



OPEN ACCESS

EDITED BY

Fushun Wang,
Nanjing University of Chinese
Medicine, China

REVIEWED BY

Alexandra Begau,
Leibniz Research Centre for Working
Environment and Human Factors
(IfADo), Germany
Wucheng Tao,
Fujian Medical University, China

*CORRESPONDENCE

Yanna Ren
yanna052267213@163.com
Zhilin Zhang
zhangzhilin@siat.ac.cn

SPECIALTY SECTION

This article was submitted to
Neurocognitive Aging and Behavior,
a section of the journal
Frontiers in Aging Neuroscience

RECEIVED 02 August 2022

ACCEPTED 11 October 2022

PUBLISHED 28 October 2022

CITATION

Yang W, Guo A, Yao H, Yang X, Li Z,
Li S, Chen J, Ren Y, Yang J, Wu J and
Zhang Z (2022) Effect of aging on
audiovisual integration: Comparison
of high- and low-intensity conditions
in a speech discrimination task.
Front. Aging Neurosci. 14:1010060.
doi: 10.3389/fnagi.2022.1010060

COPYRIGHT

© 2022 Yang, Guo, Yao, Yang, Li, Li,
Chen, Ren, Yang, Wu and Zhang. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Effect of aging on audiovisual integration: Comparison of high- and low-intensity conditions in a speech discrimination task

Weiping Yang^{1,2}, Ao Guo³, Hanyun Yao⁴, Xiangfu Yang¹,
Zimo Li¹, Shengnan Li³, Jianxin Chen¹, Yanna Ren^{5*},
Jiajia Yang³, Jinglong Wu^{3,6} and Zhilin Zhang^{6*}

¹Department of Psychology, Faculty of Education, Hubei University, Wuhan, China, ²Brain and Cognition Research Center, Faculty of Education, Hubei University, Wuhan, China, ³Applied Brain Science Lab, Faculty of Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama, Japan, ⁴Department of Neurology, Wuhan No. 1 Hospital, Wuhan, China, ⁵Department of Psychology, College of Humanities and Management, Guizhou University of Traditional Chinese Medicine, Guiyang, China, ⁶Research Center for Medical Artificial Intelligence, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China

Audiovisual integration is an essential process that influences speech perception in conversation. However, it is still debated whether older individuals benefit more from audiovisual integration than younger individuals. This ambiguity is likely due to stimulus features, such as stimulus intensity. The purpose of the current study was to explore the effect of aging on audiovisual integration, using event-related potentials (ERPs) at different stimulus intensities. The results showed greater audiovisual integration in older adults at 320–360 ms. Conversely, at 460–500 ms, older adults displayed attenuated audiovisual integration in the frontal, fronto-central, central, and centro-parietal regions compared to younger adults. In addition, we found older adults had greater audiovisual integration at 200–230 ms under the low-intensity condition compared to the high-intensity condition, suggesting inverse effectiveness occurred. However, inverse effectiveness was not found in younger adults. Taken together, the results suggested that there was age-related dissociation in audiovisual integration and inverse effectiveness, indicating that the neural mechanisms underlying audiovisual integration differed between older adults and younger adults.

KEYWORDS

audiovisual integration, aging, stimulus intensity, speech perception, inverse effectiveness

Introduction

In our daily lives, events convey information through different sensory modalities. Our brain is able to integrate different sensory information into a unique and coherent percept. For instance, when communicating with others, we recognize speech content by integrating visual and auditory speech information. This process is called audiovisual integration (Laurienti et al., 2006; Stein, 2012). The benefit of simultaneous (bimodal) presentation of visual and auditory stimuli over unimodal presentation has been repeatedly validated; for instance, the response time to audiovisual stimuli is faster than that to visual or auditory stimuli alone (Teder-Sälejärvi et al., 2005; Yang et al., 2015).

As we age, sensory systems and cognitive functions significantly decline (Bäckman et al., 2006; Liu and Yan, 2007; Grady, 2012; Mozolic et al., 2012). Audiovisual integration can bridge the gap between sensory and cognitive processing (McGovern et al., 2014). Numerous aging studies have compared audiovisual integration between younger adults and older adults. Some researchers employed a color-naming task with semantic stimuli presented under visual, auditory, and audiovisual conditions (Laurienti et al., 2006). The results showed that the reaction time to audiovisual stimuli was reduced in younger and older adults, with a significantly greater audiovisual gain in older adults. Similarly, on a simple discrimination task with meaningless stimuli, older adults also exhibited enhanced audiovisual integration (Peiffer et al., 2007). To obtain a clearer and more complete understanding of the neural changes associated with audiovisual processing during aging, subsequent event-related potential (ERP) research was conducted; the results revealed that the neural response to audiovisual stimuli in older adults was stronger than that in younger adults (Diaconescu et al., 2013). The stronger neural response in older adults was confirmed in the medial prefrontal and inferior parietal regions 100 ms after stimulus onset (Diaconescu et al., 2013). These studies supported the idea that the enhanced audiovisual integration in older adults related to the activation of specific brain regions may compensate for deficiencies in unimodal sensory processing (Laurienti et al., 2006; Davis et al., 2008; Mozolic et al., 2012; Freiherr et al., 2013). Notably, the above studies on audiovisual integration in older adults mostly presented a simple combination of auditory and visual stimuli and were based on non-speech paradigms. However, findings regarding improved audiovisual integration in older adults were equivocal when focusing on speech perception, which is more complex and requires greater cognitive resources (Freiherr et al., 2013; Stevenson and Wallace, 2013). Sommers et al. (2005) examined the age-related difference in audiovisual integration with speech stimuli by testing the contribution of visual and auditory processing to audiovisual integration. The results showed that the audiovisual gain in older adults

less than that in younger adults. The authors further proposed that the encoding of visual information was reduced in older adults, which may diminish their ability to comprehend auditory information (Sommers et al., 2005). In contrast, other researchers reported that audiovisual speech processing did not deteriorate in older adults. That is, older adults exhibited equivalent audiovisual gain with younger adults, with a larger reduction in the amplitude of the N1 component in the audiovisual condition relative to the summed amplitude in the auditory and visual conditions (Winneke and Phillips, 2011). This difference in neural alteration of older adults in audiovisual speech processing might be due to accumulated experience, suggesting that older adults would be better at extracting useful information from visual stimuli (lip movements) to predict the upcoming auditory stimuli (spoken utterance) (Winneke and Phillips, 2011). These results suggested that the discrepancy in audiovisual integration between older and younger adults during audiovisual speech perception remains controversial. It remains to be determined whether, in older adults, sensory processing mainly affects audiovisual integration leading to reduced audiovisual integration compared to younger adults or whether prior experience influences audiovisual integration leading to similar or enhanced audiovisual integration compared to younger adults. According to the multistage audiovisual integration theory, there is a distinct early stage of integration where separate visual information influences the auditory processing at sensory level and a later stage of integration where top-down process engages. The early stage of integration was governed by sensory processing efficiency and later stage of integration was influenced by top-down factors (i.e., prior experience or top-down attention). The relation between increasing age and the functioning of sensory processing and top-down processing turns to be different trend. Despite older adults undergoes age-related decline in sensory processing efficiency, there is a dissociation in top-down processing. That is older adults have top-down deficits but have more knowledgeable experience. Given the above reasons, we hypothesized that the audiovisual integration at early and later stage between older adults and younger adults were different. Therefore, we examined different stages of audiovisual integration with ERPs and compared the neural activity during the processing of audiovisual speech stimuli between age groups.

Additionally, previous researchers have found that stimulus intensity impacts audiovisual integration (Senkowski et al., 2011; Mozolic et al., 2012). Many behavioral studies have shown that the magnitude of audiovisual integration is inversely associated with stimulus intensity. That is, audiovisual integration is greater with low-intensity stimuli than high-intensity stimuli (Corneil et al., 2002). Subsequent research using ERPs demonstrated a robust audiovisual integration effect in the right anterior and left posterior regions, particularly

under low-intensity conditions compared to middle-intensity and high-intensity conditions (Senkowski et al., 2011). This phenomenon was considered as inverse effectiveness (IE) (Stein and Stanford, 2008). Numerous aging researches indicated that inverse effectiveness differs according to speech stimuli. Several behavioral studies have shown that audiovisual integration in older adults at the word and sentence levels is not enhanced in low-intensity conditions compared to high-intensity conditions, suggesting that IE in older adults is not found when processing complex speech stimuli (Gordon and Allen, 2009; Tye-Murray et al., 2010; Stevenson et al., 2015; Jansen et al., 2018). However, conflicting results were reported with relatively simple speech stimuli, at the phoneme level. Using phonemes as stimuli, Tye-Murray et al. (2011) reported that older adults only showed audiovisual enhancement when the stimuli intensity was relatively high. When the signal contrast was reduced (low-intensity stimuli), such enhancement was not found (Tye-Murray et al., 2011). This finding suggests that older adults are less able to combine information from visual and auditory modalities when stimulus intensity is lower. However, some researchers have suggested that audiovisual processing is preserved at more elementary levels of speech perception (Stevenson et al., 2015). That is, audiovisual integration in older adults increases as stimulus intensity decreases at the phoneme level (Stevenson et al., 2015). These conflicting findings indicate instability in IE during audiovisual integration at the phoneme level in older adults. In addition, a speech perception task administered during functional magnetic resonance imaging (fMRI) suggested that the superior temporal sulcus (STS) is strongly associated with IE, given the larger BOLD activations in the STS due to audiovisual speech stimuli compared to unisensory speech stimuli in the low-intensity conditions (Stevenson and James, 2009). Due to the unstable IE and the deterioration of the STS in older adults (de Dieuleveult et al., 2017), explorations of IE related to audiovisual integration at the phoneme level is necessary in older adults. Therefore, we compared the magnitude of audiovisual integration between high-intensity conditions and low-intensity conditions in older and younger adults during a speech discrimination task based on phoneme-level stimuli. Evidence showed that there is a direct link between sensory processing and higher-order cognitive processing, reflected by the factor that they may share same resources (Schneider and Pichorafuller, 2000). Speech perception at low-intensity condition is more effortful for older adults. We hypothesized that if sensory processing is effortful, more processing resources have to be devoted to sensory processing, which may lead to enhanced audiovisual integration. In return, this leads to fewer resources available for high-order cognitive processing, which may result in reduced audiovisual integration at this stage.

Materials and methods

Participants

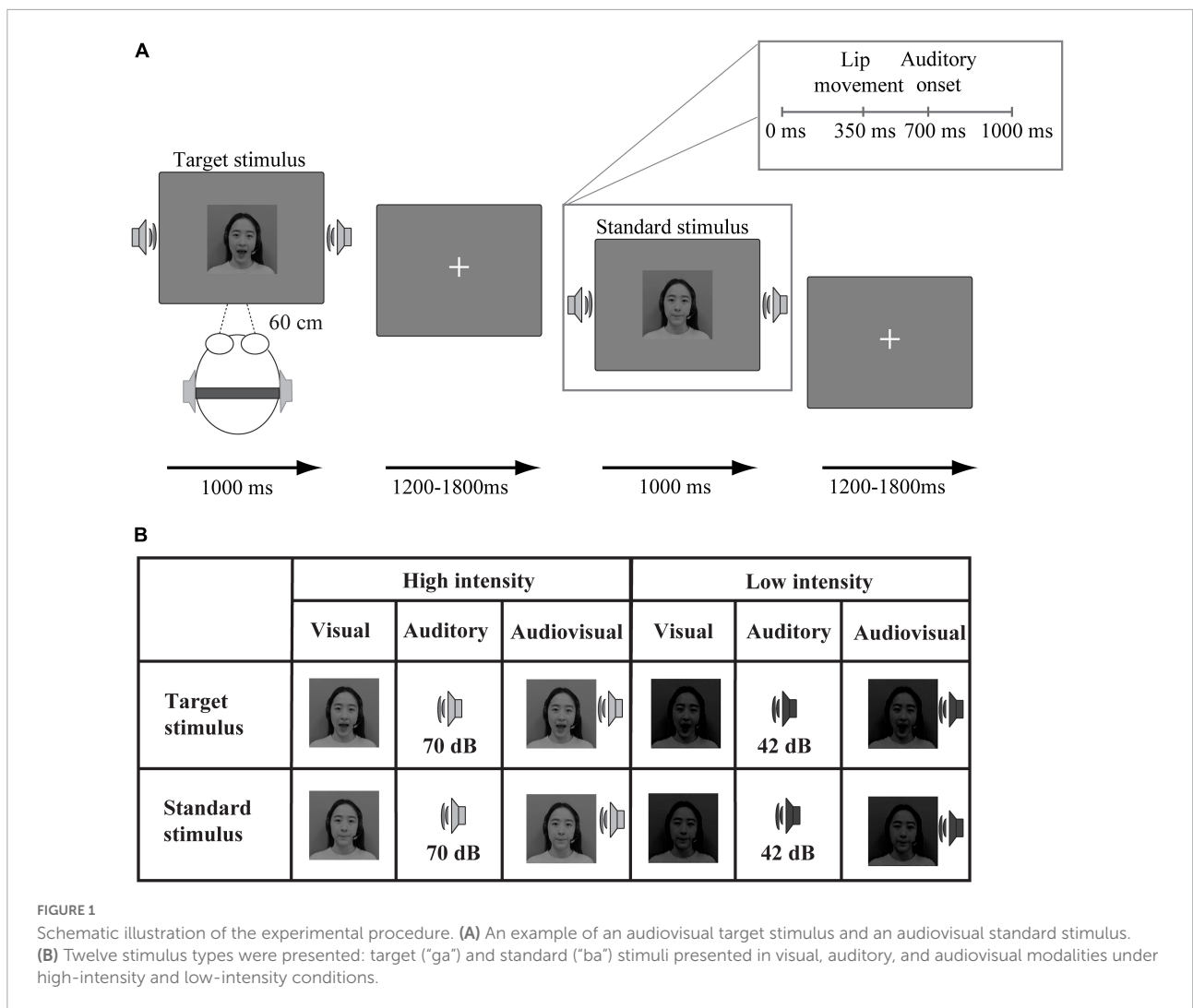
Twenty younger adults (13 females, mean age = 21.3 years, SD = 1.9) and 20 older adults (13 females, mean age = 63.5 years, SD = 3.1) were recruited for the present study. All participants were right-handed and reported having normal or corrected-to-normal vision and hearing capabilities. In addition, the Mini-Mental State Examination (MMSE) was used to screen for cognitive impairments in the older participants. This study was approved by the Ethics Committee of Hubei University, and informed consent was acquired from each participant. All participants were given a reward after the procedures were completed.

Stimuli and procedure

Audiovisual stimuli were created from audio-video recordings of a Chinese woman uttering “ga” and “ba” using Huawei Honour View 10. Unimodal stimuli were created by separating the video clips and audio tracks from the audiovisual recording using Premiere Pro Cc 2018. Under the unimodal visual condition, the video clips were presented in silence, whereas under the unimodal auditory condition, the audio clips were presented without visual information. The stimuli were presented with real audiovisual onset delay. Specifically, the auditory speech started 700 ms (“ba”) with a duration of 300 ms, 680 ms (“ga”) with a duration of 320 ms relative to the beginning of the videos. The lip movement (“ga”) started 370 ms relative to the beginning of the video and the lip movement (“ba”) started 350 ms relative to the beginning of the video. The duration of the lip movement (“ga”) was 630 ms while that of lip movement (“ba”) was 650 ms.

In addition, the visual stimuli were obtained at a resolution of 996×554 pixels and later converted to 720×480 pixels. The auditory stimuli were 32-bit and had a sampling rate of 44.1 kHz. Michelson contrast, $(\max - \min)/(\max + \min)$, was adopted to manipulate the intensity of visual stimuli. The high- and low-intensity stimuli had 90 and 10% contrast, respectively. The intensity of auditory stimuli was manipulated by the sound pressure level. Specifically, we conducted a pre-experiment (stimuli detecting task) prior to all experiments to determine the auditory stimuli intensity. The auditory stimulus at 70 dB was considered as high-intensity because the accuracy is 90% when detecting it. The auditory stimulus was decreased to 42 dB (70% accuracy), which was considered as low-intensity, see **Figure 1B**.

In the formal experiment, the stimuli were divided into two types (target and standard). The visual target stimulus was the lip movements associated with “ga,” and the standard stimulus was the lip movements associated with “ba.” The auditory target stimulus was the sound “ga,” and the standard stimulus was the



sound “ba.” Additionally, the audiovisual target stimuli were the combination of the visual and auditory target stimuli. The audiovisual standard stimuli were the combination of the visual and auditory standard stimuli.

The experiment included 3 high-intensity blocks and 3 low-intensity blocks. Each block consisted of 15 target stimuli (5 auditory, 5 visual, and 5 audiovisual) and 90 standard stimuli (30 auditory, 30 visual, and 30 audiovisual). The experiment was presented via E-prime 2.0. As shown in **Figure 1**, visual stimuli were displayed on a 19-inch Dell screen at a distance of 60 cm from the participant, and auditory stimuli were presented via earphones. The target or standard stimuli from three modalities (visual, auditory, and audiovisual) were randomly presented for 1,000 ms, see **Figure 1A**. The interstimulus interval ranged between 1,200 and 1,800 ms. The experiment was conducted in a sound-attenuated room. Participants were required to fix their eyes on the central of the monitor in which the fixation presented during presentation. The task was to press the left mouse button as quickly and accurately as possible to indicate

the target stimulus “ga” (unimodal or bimodal) but not to press the button in response to the standard stimulus “ba” (unimodal or bimodal). After completing each block, participants were allowed to take a 1-min break.

Data analysis

Behavioral data

Before analyzing the behavioral data, reactions were excluded if they were three standard deviations from the mean. The mean number of excluded trials in older adults was 0.65 (high-intensity: visual = 1, auditory = 0.56, audiovisual = 0.22; Low-intensity: visual = 0.73, auditory = 1.12, audiovisual = 0.22). The mean number of remaining trials in younger adults was 0.18 (high-intensity: visual = 0.06, auditory = 0.22, audiovisual = 0.17; low-intensity: visual = 0.22, auditory = 0.17, audiovisual = 0.22). One younger participant and two older participants were excluded because of their poor performance

detecting visual stimuli (hit rates lower than 50%). The data from 19 younger adults and 18 older adults were used for further analysis. The hit rates and mean response times to the targets were computed for both younger adults and older adults under low-intensity and high-intensity conditions. We utilized a three-factor analysis of variance (ANOVA) with age as a between-subjects factor and modality and intensity as within-subject factors (if Mauchly's sphericity was violated, Greenhouse–Geisser corrections were applied to correct the degrees of freedom). The *post-hoc* analysis used pairwise comparison with Bonferroni correction.

Event-related potential data

Electroencephalogram (EEG) signals were recorded by an EEG system (BrainAmp MR plus, Gilching, Germany) through 32 electrodes located according to the international 10–20 system using an electrode cap (Easy Cap, Herrsching Breitbrunn, Germany). All signals were referenced to the FCz electrode. Electrooculogram (EOG) data were recorded by two additional electrodes. Horizontal eye movements were recorded from an electrode placed at the outer canthus of the left eye, and vertical eye movements were recorded from an electrode placed approximately 1 cm below the right eye. The impedances of all electrodes were maintained below 5 k Ω .

Electroencephalogram data was stored digitally for offline analysis and was conducted by using Brain Vision Analyser software, including rereferencing to the average of TP9 and TP10. The EEG signals were bandpass filtered from 0.01 to 60 Hz at a sampling rate of 250 Hz. The ERPs elicited by standard stimuli were analyzed as in previous studies (Yang et al., 2021). The ERP epochs in three conditions (visual, auditory, and audiovisual) were locked to auditory onset. Then, the data were epoched from 100 ms before to 600 ms after the onset of the stimuli, and baseline corrections were made from –100 to 0 ms relative to stimulus onset. In addition to the fact that EEG epochs containing eye blinks and small artifacts were removed based on a trail-by-trail inspection of the data, epochs contaminated with large artifacts were rejected from the subsequent averaging if the voltage exceeded ± 100 μ V at any electrode site relative to baseline. Then, all the remained data without artifacts were averaged for each participant, each stimulus type (auditory, visual, and audiovisual), each stimulus intensity, following digital filtering with a bandpass filter of 0.01–30 Hz. The grand-averaged data were obtained across all the conditions and groups. The mean number of rejected epochs in older adults was 14.30 (high-intensity: visual = 13.56, auditory = 14.28, audiovisual = 13.39; Low-intensity: visual = 14.89, auditory = 14.61, audiovisual = 15.06). The mean number of rejected epochs in younger adults was 13.89 (high-intensity: visual = 16.37, auditory = 13.53, audiovisual = 13.95; Low-intensity: visual = 11.90, auditory = 12.74, audiovisual = 14.84). For remained epochs, the mean number in older adults was 75.70 (high-intensity:

visual = 76.44, auditory = 75.72, audiovisual = 76.61; Low-intensity: visual = 75.11, auditory = 75.39, audiovisual = 74.94). The mean number in younger adults was 76.11 (high-intensity: visual = 73.63, auditory = 76.47, audiovisual = 76.05; Low-intensity: visual = 78.10, auditory = 77.26, audiovisual = 75.16). Two older adults and one younger adult were excluded from further analysis due to the loss of more than 70% of the epochs in at least one stimulus type (the excluded participants were identical to those in behavioral data).

To assess the audiovisual integration in ERPs, we calculated the result of $ERP(AV) - [ERP(A) + ERP(V)]$ by subtracting the summed ERPs on the unimodal visual and auditory trials from the ERPs on the audiovisual trials at each time bin for each electrode (Giard and Peronnet, 1999). Then, one-sample *t*-tests were conducted to compare the amplitude of $AV - (A + V)$ with 0 in each time bin ranging from 0 to 600 ms. If more than 7 consecutive data points were significant ($\alpha < 0.05$), the participant was considered to display audiovisual integration (Senkowski et al., 2011). In the present data, audiovisual integration was defined as at least 7 or more consecutive data points that met an alpha criterion of 0.05, which suited a Bonferroni correction for multiple comparisons. This criterion meets the strict criteria for assessing reliable results when a large number of *t*-tests are calculated (Guthrie and Buchwald, 1991). Based on the results of the *t*-tests, three-time intervals of integration (200–230 ms, 320–360 ms, and 460–500 ms) and five regions of interest (frontal: Fz, F3, F4, F7, F8; fronto-central: FC1, FC2, FC5, FC6; central: C3, C4, Cz; centro-parietal: CP1, CP2, CP5, CP6; occipital: O1, O2, Oz) were chosen for further analysis. In addition, ERPs were averaged across electrodes within each region of interest. For each of these time intervals (200–230 ms, 320–360 ms, and 460–500 ms), we conducted a repeated-measures ANOVA with the mean amplitude of $ERP(AV) - [ERP(A) + ERP(V)]$ as the dependent variable. The ANOVA included the between-subject factor of age and the within-subject factors of stimuli intensity and region of interest (ROI). Greenhouse–Geisser corrections were applied to correct the degrees of freedom if the sphericity assumption was violated. The *post-hoc* analysis used pairwise comparison with Bonferroni correction. All statistical analyses were conducted by SPSS version of 26.0 software package.

Results

Behavioral results

Hit rates

The results showed a main effect of modality [$F(2, 70) = 4.36$, $p = 0.02$, $\eta_p^2 = 0.11$]; specifically, audiovisual targets ($M = 98\%$, $SE = 0.3\%$) detected correctly at a significantly higher rate than either unimodal visual targets ($M = 94\%$, $SE = 0.8\%$) or auditory targets ($M = 95\%$, $SE = 1.3\%$). There was also

a significant main effect of age [$F_{(1, 35)} = 6.25, p = 0.02, \eta_p^2 = 0.15$], as younger adults ($M = 98\%$, $SE = 0.8\%$) exhibited a significantly higher hit rate than older adults ($M = 94\%$, $SE = 1\%$). Additionally, we found a significant effect of intensity [$F_{(1, 35)} = 4.58, p = 0.04, \eta_p^2 = 0.15$], suggesting that high-intensity targets ($M = 97\%$, $SE = 0.5\%$) were detected correctly at a significantly higher rate than low-intensity targets ($M = 95\%$, $SE = 1\%$). However, the interaction of these three factors was not significant [$F_{(2, 70)} = 0.44, p > 0.05, \eta_p^2 = 0.01$]. Additionally, the intensity \times age, modality \times age, and intensity \times modality interactions were not significant [$F_{(1, 35)} = 0.25, p > 0.05, \eta_p^2 = 0.01$], [$F_{(2, 70)} = 0.34, p > 0.05, \eta_p^2 = 0.01$], and [$F_{(2, 70)} = 0.86, p > 0.05, \eta_p^2 = 0.02$], respectively. The detailed hit rates were demonstrated in **Table 1**.

Response times

The results revealed that the main effect of modality was significant [$F_{(2, 70)} = 212.51, p < 0.001, \eta_p^2 = 0.86$]; specifically, the response time to audiovisual stimuli ($M = 579$ ms, $SE = 25$ ms) was significantly shorter than that to visual ($M = 618$ ms, $SE = 28$ ms) and auditory stimuli ($M = 907$ ms, $SE = 23$ ms) alone. The main effect of age was also significant [$F_{(1, 35)} = 36.17, p < 0.001, \eta_p^2 = 0.51$], indicating that the response time of the older adults ($M = 840$ ms, $SE = 33$ ms) was longer than that of the younger adults ($M = 562$ ms, $SE = 32$ ms). However, there was no significant main effect of intensity [$F_{(1, 35)} = 3.54, p > 0.05, \eta_p^2 = 0.09$].

In addition, we observed a two-way interaction between age and stimulus intensity [$F_{(1, 35)} = 5.98, p = 0.02, \eta_p^2 = 0.15$]. Therefore, further *post-hoc* analyses were conducted. Specifically, regarding age, we found that older adults ($M = 844$ ms, $SE = 33$ ms) responded significantly slower than younger adults ($M = 534$ ms, $SE = 32$ ms) under the high-intensity condition [$t(35) = 6.67, p < 0.001$]. Similar results were found under the low-intensity condition, with the reaction time for older adults ($M = 837$ ms, $SE = 35$ ms) being significantly longer than that for younger adults ($M = 590$ ms, $SE = 34$ ms) [$t(35) = 4.97, p < 0.001$]. Regarding stimulus intensity factor, there was no significant difference in response times between the high-intensity ($M = 844$ ms, $SE = 33$ ms) and low-intensity conditions ($M = 837$ ms, $SE = 35$ ms) in older adults [$t(17) = 0.41, p > 0.05$]. However, younger adults responded to high-intensity stimuli ($M = 534$ ms, $SE = 32$ ms) significantly faster than to low-intensity stimuli ($M = 590$ ms, $SE = 34$ ms) [$t(18) = -3.02, p < 0.01$]. There was no three-way interaction [$F_{(2, 70)} = 2.62, p = 0.09, \eta_p^2 = 0.07$]. The detailed response times were shown in **Table 1**.

Event-related potential results

Audiovisual integration at 200–230 ms

The ANOVA results showed no main effects of intensity, ROI or age [$F_{(1, 35)} = 3.82, p > 0.05, \eta_p^2 = 0.10$], [$F_{(4, 140)} = 1.94, p > 0.05, \eta_p^2 = 0.06$], and [$F_{(1, 35)} = 1.29, p > 0.05, \eta_p^2 = 0.04$], respectively.

140) = 1.94, $p > 0.05, \eta_p^2 = 0.06$], and [$F_{(1, 35)} = 1.29, p > 0.05, \eta_p^2 = 0.04$], respectively.

Additionally, the three-way interaction was also not significant [$F_{(4, 140)} = 1.11, p > 0.05, \eta_p^2 = 0.03$]. However, the analysis showed a significant interaction between ROI and age [$F_{(4, 140)} = 3.31, p = 0.05, \eta_p^2 = 0.07$]. *Post-hoc* analysis indicated that the amplitude in the frontal region of older adults ($M = 0.68 \mu\text{V}$, $SE = 0.44 \mu\text{V}$) was significantly greater than that of younger adults ($M = -0.52 \mu\text{V}$, $SE = 0.41 \mu\text{V}$) [$t(35) = 2.01, p = 0.05$]. The interaction between intensity and age was also significant [$F_{(1, 35)} = 13.64, p < 0.001, \eta_p^2 = 0.29$]. *Post-hoc* analysis indicated that the amplitude under the low-intensity condition ($M = 1.05 \mu\text{V}$, $SE = 0.43 \mu\text{V}$) was more positive than that under the high-intensity condition ($M = -0.87 \mu\text{V}$, $SE = 0.40 \mu\text{V}$) in older adults [$t(17) = -4.61, p < 0.001$], see **Figures 2, 3**. However, such an effect was not found in younger adults (high-intensity: $M = 0.10 \mu\text{V}$, $SE = 0.37 \mu\text{V}$, low-intensity: $M = -0.82 \mu\text{V}$, $SE = 0.39 \mu\text{V}$) [$t(18) = 1.46, p > 0.05$]. Moreover, in terms of age factor, greater amplitude was also found in older adults ($M = 1.05 \mu\text{V}$, $SE = 0.43 \mu\text{V}$) under the low-intensity condition comparing to younger adults ($M = -0.82 \mu\text{V}$, $SE = 0.39 \mu\text{V}$) [$t(35) = 3.23, p < 0.01$]. There was no significant difference between younger ($M = 0.10 \mu\text{V}$, $SE = 0.37 \mu\text{V}$) and older adults ($M = -0.87 \mu\text{V}$, $SE = 0.40 \mu\text{V}$) under the high-intensity condition [$t(35) = -1.72, p > 0.05$], see **Figure 4A**.

Audiovisual integration at 320–360 ms

In this time window, we found that there was a significant main effect of age [$F_{(1, 35)} = 5.59, p = 0.02, \eta_p^2 = 0.15$], indicating that the amplitude in older adults ($M = 1.00 \mu\text{V}$, $SE = 0.48 \mu\text{V}$) was more positive than that of younger adults ($M = -0.54 \mu\text{V}$, $SE = 0.44 \mu\text{V}$), see **Figures 4B, 5**. The main effect of ROI was also significant [$F_{(4, 140)} = 9.04, p < 0.001, \eta_p^2 = 0.22$], with the amplitude of the frontal region ($M = 1.37 \mu\text{V}$, $SE = 0.56 \mu\text{V}$) being significantly more positive than that in the central ($M = -0.42 \mu\text{V}$, $SE = 0.38 \mu\text{V}$) and centro-parietal ($M = -0.37 \mu\text{V}$, $SE = 0.29 \mu\text{V}$) and fronto-central regions ($M = 0.44 \mu\text{V}$, $SE = 0.42 \mu\text{V}$). The amplitude of the fronto-central region ($M = 0.44 \mu\text{V}$, $SE = 0.42 \mu\text{V}$) was greater than that in the central ($M = -0.42 \mu\text{V}$, $SE = 0.38 \mu\text{V}$) and centro-parietal regions ($M = -0.37 \mu\text{V}$, $SE = 0.29 \mu\text{V}$). The main effect of intensity was not significant [$F_{(1, 35)} = 0.18, p > 0.05, \eta_p^2 = 0.01$].

The three-way interaction was not significant [$F_{(4, 140)} = 1.27, p > 0.05, \eta_p^2 = 0.04$]. The interaction between intensity and age was also not found [$F_{(1, 35)} = 3.89, p = 0.057, \eta_p^2 = 0.11$].

Audiovisual integration at 460–500 ms

The ANOVA results showed that the main effect of intensity was not significant [$F_{(1, 35)} = 2.91, p > 0.05, \eta_p^2 = 0.08$]. There was a significant main effect of ROI [$F_{(4, 140)} = 3.37,$

TABLE 1 Mean response times and hit rates for younger and older adults under high-intensity and low-intensity conditions.

	Low intensity			High intensity		
	Visual	Auditory	Audiovisual	Visual	Auditory	Audiovisual
Older adults						
HT (%)	92 (11)	92 (15)	96 (3)	94 (6)	93 (8)	98 (3)
RT (ms)	778 (205)	1,021 (175)	712 (158)	804 (208)	1004 (161)	724 (180)
Younger adults						
HT (%)	94 (9)	96 (5)	99 (2)	97 (4)	98 (2)	99 (1)
RT (ms)	481 (150)	821 (126)	468 (158)	409 (145)	780 (104)	412 (137)

Standard deviations are presented in parentheses.

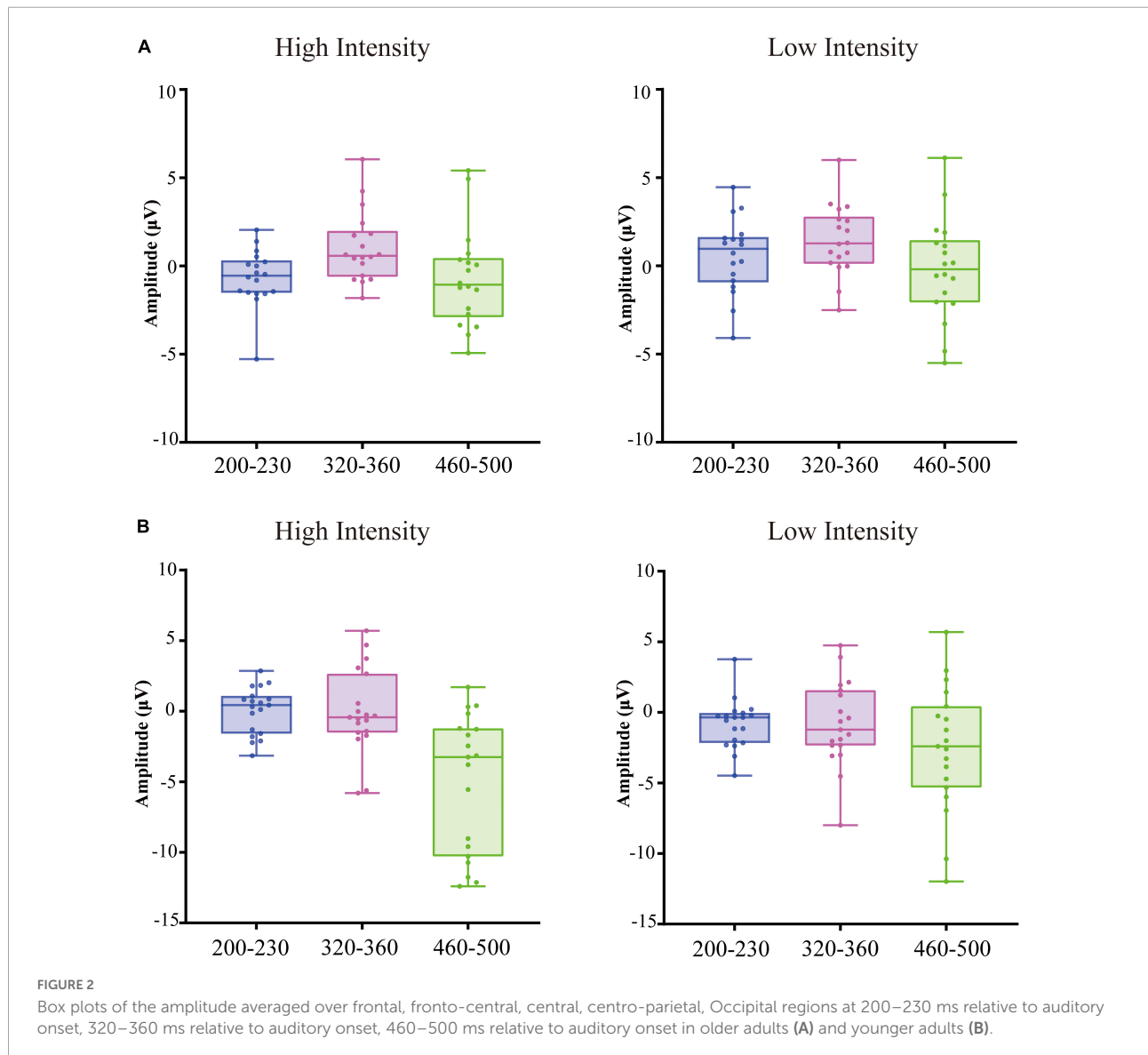
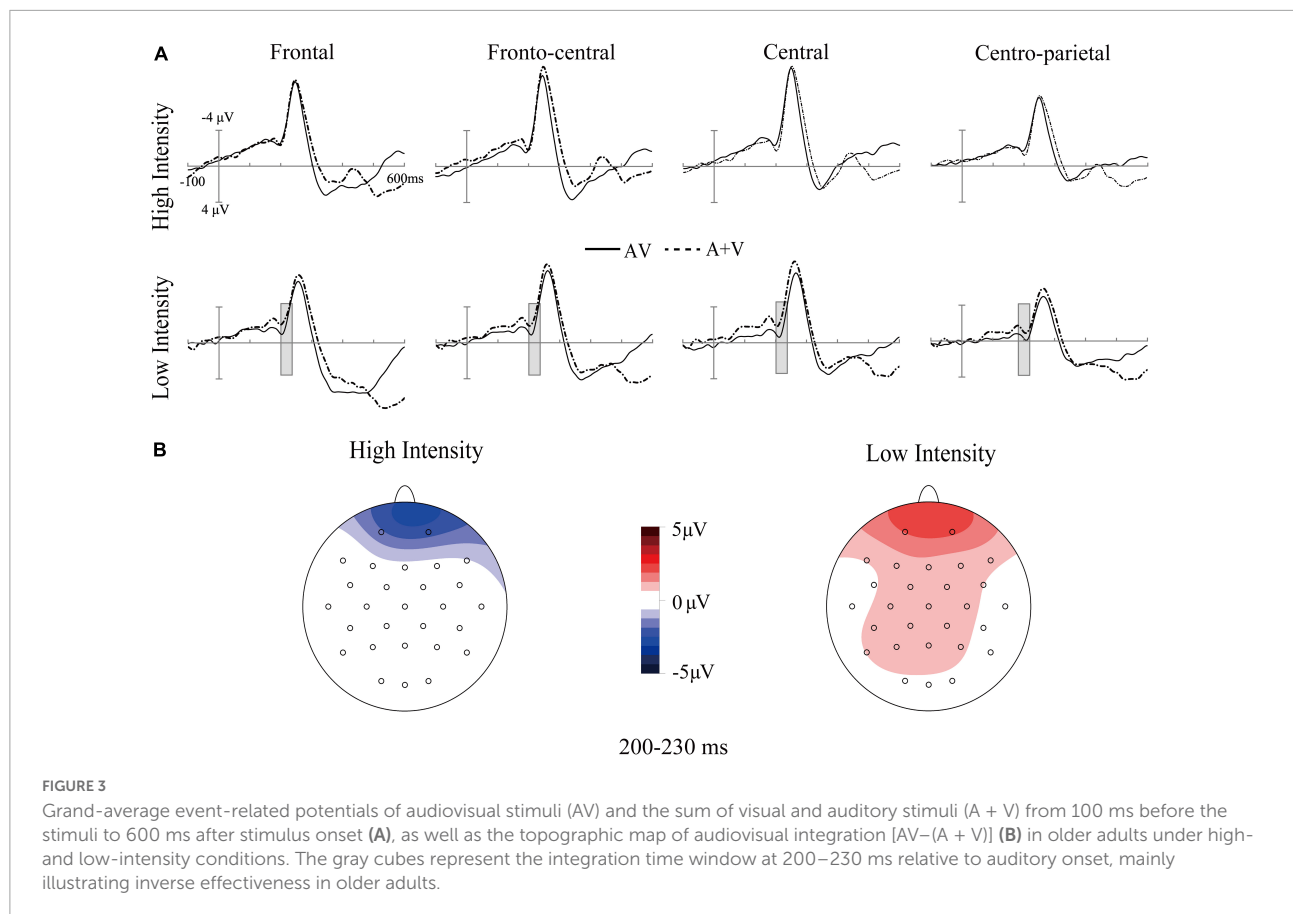


FIGURE 2 Box plots of the amplitude averaged over frontal, fronto-central, central, centro-parietal, Occipital regions at 200–230 ms relative to auditory onset, 320–360 ms relative to auditory onset, 460–500 ms relative to auditory onset in older adults (A) and younger adults (B).

$p = 0.01$, $\eta_p^2 = 0.10$], with the amplitude of central region ($M = -3.16 \mu V$, $SE = 0.65 \mu V$) being more negative than that in the centro-parietal ($M = -2.56 \mu V$, $SE = 0.49 \mu V$) and

occipital regions ($M = -1.37 \mu V$, $SE = 0.35 \mu V$); additionally, the amplitude of fronto-central ($M = -3.14 \mu V$, $SE = 0.79 \mu V$) and centro-parietal ($M = -2.56 \mu V$, $SE = 0.49 \mu V$) regions was



more negative than that in the occipital region ($M = -1.37 \mu\text{V}$, $SE = 0.35 \mu\text{V}$). A significant main effect of age was found [$F_{(1, 35)} = 6.98$, $p = 0.01$, $\eta_p^2 = 0.17$], with a more negative amplitude in younger adults ($M = -3.56 \mu\text{V}$, $SE = 0.89 \mu\text{V}$) than in older adults ($M = -1.71 \mu\text{V}$, $SE = 0.78 \mu\text{V}$), see **Figure 5**.

The three-way interaction was not significant [$F_{(4, 140)} = 2.15$, $p > 0.05$, $\eta_p^2 = 0.06$]. However, we found a significant interaction between ROI and age [$F_{(4, 140)} = 9.06$, $p < 0.001$, $\eta_p^2 = 0.21$]. *Post-hoc* analysis revealed that the amplitudes were more negative in the frontal ($M = -5.37 \mu\text{V}$, $SE = 1.44 \mu\text{V}$), fronto-central ($M = -5.19 \mu\text{V}$, $SE = 1.07 \mu\text{V}$), central ($M = -4.83 \mu\text{V}$, $SE = 0.88 \mu\text{V}$), centro-parietal ($M = -3.6 \mu\text{V}$, $SE = 0.66 \mu\text{V}$) regions of younger adults than those in older adults (frontal: $M = -0.49 \mu\text{V}$, $SE = 1.58 \mu\text{V}$, fronto-central: $M = -1.08 \mu\text{V}$, $SE = 1.16 \mu\text{V}$, central: $M = -1.49 \mu\text{V}$, $SE = 0.96 \mu\text{V}$, centro-parietal: $M = -1.25 \mu\text{V}$, $SE = 0.51 \mu\text{V}$) [$t(35) = 2.28$, $p < 0.05$; $t(35) = 2.60$, $p < 0.05$; $t(35) = 2.57$, $p < 0.05$; $t(35) = 2.13$, $p < 0.05$]; see **Figure 4C**.

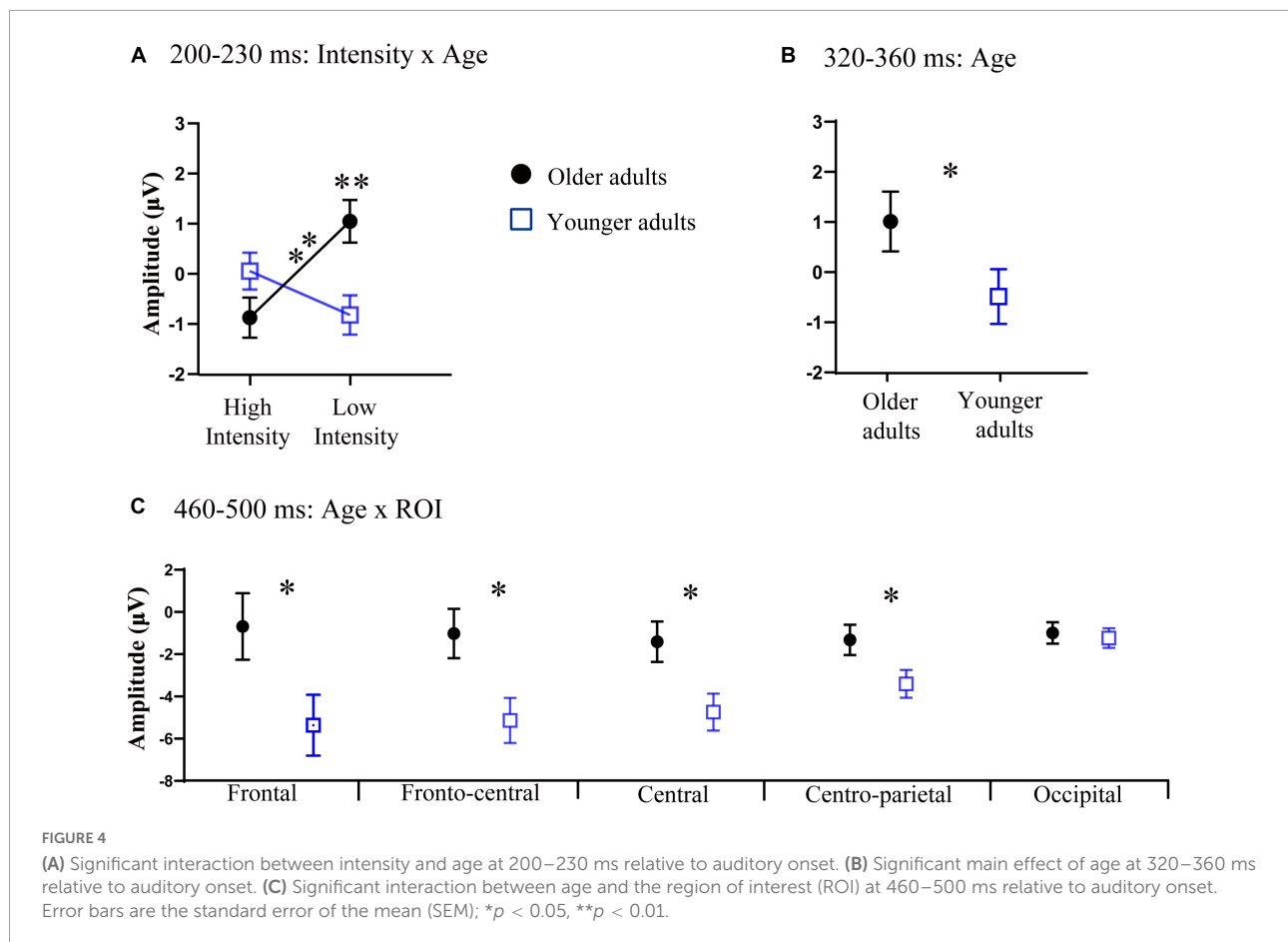
Discussion

The current study compared audiovisual integration in older and younger adults under high- and low-intensity conditions.

The results showed enhanced audiovisual integration at 320–360 ms in older adults compared to younger adults, regardless of stimulus intensity and ROI. Conversely, at a relatively later stage (460–500 ms), attenuated audiovisual integration in the frontal, fronto-central, central, and centro-parietal regions were found in the older adults compared to younger adults. In addition, in older adults, we found greater audiovisual integration at 230–300 ms under the low-intensity condition than the high-intensity condition, suggesting inverse effectiveness occurred. However, inverse effectiveness was not observed in younger adults.

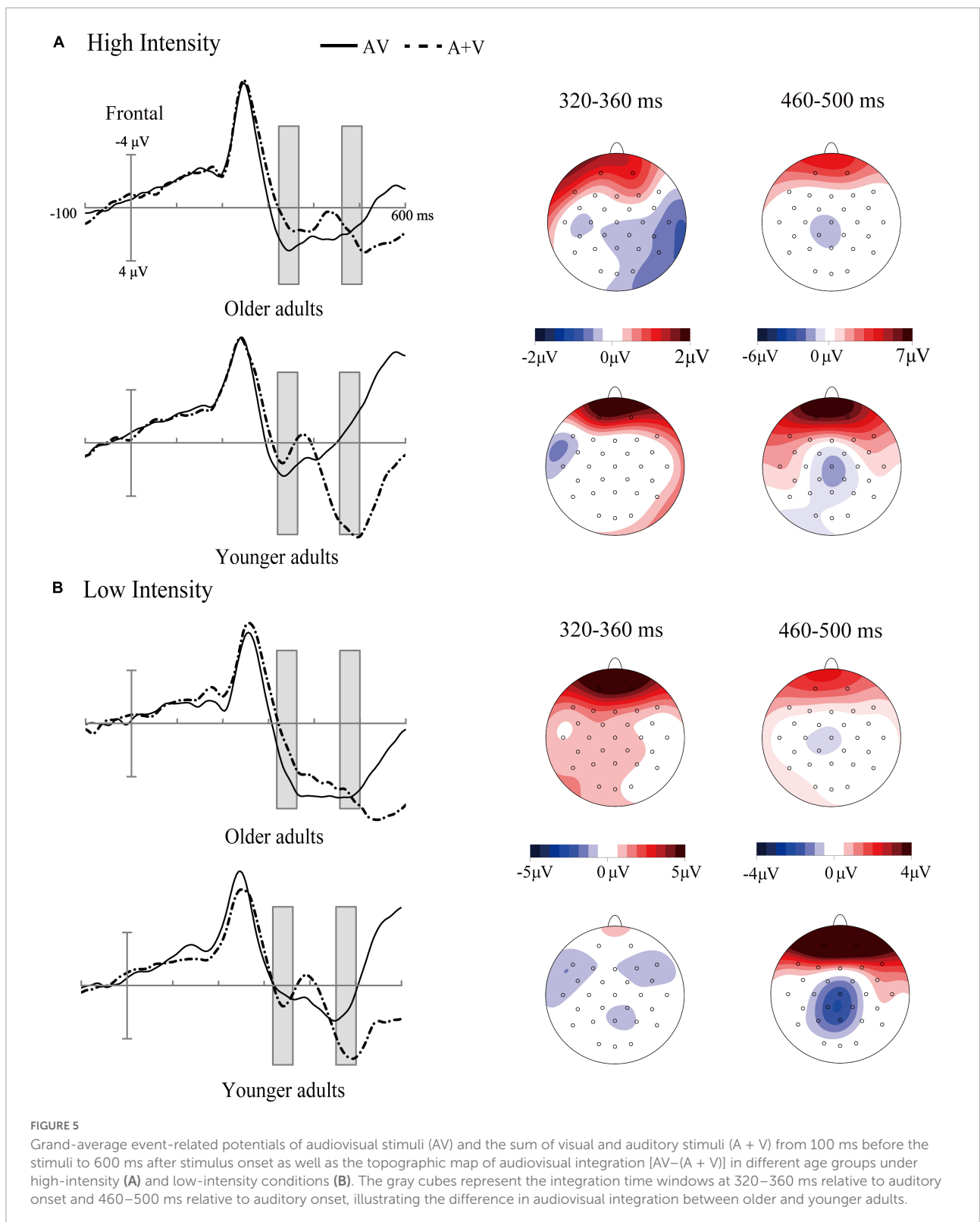
Effect of aging on audiovisual integration

At 320–360 ms, older adults exhibited greater audiovisual integration than younger adults. The weaker ability of suppressing crossmodal information in older adults may render a plausible explanation for the results (de Dieuleveult et al., 2017). Previous study utilized a cued audiovisual discrimination task to explore changes in audiovisual integration under selective and divided conditions between age groups (Hugenschmidt et al., 2009a). The results not only



showed that older adults had greater audiovisual integration under selective conditions but also that older adults exhibited greater audiovisual integration under divided conditions than younger adults (Hugenschmidt et al., 2009a). This finding might be because younger adults could ignore modality-specific irrelevant information, which suppressed audiovisual integration. In contrast, older adults may have been unable to judge whether the information was irrelevant effectively, leading to enhanced audiovisual integration. Moreover, their subsequent study found neural evidence that older adults showed greater brain activity than younger adults when they were involved in irrelevant information (Hugenschmidt et al., 2009b). From this neural evidence along with our own behavioral data indicating that older adults achieved lower accuracy (Table 1), we infer that the greater audiovisual integration of older adults reflected greater distraction from irrelevant information. This phenomenon was also found in other audiovisual perception tasks, such as the McGurk task or the sound-induced illusion task, in which the older adults had a higher possibility of being influenced by the interaction between visual modality and auditory modality and was more vulnerable to yield the audiovisual effect (DeLoss et al., 2013; Sekiyama et al., 2014).

Because overall good performance was found in both groups, indicating that older adults can achieve a level of performance comparable to that of younger adults. This could be explained by compensatory mechanisms, which help maintain good performance in older adults (Cabeza et al., 2002; Diaconescu et al., 2013; Freiherr et al., 2013). In other words, increased brain activity in some regions may compensate the older adults cognitive decline to improve their performance. Moreover, Enhanced activation in older adults was relative to younger adults. The compensatory interpretation is often invoked when older adults show more activity in a brain region than younger adults when they perform the same task (Cabeza et al., 2002). If we find enhanced activation in older adults comparing to younger adults and their behavioral performance was also preserved comparing to younger adults, the compensatory mechanisms may work (Davis et al., 2012). Some studies have further confirmed that enhanced activation in frontal and posterior parietal areas is positively associated with the performance of speech perception in older adults, suggesting a compensatory role in aiding older adults to achieve better performance (Wong et al., 2009). However, Reuter-Lorenz and Cappell (2008) proposed that the compensatory mechanism is effective at lower levels of task



demand. When task-demanding increases at high levels of the task, the cognitive resources may reach limitation, resulting in insufficient processing and poor performance in older

adults (Rypma et al., 2007; Reuter-Lorenz and Cappell, 2008). For instance, studies utilizing working memory tasks have shown greater activation in the prefrontal and parietal cortices

in older adults than younger adults under relatively low working memory loads; conversely, less activity in these areas in older adults was observed under high working memory loads (Schneider and Pichorafuller, 2000; Mattay et al., 2006). In terms of speech stimuli, the findings of the current study, which used relatively elementary speech stimuli (i.e., phonemes), were consistent with those of a previous study, which reported that audiovisual processing was intact at low levels of speech perception in older adults (Stevenson et al., 2015). In other words, the compensatory mechanism might have worked in older adults due to the elementary speech stimuli we used.

Despite well-documented increases in audiovisual integration related to speech stimuli in older adults (Cienkowski and Carney, 2002), we found attenuated audiovisual integration in the frontal, fronto-central, central, and parietal regions in older adults compared to younger adults at 460–500 ms. Xi et al. (2020) suggested that different integration time windows may reflect different stages of processing during audiovisual integration (Xi et al., 2020). These authors investigated the neural activity associated with audiovisual integration in different attention conditions. During the experiment in their studies, participants were required to focus on the stimuli on one side while ignoring the stimuli on the opposite side, which produced two types of attention conditions. Under attended condition, when the participants were asked to respond to stimuli on one side (e.g., left session), the stimuli were presented on the same side (e.g., left side). Under unattended condition, for instance, the participants were asked to respond to stimuli on right side (right session), the stimuli were presented on left side. Specifically, in the attended condition, audiovisual integration was observed in the frontal regions at 560–600 ms. In the unattended condition, audiovisual integration was observed in the centro-frontal regions at 340–360 ms. The authors suggested that audiovisual integration (340–360 ms) in the unattended condition was related to automatic processes (Xi et al., 2020). More importantly, a goal-driven process influenced by top-down attention was located at a later audiovisual integration time window (560–600 ms) in the attended condition (Xi et al., 2020). In the present study, participants were asked to attend to the speech stimuli and to carefully focus on the speaker's lip movements and utterances to discriminate which stimuli to respond to, suggesting that the task was under attended condition. Therefore, we infer that later audiovisual integration at 460–500 ms in our research may reflect the influence of top-down attention processing. Additionally, previous studies have shown that the frontal and fronto-central regions are highly associated with cognitive function, such as top-down attention (de Dieuleveult et al., 2017). We infer that deficit in top-down attention may offer an interpretation for older adults having attenuated audiovisual integration in these regions compared to younger adults.

Inverse effectiveness in audiovisual integration

Interestingly, N200 component could be observed at 200–230 ms (see Figure 3), which may reflect the auditory attentional processes during audiovisual speech processing (Tremblay et al., 2021). However, our main concern is the time window of audiovisual integration similar to previous study (Ren et al., 2018). At 200–230 ms, enhanced audiovisual integration in older adults was observed when the presented stimuli were in the low-intensity condition, which reflects the involvement of higher processing resources. The theory of cognitive resource limitations describes the mutual constraint relationship between sensory and cognitive processing (Schneider and Pichorafuller, 2000). In particular, sensory and cognitive processing share common resources from unified information-processing systems, indicating that alterations in either processing would influence the other (Schneider and Pichorafuller, 2000). When stimuli intensity decreases, older adults may not effectively perceive the degraded information delivered by the stimuli. More processing resources may have been recruited to engage in this stage at the cost of the processing resources allocated to cognitive processing, resulting in greater audiovisual integration in older adults under low-intensity conditions.

Additionally, according to predictive coding theory, our perception is determined by a trade-off between prediction based on prior experience accumulated over our life span and efficiency of sensory processing (Sohoglu et al., 2012; Alm and Behne, 2013; Chan et al., 2017). However, a Bayesian model suggests that the weight between predictions and efficiency of sensory processing changes with age (Chan et al., 2017). Some investigations using sound-induced flash tasks have attempted to explore differences in the weighting of predictions and sensory efficiency in older and younger adults. Their results showed that older adults had increased beta-band activity before processing the stimulus, indicating that older adults may rely more heavily on predictions based on prior experience (Chan et al., 2017). In the current study, visual speech information could play a crucial role in this predictive process. Because in many cases, lip movements are visible which may set up predictions about what is about to be heard. The prediction based on visual information (lip movements) may depend on the prior experience. Studies have shown that audiovisual speech integration is a skill that is obtained by experiencing sound and observing lip movements (Dick et al., 2010). These experiences could contribute to the development of audiovisual speech (Dick et al., 2010). As the amount of information accumulated in our priors increases throughout one's lifetime, older adults could make use of experience to perceive the low stimuli that delivered insufficient information.

Interestingly, some studies have reported that IE in audiovisual integration occurs at an early stage, indicating that early sensory processing may be associated with IE

(Giard and Peronnet, 1999; Senkowski et al., 2011). However, our results showed delayed IE compared with these studies. This dissociation could be explained by additional cognitive processing engaged in audiovisual speech. Previous studies have investigated audiovisual integration at different levels of complexity. At lower levels, researchers have adopted simple stimuli, such as white noise or flashing lights, to elicit audiovisual integration. These stimuli can be integrated due to the factor that they coincide with regard to time, suggesting that only interactions of sensory processing occurred. At higher levels, such as audiovisual speech, speech stimuli from different modalities may still have a temporal relationship, but may also have a connection beyond the temporal factor (Wyk et al., 2010). That is, the speech sounds and lip movements are not only related in time, but also related by phonological knowledge. The current study showed that delayed IE might reflect phonological knowledge processing rather than sensory processing.

Our discussion was based on the ERP results of audiovisual integration elicited by the approach ($ERP(AV) - [ERP(A) + ERP(V)]$). The limitations of this approach result in if anticipatory slow are elicited in association with each of the stimulus types (A, V, AV). The potentials of audiovisual integration ($ERP(AV) - [ERP(A) + ERP(V)]$) is subtracted twice, creating a deflection that might be confounded with true audiovisual integration (Teder-Sälejärvi et al., 2002). Future studies may take this effect into consideration.

Conclusion

We present results of older and younger adults on an audiovisual discrimination task with varying levels of stimulus intensity. These results confirm differences in the audiovisual integration process between the age groups. Greater audiovisual integration was found in older adults at a relatively earlier stage (320–360 ms). However, due to the deficits of top-down attention in older adults, their audiovisual integration was attenuated at a relatively later stage (460–500 ms). Additionally, the results showed different effects of stimulus intensity in older and younger adults. That is, IE occurred exclusively in older adults. Moreover, our research showed that such an effect was delayed (200–230 ms) compared that found in a previous study regarding IE; this delay might be associated with phonological knowledge processing in audiovisual speech.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Hubei University, Wuhan, China. The patients/participants provided their written informed consent to participate in this study.

Author contributions

WY and ZZ conceived and designed the experiment. XY, ZL, and SL collected the data. AG analyzed the data, wrote the manuscript, and received the comments from YR and JW. HY and JC interpreted the data. All authors contributed to the article and approved the submitted version.

Funding

This insight was supported by grants from the National Natural Science Foundation of China (31700973, 31800932, 62103404, and 32260198), the Humanity and Social Science Youth Foundation of Ministry of Education of China (16YJC190025 and 18XJC190003), and the Humanity and Social Science Youth Foundation of the Education Bureau of Hubei Province of China (16Q030), Science and Technology Planning Project of Guizhou Province (QianKeHeJiChu-ZK [2021] General 120), Shenzhen Overseas Innovation Team Project (KQTD20180413181834876), Shenzhen Basic Research Program (JCYJ20210324101402008 and JCYJ20210324115810030), and the CAS President's International Fellowship Initiative (2022VBA0031).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alm, M., and Behne, D. (2013). Audio-visual speech experience with age influences perceived audio-visual asynchrony in speech. *J. Acoust. Soc. Am.* 134, 3001–3010. doi: 10.1121/1.4820798
- Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., and Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neurosci. Biobehav. Rev.* 30, 791–807. doi: 10.1016/j.neubiorev.2006.06.005
- Cabeza, R., Anderson, N. D., Locantore, J. K., and Mcintosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Chan, J. S., Wibral, M., Wollstadt, P., Stawowsky, C., Brandl, M., Helbling, S., et al. (2017). Predictive coding over the lifespan: Increased reliance on perceptual priors in older adults—a magnetoencephalography and dynamic causal modelling study. *bioRxiv* [Preprint] 2, 80–95. doi: 10.1101/178095
- Cienkowski, K. M., and Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear Hear.* 23, 439–449. doi: 10.1097/01.AUD.0000034781.95122.15
- Corneil, B., Van Wanrooij, M., Munoz, D., and Van Opstal, A. (2002). Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J. Neurophysiol.* 88, 438–454. doi: 10.1152/jn.2002.88.1.438
- Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., and Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cereb. Cortex* 18, 1201–1209. doi: 10.1093/cercor/bhm155
- Davis, S. W., Kragel, J. E., Madden, D. J., and Cabeza, R. (2012). The architecture of cross-hemispheric communication in the aging brain: Linking behavior to functional and structural connectivity. *Cereb. Cortex* 22, 232–242. doi: 10.1093/cercor/bhr123
- de Dieuleveult, A. L., Siemonsma, P. C., van Erp, J. B., and Brouwer, A.-M. (2017). Effects of aging in multisensory integration: A systematic review. *Front. Aging Neurosci.* 9:80. doi: 10.3389/fnagi.2017.00080
- DeLoss, D. J., Pierce, R. S., and Andersen, G. J. (2013). Multisensory integration, aging, and the sound-induced flash illusion. *Psychol. Aging* 28, 802–812. doi: 10.1037/a0033289
- Diaconescu, A. O., Hasher, L., and McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *Neuroimage* 65, 152–166. doi: 10.1016/j.neuroimage.2012.09.057
- Dick, A. S., Solodkin, A., and Small, S. L. (2010). Neural development of networks for audiovisual speech comprehension. *Brain Lang.* 114, 101–114. doi: 10.1016/j.bandl.2009.08.005
- Freiherr, J., Lundstrom, J. N., Habel, U., and Retz, K. (2013). Multisensory integration mechanisms during aging. *Front. Hum. Neurosci.* 7:863. doi: 10.3389/fnhum.2013.00863
- Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/089992999563544
- Gordon, M. S., and Allen, S. (2009). Audiovisual speech in older and younger adults: Integrating a distorted visual signal with speech in noise. *Exp. Aging Res.* 35, 202–219. doi: 10.1080/03610730902720398
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505.
- Guthrie, D., and Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology* 28, 240–244. doi: 10.1111/j.1469-8986.1991.tb00417.x
- Hugenschmidt, C. E., Mozolic, J. L., and Laurienti, P. J. (2009a). Suppression of multisensory integration by modality-specific attention in aging. *Neuroreport* 20, 349–353. doi: 10.1097/WNR.0b013e328323ab07
- Hugenschmidt, C. E., Mozolic, J. L., Tan, H., Kraft, R. A., and Laurienti, P. J. (2009b). Age-related increase in cross-sensory noise in resting and steady-state cerebral perfusion. *Brain Topogr.* 21, 241–251. doi: 10.1007/s10548-009-0098-1
- Jansen, S. D., Keebler, J. R., and Chaparro, A. (2018). Shifts in maximum audiovisual integration with age. *Multisens. Res.* 31, 191–212. doi: 10.1163/22134808-00002599
- Laurienti, P. J., Burdette, J. H., Maldjian, J. A., and Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol. Aging* 27, 1155–1163. doi: 10.1016/j.neurobiolaging.2005.05.024
- Liu, X., and Yan, D. (2007). Ageing and hearing loss. *J. Pathol.* 211, 188–197. doi: 10.1002/path.2102
- Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., and Weinberger, D. R. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neurosci. Lett.* 392, 32–37. doi: 10.1016/j.neulet.2005.09.025
- McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., and Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Front. Aging Neurosci.* 6:250. doi: 10.3389/fnagi.2014.00250
- Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., and Laurienti, P. J. (2012). “Multisensory integration and aging,” in *The neural bases of multisensory processes*, eds M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press/Taylor & Francis).
- Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., and Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport* 18, 1077–1081. doi: 10.1097/WNR.0b013e3281e72ae7
- Ren, Y., Ren, Y., Yang, W., Tang, X., Wu, F., Wu, Q., et al. (2018). Comparison for younger and older adults: Stimulus temporal asynchrony modulates audiovisual integration. *Int. J. Psychophysiol.* 124, 1–11. doi: 10.1016/j.ijpsycho.2017.12.004
- Rypma, B., Eldreth, D. A., and Rebbeci, D. (2007). Age-related differences in activation-performance relations in delayed-response tasks: A multiple component analysis. *Cortex* 43, 65–76. doi: 10.1016/S0010-9452(08)70446-5
- Schneider, B. A., and Pichorafuller, M. K. (2000). “Implications of perceptual deterioration for cognitive aging research,” in *Handbook of aging and cognition*, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 155–219.
- Sekiya, K., Soshi, T., and Sakamoto, S. (2014). Enhanced audiovisual integration with aging in speech perception: A heightened McGurk effect in older adults. *Front. Psychol.* 5:323. doi: 10.3389/fpsyg.2014.00323
- Senkowski, D., Saint-Amour, D., Hofle, M., and Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage* 56, 2200–2208. doi: 10.1016/j.neuroimage.2011.03.075
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* 32, 8443–8453. doi: 10.1523/JNEUROSCI.5069-11.2012
- Sommers, M. S., Tye-Murray, N., and Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear Hear.* 26, 263–275. doi: 10.1097/00003446-200506000-00003
- Stein, B. E. (2012). *The new handbook of multisensory processing*. Cambridge, MA: MIT Press.
- Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266. doi: 10.1038/nrn2331
- Stevenson, R. A., and James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage* 44, 1210–1223. doi: 10.1016/j.neuroimage.2008.09.034
- Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Exp. Brain Res.* 227, 249–261. doi: 10.1007/s00221-013-3507-3
- Stevenson, R. A., Nelms, C., Baum, S. H., Zurkovsky, L., Barense, M., Newhouse, P. A., et al. (2015). Deficits in audiovisual speech perception in normal aging emerge at the level of whole-word recognition. *Neurobiol. Aging* 36, 283–291. doi: 10.1016/j.neurobiolaging.2014.08.003
- Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., and Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *J. Cogn. Neurosci.* 17, 1396–1409. doi: 10.1162/0898929054985383
- Teder-Sälejärvi, W., McDonald, J., Di Russo, F., and Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cogn. Brain Res.* 14, 106–114. doi: 10.1016/S0926-6410(02)00065-4
- Tremblay, P., Basirat, A., Pinto, S., and Sato, M. (2021). Visual prediction cues can facilitate behavioural and neural speech processing in young and older adults. *Neuropsychologia* 159:107949. doi: 10.1016/j.neuropsychologia.2021.107949
- Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., and Hale, S. (2010). Aging, audiovisual integration, and the principle of inverse effectiveness. *Ear Hear.* 31, 636–644. doi: 10.1097/AUD.0b013e3181dd17ff
- Tye-Murray, N., Spehar, B., Myerson, J., Sommers, M. S., and Hale, S. (2011). Cross-modal enhancement of speech detection in young and older

adults: Does signal content matter? *Ear Hear.* 32, 650–655. doi: 10.1097/AUD.0b013e31821a4578

Winneke, A. H., and 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. doi: 10.1037/a0021683

Wong, P. C., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., and Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* 47, 693–703. doi: 10.1016/j.neuropsychologia.2008.11.032

Wyk, B. C. V., Ramsay, G. J., Hudac, C. M., Jones, W., Lin, D., Klin, A., et al. (2010). Cortical integration of audio-visual speech and

non-speech stimuli. *Brain Cogn.* 74, 97–106. doi: 10.1016/j.bandc.2010.07.002

Xi, Y., Li, Q., Gao, N., Li, G., Lin, W., and Wu, J. (2020). Co-stimulation-removed audiovisual semantic integration and modulation of attention: An event-related potential study. *Int. J. Psychophysiol.* 151, 7–17. doi: 10.1016/j.ijpsycho.2020.02.009

Yang, W., Li, Z., Guo, A., Li, S., Yang, X., and Ren, Y. (2021). Effects of stimulus intensity on audiovisual integration in aging across the temporal dynamics of processing. *Int. J. Psychophysiol.* 162, 95–103. doi: 10.1016/j.ijpsycho.2021.01.017

Yang, W., Yang, J., Gao, Y., Tang, X., Ren, Y., Takahashi, S., et al. (2015). Effects of sound frequency on audiovisual integration: An event-related potential study. *PLoS One* 10:e0138296. doi: 10.1371/journal.pone.0138296