for hearing problems in noise in individuals with normal hearing thresholds (Alicea and Doherty, 2017; Badri et al., 2011), standard audiometry was complemented by subjective ratings of listening effort in different daily-life situations.

Recent work has demonstrated that the cortical tracking of speech streams in noise and audiovisual benefits on speech tracking can be studied in naturalistic settings (Crosse et al., 2015; Ding and Simon, 2013; Zion Golumbic et al., 2013). Moreover, Crosse et al. (2016b) reported that the magnitude of audiovisual enhancements in cortical speech tracking was directly related to audiovisual benefits in listening performance, as measured using an online target word detection task. Following up on these studies, we here used electroencephalography (EEG) and a speech envelope reconstruction method (Crosse et al., 2016a) to assess speech envelope tracking in older participants while they listened to continuous sequences of auditory or audiovisual speech in noise and detected rare target words. Sequences containing only visual speech served as a control for changed visual information processing in hearing-impaired listeners. If hearing-impaired individuals benefit more from added visual input than normal-hearing older listeners, we expected to observe an increasing audiovisual benefit in both target word detection accuracy and the cortical tracking of speech with increasing levels of high-frequency hearing loss and/or subjective listening effort ratings.

2. Material and methods

2.1. Participants

Forty older volunteers (18 females, age (mean ± standard deviation): 63 ± 6 years, age range: 52–77 years) with a varying degree of bilateral high-frequency hearing loss participated in the experiment. None of the hearing-impaired individuals was using a hearing aid. All participants were right-handed and reported normal or corrected to normal vision. The experimental procedures were approved by the ethics committee of the University of Oldenburg "Kommission für Forschungsfolgenabschätzung und Ethik" (Committee for research outcome assessment and ethics) and written informed consent was obtained from all participants.

2.2. Audiometric and cognitive pre-testing

A standard pure tone audiogram (frequency range: 125–8000 Hz) was obtained for each participant. Fig. 1A depicts the individual hearing thresholds, averaged over both ears. The average hearing threshold measured between 2000 and 8000 Hz, averaged over both ears, served as a measure of individual high-frequency hearing loss (Puschmann and Thiel, 2017).

The subjective listening effort in challenging daily-life situations was assessed using a questionnaire by Schulte et al. (2015). The questionnaire asks for listening effort in 17 daily-life situations; 14 of them describe settings with reduced signal quality or background noise, and with or without eye-contact to the speaker. Participants rated the listening effort according to the given situation on a scale from 0 (not exhausting at all) to 10 (extremely exhausting). The average effort rating obtained for the 14 challenging settings served as marker of individual listening effort. The 80%-speech reception threshold in noise was measured in an adaptive free-field procedure using a German matrix sentence test (Oldenburg sentence test; Wagener et al., 1999). Individual differences in

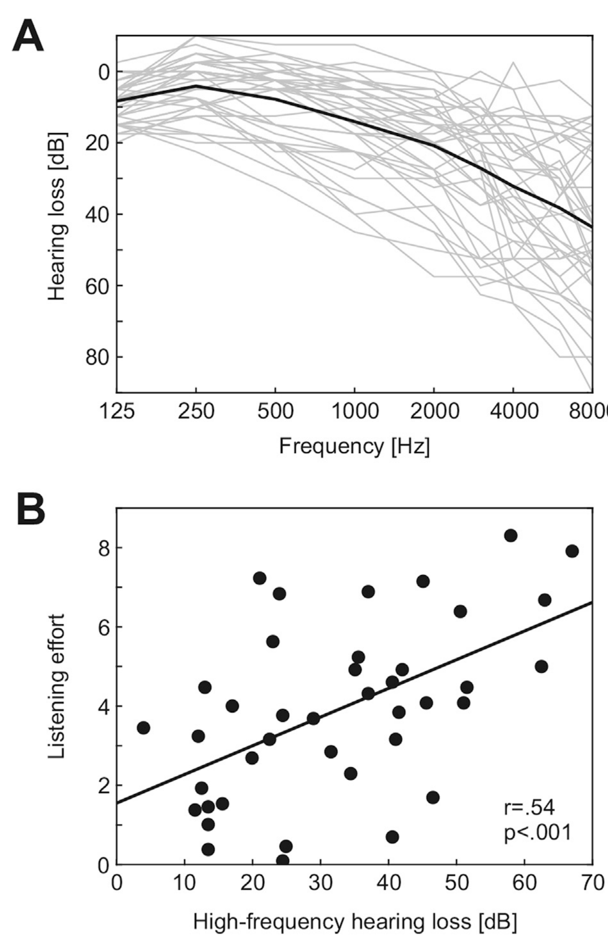

Fig. 1. Forty older volunteers with varying degree of high frequency hearing loss participated in the study. **A)** The figure depicts the individual pure tone audiogram of each participant, averaged over both ears. The group mean is shown in bold black. The mean hearing loss between 2000 and 8000 Hz served as measure of high-frequency hearing impairment. **B)** Individual high-frequency hearing loss was significantly correlated with the reported listening effort in challenging daily-life situations, which served as an additional subjective measure of hearing impairment.

overall cognitive function were assessed using a German version of the Montreal Cognitive Assessment (MoCA) test. This screening assessment for mild cognitive impairment tests for executive functions, attention, concentration, working memory, short-term memory recall, visuospatial abilities, language, as well as orientation in time and place (Nasreddine et al., 2005).

2.3. Stimuli and task

Custom-made video recordings of a male speaker performing humorous monologues served as stimulus material. The speaker spoke standard German. The videos showed the speaker's head and shoulders, with eyes directed into the camera to address the participants and to create the impression of a natural listening situation. No hand movements were used by the speaker and the video background was static. Videos were recorded at a resolution of 1920 × 1080 pixels with a rate of 30 frames per second. The audio track was sampled at 48 kHz with 16-bit resolution.

Eleven monologues with a duration of four to seven minutes were recorded. Videos were cut into consecutive 30-s trials and the root-mean-square sound intensity of each audio track was normalized. The speech tracks were presented at a fixed level of 65 dB(A). The audio tracks were overlaid with a continuous multi-speaker babble noise. The babble noise was created from the stimulus material by averaging the audio tracks of all 30-s trials used in experiment and training (100 tracks in total). The same noise signal was used in all trials. The noise level was adjusted to the individual 80%-speech-reception threshold in noise to ensure that the speech intelligibility was comparable across subjects. Auditory and visual trials were generated by extraction or removal of the audio track. In auditory trials, participants viewed a fixation cross presented centrally at the screen. Stimulus presentation was conducted using Presentation software (Version 17.0, Neurobehavioral Systems, Inc., Berkeley, CA, USA).

The experiment consisted of a training block and ten experimental blocks. Each block contained a complete monologue in correct order and was composed of audiovisual, auditory, and visual trials of 30 s duration, presented in pseudorandomized order and occurring with equal probability. Each block started with an audiovisual trial; trials of the same condition (i.e., audiovisual, auditory, visual) did not occur back to back. Subsequent trials were delayed by 125 ms to warrant exact stimulus timing, leading to a quasi-continuous presentation of the monologue. Fig. 2A depicts an exemplary trial structure. In total, the experiment contained 30 trials of each condition. Participants were instructed to listen as closely as possible to the content of the monologue and to respond via button-press to rare target words. The target word was displayed for 2 s on the video screen before the start of each block. To reduce training effects, a different target word was selected in each block. Target words could occur up to three times per 30-s trial (overall number of targets per condition: audiovisual = 31, auditory = 29, visual = 31). Participants were encouraged to lipread target words during visual trials. After each block, participants had to answer three multiple-choice questions on the content of the monologue. To minimize eye movements, participants were instructed to fixate on the speaker's face or – in auditory trials – on the fixation cross.

2.4. EEG data acquisition and preprocessing

EEG data were acquired from 64 Ag/AgCl electrodes using a BrainAmp amplifier system (BrainProducts, Gilching, Germany) and an electrode cap with an equidistant infracerebral electrode layout (Easycap, Herrsching, Germany). Data were recorded with a sampling rate of 1000 Hz using the nose tip as reference. An analogue bandpass filter (0.01–250 Hz) was applied. Electrode impedances were maintained below 10 kΩ before data acquisition.

EEG data were preprocessed using the EEGLAB toolbox (Version13.6.5b; Delorme and Makeig, 2004) for MATLAB (Version 9.0, MathWorks, Natick, MA, USA). Artifacts related to eye blinks, lateral eye movements, and heart beats were pruned from the datasets using independent component analysis. For this procedure, a copy of the EEG data was filtered from 1 to 40 Hz, segmented into consecutive 1-s epochs, artifact-pruned (by removing all epochs in which the signal amplitude exceeded three standard deviations of the mean), and submitted to the

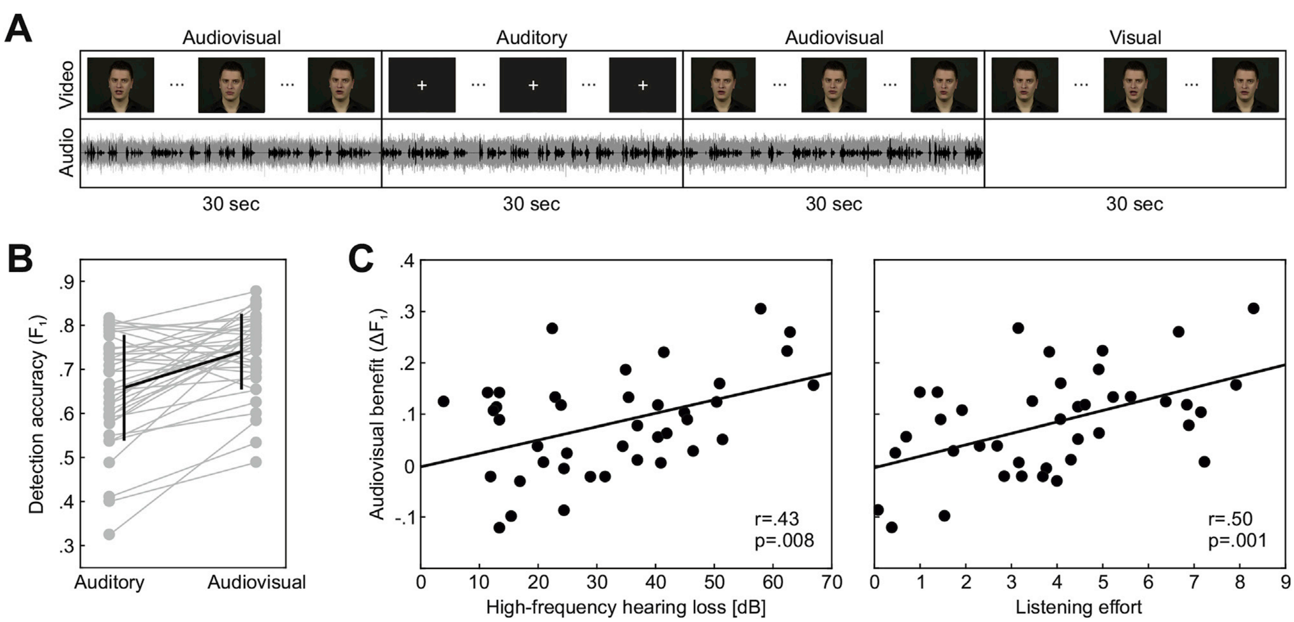

Fig. 2. Experimental design and behavioral data: **A**) During the experiment, participants listened to ten short monologues performed by a male speaker and detected rare target words. Each experimental block contained a full monologue and consisted of auditory, visual, and audiovisual trials of 30 s duration, presented in pseudorandomized order. **B**) The figure depicts the individual (grey) and mean (black; ± standard deviation) target word detection accuracy F_1 in the auditory and audiovisual task condition. Overall, the target word detection accuracy was significantly increased in the audiovisual condition. **C**) The individual audiovisual benefit was positively correlated with the participants' high frequency hearing loss (left) and the reported listening effort in challenging daily-life situations (right).

extended infomax algorithm. The unmixing matrix obtained from this procedure was applied to the original unfiltered EEG dataset for selection and rejection of artifact components. The fully-automated Eye-Catch approach (Bigdely-Shamlo et al., 2013) was used to identify independent components reflecting eye blinks and lateral eye movements. This method is based on the semi-automatic CORRMAP algorithm (Viola et al., 2009), but instead of using single user-initiated templates it uses a large database of exemplary eye scalp maps to identify eye-related artifact components. Independent components reflecting cardiac activity were identified by visual inspection of the components.

For the speech envelope reconstruction, the EEG data were offline filtered from 0.3 to 30 Hz, epoched from 0 to 30 s relative to the onset of each trial, and downsampled to 64 Hz to reduce computational demands. Filtering was performed using separate high- and low-pass finite-impulse response filters with a hamming window for a zero-phase digital filtering of the data.

2.5. Speech envelope reconstruction

The broadband envelope of the speech tracks was computed. For this, the audio signals were z-normalized and subsequently bandpass filtered into 128 logarithmically-spaced frequency bands between 100 and 6500 Hz with a gammatone filter bank (Herzke and Hohmann, 2007; Hohmann, 2002). The 100–6500 Hz range was chosen based on previous research suggesting a high temporal coherence between mouth movements and speech envelope within this frequency range (Chandrasekaran et al., 2009; Crosse et al., 2015; Grant and Seitz, 2000). Hilbert transformation was used to compute the signal envelope within each of the 128 frequency bands. The broadband envelope was then obtained by averaging the absolute Hilbert values across all bands. Similar to the EEG data, the broadband envelope was filtered from 0.3 to 30 Hz using a 3rd order Butterworth filter and downsampled to 64 Hz sampling rate for further processing.

The mTRF toolbox (Crosse et al., 2016a) was used to reconstruct the broadband envelope of the speech signals from the EEG data. The underlying methods have been described by O'Sullivan et al. (2015). In short, ridge regression was used to obtain a linear mapping between the EEG sensor data and the broadband speech envelope. A search grid and a leave-one-out cross-validation with the goal to minimize the mean-squared-error of the regression were used to optimize the ridge parameter λ (grid values: 10^{-2}, 10^{-1}, …, 10^{4}, 5×10^{4}, 10^{5}, …, 10^{9}). Speech envelope reconstruction was performed using a leave-one-out cross-validation on the subject-level to ensure that the mapping between speech envelope and EEG data does not depend on trial-specific features but represents a general relationship between speech input and neural response. This means that the speech envelope of a given trial in a given experimental condition and for a specific time lag was reconstructed using the mean regression weights obtained from all other trials of this condition at the given time lag. The reconstruction performance was evaluated by performing Pearson's correlation between the reconstructed and the original speech envelope. For statistical testing, the r values were converted to normally distributed r_z values using Fisher's z-transformation and averaged across trials.

For an initial time-resolved analysis of speech envelope tracking, single-lag models with 64 regressors (i.e., one for each EEG channel) were computed for each trial and 33 different time lags between stimulus presentation and EEG signal, ranging from 0 to 500 ms. Based on the resulting r_z time course, two time windows of interest, which encompassed the group-level peaks in the auditory and audiovisual task conditions, were defined (i.e., 75–125 ms, 200–240 ms). For further analyses of audiovisual enhancements, multi-lag models containing $64 \times$ N(lags) regressors were computed for each of these time windows and all trials. A similar combination of single-lag models for time window identification and multi-lag models for the final analysis of speech tracking accuracy was previously applied by Puschmann et al. (2017).

2.6. Data analysis

The main data analysis aimed to identify relationships between hearing impairment and audiovisual enhancements in both target word detection and the neural tracking of the speech envelope. Hearing impairment was specified in terms of high-frequency hearing thresholds and subjective listening effort in challenging daily-life situations. Given that we expected to see an increased audiovisual gain in participants with greater hearing impairments, we tested only for positive relationships between these measures.

Target word detection and false alarm rates were extracted for all three experimental conditions and F_1 scores (Van Rijsbergen, 1979) were computed as an overall measure of response accuracy. Target words were regarded as detected if a button press was registered between 0.2 and 2 s after the word onset. For participants failing to detect at least one target word in a given condition, detection rates were adjusted to 0.5/N (targets per condition) for F_1 score computation. F_1 scores obtained in the auditory and audiovisual condition were compared using a paired *t*-test. The audiovisual gain was computed as ΔF_1 between both conditions. Pearson's correlation (one-tailed) was used to test for positive linear relationships between the audiovisual gain ΔF_1 and the two measures of hearing impairment. Results are reported as statistically significant when passing a threshold of $p < .05$, corrected for multiple comparisons using a Bonferroni correction over 3 tests.

To test for audiovisual enhancements in speech envelope tracking, r_z values obtained for the auditory and audiovisual condition within the 75–125 ms and the 200–240 ms windows of interest were compared using one-tailed paired t-tests ($p < .05$; Bonferroni correction over 2 tests). Given that no audiovisual enhancement was found in the late window, only the individual audiovisual gain within the 75–125 ms window was analyzed as a function of hearing impairment using Pearson's correlation (one-sided test; $p < .05$, Bonferroni correction over 2 tests). The audiovisual gain was computed as $(r_{z,audiovisual} - r_{z,auditory})/r_{z,auditory}$ to account for subject-specific factors that may affect overall r_z values (cf. O'Sullivan et al., 2015; Presacco et al., 2016).

For a better understanding of the relationships between the different audiological measures acquired in the experiment, we performed Pearson's correlations between high-frequency hearing loss, reported listening effort in daily life settings, and speech-reception thresholds ($p < .05$, Bonferroni correction over 5 tests).

Since no eye-tracking was performed during the experiment, we used independent component time courses to control for a relationship between hearing loss and eye motion. The time course of the first principal component that reflected lateral eye movements (as identified by the Eye-Catch approach) was z-transformed by means of the standard deviation of the corresponding topographical map. The mean variance across the resulting signal served as a measure of lateral eye activity per subject. We tested for linear relationships to high-frequency hearing loss and listening effort ratings using Pearson's correlations. Participants for which no component reflecting lateral eye-movements was identified ($N = 5$) were treated as missing value. Additional control analyses were performed to assess the influence of general subject-specific factors on the behavioral and neural audiovisual gain, such as age, overall cognitive function (as indicated by MoCa scores), and speech-reading performance (as indicated by F_1 scores and r_z values obtained for the visual task condition). Also, we tested for an effect of high-frequency hearing loss and ratings of listening effort on F_1 scores obtained in the auditory, audiovisual, and visual task condition. No correction for multiple comparisons was applied for these tests to ease the identification of potentially confounding factors.

3. Results

3.1. Audiological data

The average high-frequency hearing threshold was 32.4 ± 16.4 dB (mean ± standard deviation), with individual values ranging from 4 to 67 dB. The average listening effort across a range of challenging daily-life situations, which was rated on a scale from 0 (not exhausting at all) to 10 (extremely exhausting), was 3.9 ± 2.2 (range: 0.1–8.3). The mean signal-to-noise ratio corresponding to 80%-speech-reception in noise was −0.8 ± 1.2 dB (range: −3.1 – 3.2 dB).

High-frequency hearing loss was positively correlated with the subjectively reported listening effort ($r = 0.54$, $p = .002$, Bonferroni corrected; Fig. 1B). This relationship however tended to be more reliable in individuals with pronounced high-frequency hearing loss than in participants with low high-frequency hearing thresholds (median split; sub-group with low degree of hearing loss: $r = .20$, $p = .971$, one-tailed, Bonferroni corrected; sub-group with high degree of hearing loss: $r = 0.49$, $p = .070$, one-tailed, Bonferroni corrected). High-frequency hearing loss ($r = 0.47$, $p = .013$, Bonferroni corrected), but not reported listening effort ($r = 0.32$, $p = .204$, Bonferroni corrected), was positively associated with the individual speech-reception threshold. Control analyses showed that subject age was not significantly related to any of the audiological measures (hearing loss: $r = 0.19$, $p = .244$; reported listening effort: $r = -0.24$, $p = .131$; speech-reception threshold: $r = 0.30$; $p = .064$). Likewise, no relationship to individual cognitive performance (MoCA scores: 26 ± 2, range: 22–30) was found (hearing loss: $r = 0.11$, $p = .517$; reported listening effort: $r = 0.18$, $p = .279$; speech-reception: $r = -0.09$; $p = .581$). Neither the degree of high-frequency hearing loss ($r = 0.08$, $p = .667$) nor subjective ratings of listening effort ($r = 0.10$, $p = .559$) were related to lateral eye activity during the experiment.

3.2. Target word detection

The mean target word detection rate was 53.0% ± 13.0% (range: 20.7%–72.4%) in the auditory and 63.8% ± 10.5% (range: 35.5%–80.7%) in the audiovisual condition. Most participants could not reliably detect target words based on lip-reading in the visual condition (group average: 3.2% ± 6.7%; range: 0%–29.0%; 28 of 40 participants did not detect a single target word). False alarm rates were 10.7% ± 6.7% (range: 0–27.3%) in the auditory, 10.8% ± 4.7% (range: 3.9–21.4%) in the audiovisual, and 28.8% ± 40.7% (range: 0–100%) in the visual task condition. In the visual condition, behavioral responses were registered for 19 of 40 participants, with only 4 of them responding more than five times in total. Thus, further analyses concentrated on the auditory and audiovisual task conditions.

F_1 scores served as a measure of overall response accuracy. Fig. 2B depicts F_1 scores in auditory (group average: 0.66 ± 0.12; range: 0.32–0.82) and audiovisual trials (group average: 0.74 ± 0.09; range: 0.49−0.88). Response accuracy was significantly lower in the auditory condition, which lacked any visual cues ($t(39) = -5.2$, $p < .001$, Bonferroni corrected). Note that we aimed to account for individual differences in speech intelligibility in noise by adjusting the signal-to-noise ratio in the experiment to the measured 80%-speech-reception threshold. Still, hearing impairment was negatively correlated with F_1 scores in the auditory (hearing loss: $r = -0.58$, $p < .001$; reported listening effort: $r = -0.47$, $p = .002$) but not in the audiovisual listening condition (hearing loss: $r = -0.31$, $p = .051$; reported listening effort: $r = -0.09$, $p = .583$).

The audiovisual gain ΔF_1 ranged from −0.12 to 0.30, with 32 out of 40 participants showing an accuracy increase in the audiovisual condition. Individual differences in ΔF_1 were positively associated with the degree of high-frequency hearing loss ($r = 0.43$, $p = .008$, Bonferroni corrected; one-tailed test) and with the reported listening effort in challenging daily-life settings ($r = 0.50$, $p = .001$, Bonferroni corrected; one-tailed test), suggesting that listeners with greater hearing impairment experienced an increased benefit from the additional visual input (Fig. 2C). A control analysis showed no relationship between ΔF_1 and age ($r = -0.16$, $p = .334$) or MoCA scores ($r = 0.15$, $p = .369$). Also, the individual audiovisual gain ΔF_1 was not correlated with F_1 scores obtained for the visual condition ($r = -0.16$, $p = .326$).

3.3. Speech envelope reconstruction

Fig. 3A depicts the speech envelope reconstruction accuracy r_z for the three listening conditions as a function of relative time lag between auditory input and EEG response. Two time windows of interest (i.e., 75–125 ms, 200–240 ms; indicated in grey) were defined based on the group-level peaks of envelope reconstruction accuracy in the auditory and audiovisual conditions. Paired t-tests revealed a significant audiovisual enhancement in r_z within the early time window ($t(39) = 4.9$, $p < .001$, Bonferroni corrected, one-tailed), while no enhancement was observed in the 200–240 ms window ($t(39) = -0.3$, $p = .612$). Fig. 3B shows the topography of mean decoder weights for all conditions within both windows of interest.

To investigate whether the audiovisual gain in speech envelope tracking was, similarly to the behavioral gain, directly related to individual hearing impairment, correlation analyses were performed. Given that no audiovisual enhancement was observed within the late time window of interest, we only considered the 75–125 ms interval. We found that the individual audiovisual gain within this time window was positively associated with the reported listening effort in challenging daily-life situations ($r = 0.38$, $p = .016$, Bonferroni corrected; one-tailed test; Fig. 3C). The audiovisual gain and the degree of high-frequency hearing loss were not correlated ($r = 0.16$, $p = .333$, Bonferroni corrected; one-tailed test). Similar to our behavioral data, no relationship between the reported audiovisual gain and the reconstruction accuracy obtained in the visual condition was found ($r = -.001$, $p = .995$).

For a better visualization of the observed effect, we computed the mean speech envelope reconstruction accuracy in the auditory and audiovisual task condition as a function of time lag for two subgroups of participants, those with the smallest and highest ratings of listening effort (quartile split; Fig. 3D). While participants reporting low listening effort hardly benefitted from the additional visual input, participants with high listening effort ratings showed a qualitative increase in speech reconstruction accuracy in the audiovisual condition, in particular at early time lags around the chosen time window of interest.

4. Discussion

We here investigated whether hearing-impaired older individuals benefit more from visual cues, such as seeing a speaker's face, when listening to speech in noise than age-matched listeners with normal hearing. In line with this hypothesis, we observed a positive relationship between the individual degree of hearing impairment – quantified using a standard audiogram and subjective ratings of listening effort in challenging daily-life situations – and the audiovisual benefit in target word detection accuracy. On the neural level, the audiovisual enhancement of speech envelope tracking increased with increasing levels of reported listening effort.

4.1. Audiovisual enhancement of target word detection performance

Congruent visual input can facilitate the perception of speech in background noise, both in young and older listeners, with or without hearing impairments (Baskent and Bazo, 2011; Sommers et al., 2005; Tye-Murray et al., 2007a). While audiovisual enhancements are most effective at low signal-to-noise ratios, additional visual input also facilitates speech perception at moderate noise levels, corresponding to our results (Ross et al., 2006; Sumby and Pollack, 1954). Some previous studies suggested a positive relationship between presbycusis and audiovisual enhancements in speech-in-noise perception (Altieri and Hudock, 2014; Moradi et al., 2016), whereas others did not reveal any differences between normal-hearing and hearing-impaired older listeners (Rosemann and Thiel, 2018; Tye-Murray et al., 2007a). Most previous studies however assessed the reception of individually presented

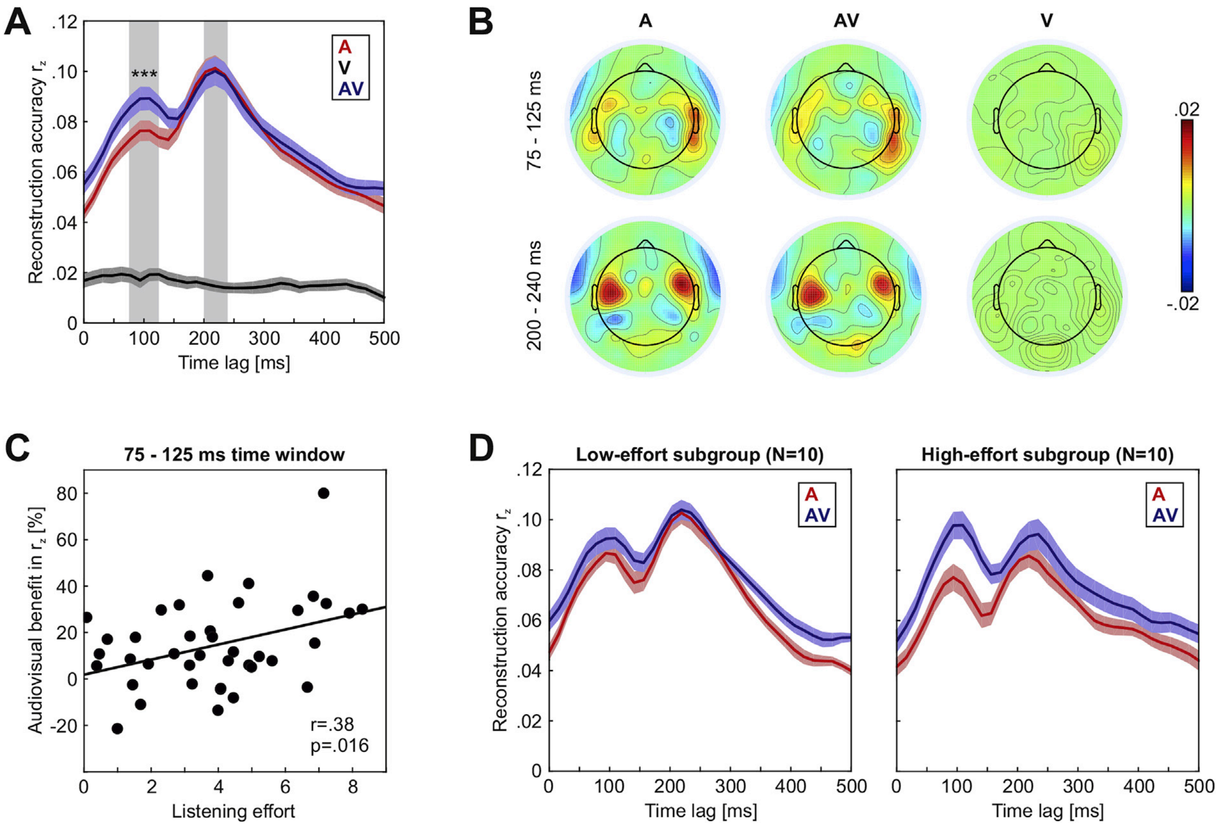

Fig. 3. A speech envelope reconstruction approach was applied to study the neural tracking of the speech envelope. **A**) The figure shows the mean speech envelope reconstruction accuracy r_z (±standard error of the mean) as a function of the time lag of the EEG signal relative to the audio input, from 0 to 500 ms. The auditory (A) and audiovisual (AV) conditions show two distinct peaks in reconstruction accuracy, around 109 and 209 ms after auditory stimulus presentation. Grey boxes highlight two time windows of interest surrounding these peaks (75–125 ms; 200–240 ms). No robust tracking of the auditory speech envelope was observed in the visual (V) condition. A statistically significant increase of r_z in the audiovisual as compared to the auditory condition was only observed in the early time window of interest (***; $p < .001$, Bonferroni corrected). **B**) The figure depicts the mean scalp topography of the regression weights used for speech envelope reconstruction in the auditory, audiovisual, and visual condition within the two time windows of interest. **C**) Within the early time window of interest, the audiovisual gain in envelope reconstruction accuracy was positively correlated with the individual listening effort in daily-life. No significant association was observed with the degree of high-frequency hearing loss. **D**) The figure shows r_z as a function of time lag for two sub-groups of participants, which reported the least and highest listening effort in challenging daily-life settings. While the low-effort subgroup shows hardly any audiovisual benefit on speech envelope reconstruction, a profound increase in early speech envelope reconstruction accuracies can be observed for the high-effort group.

phonemes, words, or short sentences in noise, whereas we here measured the detection of rare target words embedded into a continuous speech stream. The continuous stimulation resembles a natural listening situation and places additional cognitive demands – in terms of attention, vigilance, and executive control – on the listener (Johnsrude and Rodd, 2016). On the other hand, it provides rich contextual cues that can facilitate speech perception and target word detection. Previous work suggested that hearing-impaired participants may benefit more from contextual cues in speech-in-noise reception tasks (Gordon-Salant, 2005; Pichora-Fuller et al., 1995). However, subjects with high amounts of hearing-impairment still showed reduced behavioral performance and envelope tracking (cf. Fig. 3D) in the auditory condition. Given that contextual cues were comparable in the auditory and audiovisual trials, it seems unlikely that these cues affect the reported positive relationship between audiovisual gain and hearing loss.

It has been previously demonstrated that individual differences in speech-in-noise perception can be partially related to differing cognitive abilities (e.g., auditory attention, working memory), suggesting that cognitive resources support speech perception in adverse listening conditions (Anderson et al., 2013; Peelle, 2018; Zekveld et al., 2013). Hearing-impaired listeners are thought to strongly rely on cognitive mechanisms when listening to speech in noise to compensate for the degraded sensory input (Campbell and Sharma, 2013; Peelle and Wingfield, 2016). We speculate that, in the context of our experiment, hearing-impaired participants may have attributed a considerable share of cognitive resources to speech-in-noise perception, leading to reduced performance in the actual target word detection task. Congruent visual input facilitates speech-in-noise processing and may thus allow the hearing-impaired participants to reallocate resources to the target word detection. Normal-hearing participants may be less reliant on cognitive resources for perceiving speech in noise, thus limiting the audiovisual benefit. In line with this view, Rosemann and Thiel (2018) reported an increased recruitment of frontal brain regions in hearing-impaired as compared to normal-hearing older participants when listening to auditory speech in noise but no difference in activation patterns for congruent audiovisual speech, suggesting that the cognitive load of hearing-impaired individuals approaches normal levels (i.e., similar to normal-hearing listeners) in audiovisual settings.

4.2. Neural tracking of audiovisual speech

It has been argued that audiovisual enhancements in speech perception rely on the temporal coherence between the speech envelope and lip movements (Grant and Seitz, 2000). Schroeder et al. (2008) proposed that lip motion strengthens the phase-alignment of auditory cortical oscillations to the speech input. In line with this idea, previous work suggests that lip movements entrain neural activity in human auditory cortex (Besle et al., 2008; Micheli et al., 2018; Park et al., 2016).

Several studies demonstrated a facilitation of cortical speech envelope tracking for audiovisual as compared to auditory-only speech (Crosse et al., 2015; Zion Golumbic et al., 2013). This audiovisual enhancement in speech envelope tracking is most evident in the presence of background noise or when listening selectively to one out of multiple competing speakers (Crosse et al., 2016b; Zion Golumbic et al., 2013). Extending these data obtained in young and healthy listeners, we here show that an audiovisual enhancement in speech envelope tracking can be observed in older listeners. Moreover, we found that the magnitude of this effect is related to the reported listening effort in challenging daily-life situations, which served as a subjective marker of hearing impairment. Noteworthy, the audiovisual enhancement of speech envelope tracking was only found in the early time window of interest, reflecting cortical processing occurring 75–125 ms after the acoustic signal entered the listener's ear. This suggests that the congruent visual input mainly strengthened the cortical sensory processing of the speech signal.

We were not able to reconstruct the envelope of the acoustic speech signal from the visual stimulation itself (cf. Fig. 3A). This disagrees to prior findings by Crosse et al. (2015, 2016b), showing robust vision-based speech envelope reconstruction. While the studies differ in several technical aspects from our current experiment (i.e., higher number of EEG channels, longer stimulus blocks used for reconstruction), it seems unlikely that these factors can explain the differing reconstruction performance (cf. Fuglsang et al., 2017; Mirkovic et al., 2015). Instead, it may be speculated that differences in envelope reconstruction accuracy arise from study-specific differences in the actual visual stimulation (i.e., amateur vs. trained speaker, absence vs. frequent usage of hand gestures, nonrecurring vs. recurring presentation of the same stimulus material across different task conditions). Also, we cannot rule out a potential role of the tested population (i.e., older vs. young adults). Previous data show decreased lip-reading abilities in older as compared to young individuals (cf. Tye-Murray et al., 2007b). In agreement with this notion, Crosse et al. (2015, 2016b) found significantly higher visual target word detection rates than observed in our experiment, suggesting that young individuals could extract more information from the visual input. However, despite being unable to robustly extract speech information from the visual stimuli, our participants still showed a clear audiovisual enhancement of speech envelope tracking and target detection performance. This suggests that the observed audiovisual benefit does not rely on the extraction of speech information from the visual input per se, but on a vision-based facilitation of auditory sensory processing (cf. Schroeder et al., 2008).

Recent EEG data show increased auditory cortex responses to visual and somatosensory input in older listeners with moderate degrees of high-frequency hearing loss (Campbell and Sharma, 2013; Cardon and Sharma, 2018). Using functional magnetic resonance imaging, Puschmann and Thiel (2017) observed a positive relationship between the degree of high-frequency hearing loss and functional connectivity between right middle temporal area MT/V5, which is involved in visual motion processing, and secondary auditory cortex while processing congruent audiovisual input. These data suggest that the aged auditory cortex adapts to compensate for degraded auditory sensory input, possibly by integrating more information from other sensory modalities. In our view, it seems likely that the observed association between hearing impairment and the behavioral as well as the neural audiovisual gain is related to cross-modal neuroplastic changes of the auditory cortex. Opposing this idea, EEG data by Musacchia et al. (2009) showed decreased audiovisual effects on the latency and amplitude of early auditory-evoked responses in hearing-impaired as compared to normal-hearing listeners, suggesting a reduced neural integration of audiovisual information in hearing-impaired individuals. Similarly, Meredith et al. (2012) found that cross-modal plasticity in adult animals leads to a reduced multisensory integration at the level of auditory cortex.

It should be noted that no eye-tracking was performed in this study. While our data show no relationship between hearing loss and EEG-markers of eye motion, we cannot rule out that individual differences in fixation and gaze direction may have modulated the effect of visual input in speech-in-noise settings. Hence, the basis of the increased audiovisual enhancement in hearing-impaired listeners is not clear at present. Additional work is needed to investigate the nature of this effect in more detail and to better understand how older hearing-impaired listeners make use of non-auditory cues to support speech processing in challenging daily-life situations.

4.3. Summary

Our data provide evidence for a positive relationship between the degree of age-related hearing problems and the perceptual benefit from seeing a speaker's face in a noisy listening setting. On the neural level, this effect was associated with an audiovisual enhancement of early cortical speech envelope tracking, possibly reflecting a lip motion-induced entrainment of auditory cortex oscillations to speech input (Schroeder et al., 2008). We speculate that the increased effect of congruent visual input on auditory sensory processing in older hearing-impaired individuals is related to adaptive cross-modal plasticity of the auditory cortex (Campbell and Sharma, 2013; Puschmann and Thiel, 2017). Potentially, neuroplastic changes may allow hearing-impaired listeners to make more use of visual input, which may help to partially compensate for the degraded auditory sensory input.

Conflicts of interest

None.

Acknowledgments

The work was supported by the Cluster of Excellence "Hearing4all", funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; grant: EXC1077/1). Sebastian Puschmann was supported by a research fellowship of the German Research Foundation (Deutsche Forschungsgemeinschaft; grant: PU590/1-1). Mareike Daeglau was supported by a research grant of the German Research Foundation (Deutsche Forschungsgemeinschaft; grant: KR3433/3-1).

The authors thank Alexander Scholz (speaker), Isabelle Andreeßen (technical assistance), Tim Alexander Kerting (video editing), and Jan Michalsky (linguistic consultation) for supporting the stimulus preparation. The EEG data was acquired with the help of Jeanette Nischan, Marie Dewenter, and Volkan Elis.

References

Alicea, C.C.M., Doherty, K.A., 2017. Motivation to address self-reported hearing problems in adults with normal hearing thresholds. J. Speech Lang. Hear. Res. 60, 3642–3655.

Altieri, N., Hudock, D., 2014. Assessing variability in audiovisual speech integration skills using capacity and accuracy measures. Int. J. Audiol. 53, 710–718.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., Kraus, N., 2013. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. Hear. Res. 300, 18–32.

Arnal, L.H., Morillon, B., Kell, C.A., Giraud, A.L., 2009. Dual neural routing of visual facilitation in speech processing. J. Neurosci. 29, 13445–13453.

Atilgan, H., Town, S.M., Wood, K.C., Jones, G.P., Maddox, R.K., Lee, A.K.C., Bizley, J.K., 2018. Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. Neuron 97, 640-655 e644.

Badri, R., Siegel, J.H., Wright, B.A., 2011. Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. J. Acoust. Soc. Am. 129, 852–863.

Baskent, D., Bazo, D., 2011. Audiovisual asynchrony detection and speech intelligibility in noise with moderate to severe sensorineural hearing impairment. Ear Hear. 32, 582–592.

Besle, J., Fischer, C., Bidet-Caulet, A., Lecaignard, F., Bertrand, O., Giard, M.H., 2008. Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. J. Neurosci. 28, 14301–14310.

Bigdely-Shamlo, N., Kreutz-Delgado, K., Kothe, C., Makeig, S., 2013. EyeCatch: data-mining over half a million EEG indipendent components to construct a fully-automated eye-component detector. In: Conf Proc IEEE Eng Med Biol Soc, 2013.

Butler, B.E., Lomber, S.G., 2013. Functional and structural changes throughout the auditory system following congenital and early-onset deafness: implications for hearing restoration. Front. Syst. Neurosci. 7, 92.

Campbell, J., Sharma, A., 2013. Compensatory changes in cortical resource allocation in adults with hearing loss. Front. Syst. Neurosci. 7, 71.

Cardon, G., Sharma, A., 2018. Somatosensory cross-modal reorganization in adults with age-related, early-stage hearing loss. Front. Hum. Neurosci. 12, 172.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., Ghazanfar, A.A., 2009. The natural statistics of audiovisual speech. PLoS Comput. Biol. 5 e1000436.

Crosse, M.J., Butler, J.S., Lalor, E.C., 2015. Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. J. Neurosci. 35, 14195–14204.

Crosse, M.J., Di Liberto, G.M., Bednar, A., Lalor, E.C., 2016a. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10, 604.

Crosse, M.J., Di Liberto, G.M., Lalor, E.C., 2016b. Eye can hear clearly now: inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. J. Neurosci. 36, 9888–9895.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Ding, N., Simon, J.Z., 2013. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. J. Neurosci. 33, 5728–5735.

Fuglsang, S.A., Dau, T., Hjortkjaer, J., 2017. Noise-robust cortical tracking of attended speech in real-world acoustic scenes. Neuroimage 156, 435–444.

Gordon-Salant, S., 2005. Hearing loss and aging: new research findings and clinical implications. J. Rehabil. Res. Dev. 42, 9–24.

Grant, K.W., Seitz, P.-F., 2000. The use of visible speech cues for improving auditory detection of spoken sentences. J. Acoust. Soc. Am. 108, 1197–1208.

Guerreiro, M.J.S., Van Gerven, P.W.M., 2017. Disregarding hearing loss leads to overestimation of age-related cognitive decline. Neurobiol. Aging 56, 180–189.

Hallam, R.S., Corney, R., 2014. Conversation tactics in persons with normal hearing and hearing-impairment. Int. J. Audiol. 53, 174–181.

Hauthal, N., Sandmann, P., Debener, S., Thorne, J.D., 2013. Visual movement perception in deaf and hearing individuals. Adv. Cognit. Psychol. 9, 53–61.

Herzke, T., Hohmann, V., 2007. Improved numerical methods for gammatone filterbank analysis and synthesis. Acta Acust. United Ac. 93, 498–500.

Hohmann, V., 2002. Frequency analysis and synthesis using a Gammatone filterbank. Acta Ac. United Ac. 88, 433–442.

Johnsrude, I.S., Rodd, J.M., 2016. Factors that increase processing demands when listening to peech. In: Hickok, G., Small, S. (Eds.), Neurobiology of Language. Academic Press, London, UK, pp. 491–502.

Kayser, C., Logothetis, N.K., Panzeri, S., 2010. Visual enhancement of the information representation in auditory cortex. Curr. Biol. 20, 19–24.

Kral, A., 2007. Unimodal and cross-modal plasticity in the 'deaf' auditory cortex. Int. J. Audiol. 46, 479–493.

Meredith, M.A., Keniston, L.P., Allman, B.L., 2012. Multisensory dysfunction accompanies crossmodal plasticity following adult hearing impairment. Neuroscience 214, 136–148.

Micheli, C., Schepers, I.M., Ozker, M., Yoshor, D., Beauchamp, M.S., Rieger, J.W., 2018. Electrocorticography reveals continuous auditory and visual speech tracking in temporal and occipital cortex. Eur. J. Neurosci. https://doi.org/10.1111/ejn.13992 [Epub ahead of print].

Mirkovic, B., Debener, S., Jaeger, M., De Vos, M., 2015. Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. J. Neural Eng. 12, 046007.

Moradi, S., Lidestam, B., Rönnberg, J., 2016. Comparison of gated audiovisual speech identification in elderly hearing aid users and elderly normal-hearing individuals: effects of adding visual cues to auditory speech stimuli. Trends Hear 20, 1–15.

Musacchia, G., Arum, L., Nicol, T., Garstecki, D., Kraus, N., 2009. Audiovisual deficits in older adults with hearing loss: biological evidence. Ear Hear. 30, 505–514.

Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H., 2005. The Montreal Cognitive Assessment, MoCa: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Sco. 53, 695–699.

O'Sullivan, J.A., Power, A.J., Mesgarani, N., Rajaram, S., Foxe, J.J., Shinn-Cunningham, B.G., Slaney, M., Shamma, S.A., Lalor, E.C., 2015. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. Cerebr. Cortex 25, 1697–1706.

Park, H., Kayser, C., Thut, G., Gross, J., 2016. Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. Elife 5 e14521.

Peelle, J.E., 2018. Listening effort: how the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear Hear. 39, 204–214.

Peelle, J.E., Wingfield, A., 2016. The neural consequences of age-related hearing loss. Trends Neurosci. 39, 486–497.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593–608.

Presacco, A., Simon, J.Z., Anderson, S., 2016. Effect of informational content of noise on speech representation in the aging midbrain and cortex. J. Neurophysiol. 116, 2356–2367.

Puschmann, S., Sandmann, P., Bendixen, A., Thiel, C.M., 2014. Age-related hearing loss increases cross-modal distractibility. Hear. Res. 316, 28–36.

Puschmann, S., Steinkamp, S., Gillich, I., Mirkovic, B., Debener, S., Thiel, C.M., 2017. The right temporoparietal junction supports speech tracking during selective listening: evidence from concurrent EEG-fMRI. J. Neurosci. 37, 11505–11516.

Puschmann, S., Thiel, C.M., 2017. Changed crossmodal functional connectivity in older adults with hearing loss. Cortex 86, 109–122.

Rosemann, S., Thiel, C.M., 2018. Audio-visual speech processing in age-related hearing loss: stronger integration and increased frontal lobe recruitment. Neuroimage 175, 425–437.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., Foxe, J.J., 2006. Do you see what I Am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebr. Cortex 17, 1147–1153.

Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R.D., Marcar, V.L., Jäncke, L., Debener, S., 2012. Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. Brain 135, 555–568.

Schormans, A.L., Typlt, M., Allman, B.L., 2017. Crossmodal plasticity in auditory, visual and multisensory cortical areas following noise-induced hearing loss in adulthood. Hear. Res. 343, 92–107.

Schroeder, C.E., Lakatos, P., Kajikawa, Y., Partan, S., Puce, A., 2008. Neuronal oscillations and visual amplification of speech. Trends Cognit. Sci. 12, 106–113.

Schulte, M., Meis, M., Wagener, K., 2015. Der Höranstrengungsfragebogen [The listening effort questionnaire]. 18. Jahrestagung der deutschen Gesellschaft für Audiologie., March 4-7. Bochum, Germany.

Shiell, M.M., Champoux, F., Zatorre, R.J., 2014. Enhancement of visual motion detection thresholds in early deaf people. PLoS One 9 e90498.

Sommers, M.S., Tye-Murray, N., Spehar, B., 2005. Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. Ear Hear. 26, 263–275.

Stropahl, M., Debener, S., 2017. Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. Neuroimage Clin 16, 514–523.

Stropahl, M., Plotz, K., Schonfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., Debener, S., 2015. Cross-modal reorganization in cochlear implant users: auditory cortex contributes to visual face processing. Neuroimage 121, 159–170.

Sumby, W.H., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. 26, 212–215.

Tye-Murray, N., Sommers, M.S., Spehar, B., 2007a. Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. Ear Hear. 28, 656–668.

Tye-Murray, N., Sommers, M.S., Spehar, B., 2007b. The effect of age and gender on lipreading abilities. J. Am. Acad. Audiol. 18, 883–892.

Van Rijsbergen, C.J., 1979. Information Retrieval, second ed. Butterworth, Newton, MA.

Van Wassenhove, V., Grant, K.W., Poeppel, D., 2005. Visual speech speeds up the neural processing of auditory speech. Proc. Natl. Acad. Sci. U. S. A. 102, 1181–1186.

Viola, F.C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., Debener, S., 2009. Semi-automatic identification of independent components representing EEG artifact. Clin. Neurophysiol. 120, 868–877.

Wagener, K., Kühnel, V., Kollmeier, B., 1999. Entwicklung und Evaluation eines Satztests in deutscher Sprache I: Design des Oldenburg Satztests [Development and evaluation of a German sentence test I: design of the Oldenburg sentence test]. Z. Audiol. 38, 4–15.

Winneke, A.H., Phillips, N.A., 2011. Does audiovisual speech offer a fountain of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception. Psychol. Aging 26, 427–438.

Zekveld, A., Rudner, M., Johnsrude, I.S., Rönnberg, J., 2013. The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. J. Acoust. Soc. Am. 134, 2225–2234.

Zion Golumbic, E., Cogan, G.B., Schroeder, C.E., Poeppel, D., 2013. Visual input enhances selective speech envelope tracking in auditory cortex at a "cocktail party". J. Neurosci. 33, 1417–1426.