HAL open science

Visual prediction cues can facilitate behavioural and neural speech processing in young and older adults

Pascale Tremblay, Anahita Basirat, Pinto Serge, Marc Sato

To cite this version:

Pascale Tremblay, Anahita Basirat, Pinto Serge, Marc Sato. Visual prediction cues can facilitate behavioural and neural speech processing in young and older adults. Neuropsychologia, 2021, 159, pp.107949. 10.1016/j.neuropsychologia.2021.107949 . hal-03371896

HAL Id: hal-03371896

https://hal.science/hal-03371896v1

Submitted on 11 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Highlights

1. This study examined age differences in the neural processing of audiovisual speech.
2. Lip reading capacity was reduced in older adults.
3. Visual predictive cues facilitated speech recognition in younger and older adults.
4. Audiovisual facilitation on P2 amplitude was lower in older adults.
5. Prediction facilitation on N2 latency was lower in older adults.

REVISED Manuscript (text UNmarked) Click here to view linked References

Visual Prediction cues can facilitate behavioural and neural speech processing in young and older adults

Pascale Tremblay[1,2], Anahita Basirat[3], Serge Pinto[4], Marc Sato[4]

1 Département de Réadaptation, Faculté de Médecine, Université Laval, Quebec City, Canada

2 Cervo Brain Research Centre, Quebec City, Canada

3 Univ. Lille, CNRS, UMR 9193 – SCALab – Sciences Cognitives et Sciences Affectives, Lille, France

4 France Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

Corresponding author:

Pascale Tremblay, Ph. D.

Département de réadaptation, Université Laval

1050 avenue de la Médecine, Québec (QC)

CANADA G1V 0A6

Centre de recherche CERVO

Email: Pascale.Tremblay@fmed.ulaval.ca

Phone: +001 418 663–5000 ext. 8717

Fax: +001 418 656–5476

Abstract

The ability to process speech evolves over the course of the lifespan. Understanding speech at low acoustic intensity and in the presence of background noise becomes harder, and the ability for older adults to benefit from audiovisual speech also appears to decline. These difficulties can have important consequences on quality of life. Yet, a consensus on the cause of these difficulties is still lacking. The objective of this study was to examine the processing of speech in young and older adults under different modalities (i.e. auditory [A], visual [V], audiovisual [AV]) and in the presence of different visual prediction cues (i.e., no predictive cue (control), temporal predictive cue, phonetic predictive cue, and combined temporal and phonetic predictive cues.). We focused on recognition accuracy and four auditory evoked potential (AEP) components: P1-N1-P2 and N2. Thirty-four right-handed French-speaking adults were recruited, including 17 younger adults (28 ± 2 years; 20–42 years) and 17 older adults (67 ± 3.77 years; 60–73 years). Participants completed a forced-choice speech identification task. The main findings of the study are: (1) The faciliatory effect of visual information was reduced, but present, in older compared to younger adults, (2) visual predictive cues facilitated speech recognition in younger and older adults alike, (3) age differences in AEPs were localized to later components (P2 and N2), suggesting that aging predominantly affects higher-order cortical processes related to speech processing rather than lower-level *auditory* processes. (4) Specifically, AV facilitation on P2 amplitude was lower in older adults, there was a reduced effect of the temporal predictive cue on N2 amplitude for older compared to younger adults, and P2 and N2 latencies were longer for older adults, and finally (5) behavioural performance was associated with P2 amplitude in older adults. Our results indicate that aging affects speech processing at multiple levels, including audiovisual integration (P2) and auditory attentional processes (N2). These findings have important implications for understanding barriers to communication in older ages, as well as for the development of compensation strategies for those with speech processing difficulties.

Keywords: Event-related potentials, Speech perception, Aging, Audiovisual enhancement, Predictive coding, Auditory processing

1. Introduction

One of the most common complaints of older adults is difficulty in understanding speech at low acoustic intensity and in the presence of background noise (CHABA, 1988; Pichora-Fuller, 1997). These difficulties are associated with increased self-consciousness about one's communication competences, reduced self-confidence and disengagement from social activities, which can lead to isolation (Tobias, 1977). Yet, a unified theory of the etiology of these difficulties is still lacking and so are effective treatment options and prevention strategies (El-Assal & El-Gharib, 2019; Humes, Wilson, Barlow, & Garner, 2002).

Importantly, speech processing difficulty can arise during face-to-face conversations even when articulatory (visual) information from the speaker complements the acoustic speech signal. The influence of visual information on speech perception is well established. The most well-known example of the influence of visual cues on perception of speech sounds is the McGurk effect, in which, when presented with incongruent auditory and visual speech stimuli, most people report hearing a fusion of the two syllables presented (auditory and visual) (McGurk & MacDonald, 1976). In young adults, several studies have shown that audiovisual speech (AV) can be associated with a speech recognition gain compared to auditory (A) alone (e.g. Erber, 1969; Sumby & Pollack, 1954), a phenomenon that is often referred to as the *audiovisual speech advantage* or, more generally, *multimodal enhancement*. Importantly, in accordance with the principle of inverse effectiveness (i.e. multisensory enhancement is greatest when unimodal stimuli are least effective), AV speech recognition is more resistant to noise than A-only (Erber, 1969). These effects indicate that adding articulatory (visual) information helps reduce uncertainty, which facilitates auditory speech recognition, especially in noisy auditory conditions. A decline in the ability to extract and to use articulatory (visual) information to disambiguate speech could be contributing to the speech processing difficulties experienced by older adults. Several studies have shown that, compared to younger adults, middle-age and older adults exhibit *lower* performance enhancement for AV speech compared to A

speech (Tye-Murray, Sommers, Spehar, Myerson, & Hale, 2010; Yang & Ren, 2018). In contrast, other studies have found comparable (Avivi-Reich, Puka, & Schneider, 2018; Cienkowski & Carney, 2002; Sommers, Tye-Murray, & Spehar, 2005; Winneke & Phillips, 2011) or *enhanced* AV integration and binding abilities for older compared to younger adults (Ganesh, Berthommier, & Schwartz, 2017; Laurienti, Burdette, Maldjian, & Wallace, 2006; Sekiyama, Soshi, & Sakamoto, 2014). In an attempt to resolve these inconsistent findings, it has been suggested that there may be no overall decline in the capacity to benefit from audiovisual information in aging, but instead, that the conditions needed to benefit from it may change over the course of the lifespan, with older adults benefiting less when the (acoustic) signal-to-noise ratio is low (Jansen, Keebler, & Chaparro, 2018; Stevenson, Nelms, Baum, Zurkovsky, Barense, Newhouse, & Wallace, 2015) and when the visual information is degraded (Gordon & Allen, 2009). In sum, the ability for older adults to benefit from an audiovisual speech signal appears to evolve with age and to depend upon the clarity of the acoustic and/or visual sources.

While there have been several behavioural studies focusing on AV speech perception in aging, the neural mechanisms underlying AV speech perception performance and multimodal enhancement in aging, and, more generally, speech processing capacities, are still unclear. Neurophysiological studies have shown that, in young adults, prior knowledge of a speech input facilitate neural processing of deteriorated or missing speech (Cervantes Constantino & Simon, 2018), and that adding visual articulatory information to auditory speech modulates activity in primary auditory and associative regions of the temporal cortex including the superior temporal sulcus (Arnal, Morillon, Kell, & Giraud, 2009), affecting the P1-N1-P2 complex (e.g. Stekelenburg & Vroomen, 2007; van Wassenhove, Grant, & Poeppel, 2005). The P1-N1-P2 complex is a series of co-occurring auditory evoked potentials (AEPs), which indicate that a sound has reached the auditory cortex and that initial cortical acoustic-phonetic processing has begun. As such, the P1/N1/P2 complex indexes the capacity for speech sound processing and discrimination. It consists of positive and negative voltage deflections peaking around 50 ms (P1), 100 ms (N1), and 200 ms (P2) after stimulus onset. AEPs components such as the P1-N1-

P2 complex are typically identified by their polarity (positive or negative), latency (peak occurrence after stimulus onset, measured in milliseconds) and amplitude (in µV). Specifically, it is well established that, compared to unimodal auditory perception, adding visual articulatory information to auditory speech leads to an attenuated amplitude and earlier latency of the N1/P2 complex (e.g., Besle, Fort, Delpuech, & Giard, 2004; Klucharev, Mottonen, & Sams, 2003; Treille, Cordeboeuf, Vilain, & Sato, 2014a; Treille, Vilain, Kandel, & Sato, 2017; Treille, Vilain, & Sato, 2014b; Treille, Vilain, Schwartz, Hueber, & Sato, 2018; van Wassenhove et al., 2005). Hence, this complex is usually seen as a reliable marker of AV integration.

Winneke & Phillips were the firsts to investigate the behavioural benefit of and the neural processes (P1/N1/P2 complex) associated with AV perception of spoken words presented in babbling noise in young and older adults (Winneke & Phillips, 2011). Behaviourally, their results showed no difference in the ability for younger and older adults to benefit from audiovisual information. Yet, EEG results revealed an age-related difference in P1 amplitude, with an increased reduction from A to AV speech (i.e. multisensory integration) for older compared to younger adults. There was also an age difference in N1 latency with a more pronounced latency shift (i.e. facilitation) for older compared to younger adults when comparing A to AV speech. The authors concluded that older adults, compared with younger adults, "*are not better lip readers per se but rather are better "lip/speech integrators*." (p. 436), perhaps as a compensation for more laborious auditory processing in older adults. Though these results critically await replication, abnormal cortical response patterns have been shown during speech and speech in noise tasks in older adults in several studies (e.g. Brodbeck, Presacco, Anderson, & Simon, 2018; Presacco, Simon, & Anderson, 2016a, 2016b).

During audiovisual speech processing, perceptual experience is aided by prior crossmodal associations and integration mechanisms, which can help reduce sensory uncertainty. Predictive coding theories postulates that the brain actively predicts upcoming sensory input rather than simply registering it. According to this view, bottom-up sensory information is compared with top-down

predictions from higher levels to estimate prediction errors (Clark, 2013; Friston, 2005, 2010; Rao & Ballard, 1999). For speech, audiovisual integration operates through temporal expectations and speech-specific predictions. From a Bayesian perspective, perceptual experience derives from the processing and integration of multisensory (AV) inputs based on their predictability and joint probability (Massaro, 1998; van Wassenhove, 2013). Therefore, while visual (articulatory) information can facilitate speech recognition and modulate the N1/P2 complex, other types of predictive cues could also have a facilitatory impact on speech perception and AEPs and could potentially facilitate speech processing in older adults. Notably, we showed in a previous EEG study that visual temporal and phonetic predictions can attenuate the amplitude of the N1/P2 components during *auditory* speech perception, and that P2 (amplitude and latency), but not N1, is modulated (facilitated) by phonetic prediction during *audiovisual* speech perception (Pinto, Tremblay, Basirat, & Sato, 2019), demonstrating an enhanced sensitivity for P2 to phonetic cues, consistent with previous work (e.g. Baart, Stekelenburg, & Vroomen, 2014). The effect of age on the ability to process different kinds of visual predictive cues, however, remains unclear.

The first objective of the present study was to investigate the effect of age and audiovisual integration on speech perception performance (Objective 1). Specifically, we wanted to determine whether audiovisual integration (objective 1a: adding natural speech movements through video) and/or other visual prediction cues (objective 1b: temporal onset and/or visuo-orthographic (e.g. /pa/) cues) facilitate auditory speech syllable perception similarly for younger and older adults. The second objective of the study was to compare the neurophysiological response to speech in younger and older adults focusing on AEPs (P1-N1-P2 and N2). First, we compared audiovisual speech integration in younger and older adults using an additive model (objective 2a; AV vs. A+V). This analysis focused strictly on visual articulatory movements in order to confirm and extend previous EEG findings (Brodbeck et al., 2018; Presacco et al., 2016a, 2016b; Winneke & Phillips, 2011). Next, we tested whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural

processing of speech similarly for younger and older adults (Objective 2b; AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). Finally, the third objective of the study was to examine the relationship between speech perception performance and AEPs to shed new lights on brain aging and its impact on human behaviour.

To achieve these goals, the predictability of auditory syllables was manipulated experimentally by adding unnatural visual information indicative of their temporal unfolding (*when*) and phonetic content (*what*) (Pinto et al., 2019). The presentation modality (auditory (A), visual (V) and audiovisual (AV)) was manipulated to examine whether the visual predictive cues would be processed similarly when the speech signal was audio and audiovisual.

Based on prior studies, we hypothesized that older adults would show similar or enhanced audiovisual and predictive gain. Specifically, we expected a facilitation of AEPs (shorter latency and/or reduced amplitude) during AV speech compared to A speech, as well as during the processing of temporal and phonetic predictive cues. We expected those patterns to either be identical for the young and older adults, or to be enhanced in older adults, based on their lifetime of experience processing speech. Indeed, predictive coding for speech could be heightened with age and serve as a compensation strategy to overcome declining unisensory processing. Alternatively, predictive coding could, like other higher-order cognitive functions, decline with age and this decline could be associated with age-related speech processing difficulties. Given the well-established hearing loss and cognitive decline that occurs in aging, and the known association between these two factors (e.g.Humes, Busey, Craig, & Kewley-Port, 2013; Humes, Kidd, & Lentz, 2013; Lin, 2011; Lin, Ferrucci, Metter, An, Zonderman, & Resnick, 2011; Lin, Yaffe, Xia, Xue, Harris, Purchase-Helzner, Satterfield, Ayonayon, Ferrucci, Simonsick, & Health, 2013; Wayne & Johnsrude, 2015), we expected that the auditory N2, which indexes late cognitive processes, such as executive functions and attention (Fritz, Elhilali, David, & Shamma, 2007), would show either an increased amplitude, reflecting compensatory activity, or perhaps a decline, reflecting a disruption in processing.

2. Material and Methods

2.1. Participants

Thirty-four right-handed French-speaking adults participated in the study after giving informed consent. All participants reported normal or corrected-to-normal vision and were allowed to wear their glasses or lenses during the experiment. Participants reported no history of hearing, speaking, language, neurological and/or neuropsychological disorders. The cognitive functioning of all participants was evaluated using the Montreal Cognitive Assessment scale (MoCA) (Nasreddine, Chertkow, Phillips, Bergman, & Whitehead, 2003; Nasreddine, Phillips, Bedirian, Charbonneau, Whitehead, Collin, Cummings, & Chertkow, 2005). None of the participant had mild cognitive decline using the criteria by Larouche et al. (Larouche, Tremblay, Potvin, Laforest, Bergeron, Laforce, Monetta, Boucher, Tremblay, Belleville, Lorrain, Gagnon, Gosselin, Castellano, Cunnane, Macoir, & Hudon, 2016).

Participants were divided into a younger and an older group. The younger group included 17 adults (14 females), with a mean age of 28 ± 2 years (20–42 years) and an average of 15.6 ± 2.61 years of education (range: 11–20 years). The data from the younger group were published in Pinto et al. (2019). The older group included 17 adults (8 females), with a mean age of 67 ± 3.77 years (60–73 years) and an average of 15.25 ± 3.55 years of education (range: 9–20 years). Participants' characteristics are detailed in Table 1. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (World Medical, 2013) and participants were compensated for the time spent in the study.

Table 1. Participants' characteristics

	Young (N = 17; 13 W)		**Older** (N = 17; 8 W)	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**
Age	28.20 (6.87)	20–42	67 (3.88)	60–73
MoCA (/30)	29.07 (1.38)	25–30	26.38 (1.66)	23–29

Handedness	84.8 (.15)	55–100	95.9 (11.5)	58–100
Right ear PTA	4.77 (5.22)	-5 – 15	13.699 (6.41)	4.16 – 26.6
Left ear PTA	5.55 (4.3)	-3.33 – 13.33	14.79 (6.26)	3.33 – 25
Education	15.6 (2.61)	11 – 20	15.25 (3.55)	9 – 20

Table 1 Participants' characteristics. SD = standard deviation. W = women. PTA = Pure tone average.

2.2. Hearing Assessment

To ensure that participants had normal hearing, pure tone audiometry was performed using a clinical audiometer (Resonance R17A, MRS, Italy) for each ear separately, at the following frequencies: .5, 1, 2 kHz. For each participant, a standard pure tone average (PTA: average of thresholds at .5, 1 and 2 kHz) was computed for the left and right ear. The result of the hearing assessment is provided in Table 1. Figure 1 illustrates the average threshold at each frequency tested separately for each Age Group.

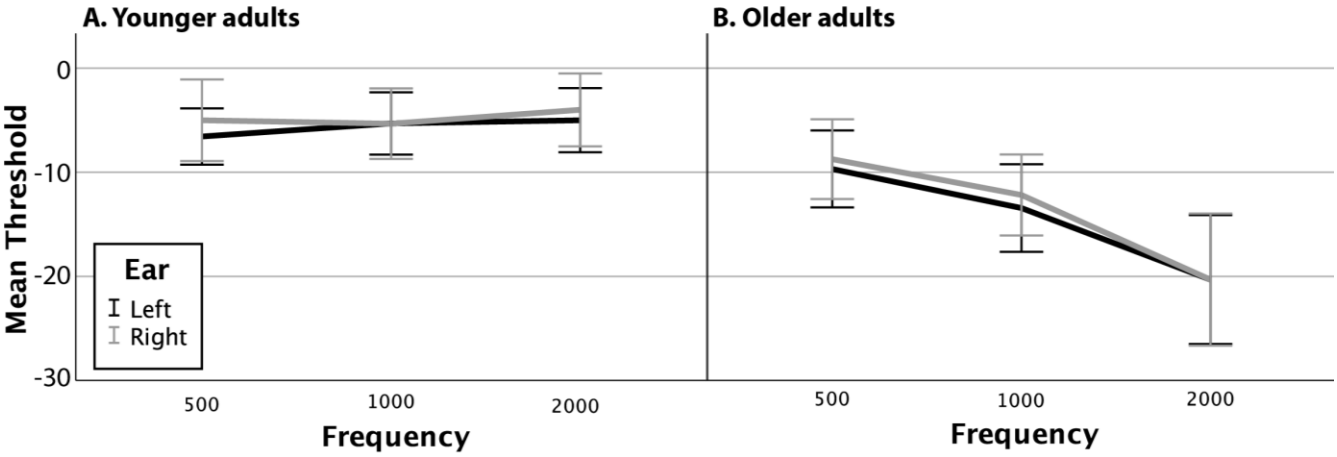

Figure 1 . Hearing thresholds. The line charts present an overview of participants' hearing thresholds, separately for each ear, and the younger (A) and the older adults. Each line in the figure represents the average threshold for a group of participants. Error bars represent the confidence interval of the means.

A linear mixed model (LMM) analysis was conducted using the r package version 4.0.3 for Mac (Team, 2019), with Age Group (younger, older) as the between-subject factor, Ear (Left, Right), and Frequency (500, 1000, 2000 Hz) as the within-subject factors. The random intercepts for participants were included in the model. The LMM results show main effects of Age Group ($p \leq .001$)

and Frequency ($p \leq .001$), as well as a 2-way interaction between Age Group and Frequency ($p \leq .001$). Tukey-corrected post hoc contrasts indicated that thresholds were higher in the older adults at 1000 Hz (p = .0022, d = -1.230) and 2000 Hz ($p \leq .001$, d = -2.588). The descriptive statistics and the detailed results of the LMM analyses as well as the pairwise contrasts are provided in Supplementary Material 1. Because of these (expected) Group differences, Hearing (average PTA) was included in all analyses as a covariate.

2.3. Speech Perception Assessment

All procedures were carried out in a sound-attenuated room at the LPL Lab in Aix-en-Provence. Participants sat in front of a computer monitor at a distance of approximately 50 cm. The acoustic stimuli were presented through loudspeakers located on each side of a computer monitor. Stimuli were presented using Presentation software (Neurobehavioral Systems, Albany, USA), which was also used to record participants' behavioural responses. During the task, participants were asked to complete a forced-choice speech identification task. On each trial, they identified one syllable by pressing one of three keys on a keyboard with their left hand. No feedback was provided. The stimuli were the syllables /pa/, /ta/, and/ka/. All stimuli were presented in quiet. The response key designation was counterbalanced. To dissociate sensory/perceptual from motor responses on the EEG recordings, each stimulus was followed, after 600 ms, by a brief auditory tone and question mark (?), which served as "GO" cues. The inter-trial interval was 3s. The experiment lasted approximately 45 min and was divided in four sessions of ~11 minutes with short breaks in between sessions.

The syllables were presented in three modalities (auditory [A], visual [V], audiovisual [AV]) and under 4 different cue conditions (control [no cue], *when*, *what*, *what-when*), which resulted in 12 experimental conditions: control (A, V, AV), when (A_{when}, V_{when}, AV_{when}), what (A_{what}, V_{what}, AV_{what}) and what-when ($A_{what\text{-}when}$, $V_{what\text{-}when}$, $AV_{what\text{-}when}$). The experiment consisted of 864 trials presented in a pseudo-randomized order, including 72 trials in each of the 12 experimental conditions.

A full description of the stimuli can be found in Pinto et al. (2019). Briefly, the auditory stimuli consisted of an acoustic syllable dubbed on a static image of a neutral mid-open mouth position of a speaker. The visual stimuli consisted of the visual speech movements displayed without any sound. The audiovisual stimuli started with an initial neutral mid-open mouth position followed by visual speech movements (30 frames, 1200 ms) before the acoustic consonantal burst and the syllable (5 frames, 200 ms). For all stimuli, the auditory signal intensity was normalized using a common maximal amplitude criterion. Importantly, the audiovisual stimuli were first created. The visual and auditory-only stimuli were created from the audiovisual stimuli, by removing the acoustic signal (visual stimuli) or by replacing the visual speech movements by a static face (auditory stimuli). Participants were informed that visual cues were always coherent with the auditory syllable.

In all conditions, "##" orthographic symbols and a static timeline were visually presented during the first 15 frames (0 to 600 ms). In the *when* conditions, a moving visual timeline indicative of the temporal consonantal onset of the acoustic syllable replaced the static timeline during the subsequent 15 ± 2 frames (600 ± 80 to 1200 ms). In the *what* conditions, a visuo-orthographic cue indicative of the syllable (/pa/, /ta/or/ka/) replaced the "##" symbols during the subsequent 15 ± 2 frames (600 ± 80 to 1200 ms). In the *what-when* conditions, both the visual timeline and visuo-orthographic cues were presented. Finally, in the control conditions, speech signals were presented only with the "##" orthographic symbols and the static timeline during 15 ± 2 frames (600 ± 80 to 1200 ms). The Prediction conditions are illustrated in Figure 2.

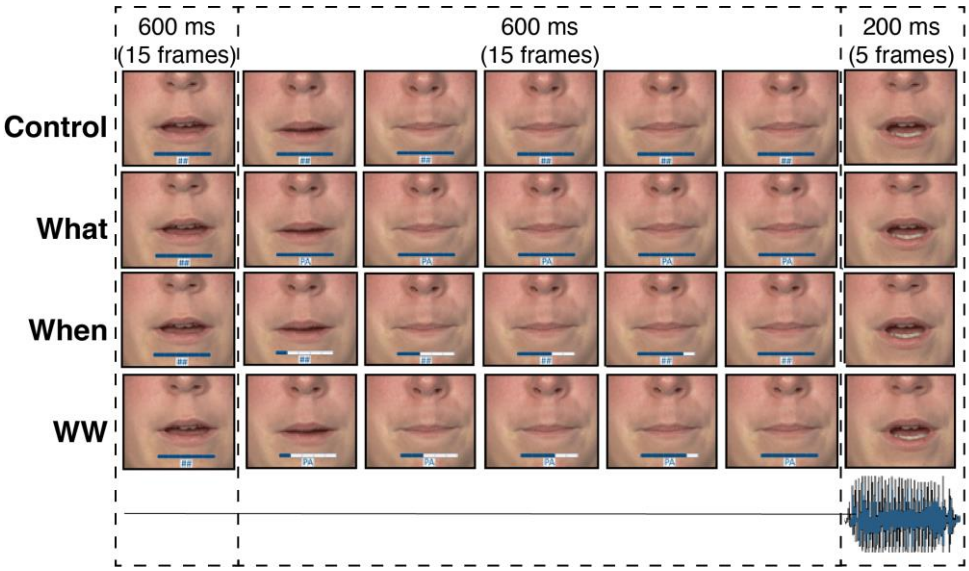

Figure 2 Prediction Cue conditions. WW = What+When condition. The sound wave at the bottom right represents the auditory speech signal.

2.4. EEG Recordings and Pre-Processing

EEG data were recorded continuously from 9 scalp electrodes (Electro-Cap International, INC, according to the international 10–20 system) using the Biosemi Active Two AD-box EEG system operating at a 512 Hz sampling rate. Because N1/P2 AEPs have maximal response over central sites on the scalp (Naatanen & Picton, 1987; Scherg & Von Cramon, 1986), EEG was only collected from fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) (Pinto et al., 2019). This minimal recording procedure has been used in several EEG studies on audiovisual speech integration and auditory evoked responses that showed classical audiovisual speech interactions on N1/P2 AEPs (Stekelenburg & Vroomen, 2007; Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; Vroomen & Stekelenburg, 2010).

Two additional electrodes were used as ground electrodes (Common Mode Sense [CMS] active and Driven Right Leg [DRL] passive electrodes). In addition, one external reference electrode was set at the top of the nose. Horizontal (HEOG) and vertical (VEOG) eye movements were recorded using electrodes positioned at the outer canthus of each eye, as well as above and below the right eye. Before the experiment, the impedance of all electrodes was adjusted to get low offset voltages and stable DC.

EEG data were processed using EEGLAB (Delorme & Makeig, 2004) running on Matlab (Mathworks, Natick, USA). EEG data were first re-referenced offline to the nose channel and band-pass filtered using a 2-way least square FIR filtering (3–30 Hz) to reduce slow drifts and high frequency noise (see Pinto et al. (2019). Next, the data was segmented into 500 ms epochs including a 100 ms prestimulus baseline (from −100 to −0 ms relative to the acoustic syllable onset). Epochs with an amplitude change exceeding ±100 uV at any channel (including HEOG and VEOG channels) were rejected (mean (±SD): 2% (±2%) trials). Responses from/pa/, /ta/and/ka/syllables were first averaged together in order to provide 72 trials per condition. For each participant and each condition (i.e., A, V, AV, A_{when}, V_{when}, AV_{when}, A_{what}, V_{what}, AV_{what}, $A_{what\text{-}when}$, $V_{what\text{-}when}$, $AV_{what\text{-}when}$), data were then averaged over the nine electrodes. Finally, the maximal amplitude and peak latency of the N1-P2-N2 complex were determined using a fixed temporal window for each component (N1: 70–150 ms; P2: 150–250 ms; N2: 180–325 ms). For P1, instead of extracting the maximal amplitude and peak latency, we computed the area under the curve for amplitude (10–100 ms), and the 50% area latency. The V condition was included in the study design to examine behaviour in this condition, as well as to compare the neural response to AV with the A+V signal (Objective 2a, see section 2.5.3).

2.5. Statistical analyses

2.5.1. Behavioural data

Behavioural data analyses were conducted to address the first objective of the project, to investigate the effect of age and audiovisual integration on speech perception performance. To achieve this goal, three complementary behavioural indexes of performance were computed to characterize speech perception performance: a classical percentage of correct responses, and two difference scores: a visual (VE) effect scores and predictive cue effect (PE) scores.

First, a classical measure of accuracy was computed and used to compare perception accuracy in young and older adults during speech perception. The percentage of correct responses was

determined for each participant and condition (note that RTs were not recorded due to the 600 ms delay between the stimuli and the "GO" cues for the manual responses). A linear mixed model (LMM) analysis was conducted using r version 4.0.3 for Mac (Team, 2019), using the nlme package, with Age Group (younger, older) as the between-subject factor, Modality (A, V, AV) and Predictive cue (Control, What, When, WW) as the within-subject factors, and hearing (PTA) as between-subject continuous fixed factor. The random intercepts for participants were also included in the model. Model selection included testing models, using likelihood ratio tests, with and without PTA as well as with different random effect structures (with or without slopes for either Modality or Predictive cue condition) and covariance structure for the residuals (homogeneous vs. heterogeneous variance across levels of Modality or Predictive Cue). The model with the best fit, assessed using likelihood ratio tests, was kept. When fit was similar across one model or more, the simplest model was kept. The same procedure was used for all LMM analyses (accuracy, VE, PE and all EEG analyses) and will not be repeated hereafter. For each analysis (accuracy, VE, PE and all EEG analyses), the final model is provided in the corresponding supplementary materials.

Next, visual effect (VE) scores were calculated to determine whether adding natural speech movements through video to auditory speech facilitates auditory speech perception through audiovisual integration similarly in young and older adults (objective 1a). These scores were derived from the percentage correct scores for each participant and each of the cue condition (Control, What, When, WW) with the following formula: VE = (AV—A) / (100-A). VE measures have been used in several AV studies as they circumvent the bias inherent to calculating the difference between AV and A, in which higher values of A necessarily lead to lower values of VE (Sommers et al., 2005). To examine Age Group differences in VE, a LMM analysis was conducted on the VE scores, with Age Group (younger, older) as the between-subject factor, Prediction Cue condition (Control, What, When, WW) as the within-subject factor, and hearing (PTA) as a between-subject continuous fixed factor. The random intercepts for participants were also included in the model. The model selection procedure was

the same as described for accuracy. The VE value for 8 participants (3 young and 5 elderly) that had perfect score in the auditory modality could not be included in the analysis because the equation resulted in an error when attempting to divide by zero. Two outliers were removed from the final analyses (one young and one older adult).

Finally, predictive cue effect (PE) scores were calculated to determine whether adding other temporal and/or phonetic visual cues to the speech signal would facilitate speech perception similarly in young and older adults (objective 1b). The PE scores were inspired by the VE scores and used to circumvent the bias inherent to calculating the difference between the score in Prediction and Control conditions, in which higher values of Prediction necessarily lead to lower values of PE. A PE score was calculated for each prediction cue (What, When, WW) and each modality. The What effect (WhatE) score was calculated from the percentage correct scores for each participant. WhatE = (What—Control)/(100-Control). The formula for the When effect (WhenE) (WhenE = (When—Control)/(100-Control)) and the What+When effect (WWE) scores were identical: WWE = (WWE—Control)/(100-Control). To examine Age Group differences in prediction effect, a LMM analysis was conducted on the PE scores, with Age Group (younger, older) as the between-subject factor, Modality (A, V, AV), and Predictive cue condition (What, When, WW) as the within-subject factors, and hearing (PTA) as the between-subject continuous covariate. The random intercepts for participants were also included in the model. The model selection procedure was the same as described for accuracy.

2.5.2. EEG Analyzes: Audiovisual Integration (additive model)

The first set of analyses focused on objective 2a of the study, to compare audiovisual speech integration in younger and older adults using an additive model (AV vs. A+V). To address this aim, we used an additive model (e.g. Pilling, 2009; Treille et al., 2014a; Treille et al., 2014b; van Wassenhove et al., 2005) in which the bimodal audiovisual EEG signal was compared to the sum of auditory and visual unimodal EEG signals (AV ≠ A + V). The resulting signal may contain task-related neural

activity common to all modalities unrelated to crossmodal integration, such as movement-related cortical potentials, characterized by a slow negative deflection on fronto-central sites starting around 1000 ms prior to the manual response (e.g. Kornhuber & Deecke, 1965; Libet, Gleason, Wright, & Pearl, 1983), as well as slow anticipatory potentials that precedes perceptual decisions and discriminative responses, characterized by a slow positive deflection on fronto-central sites (e.g. Teder-Salejarvi, McDonald, Di Russo, & Hillyard, 2002). In order to minimize the temporally contingent influence of movement-related cortical potentials on auditory evoked potentials, an 800 ms delay was introduced between the acoustic consonantal burst of all syllables and the "GO" cues. Moreover, a 3 Hz high-pass filter was applied on the EEG data to minimize the contribution of movement-related as well as slow anticipatory potentials (Teder-Salejarvi et al., 2002).

A series of LMM analysis was conducted for each dependent measure: P1 peak amplitude and peak latency, N1 peak amplitude and peak latency, P2 peak amplitude and peak latency, and N2 peak amplitude and peak latency. In all analyses, the within-subject (repeated) fixed factors were Modality (AV, A+V), and the between-subject factor was Age Group (younger, older). Hearing (PTA) was included as a between-subject continuous fixed factor (covariate). Predictive cues were not included in this analysis because here we meant to assess audiovisual integration. The analyses were conducted on the signal from the control (Control) condition). The random intercepts for participants were also included in the model. The model selection procedure was the same as described for accuracy.

2.5.3. EEG Analyzes: Temporal and Phonetic Visual Predictive Cues

This second set of EEG analyses focused on objective 2b of the study, to determine whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural processing of speech similarly for younger and older adults (Objective 2b; AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). To address this aim, a series of linear mixed model (LMM) analysis was run to examine age differences on the neural processing of

Predictive Cues under different modalities (A, AV). One analysis was conducted for each dependent measure: P1 peak amplitude and peak latency, N1 peak amplitude and peak latency, P2 peak amplitude and peak latency, and N2 peak amplitude and peak latency. In all analyses, the within-subject (repeated) fixed factors were Modality (A, AV) and Predictive Cue (Control, What, When, WW), the between-subject factor was Age Group (younger, older), and hearing (PTA) was included as between-subject continuous fixed factors (covariate). The random intercepts for participants were included in the model. The model selection procedure was the same as described for accuracy.

2.5.4. Electrophysiology-Behaviour Relationship

The third objective of the study was to examine the relationship between performance in young and older adults and AEPs. To achieve this goal, a series of simple mediation analyses were conducted using the r packages NLME and Mediation, a package to conduct causal mediation analyses (Tingley D, 2013). All analyses were run with the same seed (2021). In all analyses, Age Group was used as the categorical predictor variable (X), one AEP component was included as continuous mediators (M) in separate analyses, and hearing (PTA) was used as a continuous between-subject covariate. The dependent variables (Y) were those that showed Age Group differences or interactions. A quasi-Bayesian approximation based on normal approximation (Imai, Keele, & Tingley, 2010) was used to compute percentile confidence intervals of the mediation effects with 1000 Monte Carlo draws. Robust (heteroskedasticity-consistent) standard errors were computed for the quasi-Bayesian simulations for the mediation effects. Because each analysis aimed at relating one set of Tukey-corrected behavioural findings to relevant AEP to get at underlying mechanism, these analyses were not corrected for multiple comparisons.

3. Results

3.1. Behavioural Results

First, we report the results for accuracy, then we report the results for the VE and PE scores. A summary of all behavioural results is provided in Table 2.

3.1.1. Accuracy

This analysis focuses on the comparison of the facilitation effect of audiovisual integration (objective 1a) and other visual predictive cues (objective 1b) on speech perception in younger and older adults.

Results of the LMM analyses (marginal fixed effects) revealed a main effect of Modality ($F_{(2,330)} = 194.503$, $p < .0001$). Tukey-corrected post hoc contrasts indicated that accuracy was higher in A compared to V ($\beta = 13.78$, SE = 0.797, $p < .0001$) and in AV compared to V ($\beta = 3.651$, SE = 0.206, $p = < .0001$). There was also an interaction between Age Group and Modality ($F_{(2,330)} = 6.215$, $p = .002$). Tukey-corrected post hoc contrasts indicated that accuracy in the V condition was lower for older compared to younger adults ($\beta = 8.350$, SE =2.24, $p = .001$) (Figure 3A). Finally, the LMM analysis also revealed an interaction between Prediction Cue and Modality ($F_{(6,330)} = 39.422$, $p < .0001$) (Figure 3B). Tukey-corrected post hoc contrasts indicated that, in the A condition, accuracy was higher in the WW condition compared to the Control condition ($\beta = -2.99$, SE = .996, $p = .015$). In the V condition, accuracy was higher in What compared to the Control condition ($\beta = -23.9$, SE = .996, $p < .0001$), in the When compared to the Control condition ($\beta = -3.60$, SE = .996, $p = .002$), and in WW compared to the Control condition ($\beta = 23.9$, SE = .996, $p < .0001$). Of all the cues, the When cue had the smallest impact on performance. Finally, in the AV condition, the cues did not affect accuracy. The descriptive statistics and the detailed results of the LMM analyses as well as the pairwise contrasts are provided in Supplementary Material 2, along with a figure showing accuracy in all experimental conditions separately.

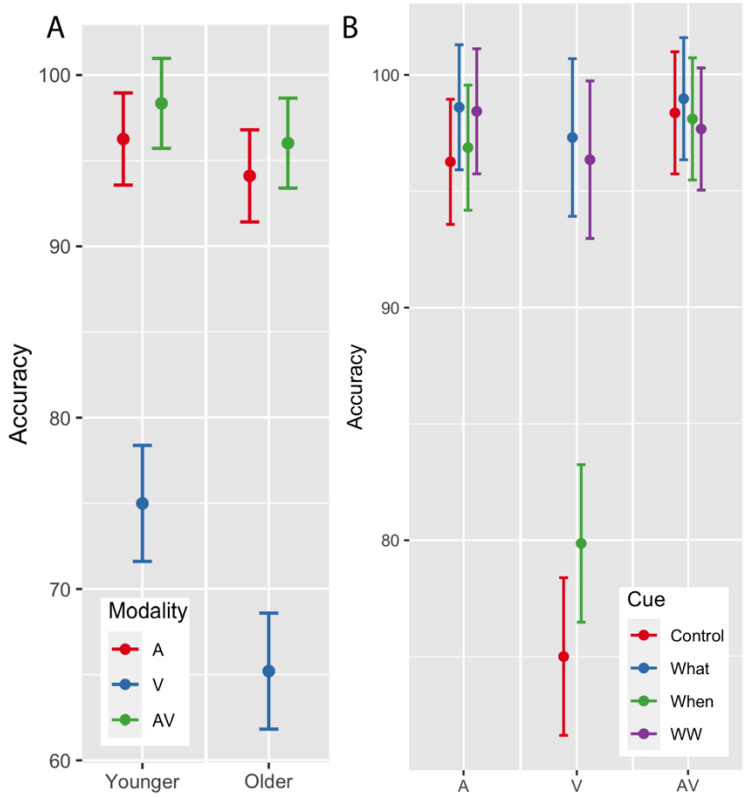

Figure 3 Accuracy results (marginal means). A. The plots display Accuracy as a function of Modality, separately for each group. B. The plots display Accuracy as a function of Modality, separately for each cue. Refer to the text for the list of the significant contrasts. The error bars represent the confidence intervals of the marginal means.

3.1.2. VE Scores

The analysis of the VE scores aimed to determine whether adding natural speech movements to auditory speech facilitates auditory speech perception in older adults through audiovisual integration (objective 1a). Results of the LMM analyses (marginal fixed effects) revealed a main effect of Age Group (Younger > Older) ($F_{(1,28)} = 8.469$, $p = .007$) (Figure 4A).

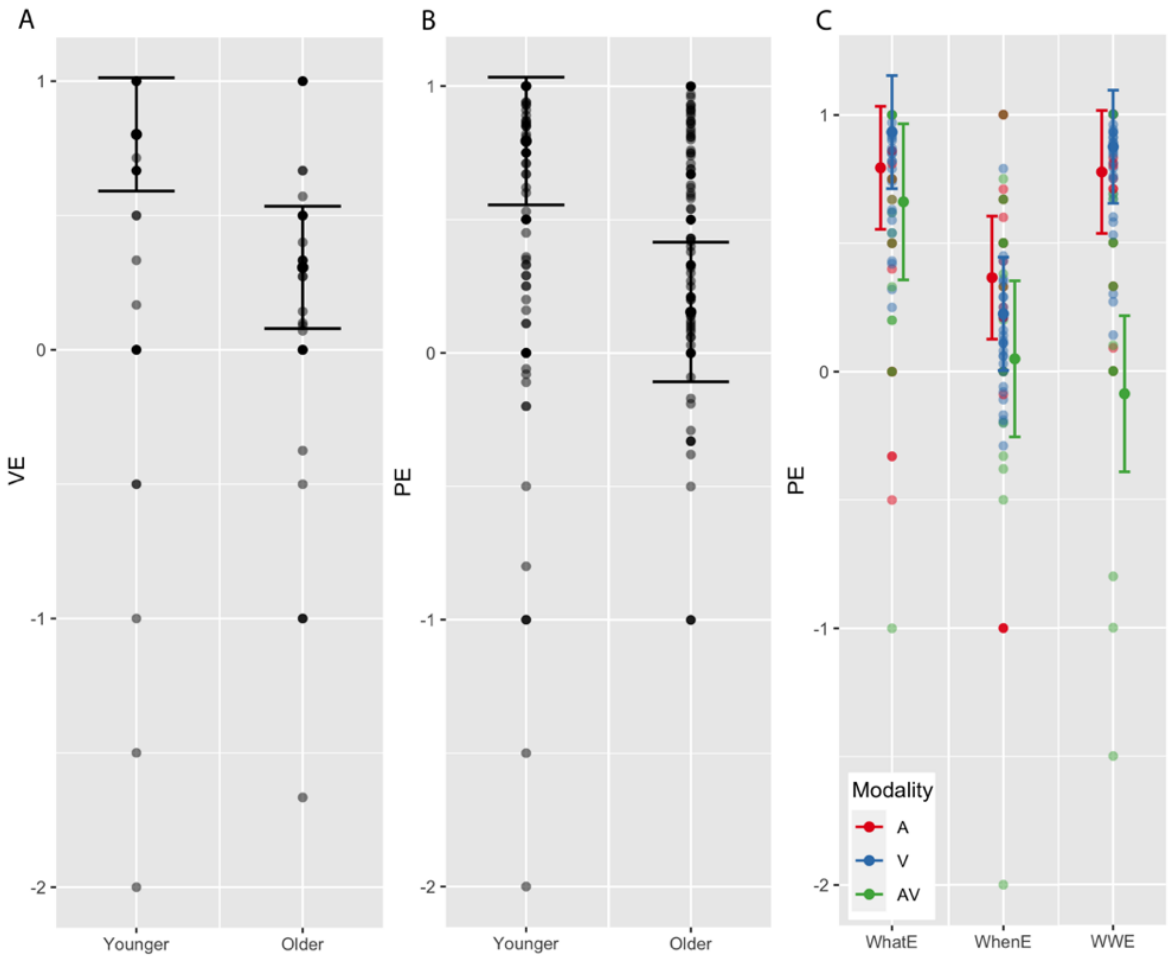

Figure 4 VE and PE results (marginal means). A. Age Group differences in VE. The groups are represented on the X-axis while VE is displayed in the y-axis. B. Age Group differences in PE. C. Cue by Modality interaction on PE. The Cue conditions are represented on the X-axis while PE is displayed in the y-axis. Modalities are colour coded. In all graphs, the error bars represent standard errors of the marginal means. Each dot is a participant.

3.1.3. PE Scores

The PE scored aimed to determine whether adding other visual Prediction Cues to the speech signal would facilitate speech perception similarly in young and older adults (objective 1b). The LMM results (marginal means) revealed a main effect of Age Group ($F_{(1,29)} = 11.822$, $p = .002$), with overall higher PE values in the younger adults (Figure 4B). The analysis also showed an interaction between Modality and Prediction Cue ($F_{(4,173)} = 3.646$, $p = .007$) (Figure 4C). This interaction revealed that PE scores did not differ as a function of Modality in the What and When conditions, but only in the WW condition. The condition with the most information (WW_{AV}) showed the lowest PE. Tukey-corrected

post hoc tests revealed that in the WW condition, the effect of adding a prediction cue on performance was stronger for V then AV (β = -0.647 SE = 0.122, p = .001) and for A compared to AV (β = 0.555, SE = 0.129, p = .001), while the effect was similar for the unisensory conditions A and V (β = -0.092, SE = 0.109, p = .995). The descriptive statistics and the detailed results of the LMM analyses and are provided in Supplementary Material 4.

Table 2. Type III Tests of Fixed Effects (marginal) for the behavioural analyses

Effects and interactions	DF		F	p
A. Accuracy				
Age Group	1	29	1.095	0.304
Prediction cue	3	330	1.348	0.2587
Modality	2	330	109.03	**<.0001**
Hearing	1	29	0.266	0.6101
Age Group * Prediction cue	3	330	0.373	0.7727
Age Group * Modality	2	330	6.215	**0.0022**
Prediction cue * Modality	6	330	39.422	**<.0001**
Age Group * Prediction cue * Modality	6	330	1.559	0.1583
B. VE				
Age Group	1	28	8.469	**0.007**
Prediction cue	3	47	1.872	0.147
Hearing	1	28	1.145	0.294
Age Group * Prediction cue	3	47	0.871	0.463
C. PE				
Age Group	1	29	11.822	**0.002**
Prediction cue	2	173	4.741	**0.010**
Modality	2	173	1.257	0.287
Hearing	1	29	0.540	0.468
Age Group * Prediction cue	2	173	2.366	0.097
Age Group * Modality	2	173	1.699	0.186
Prediction cue * Modality	4	173	3.646	**0.007**
Age Group * Prediction cue * Modality	4	173	1.042	0.387

3.2. EEG Analyzes: Audiovisual Integration (additive model)

The first set of EEG analyses addresses objective 2a, to compare audiovisual speech integration in younger and older adults using an additive model (AV vs. A+V). To achieve this goal, we compared the AV signal to the sum of the unisensory signals (A+V) in young and older adults. Figure 5 shows the average neurophysiological response for each Age Group and each modality (AV, A+V). The descriptive and inferential statistics are presented in Supplementary Materials 5-6 (P1), 7-8 (N1), 9-10 (P2) and 11-12 (N2). In line with previous EEG studies (e.g., Besle et al., 2004; Klucharev et al., 2003;

Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; van Wassenhove et al., 2005), the difference between AV and A+V was significant on the N1/P2/N2 complex in both younger (Figure 5A) and older adults (Figure 5B). As can be seen in the figure, overall, the amplitude of N1 and N2 was more negative for A+V than for AV, and the amplitude of P2 was more positive for A+V than AV. In terms of latency, the A+V signal had a longer N1 and P2 peak latency. As detailed in the following paragraphs and illustrated in Table 3, Age Group differences were found on P2 and N2 (latency).

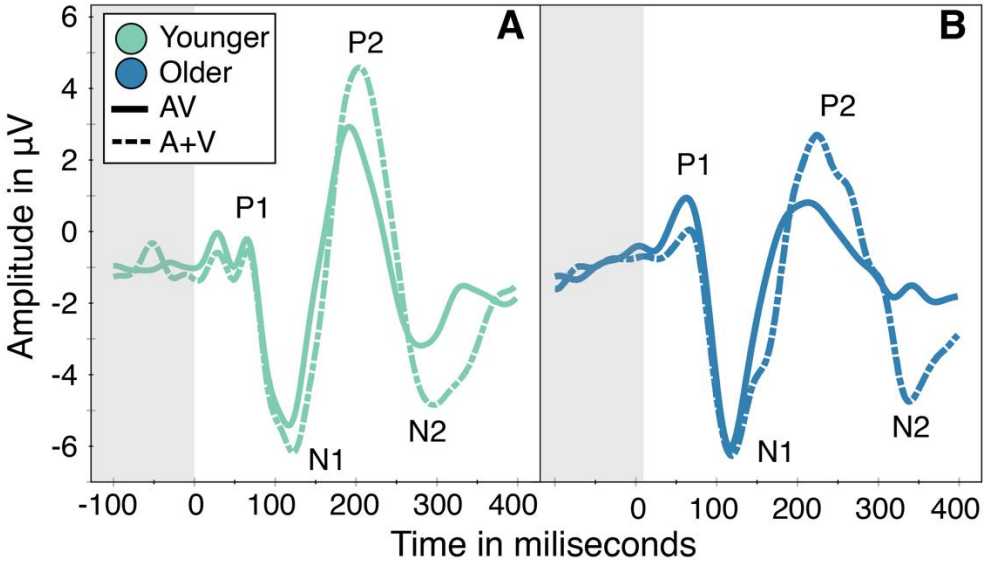

Figure 5. Group average electrophysiological responses for the younger adults (A) and older adults (B). In both panels, the response in the AV condition is represented as a plain line while the A+V is represented as a dotted line. In each plot, time is displayed in the X-axis while amplitude in µV is displayed in the y-axis. The 0 ms corresponds to the time of stimulus presentation. The shaded area represents the baseline period. The 4 components of interest are identified in each plot: P1, N1, P2 and N2.

The LMM analyses (marginal means) revealed no effect on P1. For **N1 amplitude**, there was a marginally significant effect of Modality ($F_{(1,30)} = 4.198$, $p = .049$), with a more negative signal amplitude for A+V compared to AV (Figure 6A). For **N1 latency**, there was also a significant effect of Modality ($F_{(1,30)} = 7.169$, $p = .012$), with a longer N1 latency for A+V compared to AV (Figure 6B). For **P2 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,30)} = 23.039$, $p < .001$), with higher P2 amplitude for A+V compared to AV (Figure 6C). For **P2 latency**, the LMM analyses revealed a main effect of Age Group ($F_{(1,30)} = 4.487$, $p = .043$), with P2 peaking later in the older

compared to the younger group (Figure 6D). There was also a main effect of Modality on P2 latency ($F_{(1,30)}$ = 4.983, p = .033), with a longer latency for A+V compared to AV (Figure 6E).

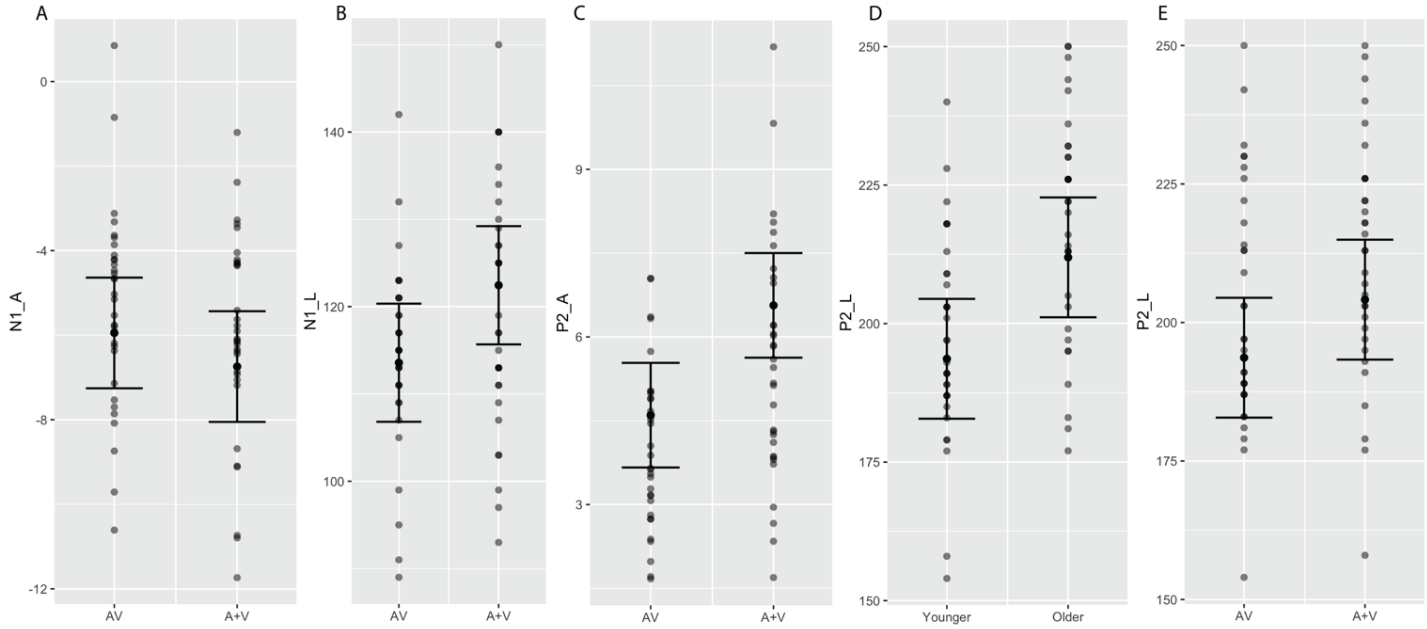

Figure 6 Results for N1 and P2 for the integration analyses (additive model). A. Modality difference in N1 amplitude. B. Modality difference in N1 latency (N1-L). C. Modality difference in P2 amplitude (P2_A). D. Age difference in P2 latency (P2_L). E. Modality difference in P2 latency (P2_L). In all figures, N1/P2 amplitude/latency is displayed in the y-axis. Each dot is a participant. The error bars represent standard errors of the marginal means.

For **N2 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,30)}$ = 22.341, p <.001), with a more negative N2 amplitude for A+V compared to AV (Figure 7A). For **N2 latency**, the LMM analyses revealed a main effect of Age Group ($F_{(1,29)}$ = 5.872, p = .022), with N2 peaking later in the older compared to the younger group (Figure 7B).

Table 3 summarizes all effects for this set of analyses.

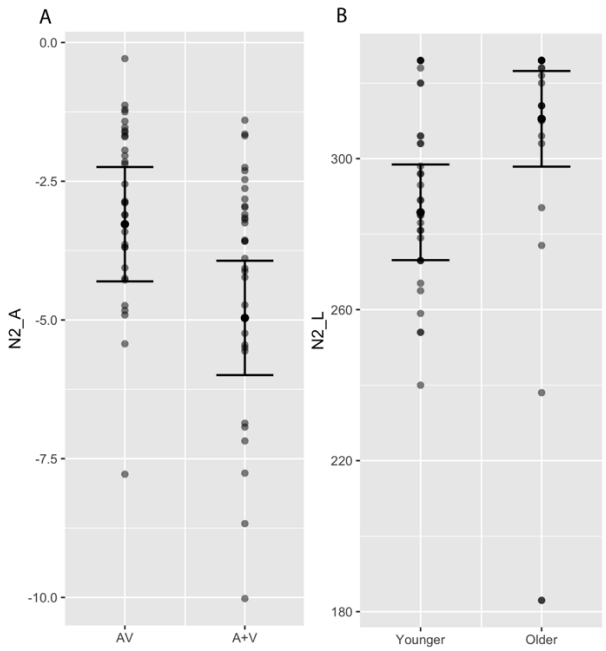

Figure 7 Results for N2 for the integration analyses (additive model). A. Modality difference in N2 amplitude. B. Age difference in N2 latency. D. Modality difference in P2 latency. N2 amplitude/latency is displayed in the y-axis. Each dot is a participant. In all graphs, the error bars represent standard errors of the marginal means.

Table 3. Type III Tests of Fixed Effects (marginal) for the EEG integration analyses (additive model)

Effects and interactions	DF		F	p
A. P1 amplitude				
Age Group	1	29	2.176	0.151
Modality	1	29	0.090	0.767
Hearing	1	29	0.030	0.865
Age Group * Modality	1	29	0.202	0.656
B. P1 latency				
Age Group	1	28	0.968	0.334
Modality	1	25	0.055	0.816
Hearing	1	28	0.005	0.943
Age Group * Modality	1	25	0.566	0.459
C. N1 amplitude				
Age Group	1	29	1.155	0.291
Modality	1	30	4.198	**0.049**
Hearing	1	29	4.605	**0.040**
Age Group * Modality	1	30	0.182	0.672
D. N1 latency				
Age Group	1	29	0.075	0.787
Modality	1	30	7.169	**0.012**
Hearing	1	29	2.555	0.121
Age Group * Modality	1	30	1.923	0.176
E. P2 amplitude				
Age Group	1	29	2.348	0.136
Modality	1	30	23.039	**<.0001**
Hearing	1	29	0.000	0.995
Age Group * Modality	1	30	2.147	0.153
F. P2 latency				
Age Group	1	29	4.487	**0.043**
Modality	1	30	4.983	**0.033**
Hearing	1	29	0.499	0.485
Age Group * Modality	1	30	0.248	0.622

		G. N2 amplitude		
Age Group	1	29	0.680	0.416
Modality	1	30	22.341	**0.0001**
Hearing	1	29	0.017	0.898
Age Group * Modality	1	30	1.897	0.179
		H. N2 latency		
Age Group	1	29	5.872	**0.022**
Modality	1	30	1.326	0.259
Hearing	1	29	0.080	0.780
Age Group * Modality	1	30	1.618	0.213

3.3. EEG Analyzes: Temporal and Phonetic Visual Predictive Cues

The second set of EEG analyses addresses objective 2b, to test whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural processing of speech similarly for younger and older adults (AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). To address this objective, first, we report the average response for each Age Group (younger, older) and each Predictive cue condition (Control, What, When, WW), separately for each modality (A, V, AV), in Figure 8. Though the visual condition was not analyzed, we present it in the figure for the sake of transparency and completeness. As can be seen in the Figure, the average response patterns were similar for the A and AV conditions, with reduced responses in the AV compared to the A condition, for both younger and older adults.

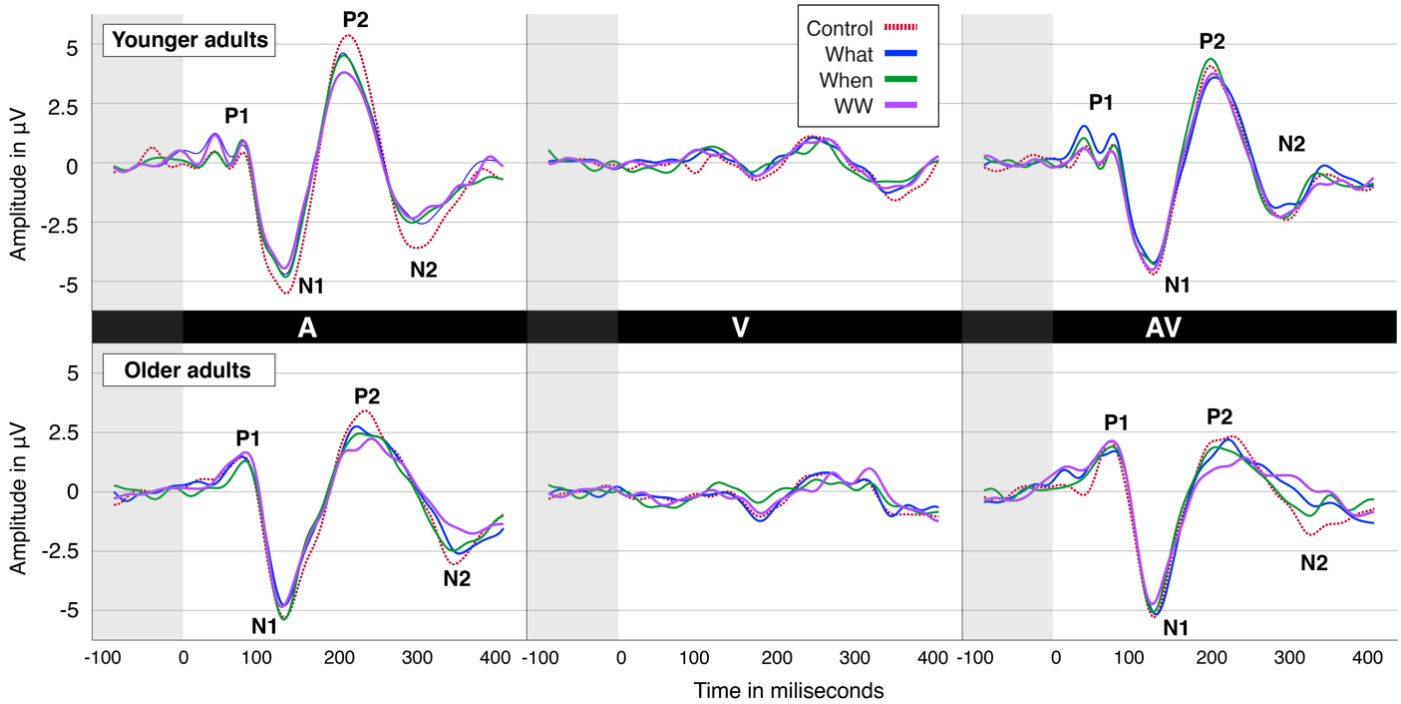

Figure 8 Group average electrophysiological responses for the younger adults (top panels) and older adults (bottom panels), separately for the A condition (left panels), V condition (central panels) and AV condition (right panel). In each plot, time is displayed in the X-axis while amplitude in µV is displayed in the y-axis. The 0 ms corresponds to the time of stimulus presentation. The shaded area represents the baseline period. The 4 components of interest are identified in each plot: P1, N1, P2 and N2.

For each component (P1, N1, P2 and N2), Q-Q plot and histograms were computed, which revealed that the residuals followed a normal or fairly normal distribution. The descriptive statistics for each dependent variable and the detailed results of the LMM analyses are provided in Supplementary Materials 13-14 (P1), 15-16 (N1), 17-18 (P2) and 19-20 (N2). As detailed in the following paragraphs and illustrated in Table 4, Age Group differences were found on P2 (both amplitude and latency) and N2 (latency).

For **P1 and N1 latency,** the marginal fixed effect tests revealed no main effects and no interaction. For **N1 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,210)} = 6.451$, $p = .012$). Tukey-corrected post hoc tests showed facilitation for AV compared to A ($\beta = -0.344$, SE = 0.26, p = .0324) on N1 amplitude (Figure 9A). There was also a main effect of Prediction Cue on N1

amplitude ($F_{(3,210)} = 4.913$, $p = .003$). Tukey-corrected post hoc tests were conducted to decompose the effect of Cue on N1 amplitude. The contrasts showed facilitation for What compared to the control condition ($\beta = -0.5469$, SE = 0.172, p = .0092), and for WW compared to the Control condition ($\beta = 0.344$, SE = 0.172, p = .0063) (Figure 9B).

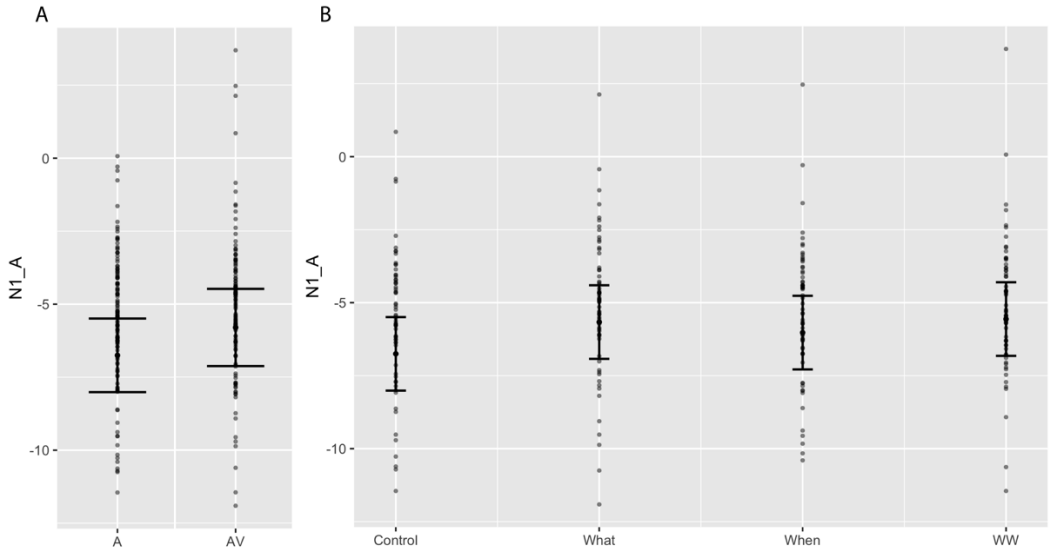

Figure 9 N1 amplitude (marginal means). A. Modality differences in N1 amplitude (µV; N1_A). Each modality is represented on the X-axis while amplitude is displayed in the y-axis. Each dot is a participant. B. Prediction Cue differences in N1 amplitude (µV; N1_A). The Cue conditions are represented on the X-axis while amplitude is displayed in the y-axis. Modalities are colour coded. Il all graphs, the error bars represent standard errors of the marginal means.

For **P2 amplitude**, the LMM analyses (marginal fixed effects) revealed a main effect of Age Group ($F_{(1,29)} = 10.958$, $p = .0003$) (Younger > Older) as well as a main effect of Modality ($F_{(1,210)} = 20.25$, $p < .001$) (A > AV), and Prediction Cue ($F_{(3,210)} = 10.30$, $p < .001$). There was also a 2-way interaction between Age Group and Modality ($F_{(1,210)} = 5.642$, $p = .0184$) and between Cue and Modality ($F_{(3,210)} = 6.12$, $p = .001$) on P2 amplitude. Finally, the LMM analyses also revealed a 3-way interaction between Age Group, Modality and Prediction Cue ($F_{(3,210)} = 3.468$, $p = .0171$) on P2 amplitude. Tukey-corrected post hoc tests were conducted to decompose the 3-way interaction on P2 amplitude. These tests showed P2 amplitude facilitation for AV compared to A in the Control condition

(β = 1.612, SE = 0.358, p = .0003), and a similar trend in the What condition (β = 1.049, SE = 0.358, p = 072), for younger adults, but no such P2 amplitude facilitation in older adults (Figure 10A). For **P2 latency**, we found a main effect of Age Group ($F_{(1,29)}$ = 8.439, *p* = .007), with longer P2 peak latency in the older group (Figure 10B).

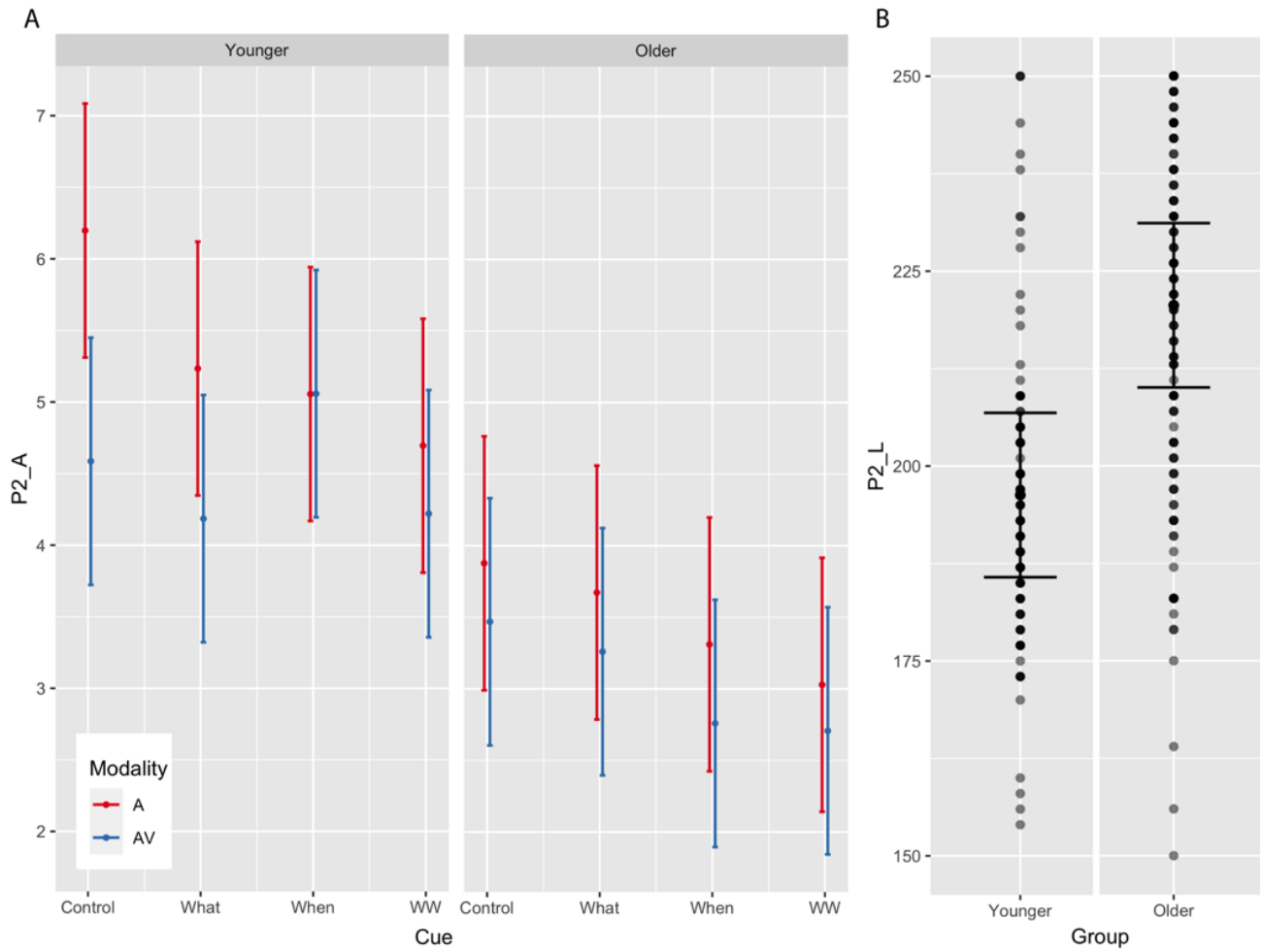

Figure 10. P2 results (marginal means). A. 3-way interaction between Age Group, Cue and Modality on P2 amplitude (P2_A). B. Group difference in P2 latency (P2_L). In all graphs, P2 amplitude/latency is displayed in the y-axis. The error bars represent standard errors of the marginal means. Each dot is a participant.

For **N2 amplitude**, the LMM analyses (marginal fixed effects) revealed a main effect of Prediction Cue ($F_{(3,210)}$ = 7.503, *p* = .0001) and a main effect of Modality ($F_{(1,210)}$ = 9.170, *p* = .003). Tukey-corrected post hoc tests were conducted to decompose the effect of Modality. These contrasts showed N2 amplitude facilitation for AV compared to A (β = -0.661, SE = 0.237, p = .0058) (Figure

11A). There was also as a marginally significant interaction between Age Group and Cue on N2 amplitude ($F_{(3,210)} = 2.6742$, $p = .048$) (Figure 11B). Tukey-corrected post hoc tests were conducted to decompose the interaction between Age Group and Cue on N2 amplitude. The tests revealed that, in younger adults, there was a difference in N2 amplitude between the Control and the What (β = -0.697, SE = 0.215, p = .03), When (β = -0.746, SE = 0.215, p = .014), and WW conditions (β = -0.831, SE = 0.215, p = .004). In the older adults, there was a difference in N2 amplitude between the Control and the What (β = -0.653, SE = 0.215, p = .054) and WW conditions (β = -0.779, SE = 0.215, p = .009), but not between the Control condition and the When condition (β = -0.312, SE = 0.215, p = .833). For **N2 latency**, the LMM analyses (marginal fixed effects) revealed a main effect of Age Group ($F_{(3,210)} = 4.1524$, $p = .007$), with longer latency for the older compared to the younger adults (Figure 11C).

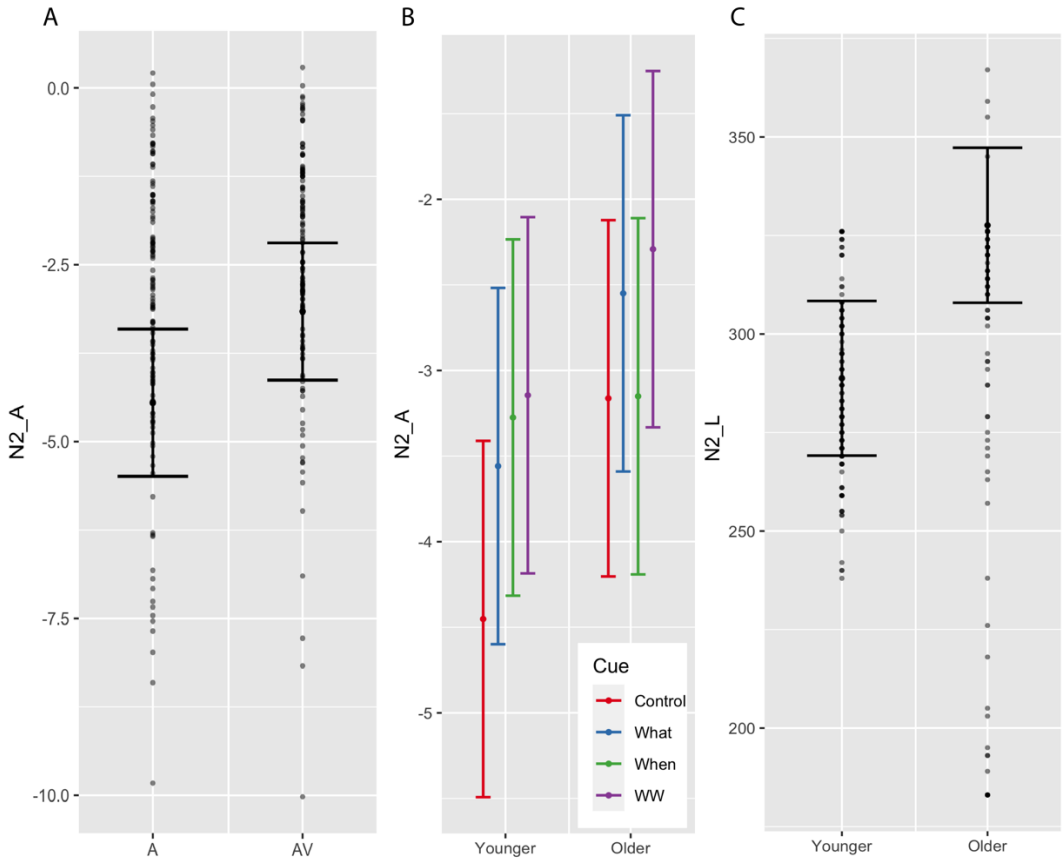

Figure 11. N2 results (marginal means). A. Main effect of Modality on N2 amplitude (N2_A). B. 2-way interaction between Age Group and Prediction Cue on N2 Amplitude (N2_A). C. Group difference in N2 latency (N2_L). In all graphs, the error bars represent standard errors of the marginal means. Each dot is a participant.

Table 4. Type III Tests of Fixed Effects (marginal) for the EEG analyses

Effects and interactions	DF		F	p
B. P1 amplitude				
Age Group	1	29	2.753	0.108
Prediction cue	3	210	0.703	0.551
Modality	1	210	0.412	0.522
Hearing	1	29	0.034	0.854
Age Group * Prediction cue	3	210	0.469	0.704
Age Group * Modality	1	210	0.086	0.770
Prediction cue * Modality	3	210	0.027	0.994
Age Group * Prediction cue * Modality	3	210	0.631	0.596
B. P1 latency				
Age Group	1	29	0.017	0.896
Prediction cue	3	198	1.579	0.196
Modality	1	198	0.150	0.699
Hearing	1	29	0.075	0.786
Age Group * Prediction cue	3	198	1.540	0.205
Age Group * Modality	1	198	2.223	0.138
Prediction cue * Modality	3	198	0.356	0.785
Age Group * Prediction cue * Modality	3	198	1.448	0.230
I. N1 amplitude				
Age Group	1	29	2.417	0.131
Prediction cue	3	210	4.913	**0.003**
Modality	1	210	6.451	**0.012**
Hearing	1	29	3.251	0.082
Age Group * Prediction cue	3	210	1.107	0.347
Age Group * Modality	1	210	1.904	0.169
Prediction cue * Modality	3	210	2.163	0.094
Age Group * Prediction cue * Modality	3	210	1.353	0.258
J. N1 latency				
Age Group	1	29	0.011	0.917
Prediction cue	3	210	0.326	0.807
Modality	1	210	2.703	0.102
Hearing	1	29	2.058	0.162
Age Group * Prediction cue	3	210	1.336	0.264
Age Group * Modality	1	210	0.338	0.562
Prediction cue * Modality	3	210	0.414	0.743
Age Group * Prediction cue * Modality	3	210	0.016	0.997
K. P2 amplitude				
Age Group	1	29	10.96	**0.003**
Prediction cue	3	210	10.30	**<.0001**
Modality	1	210	20.25	**<.0001**
Hearing	1	29	0.00	0.955
Age Group * Prediction cue	3	210	1.45	0.231
Age Group * Modality	1	210	5.64	**0.018**
Prediction cue * Modality	3	210	6.12	**0.001**
Age Group * Prediction cue * Modality	3	210	3.47	**0.017**
L. P2 latency				
Age Group	1	29	8.439	**0.007**
Prediction cue	3	210	0.322	0.810
Modality	1	210	0.441	0.507
Hearing	1	29	0.271	0.607
Age Group * Prediction cue	3	210	1.046	0.373
Age Group * Modality	1	210	0.529	0.468
Prediction cue * Modality	3	210	0.511	0.675
Age Group * Prediction cue * Modality	3	210	1.576	0.196
M. N2 amplitude				
Age Group	1	29	2.476	0.127
Prediction cue	3	210	7.503	**0.0001**
Modality	1	210	9.170	**0.003**
Hearing	1	29	0.402	0.531
Age Group * Prediction cue	3	210	2.674	**0.048**
Age Group * Modality	1	210	1.906	0.169
Prediction cue * Modality	3	210	2.117	0.099
Age Group * Prediction cue * Modality	3	210	2.029	0.111

N. N2 latency				
Age Group	1	29	7.476	**0.011**
Prediction cue	3	210	0.104	0.958
Modality	1	210	0.117	0.733
Hearing	1	29	0.246	0.624
Age Group * Prediction cue	3	210	1.456	0.228
Age Group * Modality	1	210	0.585	0.445
Prediction cue * Modality	3	210	0.355	0.786
Age Group * Prediction cue * Modality	3	210	0.065	0.979

3.4. Electrophysiology-Behaviour Relationship

This final set of analysis addresses the third objective of the study, to examine the relationship between speech perception performance and AEPs in young and older adults to shed new lights on brain aging and its impact on human behaviour. This analysis focused on overall VE and overall PE scores, and on AEPs (P2A, P2L, N2A and N2L) measures that showed Age Group effects or interactions in the statistical analyses detailed in the previous subsections.

For the analysis with overall **VE as dependent variable**, the predictor variables were the AEP components that showed an effect of Age Group or an interaction with Age Group: P2 amplitude average (basic analysis), P2 latency average (basic and integration analyses), N2 amplitude average (basic analysis) and N2 average latency (integration analysis). There were thus 5 analyses, which are detailed in **Supplementary Material 21**. The analyses revealed a significant mediation effect of P2 amplitude on overall VE for the older adults. As shown in Figure 12B and C, in older adults with higher overall VE, the amplitude of P2 was less positive.

For the analysis with overall **PE as dependent variable**, the predictor variables were the AEP components that showed an effect of Age Group or an interaction with Age Group: P2 amplitude average (basic analysis), P2 latency average (basic and integration analyses), N2 amplitude average (basic analysis) and N2 average latency (integration analysis). There were thus 5 analyses, which are detailed in **Supplementary Material 22**. There were no significant mediation effects of AEPs and no direct effect of AEPs on overall VE scores.

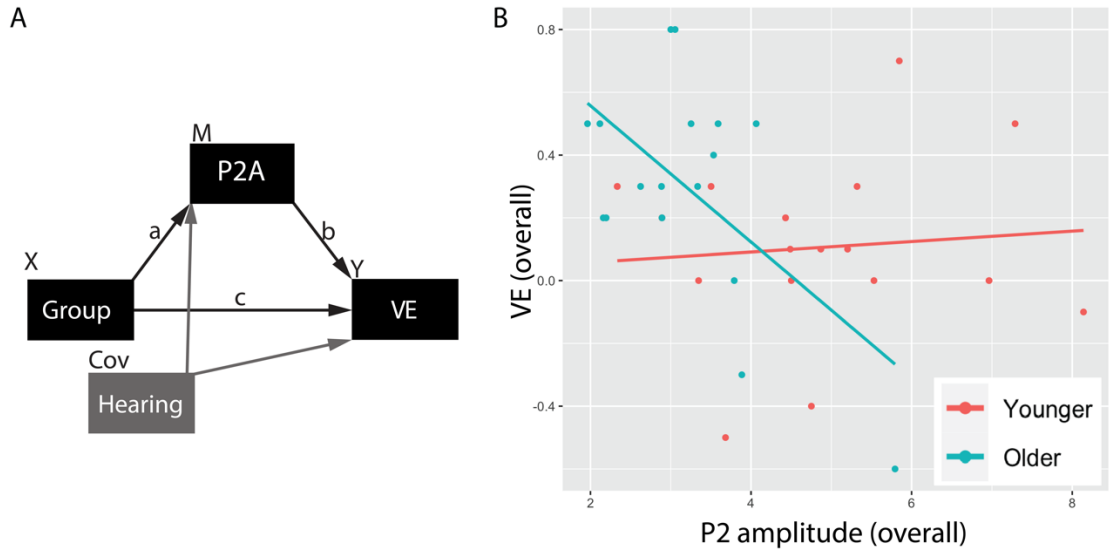

Figure 12 Significant mediation results illustrating the relationship between VE and P2 amplitude for younger and older adults. A. Mediation model for the significant mediation. X = the predictor variable; M = the mediator variable; Y = the dependent variable; Cov = covariate; a represents the effect of X on M; b represents the effect of M on Y. B. The scatterplot displays the mediating effect of Age Group on VE through P2 amplitude. In both the scatterplots, each dot represents a participant. The red dots represent the younger participants while the blue dots represent the older participants.

4. Discussion

The general objective of the present study was to fill a knowledge gap about the neural processing of auditory speech in aging under different levels of prediction. By measuring multiple early and late auditory evoked potentials (P1-N1-P2 and N2), we aimed to shed new lights on the locus of age-related differences in neural speech processing. To examine the effect of prediction, we compared the processing of auditory and audiovisual speech, and we manipulated prior knowledge on auditory syllables by presenting participants with visual information indicative of the temporal unfolding (*when*) or the phonetic content (*what*) of auditory syllables. Based on Winneke and Phillips (2011), we hypothesized that, compared to younger adults, older adults would show similar or enhanced audiovisual and predictive effects. Specifically, we expected an enhanced facilitation of AEPs (shorter latency and/or reduced amplitude) during AV compared to A speech, as well as during the processing of temporal and phonetic predictive cues, reflecting a lifetime of experience with speech, or perhaps a

compensation strategy to overcome possible hearing loss. Given the well-established cognitive decline that occurs in aging (e.g., Park, Lautenschlager, Hedden, Davidson, Smith, & Smith, 2002; Salthouse, 1996; Salthouse, 2009), we expected that the auditory N2, which indexes cognitive and executive processes (e.g. Czigler, Csibra, & Ambró, 1997; Falkenstein, Hoormann, & Hohnsbein, 1999; Folstein & Van Petten, 2008), would show either an increased amplitude, reflecting compensatory activity, or perhaps a decline, reflecting a disruption in processing.

The main findings of the study are as follows: the behavioural results show that (1) there was a VE effect in both groups (as shown in Figure 4 and detailed in Table 2B), but, contrary to our prediction, it was lower in older compared to younger adults, (2) predictive cues facilitated speech recognition in younger and older adults alike (Table 2C). The EEG results showed that (3) age differences in AEPs are localized to later components (P2 and N2, see Tables 3 and 4), suggesting that aging predominantly affects higher-order cortical processes related to speech processing rather than lower-level auditory processes. (4) Specifically, the latency of the P2 and N2 components were delayed in older adults (Tables 3 and 4). Moreover, there was reduced AV facilitation on P2 amplitude in older adults (Figure 10), while there was a reduced effect of the When cue on N2 amplitude for older compared to younger adults (Figure 11). Finally, the mediation analyzes on EEG-behaviour relationship revealed that (5) overall VE scores were associated with P2 amplitude (Figure 12).

4.1. Prediction Facilitation as a Function of Age

The main objective of the study was to examine whether adding visual cues (natural or unnatural) to the speech signal would facilitate speech perception similarly in younger and older adults, behaviourally (objective 1) and at the neural level (objective 2).

First, our behavioural results revealed that overall speech recognition accuracy in a straightforward syllable identification task performed in quiet was lower for older adults in the V condition, suggestive of poorer lipreading ability in older compared to younger adults. Here,

participants reported normal or corrected-to-normal vision, and they were allowed to wear their glasses or contact lenses during the experiment, although their visual acuity was not assessed. The low accuracy in the V modality could therefore reflect degraded sensory input or reduced visual processing efficiency. While the clinical relevance of reduced lipreading should be investigated in future work, this finding is consistent with previous studies showing that, compared to younger adults, middle-aged and/or older adults exhibit reduced lipreading abilities (Cienkowski & Carney, 2002; Sommers et al., 2005; Tye-Murray et al., 2010; Winneke & Phillips, 2011). Future studies need to determine if adequate correction for age-related visual impairment can restore lip-reading ability.

Despite a lower performance in the V condition, however, our study is the first to show that visual cues can improve speech processing accuracy in older adults. Accuracy in the visual modality was enhanced for the young and older adults when visual predictive cues were provided (especially the What and the WW cues). Moreover, the prediction effect (PE) scores did not show any effect of age groups. In the V modality, older adults improved maximally upon presentation of the dual (WW) cue, as did younger adults. This suggests that the visual cues were processed and that they were helpful to older adults. Given that visual speech information can be processed, even at the categorical level, by the human brain (O'Sullivan, Crosse, Di Liberto, & Lalor, 2016), the finding that prediction cues can improve lipreading in older adults has potential implications for the care and rehabilitation of older adults with communication difficulties, especially for those with significant hearing loss who rely more heavily on visual speech and audiovisual integration (Puschmann, Daeglau, Stropahl, Mirkovic, Rosemann, Thiel, & Debener, 2019). Identifying strategies to improve visual speech processing is of key importance for this population, but additional empirical evidence is needed to determine if visual cues such as the ones used in the present study have beneficial impacts on elderly populations with hearing impairments.

Though our results show that older adults are capable of integrating AV information to decipher speech, older participants in the present study exhibited poorer VE compared to younger ones, which

suggests that, though preserved, the ability to benefit from natural visual information is reduced in aging. Importantly, VE was unaffected by hearing. This result is in line with studies demonstrating that performance enhancement for AV speech compared to A speech is lower for older compared to younger adults (Tye-Murray et al., 2010; Yang & Ren, 2018), but it is at odds with other studies that did not find a decline in enhancement in aging (Avivi-Reich et al., 2018; Cienkowski & Carney, 2002; Ganesh et al., 2017; Sommers et al., 2005; Winneke & Phillips, 2011). Recent investigations have suggested an age-related change in the conditions needed for older adults to benefit from AV speech, rather than an absolute incapacity to benefit from an audiovisual signal, with older adults benefiting less when signal-to-noise ratio is low (Jansen et al., 2018; Stevenson et al., 2015) and when the visual information is degraded (Gordon & Allen, 2009). Crucially, in the present study, an age difference was observed even in quiet, using non-degraded stimuli. In addition to the context, it is possible that changes in AV integration are progressive. Indeed, in several previous studies, participants classified as "older" were, on average, younger than 60 (Jansen et al., 2018; Stevenson et al., 2015), which is young. In the present study, the older group was older (average of 67 years), and a group difference was observed in the absence of noise. Additionally, one cannot exclude that effects of dual-tasking and attention-sharing might be present in our experimental tasks, preventing older participants from maximally benefiting from the visual cues. Additional studies are needed to investigate the environmental conditions and participants characteristic (including visual acuity and cognitive capabilities) that affect audiovisual facilitation.

Together, these findings suggest that predictive coding capabilities, in general, may be slightly reduced in older adults. Predictive coding theories suggest that our perceptual experience is determined by a fine balance between internal predictions based on priors acquired over the course of our lifetimes and incoming sensory evidence (Clark, 2013; Friston, 2005, 2010; Rao & Ballard, 1999). Sensory evidence and priors are thought to be fused in a Bayesian way, leading to a prediction about the state of the world. Aging offers a unique opportunity to probe this notion, as the amount of phonetic

information accumulated, especially about talker variability, increases over the lifetime, as well as knowledge about communication contexts, while at the same time predictive coding mechanisms probability become more fine-tuned and efficient. Importantly, however, the precision of sensory processing degrades with age, in both the auditory and visual modalities. While the first factor (accumulated knowledge) strengthens the influence of predictions, the second (reduced sensory processing) lowers the influence of sensory evidence, especially unisensory signals. One possibility is that experience plays a stronger role than sensory decline, but our results suggest that impoverished sensory processing may be playing a stronger part in the interplay between these opposing forces. Additional studies are needed to compare predictive coding in younger and older adults with various levels of visual and auditory acuity to determine the threshold at which sensory processing becomes the dominant mechanisms driving predictive coding.

4.2. Age Differences in Electrophysiological Responses Are Localized to P2 and N2

A central objective of the present study was to shed new lights on the locus of age differences in the neural processing of speech sounds, which has major implications for understanding speech processing difficulties and to guide rehabilitation research and practice. This was achieved by investigating several ERP components: the P1-N1-P2 complex as well as the later N2 component.

Our results show that, controlling for hearing thresholds in the low frequency range (PTA), age differences in electrophysiological response to speech were predominantly located to P2 and N2. This suggests that speech processing difficulties may result from changes to higher-order cortical processes rather than lower-level auditory ones.

While several studies have reported age differences in the neural responses to speech in older adults, and more generally, the neural responses to sounds (e.g. Anderson & Karawani, 2020), uncertainty remains regarding the nature of these changes and how they affect speech perception performance. Previous AV speech studies investigating AEPs in aging have largely focused on the P1-

N1-P2 complex. The P1-N1-P2 complex represents the obligatory sensory-evoked response of the auditory system (Naatanen & Picton, 1987). The P1-N1-P2 complex can be used to indirectly assess the integrity of the central auditory system, up to the cortex. The earlier P1 primarily originates from the primary auditory cortex, and is thought to have subcortical contributions, including the reticular activating system (e.g., Erwin & Buchwald, 1987). The N1 component has been localized to the primary and secondary auditory associative regions (e.g., Eggermont & Ponton, 2002; Naatanen & Picton, 1987). It is related to stimulus detection and the encoding of auditory stimulus properties. Both components are obligatory auditory responses unaffected by attention. In early studies of syllable discrimination, Tremblay et al. reported evidence of neural disruptions in older adults in the form of delayed N1 latency, potentially reflecting age-related changes in neural synchrony (Tremblay, Piskosz, & Souza, 2002, 2003). A similar age-related delayed N1 latency was reported by Bidelman et al. during a vowel categorization task (Bidelman, Villafuerte, Moreno, & Alain, 2014). However, Soros and colleagues, using MEG, found no delays in P1 or N1m either during passive listening of rapid sequences of speech sounds, but instead stronger amplitude for older adults (Soros, Teismann, Manemann, & Lutkenhoner, 2009). Likewise, Roque et al. reported age differences in P1 during an auditory word identification task performed in quiet, with older adults exhibiting earlier P1 peak latency and larger amplitude compared to younger adults; but no difference on N1 was found (Roque, Karawani, Gordon-Salant, & Anderson, 2019). In the present study, the P1 and N1 showed no signs of age-related disruptions. This suggests that basic auditory processing during a simple syllable recognition in quiet task, at least in our sample, was largely unchanged with age.

In contrast to early components, the later auditory components—P2 and N2—showed evidence of age-related differences, with overall lower amplitude and longer latencies. Latency generally reflects the time point of the peak neurophysiological response peak relative to the eliciting stimulus. Latency is related to neural conduction time and site of excitation: the time it takes for the sound to travel through the peripheral auditory system to the place of excitation in the central nervous system (Alain &

Tremblay, 2007). Earlier peaks are therefore presumed to reflect the output of a process that is completed earlier in time. Several studies have shown an age-related increase in P2 latency (e.g., Billings, Penman, McMillan, & Ellis, 2015; Czigler, Csibra, & Csontos, 1992; Goodin, Squires, Henderson, & Starr, 1978; Iragui, Kutas, Mitchiner, & Hillyard, 1993; Tremblay, Billings, & Rohila, 2004; Tremblay et al., 2002, 2003). An increase in latency could reflect increased neural conduction time with normal aging, or more laborious neural processing which could be related to decline in the structure of the cerebral cortex, for example.

The auditory P2 is thought to reflect synchronous neural activation in the thalamic-cortical segment of the central nervous system, mainly originating from the supratemporal plane of the auditory cortex (e.g., Naatanen & Picton, 1987). Previous neurophysiological studies, as well as the present one, have shown that the N1/P2 complex occurs earlier and its amplitude is lower for AV compared to unimodal (A) speech processing (e.g., Besle et al., 2004; Klucharev et al., 2003; Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; van Wassenhove et al., 2005). P2 is often considered as an index of AV integration. Here, while P2 latency showed an age-related delay, it showed no evidence of interaction between group and modality, either in the basic analysis or in the integration analyses, suggesting that AV integration is taking place in older adults in a manner that is similar to younger adults. However, in terms of amplitude, there was some evidence of reduced P2 facilitation in older adults compared to younger adults (i.e. in the Control and the What conditions). These findings suggest that P2's sensibility to prediction is slightly reduced in aging. Interestingly, the brain-behaviour analyses revealed a relationship between P2 amplitude difference and overall VE scores. Older adults who exhibited facilitation (overall lower P2 amplitude) also exhibited a higher overall VE score, suggesting that a lowering of P2 amplitude with age may be normal and even beneficial to AV speech processing, perhaps reflecting increased experience and reduced processing need. Together, these results suggest that normal age-related changes to P2 affect speech processing skills in older adults, with those exhibiting a more efficient audiovisual integration process, as reflected

by lower P2 amplitude, maintaining better AV speech skills. This notion is consistent with a prior study that found that increased N1-P2 signal was negatively correlated with speech classification performance in young and older adults (Bidelman et al., 2014). The finding of reduced amplitude and delayed P2 in older adults suggest that audiovisual integration may be less energy consuming in older adults. Interestingly, Anderson and Karawani (2020) have suggested that an imprecise stimulus representation associated with age could lead to delays in P2. Future studies with larger sample sizes and using more challenging speech tasks will help clarify the mechanism that underlies age-related changes in P2, and more generally, in audiovisual integration for speech. Critically, such analyses will need to take into account not just hearing but also visual capabilities.

In addition to age differences in the auditory P2, we also found significant differences in the auditory N2. Similar to P2, N2 was generally delayed in older adults, but it also showed a decline in facilitation effects compared to younger adults. The N2 is a broad, slow negativity that occurs in the 140–300 ms latency range after stimulus onset. It is thought to index attentional discrimination process and response monitoring. Previous studies have shown delayed N2 latency for older compared to younger adults (Dushanova & Christov, 2013; Goodin et al., 1978; Schiff, Valenti, Andrea, Lot, Bisiacchi, Gatta, & Amodio, 2008) and lower N2 response amplitudes in older compared to younger adults during a speech perception in noise task (Billings et al., 2015). In a series of studies, Czigler et al. showed that attention-related processes in the visual modality, as indexed by event-related potentials—have longer latency in the elderly (Czigler et al., 1997). This delay in the posterior N2 (selection negativity) was considered as evidence for age-related delay in stimulus evaluation. As a consequence of the slowing down of attentional processes, as indicated by the slowing of the auditory N2, the use of prediction in speech perception may become less efficient in older adults, however, we did not find significant relationship between N2 and behaviour.

5. Conclusion

Our results indicate that aging affect speech processing at multiple levels, including sensory (visual) processing, multimodal integration and auditory attentional processes. Despite impaired lipreading abilities, older adults were able to use visual prediction cues to help decipher speech, suggesting preserved AV integration capabilities. These findings have important implications for rehabilitation research and interventions by suggesting that improving communication in older ages may rest on the need to enhance sensory processing, but also information integration and attention processes.

6. Acknowledgements

PT holds a Career award from the Fonds de la Recherche en Santé du Quebec (FRQ-S, #35016).

We thank all participants.

7. CRediT author statement

Pascale Tremblay: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing- Original draft preparation. **Serge Pinto**: Conceptualization, Investigation, Writing—Reviewing and Editing. **Anahita Basirat**: Conceptualization, Writing—Reviewing and Editing. **Marc Sato**: Conceptualization, Methodology, Investigation, Resources, Project administration, Writing—Reviewing and Editing.

8. References

Alain, C., & Tremblay, K. (2007). The role of event-related brain potentials in assessing central auditory processing. *Journal of the American Academy of Audiology, 18*(7), 573-589.

Anderson, S., & Karawani, H. (2020). Objective evidence of temporal processing deficits in older adults. *Hear Res, 397,* 108053. doi:10.1016/j.heares.2020.108053

Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A. L. (2009). Dual neural routing of visual facilitation in speech processing. *J Neurosci, 29*(43), 13445-13453. doi:10.1523/JNEUROSCI.3194-09.2009

Avivi-Reich, M., Puka, K., & Schneider, B. A. (2018). Do age and linguistic background alter the audiovisual advantage when listening to speech in the presence of energetic and informational masking? *Atten Percept Psychophys, 80*(1), 242-261. doi:10.3758/s13414-017-1423-5

Baart, M., Stekelenburg, J. J., & Vroomen, J. (2014). Electrophysiological evidence for speech-specific audiovisual integration. *Neuropsychologia, 53*, 115-121. doi:10.1016/j.neuropsychologia.2013.11.011

Besle, J., Fort, A., Delpuech, C., & Giard, M.-H. (2004). Bimodal speech: Early suppressive visual effects in human auditory cortex. *European Journal of Neuroscience, 20*, 2225-2234.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging, 35*(11), 2526-2540. doi:10.1016/j.neurobiolaging.2014.05.006

Billings, C. J., Penman, T. M., McMillan, G. P., & Ellis, E. M. (2015). Electrophysiology and Perception of Speech in Noise in Older Listeners: Effects of Hearing Impairment and Age. *Ear and hearing, 36*(6), 710-722. doi:10.1097/AUD.0000000000000191

Brodbeck, C., Presacco, A., Anderson, S., & Simon, J. Z. (2018). Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust United Acust, 104*(5), 774-777. doi:10.3813/AAA.919221

Cervantes Constantino, F., & Simon, J. Z. (2018). Restoration and Efficiency of the Neural Processing of Continuous Speech Are Promoted by Prior Knowledge. *Frontiers in systems neuroscience, 12*, 56. doi:10.3389/fnsys.2018.00056

CHABA. (1988). Speech understanding and aging. Working Group on Speech Understanding and Aging. Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council. *J Acoust Soc Am, 83*(3), 859-895.

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear Hear, 23*(5), 439-449. doi:10.1097/00003446-200210000-00006

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci, 36*(3), 181-204. doi:10.1017/S0140525X12000477

Czigler, I., Csibra, G., & Ambró, A. (1997). Age and Information Processing. *European Psychologist*, 247-257. doi:https://doi.org/10.1027/1016-9040.2.3.247

Czigler, I., Csibra, G., & Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biological psychology, 33*(2-3), 195-206. doi:10.1016/0301-0511(92)90031-o

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21. doi:10.1016/j.jneumeth.2003.10.009

Dushanova, J., & Christov, M. (2013). Auditory event-related brain potentials for an early discrimination between normal and pathological brain aging. *Neural Regen Res, 8*(15), 1390-1399. doi:10.3969/j.issn.1673-5374.2013.15.006

Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: from single units to evoked potentials. *Audiology & neuro-otology, 7*(2), 71-99. doi:10.1159/000057656

El-Assal, H. A., & El-Gharib, A. M. (2019). Encoding of speech in noise in adults using hearing aids: effect of noise reduction algorithm. *Hearing, Balance and Communication* doi:https://doi.org/10.1080/21695717.2019.1667687

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *J Speech Hear Res, 12*(2), 423-425. doi:10.1044/jshr.1202.423

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol (Amst), 101*(2-3), 267-291. doi:10.1016/s0001-6918(99)00008-6

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170. doi:10.1111/j.1469-8986.2007.00602.x

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 360*(1456), 815-836. doi:W5T4QMCP8T4K0UP8 [pii]

10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci, 11*(2), 127-138. doi:10.1038/nrn2787

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention--focusing the searchlight on sound. *Curr Opin Neurobiol, 17*(4), 437-455. doi:10.1016/j.conb.2007.07.011

Ganesh, A. C., Berthommier, F., & Schwartz, J.-L. (2017). Audiovisual Binding for Speech Perception in Noise and in Aging. *Laanguage learning, 68*, 193-220. doi:https://doi.org/10.1111/lang.12271

Goodin, D. S., Squires, K. C., Henderson, B. H., & Starr, A. (1978). Age-related variations in evoked potentials to auditory stimuli in normal human subjects. *Electroencephalography and clinical neurophysiology, 44*(4), 447-458. doi:10.1016/0013-4694(78)90029-9

Gordon, M. S., & Allen, S. (2009). Audiovisual speech in older and younger adults: integrating a distorted visual signal with speech in noise. *Exp Aging Res, 35*(2), 202-219. doi:10.1080/03610730902720398

Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten Percept Psychophys, 75*(3), 508-524. doi:10.3758/s13414-012-0406-9

Humes, L. E., Kidd, G. R., & Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Frontiers in systems neuroscience, 7*, 55. doi:10.3389/fnsys.2013.00055

Humes, L. E., Wilson, D. L., Barlow, N. N., & Garner, C. (2002). Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults. *J Speech Lang Hear Res, 45*(4), 772-782. doi:10.1044/1092-4388(2002/062)

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychol Methods, 15*(4), 309-334. doi:10.1037/a0020761

Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology, 30*(1), 10-22. doi:10.1111/j.1469-8986.1993.tb03200.x

Jansen, S. D., Keebler, J. R., & Chaparro, A. (2018). Shifts in Maximum Audiovisual Integration with Age. *Multisens Res, 31*(3-4), 191-212. doi:10.1163/22134808-00002599

Klucharev, V., Mottonen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Brain Res Cogn Brain Res, 18*(1), 65-75. doi:10.1016/j.cogbrainres.2003.09.004

Kornhuber, H. H., & Deecke, L. (1965). [Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials]. *Pflugers Arch Gesamte Physiol Menschen Tiere, 284*, 1-17.

Larouche, E., Tremblay, M. P., Potvin, O., Laforest, S., Bergeron, D., Laforce, R., . . . Hudon, C. (2016). Normative Data for the Montreal Cognitive Assessment in Middle-Aged and Elderly Quebec-French People. *Arch Clin Neuropsychol*. doi:10.1093/arclin/acw076

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol Aging, 27*(8), 1155-1163. doi:10.1016/j.neurobiolaging.2005.05.024

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain, 106 (Pt 3)*, 623-642. doi:10.1093/brain/106.3.623

Lin, F. R. (2011). Hearing loss and cognition among older adults in the United States. *J Gerontol A Biol Sci Med Sci, 66*(10), 1131-1136. doi:10.1093/gerona/glr115

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25*(6), 763-770. doi:10.1037/a0024238

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., . . . Health, A. B. C. S. G. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern Med, 173*(4), 293-299. doi:10.1001/jamainternmed.2013.1868

Massaro, D. W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. Cambridge, London: The MIT Press.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

Naatanen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425. doi:10.1111/j.1469-8986.1987.tb00311.x

Nasreddine, Z. S., Chertkow, H., Phillips, N., Bergman, H., & Whitehead, V. (2003). Sensitivity and Specificity of The Montreal Cognitive Assessment (MoCA) for Detection of Mild Cognitive Deficits. *Can J Neurol Sci 30*(30).

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc, 53*(4), 695-699. doi:10.1111/j.1532-5415.2005.53221.x

O'Sullivan, A. E., Crosse, M. J., Di Liberto, G. M., & Lalor, E. C. (2016). Visual Cortical Entrainment to Motion and Categorical Speech Features during Silent Lipreading. *Front Hum Neurosci, 10*, 679. doi:10.3389/fnhum.2016.00679

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and aging*, *17*(2), 299-320.

Pichora-Fuller, M. K. (1997). Language comprehension in older listeners. *Journal of Speech-Language Pathology and Audiology*, *21*(2), 125–142.

Pilling, M. (2009). Auditory event-related potentials (ERPs) in audiovisual speech perception. *J Speech Lang Hear Res*, *52*(4), 1073-1081. doi:10.1044/1092-4388(2009/07-0276)

Pinto, S., Tremblay, P., Basirat, A., & Sato, M. (2019). The impact of when, what and how predictions on auditory speech perception. *Exp Brain Res*. doi:10.1007/s00221-019-05661-5

Presacco, A., Simon, J. Z., & Anderson, S. (2016a). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol*, *116*(5), 2356-2367. doi:10.1152/jn.00373.2016

Presacco, A., Simon, J. Z., & Anderson, S. (2016b). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol*, *116*(5), 2346-2355. doi:10.1152/jn.00372.2016

Puschmann, S., Daeglau, M., Stropahl, M., Mirkovic, B., Rosemann, S., Thiel, C. M., & Debener, S. (2019). Hearing-impaired listeners show increased audiovisual benefit when listening to speech in noise. *NeuroImage*, *196*, 261-268. doi:10.1016/j.neuroimage.2019.04.017

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79-87. doi:10.1038/4580

Roque, L., Karawani, H., Gordon-Salant, S., & Anderson, S. (2019). Effects of Age, Cognition, and Neural Encoding on the Perception of Temporal Speech Cues. *Front Neurosci*, *13*, 749. doi:10.3389/fnins.2019.00749

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol Rev*, *103*(3), 403-428.

Salthouse, T. A. (2009). Decomposing age correlations on neuropsychological and cognitive variables. *J Int Neuropsychol Soc, 15*(5), 650-661. doi:10.1017/S1355617709990385

Scherg, M., & Von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and clinical neurophysiology, 65*(5), 344-360. doi:10.1016/0168-5597(86)90014-6

Schiff, S., Valenti, P., Andrea, P., Lot, M., Bisiacchi, P., Gatta, A., & Amodio, P. (2008). The effect of aging on auditory components of event-related brain potentials. *Clin Neurophysiol, 119*(8), 1795-1802. doi:10.1016/j.clinph.2008.04.007

Sekiyama, K., Soshi, T., & 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

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*(3), 263-275.

Soros, P., Teismann, I. K., Manemann, E., & Lutkenhoner, B. (2009). Auditory temporal processing in healthy aging: a magnetoencephalographic study. *BMC Neurosci, 10*, 34. doi:10.1186/1471-2202-10-34

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci, 19*(12), 1964-1973. doi:10.1162/jocn.2007.19.12.1964

Stevenson, R. A., Nelms, C. E., Baum, S. H., Zurkovsky, L., Barense, M. D., Newhouse, P. A., & Wallace, M. T. (2015). Deficits in audiovisual speech perception in normal aging emerge at the level of whole-word recognition. *Neurobiol Aging, 36*(1), 283-291. doi:10.1016/j.neurobiolaging.2014.08.003

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(212-215).

Team, R. C. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Teder-Salejarvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res Cogn Brain Res, 14*(1), 106-114. doi:10.1016/s0926-6410(02)00065-4

Tingley D, Y. T., Hirose K, Keele L, Imai K. (2013). mediation: R Package for Causal Mediation Analysis. (Version R package version 4.4.2). Retrieved from http://CRAN.R-project.org/

Tobias, J. V. (1977). The need for preventive audiology. *Audiol. Hear. Educ. , 3*(2), 37-38.

Treille, A., Cordeboeuf, C., Vilain, C., & Sato, M. (2014a). Haptic and visual information speed up the neural processing of auditory speech in live dyadic interactions. *Neuropsychologia, 57*, 71-77. doi:10.1016/j.neuropsychologia.2014.02.004

Treille, A., Vilain, C., Kandel, S., & Sato, M. (2017). Electrophysiological evidence for a self-processing advantage during audiovisual speech integration. *Exp Brain Res, 235*(9), 2867-2876. doi:10.1007/s00221-017-5018-0

Treille, A., Vilain, C., & Sato, M. (2014b). The sound of your lips: electrophysiological cross-modal interactions during hand-to-face and face-to-face speech perception. *Front Psychol, 5*, 420. doi:10.3389/fpsyg.2014.00420

Treille, A., Vilain, C., Schwartz, J. L., Hueber, T., & Sato, M. (2018). Electrophysiological evidence for Audio-visuo-lingual speech integration. *Neuropsychologia, 109*, 126-133. doi:10.1016/j.neuropsychologia.2017.12.024

Tremblay, K. L., Billings, C., & Rohila, N. (2004). Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol, 15*(3), 226-237; quiz 264. doi:10.3766/jaaa.15.3.5

Tremblay, K. L., Piskosz, M., & Souza, P. (2002). Aging alters the neural representation of speech cues. *Neuroreport, 13*(15), 1865-1870.

Tremblay, K. L., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol, 114*(7), 1332-1343.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., & Hale, S. (2010). Aging, audiovisual integration, and the principle of inverse effectiveness. *Ear Hear, 31*(5), 636-644. doi:10.1097/AUD.0b013e3181ddf7ff

van Wassenhove, V. (2013). Speech through ears and eyes: interfacing the senses with the supramodal brain. *Front Psychol, 4*, 388. doi:10.3389/fpsyg.2013.00388

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, 102*(4), 1181-1186.

Vroomen, J., & Stekelenburg, J. J. (2010). Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J Cogn Neurosci, 22*(7), 1583-1596. doi:10.1162/jocn.2009.21308

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res Rev, 23*(Pt B), 154-166. doi:10.1016/j.arr.2015.06.002

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*(2), 427-438. doi:10.1037/a0021683

World Medical, A. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191-2194. doi:10.1001/jama.2013.281053

Yang, W., & Ren, Y. (2018). Attenuated audiovisual integration in middle-aged adults in a discrimination task. *Cogn Process, 19*(1), 41-45. doi:10.1007/s10339-017-0838-1

Supplementary Material

Click here to access/download;Supplementary Material;SupplementaryMaterial_R3.docx

1. Hearing (pure tone average in the low frequency range)

Table 1.1 Estimated marginal means for each group at each frequency tested

Group	Frequency	M	SE	df	95% CI LL	95% CI UL
Younger	500 Hz	5.800	1.600	31	2.540	9.060
Older		9.220	1.580	30	5.980	12.450
Younger	1000 Hz	5.310	1.580	31	2.080	8.540
Older		12.810	1.580	30	9.580	16.050
Younger	2000 Hz	4.530	1.580	31	1.300	7.760
Older		20.310	1.580	30	17.080	23.550

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

Table 1.2 Type III Tests of Fixed Effects for hearing thresholds (marginal)

Effects	DF		F	p
(Intercept)	1	149	11.7382	0.0008
Age Group	1	30	1.33086	0.2578
Ear	1	149	0.48229	0.4885
Frequency	2	149	0.29426	0.7455
Age group * Ear	1	149	0.03649	0.8488
Age group * Frequency	2	149	8.07806	**0.0005**
Ear * Frequency	2	149	0.12539	0.8822
Age group * Ear * Frequency	2	149	0.14835	0.8623

Note. DF = degrees of freedom. The r model was: Threshold ~ Age group * Ear * Frequency, random = ~1 | Subject

Table 1.3 Coefficients, standard errors and CI for the fixed and random effects for hearing

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	6.56	1.92	2.78 – 10.35	**0.001**
Age group [1]	3.12	2.71	-2.41 – 8.66	0.258
Ear [Right]	-1.53	2.2	-5.87 – 2.81	0.488
Frequency [1000]	-1.25	2.16	-5.51 – 3.01	0.563
Frequency [2000]	-1.56	2.16	-5.82 – 2.70	0.47
Age group [1] * Ear [Right]	0.59	3.08	-5.49 – 6.67	0.849
Age group [1] * Frequency [1000]	5	3.05	-1.02 – 11.02	0.103
Age group [1] * Frequency[2000]	12.19	3.05	6.16 – 18.21	**<0.001**
Ear [Right] * Frequency [1000]	1.53	3.08	-4.56 – 7.61	0.621
Ear [Right] * Frequency[2000]	0.59	3.08	-5.49 – 6.67	0.849
(Age group [1] * Ear [Right]) *Frequency [1000]	-1.84	4.33	-10.40 – 6.72	0.672
(Age group [1] * Ear [Right]) *Frequency [2000]	0.35	4.33	-8.21 – 8.91	0.936
Random effects				
σ^2	37.17			
τ_{00}	21.53			
Marginal R^2 / Conditional R^2	0.455	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 1.4 Tukey-adjusted Group difference for each Frequency tested

Frequency	β	SE	df	T	P	d
500 Hz	-3.420	2.25	30	-1.52	0.139	-0.561
1000 Hz	-7.500	2.24	30	-3.349	0.0022	-1.230
2000 Hz	-15.780	2.24	30	-7.047	<.0001	-2.588

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

2. Accuracy

Table 2.1 Estimated marginal means for Accuracy

Cue	Modality	Everyone					Younger					Older				
		MM	SE	df	95% CI		MM	SE	df	95% CI		MM	SE	df	95% CI	
					LL	UL				LL	UL				LL	UL
Control	A	95.2	0.9	29	93.3	97	96.3	1.37	29	93.5	99.1	94.1	1.37	29	91.3	96.9
	V	70.1	1.167	29	67.7	72.5	75	1.73	29	71.5	78.5	65.2	1.73	29	61.7	68.7
	AV	97.2	0.876	29	95.4	99	98.3	1.34	29	95.6	101.1	96	1.34	29	93.3	98.7
What	A	97.4	0.9	29	95.6	99.2	98.6	1.37	29	95.8	101.4	96.2	1.37	29	93.4	99
	V	94	1.167	29	91.6	96.4	97.3	1.73	29	93.8	100.8	90.7	1.73	29	87.2	94.2
	AV	98.1	0.876	29	96.3	99.9	99	1.34	29	96.2	101.7	97.3	1.34	29	94.6	100.1
When	A	96.2	0.9	29	94.3	98	96.9	1.37	29	94.1	99.7	95.5	1.37	29	92.7	98.3
	V	73.7	1.167	29	71.3	76.1	79.9	1.73	29	76.3	83.4	67.5	1.73	29	64	71.1
	AV	97.2	0.876	29	95.4	99	98.1	1.34	29	95.3	100.8	96.3	1.34	29	93.5	99
WW	A	98.2	0.9	29	96.3	100	98.4	1.37	29	95.6	101.2	97.9	1.37	29	95.1	100.7
	V	94	1.167	29	91.6	96.4	96.4	1.73	29	92.8	99.9	91.7	1.73	29	88.1	95.2
	AV	97.5	0.876	29	95.7	99.3	97.7	1.34	29	94.9	100.4	97.4	1.34	29	94.7	100.1

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 2.2 Type III Tests of Fixed Effects for Accuracy (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	330	4598.65	**<.0001**
Age group	1	29	1.095	0.304
Prediction cue	3	330	1.348	0.2587
Modality	2	330	109.03	**<.0001**
Hearing	1	29	0.266	0.6101
Age group * Prediction cue	3	330	0.373	0.7727
Age group * Modality	2	330	6.215	**0.0022**
Prediction cue * Modality	6	330	39.422	**<.0001**
Age group * Prediction cue * Modality	6	330	1.559	0.1583

Note. DF = degrees of freedom. The r model was: Accuracy ~ Age group * Prediction Cue * Modality + PTA, random = ~ Modality | Subject.

Table 2.3 Coefficients, standard errors and CI for the fixed and random effects for Accuracy

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	96.86	1.43	94.05 – 99.67	**<0.001**
Age Group[Older]	-2.17	2.07	-6.40 – 2.07	0.304
Prediction Cue [What]	2.34	1.41	-0.43 – 5.12	0.097
Prediction Cue [When]	0.61	1.41	-2.16 – 3.38	0.667
Prediction Cue [WW]	2.17	1.41	-0.60 – 4.94	0.124
Modality[V]	-21.27	1.66	-24.54 – -18.00	**<0.001**
Modality[AV]	2.08	1.41	-0.69 – 4.86	0.141
Hearing	-0.06	0.12	-0.29 – 0.18	0.61
Age Group[Older] *Prediction Cue [What]	-0.26	1.99	-4.18 – 3.66	0.896
Age Group[Older] *Prediction Cue [When]	0.78	1.99	-3.14 – 4.70	0.695
Age Group[Older] *Prediction Cue [WW]	1.65	1.99	-2.27 – 5.57	0.409
Age Group[Older] *Modality[V]	-7.64	2.35	-12.26 – -3.02	**0.001**
Age Group[Older] *Modality[AV]	-0.17	2	-4.10 – 3.75	0.931
Prediction Cue [What] * Modality[V]	19.96	1.99	16.04 – 23.89	**<0.001**
Prediction Cue [When] * Modality[V]	4.25	1.99	0.33 – 8.17	**0.034**
Prediction Cue [WW] * Modality[V]	19.18	1.99	15.26 – 23.10	**<0.001**
Prediction Cue [What] * Modality[AV]	-1.74	1.99	-5.66 – 2.18	0.384
Prediction Cue [When] * Modality[AV]	-0.87	1.99	-4.79 – 3.05	0.663
Prediction Cue [WW] * Modality[AV]	-2.86	1.99	-6.78 – 1.06	0.152
(Age Group[Older] *Prediction Cue [What]) *Modality[V]	3.47	2.82	-2.07 – 9.02	0.219

(Age Group[Older] *Prediction Cue [When]) *Modality[V]	-3.3	2.82	-8.84 – 2.25	0.243
(Age Group[Older] *Prediction Cue [WW]) * Modality[V]	3.47	2.82	-2.07 – 9.02	0.219
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	0.96	2.82	-4.59 – 6.50	0.735
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	-0.26	2.82	-5.80 – 5.28	0.927
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	0.43	2.82	-5.11 – 5.98	0.878
Random effects				
σ^2	15.88			
τ_{00}	10.05			
$\tau_{11\ AV}$	0.07			
$\tau_{11\ V}$	12.39			
Marginal R^2 / Conditional R^2	0.785	0.874		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 2.4. Tukey-corrected contrasts for Accuracy

Contrast		β	SE	df	T	P	d
A vs. AV		-8.840	3.330	30	-2.659	**0.013**	-0.665
Group difference	**A**	1.62	1.67	29	0.971	0.3394	0.407
	V	8.35	2.24	29	3.734	**0.0008**	2.095
	AV	1.51	1.62	29	0.936	0.3571	0.38
A	**Control vs What**	-2.210	0.996	330	-2.222	0.120	-0.0441
	Control vs When	-0.998	0.996	330	-1.002	0.748	0.2608
	Control vs WW	-2.990	0.996	330	-3.006	**0.015**	-0.2401
	What vs. When	1.220	0.996	330	1.220	0.615	0.8162
	What vs. WW	-0.781	0.996	330	-0.784	0.862	0.3153
	When vs. WW	-2.000	0.996	330	-2.004	0.189	0.0104
V	**Control vs What**	-23.900	0.996	330	-24.002	**<.0001**	-5.4891
	Control vs When	-3.600	0.996	330	-3.615	**0.002**	-0.3926
	Control vs WW	-23.900	0.996	330	-24.002	**<.0001**	-5.4891
	What vs. When	20.300	0.996	330	20.386	**<.0001**	5.6078
	What vs. WW	0.000	0.996	330	0.000	1.000	0.5113
	When vs. WW	-20.300	0.996	330	-20.386	**<.0001**	-4.5852
AV	**Control vs What**	-0.955	0.996	330	-0.958	0.773	0.2717
	Control vs When	0.000	0.996	330	0.000	1.000	0.5113
	Control vs WW	-0.347	0.996	330	-0.349	0.985	0.4242
	What vs. When	0.955	0.996	330	0.958	0.773	0.7509
	What vs. WW	0.608	0.996	330	0.610	0.929	0.6638
	When vs. WW	-0.347	0.996	330	-0.348	0.985	0.4242

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

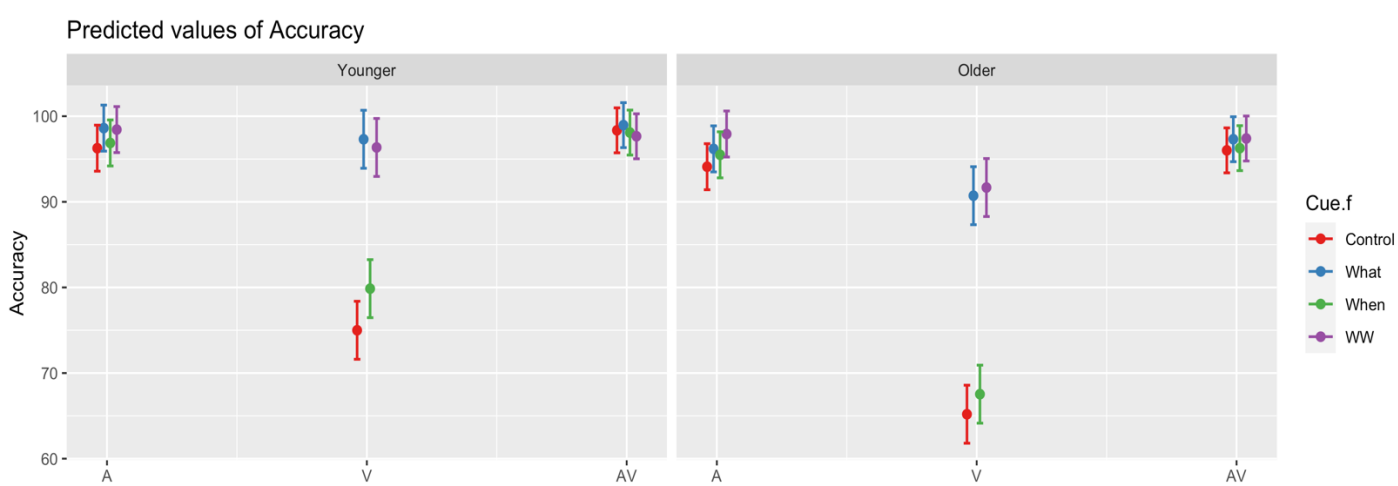

Figure 1 Accuracy Marginal Means

3. Visual Effect Scores (VE)

Table 3.1 Estimated marginal means for VE

Cue	Everyone					Younger					Older				
	MM	SE	df	95% CI		MM	SE	df	95% CI		MM	SE	df	95% CI	
				LL	UL				LL	UL				LL	UL
Control	0.554	0.073	28	0.405	0.703	0.802	0.108	28	0.581	1.022	0.307	0.116	28	0.069	0.545
What	0.295	0.127	28	0.035	0.555	0.380	0.208	28	-0.047	0.806	0.210	0.160	28	-0.118	0.538
When	0.420	0.129	28	0.157	0.683	0.517	0.187	28	0.135	0.899	0.322	0.190	28	-0.068	0.712
WW	-0.167	0.264	28	-0.707	0.374	0.018	0.388	28	-0.776	0.812	-0.351	0.366	28	-1.101	0.399

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 3.2 Type III Tests of Fixed Effects for VE (marginal)

Effects	DF		F	p
(Intercept)	1	47	31.662	<.0001
Age group	1	28	8.469	**0.007**
Prediction cue	3	47	1.872	0.147
Hearing	1	28	1.145	0.294
Age group * Prediction cue	3	47	0.871	0.463

Note. DF = degrees of freedom. The r model was: VE ~ Age group * Prediction Prediction Cue + Hearing, random = ~ Prediction Prediction Cue | Subject.

Table 3.3 Coefficients, standard errors and CI for the fixed and random effects for VE

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	0.68	0.12	0.44 – 0.92	**<0.001**
Age Group[Older]	-0.49	0.17	-0.84 – -0.15	**0.007**
Prediction Cue [What]	-0.42	0.23	-0.88 – 0.04	0.07
Prediction Cue [When]	-0.28	0.19	-0.67 – 0.10	0.145
Prediction Cue [WW]	-0.78	0.4	-1.60 – 0.03	0.059
Hearing	0.01	0.01	-0.01 – 0.03	0.294
Age Group[Older] * Cue[What]	0.33	0.3	-0.28 – 0.93	0.281
Age Group[Older] * Cue[When]	0.3	0.28	-0.26 – 0.85	0.283
Age Group[Older] * Cue[WW]	0.13	0.56	-1.01 – 1.26	0.825
Random effects				
σ^2	0.02			
τ_{00}	0.1			
$\tau_{11\ What}$	0.44			
$\tau_{11\ When}$	0.44			
$\tau_{11\ WW}$	1.7			
Marginal R^2 / Conditional R^2	0.149	0.959		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 3.4 Tukey-adjusted contrasts for VE

Contrast	β	SE	df	T	P	d
Younger vs. Older	0.458	0.293	28	1.562	0.130	2.44

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

4. Prediction Effect Scores (PE)

Table 4.1 Estimated marginal means for PE

Cue	Modality	Everyone					Younger					Older				
		MM	SE	df	95% CI		MM	SE	df	95% CI		MM	SE	df	95% CI	
					LL	UL				LL	UL				LL	UL
What	A	0.474	0.088	29	0.294	0.653	0.794	0.122	29	0.544	1.044	0.154	0.133	29	-0	0.427
	V	0.803	0.076	29	0.647	0.959	0.933	0.112	29	0.703	1.163	0.673	0.113	29	0	0.904
	AV	0.519	0.103	29	0.307	0.73	0.661	0.155	29	0.344	0.979	0.376	0.146	29	0	0.674
When	A	0.225	0.088	29	0.046	0.405	0.366	0.122	29	0.116	0.615	0.085	0.133	29	-0	0.358
	V	0.131	0.076	29	-0.024	0.287	0.224	0.112	29	-0.006	0.454	0.038	0.113	29	-0	0.269
	AV	0.139	0.103	29	-0.072	0.35	0.049	0.155	29	-0.268	0.366	0.229	0.146	29	-0	0.527
WW	A	0.699	0.088	29	0.52	0.879	0.776	0.122	29	0.526	1.025	0.623	0.133	29	0	0.896
	V	0.791	0.076	29	0.636	0.947	0.874	0.112	29	0.644	1.104	0.709	0.113	29	0	0.94
	AV	0.144	0.103	29	-0.067	0.356	-0.089	0.155	29	-0.406	0.229	0.377	0.146	29	0	0.675

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 4.2 Type III Tests of Fixed Effects for PE (marginal)

Predictors	DF		F	p
(Intercept)	1	173	33.741	**<.0001**
Age group	1	29	11.822	**0.002**
Prediction cue	2	173	4.741	**0.010**
Modality	2	173	1.257	0.287
Hearing	1	29	0.540	0.468
Age group * Prediction cue	2	173	2.366	0.097
Age group * Modality	2	173	1.699	0.186
Prediction cue * Modality	4	173	3.646	**0.007**
Age group * Prediction cue * Modality	4	173	1.042	0.387

Note. DF = degrees of freedom. The r model was: PE ~ Age group * Prediction Prediction Cue + PTA, random = ~1 | SID.

Table 4.3 Coefficients, standard errors and CI for the fixed and random effects for PE

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	0.74	0.13	0.49 – 0.99	**<0.001**
Age Group[Older]	-0.64	0.19	-1.02 – -0.26	**0.002**
Prediction Cue [WhenE]	-0.43	0.16	-0.74 – -0.12	**0.007**
Prediction Cue [WWE]	-0.02	0.16	-0.33 – 0.29	0.907
Modality[V]	0.14	0.15	-0.16 – 0.44	0.357
Modality[AV]	-0.13	0.18	-0.50 – 0.23	0.472
Hearing	0.01	0.01	-0.01 – 0.02	0.468
Age Group[Older] * Cue[WhenE]	0.36	0.23	-0.10 – 0.82	0.124
Age Group[Older] * Cue[WWE]	0.49	0.23	0.03 – 0.95	**0.037**
Age Group[Older] *Modality[V]	0.38	0.22	-0.05 – 0.81	0.084
Age Group[Older] *Modality[AV]	0.35	0.26	-0.15 – 0.86	0.171
Prediction Cue [WhenE] *Modality[V]	-0.28	0.21	-0.70 – 0.14	0.188
Prediction Cue [WWE] * Modality[V]	-0.04	0.21	-0.46 – 0.38	0.847
Prediction Cue [WhenE] *Modality[AV]	-0.18	0.25	-0.69 – 0.32	0.471
Prediction Cue [WWE] * Modality[AV]	-0.73	0.25	-1.23 – -0.23	**0.005**
(Age Group[Older] * Cue[WhenE]) * Modality[V]	-0.29	0.31	-0.89 – 0.32	0.353
(Age Group[Older] * Cue[WWE]) * Modality[V]	-0.39	0.31	-1.00 – 0.21	0.203
(Age Group[Older] * Cue[WhenE]) * Modality[AV]	0.11	0.36	-0.61 – 0.82	0.768

(Age Group[Older] * Cue[WWE]) * Modality[AV]	0.26	0.36	-0.45 – 0.98	0.466
Random effects				
σ^2	0.16			
τ_{00}	0.02			
Marginal R^2 / Conditional R^2	0.349	0.436		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 4.4. Tukey-corrected contrasts for PE

	Contrast	β	SE	df	T	P	d
Younger vs. Older		0.147	0.105	29	1.399	0.173	0.366
	A vs. V	-0.109	0.064	171	-1.704	0.207	-0.272
	A vs. AV	0.199	0.076	171	2.606	**0.027**	0.496
	V vs. AV	0.308	0.073	171	4.248	**0.000**	0.768
What	**A vs. V**	-0.329	0.109	171	-3.011	0.072	-0.821
	A vs. AV	-0.045	0.129	171	-0.347	1.000	-0.112
	V vs AV	0.285	0.122	171	2.328	0.332	0.709
When	**A vs. V**	0.094	0.109	171	0.858	0.995	0.234
	A vs. AV	0.086	0.129	171	0.667	0.999	-1.441
	V vs AV	-0.008	0.122	171	-0.064	1.000	-0.019
WW	**A vs. V**	-0.092	0.109	171	-0.842	0.995	-0.230
	A vs. AV	0.555	0.129	171	4.301	**0.001**	1.384
	V vs AV	0.647	0.122	171	5.294	**<.0001**	1.613

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

5. Integration analyses (additive model): P1 amplitude

Table 5.1 Estimated marginal means for P1 amplitude

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	77.7	10.7	29	56	99.5	58.4	17.1	29	23.4	93.4	97.1	16.7	29	63	131.2
A+V	68.2	10.5	29	46.7	89.7	53.7	16.9	29	19.1	88.4	82.6	16.7	29	48.5	116.7

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 5.2 Type III Tests of Fixed Effects for P1 Amplitude (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	29	9.041	**0.005**
Age group	1	29	2.176	0.151
Modality	1	29	0.090	0.767
Hearing	1	29	0.030	0.865
Age group * Modality	1	29	0.202	0.656

Note. DF = degrees of freedom. The r model was: P1_A ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 5.3 Coefficients, standard errors and CI for the fixed and random effects for P1 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	55.3	18.39	17.68 – 92.92	**0.005**
Age Group[Older]	38.68	26.22	-14.95 – 92.31	0.151
Modality[A+V]	-4.67	15.58	-36.53 – 27.19	0.767
Hearing	0.3	1.77	-3.31 – 3.91	0.865
Age Group[Older] *Modality[A+V]	-9.78	21.77	-54.31 – 34.74	0.656
Random effects				
σ^2	1849			
τ_{00}	1690			
Marginal R^2 / Conditional R^2	0.167	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

6. Integration analyses (additive model): P1 latency

Table 6.1 Estimated marginal means for P1 latency

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	51.5	3.83	28	43.6	59.3	58.4	17.1	29	23.4	93.4	97.1	16.7	29	63	131.2
A+V	46.3	3.71	25	38.7	54	53.7	16.9	29	19.1	88.4	82.6	16.7	29	48.5	116.7

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 6.1 Type III Tests of Fixed Effects for P1 latency (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	28	52.946	**<.0001**
Age group	1	28	0.968	0.334
Modality	1	25	0.055	0.816
Hearing	1	28	0.005	0.943
Age group * Modality	1	25	0.566	0.459

Note. DF = degrees of freedom. The r model was: P1_A ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 6.3 Coefficients, standard errors and CI for the fixed and random effects for P1 latency

Predictors	β	SE	95% CI	p
(Intercept)	47.39	6.51	34.05 – 60.74	**<0.001**
Age Group[Older]	9.03	9.18	-9.77 – 27.83	0.334
Modality[A+V]	-1.61	6.87	-15.75 – 12.53	0.816
Hearing	-0.04	0.56	-1.20 – 1.12	0.942
Age Group[Older] *Modality[A+V]	-7.03	9.35	-26.29 – 12.22	0.459
Random effects				
σ2	309.04			
τ00	101.49			
Marginal R2 / Conditional R2	0.052	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

7. Integration analyses (additive model): N1 amplitude

Table 7.1 Estimated marginal means for N1 amplitude

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	-5.38	0.41	29	-6.22	-4.54	-5.95	0.668	29	-7.31	-4.58	-4.81	0.668	29	-6.18	-3.45
A+V	-6.05	0.41	29	-6.89	-5.22	-6.74	0.668	29	-8.1	-5.37	-5.37	0.668	29	-6.74	-4.00

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 7.2 Type III Tests of Fixed Effects for N1 Amplitude (marginal)

	DF		F	p
(Intercept)	1	30	36.466	<.0001
Age group	1	29	1.155	0.291
Modality	1	30	4.198	**0.049**
Hearing	1	29	4.605	**0.040**
Age group * Modality	1	30	0.182	0.672

Note. DF = degrees of freedom. The r model was: N1_A ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 7.3 Coefficients, standard errors and CI for the fixed and random effects for N1 amplitude

Predictors	**β**	**SE**	**95% CI**	**p**
Fixed effects				
(Intercept)	-4.33	0.72	-5.79 – -2.86	**<0.001**
Age Group[Older]	1.13	1.05	-1.02 – 3.29	0.291
Modality[A+V]	-0.79	0.38	-1.57 – -0.00	**0.049**
Hearing	-0.16	0.08	-0.31 – -0.01	**0.040**
Age Group[Older] *Modality[A+V]	0.23	0.54	-0.88 – 1.34	0.672
Random effects				
σ^2	1.19			
τ_{00}	4.18			
Marginal R^2 / Conditional R^2	0.41	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 7.4 Decomposition of the Modality main effect for N1 amplitude

Contrast	**β**	**SE**	**df**	**T**	**P**	**d**
AV vs. A+V	0.672	0.272	30	2.47	0.0194	0.618

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

8. Integration EEG analyses (additive model): N1 latency

Table 8.1 Estimated marginal means for N1 latency

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	114	2.17	29	110	119	114	3.45	29	107	121	115	3.45	29	108	122
A+V	120	2.17	29	115	124	122	3.45	29	115	130	117	3.45	29	110	124

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 8.2 Type III Tests of Fixed Effects for N1 latency (marginal)

	DF		F	p
(Intercept)	1	30	865.232	**<.0001**
Age group	1	29	0.075	0.787
Modality	1	30	7.169	**0.012**
Hearing	1	29	2.555	0.121
Age group * Modality	1	30	1.923	0.176

Note. DF = degrees of freedom. The r model was: N1_L ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 8.3 Coefficients, standard errors and CI for the fixed and random effects for N1 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	107.86	3.67	100.38 – 115.35	**<0.001**
Age Group[Older]	1.47	5.37	-9.51 – 12.44	0.787
Modality[A+V]	8.87	3.31	2.11 – 15.64	**0.012**
Hearing	0.57	0.36	-0.16 – 1.30	0.121
GroupName [Younger] *Modality [AV]	-6.5	4.69	-16.07 – 3.07	0.176
Random effects				
σ^2	87.89			
τ_{00}	63.23			
Marginal R^2 / Conditional R^2	0.201	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 8.4. Decomposition of the Modality main effect for N1 latency

Contrast	β	SE	df	T	P	d
AV vs. A+V	-5.62	2.34	30	-2.4	0.0228	-0.6

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

9. Integration EEG analyses (additive model): P2 amplitude

Table 9.1 Estimated marginal means for P2 amplitude

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	4.03	0.298	29	3.42	4.64	4.6	0.478	29	3.62	5.58	3.46	0.478	29	2.48	4.43
A+V	5.57	0.298	29	4.96	6.18	6.56	0.478	29	5.59	7.54	4.57	0.478	29	3.59	5.55

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 9.2 Type III Tests of Fixed Effects for P2 amplitude (marginal)

	DF		F	p
(Intercept)	1	30	81.465	**<.0001**
Age group	1	29	2.348	0.136
Modality	1	30	23.039	**<.0001**
Hearing	1	29	0.000	0.995
Age group * Modality	1	30	2.147	0.153

Note. DF = degrees of freedom. The r model was: P2_A ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 9.3 Coefficients, standard errors and CI for the fixed and random effects for P2 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	4.60	0.51	3.56 – 5.64	**<0.001**
Age Group[Older]	-1.14	0.75	-2.67 – 0.38	0.136
Modality[A+V]	1.96	0.41	1.13 – 2.80	**<0.001**
Hearing	0.00	0.05	-0.10 – 0.10	1.00
Age Group[Older] *Modality[A+V]	-0.85	0.58	-2.03 – 0.33	0.15
Random effects				
σ^2	1.34			
τ_{00}	1.51			
Marginal R^2 / Conditional R^2	0.308	0.675		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 9.4 Decomposition of the Modality main effect for P2 amplitude

Contrast	β	SE	df	T	P	d
AV vs. A+V	-1.54	0.289	30	-5.323	**<.0001**	-1.33

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

10. Integration EEG analyses (additive model): P2 latency

Table 10.1 Estimated marginal means for P2 latency

	Everyone					Younger					Older				
Modality	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
				LL	**UL**				**LL**	**UL**				**LL**	**UL**
AV	203	3.44	29	196	210	194	5.52	29	182	205	212	5.52	29	201	223
A+V	212	3.44	29	205	219	204	5.52	29	193	215	219	5.52	29	208	230

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 10.2 Type III Tests of Fixed Effects for P2 latency (marginal)

	DF		F	p
(Intercept)	1	30	1037.684	**<.0001**
Age group	1	29	4.487	**0.043**
Modality	1	30	4.983	**0.033**
Hearing	1	29	0.499	0.485
Age group * Modality	1	30	0.248	0.622

Note. DF = degrees of freedom. The r model was: P2_L ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 10.3 Coefficients, standard errors and CI for the fixed and random effects for P2 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	189.48	5.88	177.47 – 201.50	**<0.001**
Age Group[Older]	18.27	8.62	0.63 – 35.91	**0.043**
Modality[A+V]	10.5	4.7	0.89 – 20.11	**0.033**
Hearing	0.41	0.59	-0.78 – 1.61	0.485
Age Group[Older] *Modality[A+V]	-3.31	6.65	-16.90 – 10.27	0.622
Random effects				
σ^2	177.01			
τ_{00}	202.65			
Marginal R^2 / Conditional R^2	0.423	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 10.4 Decomposition of the Modality main effect for P2 latency

Contrast	β	SE	df	T	P	d
AV vs. A+V	-8.840	3.330	30	-2.659	0.013	-0.665
Younger vs. Older	-16.600	7.960	29	-2.088	0.046	-1.250

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

11. Integration EEG analyses (additive model): N2 amplitude

Table 11.1 Estimated marginal means for N2 amplitude

Modality	Everyone					Younger					Older				
	MM	SE	df	95% CI		MM	SE	df	95% CI		MM	SE	df	95% CI	
				LL	UL				LL	UL				LL	UL
AV	-2.93	0.324	29	-3.6	-2.27	-3.27	0.526	29	-4.35	-2.2	-2.59	0.526	29	-3.67	-1.52
A+V	-4.27	0.324	29	-4.94	-3.61	-4.96	0.526	29	-6.04	-3.89	-3.58	0.526	29	-4.66	-2.51

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 11.2 Type III Tests of Fixed Effects for N2 amplitude (marginal)

	DF		F	p
(Intercept)	1	30	35.387	**<.0001**
Age group	1	29	0.680	0.416
Modality	1	30	22.341	**0.0001**
Hearing	1	29	0.017	0.898
Age group * Modality	1	30	1.897	0.179

Note. DF = degrees of freedom. The r model was: N2_A ~ Age group * Modality + PTA, random = ~ 1 | Subject.

Table 11.3 Coefficients, standard errors and CI for the fixed and random effects for N2 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	-3.35	0.56	-4.50 – -2.20	**<0.001**
Age Group[Older]	0.68	0.83	-1.01 – 2.37	0.416
Modality[A+V]	-1.69	0.36	-2.42 – -0.96	**<0.001**
Hearing	0.01	0.06	-0.11 – 0.13	0.898
Age Group[Older] *Modality[A+V]	0.7	0.51	-0.34 – 1.73	0.179
Random effects				
σ^2	1.02			
τ_{00}	2.34			
Marginal R^2 / Conditional R^2	0.437	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 11.4 Decomposition of the Modality main effect for N2 amplitude

Contrast	β	SE	df	T	P	d
AV vs. A+V	1.340	0.253	30	5.307	<.0001	1.330

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

12. Integration EEG analyses (additive model): N2 latency

Table 12.1 Estimated marginal means for N2 latency

Modality	Everyone					Younger					Older				
	MM	SE	df	95% CI		MM	SE	df	95% CI		MM	SE	df	95% CI	
				LL	UL				LL	UL				LL	UL
AV	298	3.96	29	290	306	286	6.46	29	273	299	311	6.46	29	297	324
A+V	301	6.91	29	287	315	297	10.28	29	276	318	304	10.28	29	283	325

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 12.2 Type III Tests of Fixed Effects for N2 latency (marginal)

	DF		F	p
(Intercept)	1	30	1672.254	**<.0001**
Age group	1	29	5.872	**0.022**
Modality	1	30	1.326	0.259
Hearing	1	29	0.080	0.780
Age group * Modality	1	30	1.618	0.213

Note. DF = degrees of freedom. The r model was: N2_L ~ Age group * Modality + PTA, random = ~ Modality | Subject.

Table 12.3 Coefficients, standard errors and CI for the fixed and random effects for N2 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	283.72	6.94	269.55 – 297.89	**<0.001**
Age Group[Older]	24.74	10.21	3.86 – 45.63	**0.022**
Modality[A+V]	11.56	10.04	-8.94 – 32.07	0.259
Hearing	0.21	0.73	-1.28 – 1.69	0.78
Age Group[Older] *Modality[A+V]	-18.06	14.2	-47.06 – 10.94	0.213
Random effects				
σ^2	141.85			
τ_{00}	360.45			
$\tau_{11\ A+V}$	1329.49			
Marginal R^2 / Conditional R^2	0.091			

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 12.4 Decomposition of the Modality main effect for N2 latency

Contrast	β	SE	df	T	P	d
Younger vs. Older	-15.7	10.9	29	-1.448	0.1584	-1.32

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

13. Temporal and Phonetic Visual Predictive Cue analysis: P1 Amplitude

Table 13.1 Estimated marginal means for P1 amplitude

Conditions		Everyone					Younger					Older				
					95% CI					95% CI					95% CI	
Cue	Modality	MM	SE	df	LL	UL	MM	SE	df	LL	UL	MM	SE	df	LL	UL
Control	A	72.7	11.8	29	48.5	96.8	49.4	18.3	29	12	86.9	95.9	18.3	29	58	133.4
	AV	83.4	11.8	29	59.2	107.5	60.3	18.3	29	22.9	97.8	99.8	18.3	29	62	137.2
What	A	62.4	11.8	29	38.3	86.6	71.3	18.3	29	33.8	108.7	95.4	18.3	29	58	132.9
	AV	79.8	11.8	29	55.7	104	79.3	18.3	29	41.8	116.7	131.2	18.3	29	94	168.6
When	A	80	11.8	29	55.9	104.2	51.1	18.3	29	13.6	88.5	73.7	18.3	29	36	111.2
	AV	105.2	11.8	29	81.1	129.4	58.3	18.3	29	20.8	95.7	106.8	18.3	29	69	144.2
WW	A	82.5	11.8	29	58.4	106.7	60.3	18.3	29	22.9	97.8	99.4	18.3	29	62	136.8
	AV	102.1	11.8	29	77.9	126.2	64.4	18.3	29	27	101.9	139.7	18.3	29	102	177.2

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 13.2 Type III Tests of Fixed Effects for P1 Amplitude (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	210	*7.478*	**0.007**
Age group	1	29	2.753	0.108
Prediction cue	3	210	0.703	0.551
Modality	1	210	0.412	0.522
Hearing	1	29	0.034	0.854
Age group * Prediction cue	3	210	0.469	0.704
Age group * Modality	1	210	0.086	0.770
Prediction cue * Modality	3	210	0.027	0.994
Age group * Prediction cue * Modality	3	210	0.631	0.596

Note. DF = degrees of freedom. The r model was: P1_A ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~1 | SID.

Table 13.3 Coefficients, standard errors and CI for the fixed and random effects for P1 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	52.62	19.24	14.69 – 90.55	**0.007**
Age Group[Older]	46.46	28	-10.81 – 103.74	0.108
Prediction Cue [What]	21.84	16.94	-11.56 – 55.25	0.199
Prediction Cue [When]	1.64	16.94	-31.76 – 35.04	0.923
Prediction Cue [WW]	10.87	16.94	-22.54 – 44.27	0.522
Modality[AV]	10.88	16.94	-22.53 – 44.28	0.522
Hearing	-0.32	1.70	-3.79 – 3.16	0.854
Age Group[Older] *Prediction Cue [What]	-22.33	23.96	-69.57 – 24.91	0.352
Age Group[Older] *Prediction Cue [When]	-23.82	23.96	-71.06 – 23.42	0.321
Age Group[Older] *Prediction Cue [WW]	-7.41	23.96	-54.65 – 39.83	0.757
Age Group[Older] *Modality[AV]	-7.01	23.96	-54.25 – 40.23	0.770
Prediction Cue [What] * Modality[AV]	-2.91	23.96	-50.15 – 44.33	0.904
Prediction Cue [When] * Modality[AV]	-3.71	23.96	-50.95 – 43.53	0.877
Prediction Cue [WW] * Modality[AV]	-6.78	23.96	-54.02 – 40.46	0.778
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	34.77	33.89	-32.03 – 101.58	0.306
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	32.9	33.89	-33.91 – 99.71	0.333
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	43.26	33.89	-23.54 – 110.07	0.203
Random effects				
σ^2	2296.96			
τ_{00}	2163.36			
Marginal R^2 / Conditional R^2	0.222	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

14. Temporal and Phonetic Visual Predictive Cue analysis: P1 latency

Table 14.1 Estimated marginal means for P1 latency

Conditions		Everyone					Younger					Older				
Cue	Modality	MM	SE	df	95% CI LL	95% CI UL	MM	SE	df	95% CI LL	95% CI UL	MM	SE	df	95% CI LL	95% CI UL
Control	A	45.8	3.57	29	38.5	53.1	45.2	5.65	29	33.7	56.8	46.4	5.54	29	35	57.7
	AV	49	3.63	29	41.6	56.4	43.3	5.73	29	31.6	55	54.7	5.63	29	43.1	66.2
What	A	42.4	3.57	29	35.1	49.7	35.1	5.65	29	23.5	46.6	49.8	5.54	29	38.4	61.1
	AV	42.2	3.57	29	34.9	49.5	39.6	5.65	29	28.1	51.2	44.8	5.54	29	33.5	56.1
When	A	45.5	3.67	29	37.9	53	41.7	5.65	29	30.1	53.3	49.2	5.78	29	37.4	61
	AV	44.9	3.57	29	37.6	52.2	40.8	5.65	29	29.3	52.4	49	5.54	29	37.7	60.4
WW	A	45.1	3.57	29	37.8	52.4	39.2	5.58	29	27.8	50.7	51	5.63	29	39.5	62.5
	AV	43.6	3.54	29	36.3	50.8	38.8	5.58	29	27.4	50.2	48.4	5.54	29	37.1	59.7

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 14.2 Type III Tests of Fixed Effects for P1 latency (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	198	61.798	**<.0001**
Age group	1	29	0.017	0.896
Prediction cue	3	198	1.579	0.196
Modality	1	198	0.150	0.699
Hearing	1	29	0.075	0.786
Age group * Prediction cue	3	198	1.540	0.205
Age group * Modality	1	198	2.223	0.138
Prediction cue * Modality	3	198	0.356	0.785
Age group * Prediction cue * Modality	3	198	1.448	0.230

Note. DF = degrees of freedom. The r model was: P1_L ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~1 | SID.

Table 14.3 Coefficients, standard errors and CI for the fixed and random effects for P1 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	46.74	5.95	35.01 – 58.46	**<0.001**
Age Group[Older]	1.14	8.61	-16.48 – 18.76	0.896
Prediction Cue [What]	-10.16	4.81	-19.64 – -0.67	**0.036**
Prediction Cue [When]	-3.52	4.81	-13.00 – 5.97	0.465
Prediction Cue [WW]	-5.98	4.75	-15.34 – 3.39	0.21
Modality[AV]	-1.91	4.94	-11.65 – 7.83	0.699
Hearing	-0.15	0.54	-1.26 – 0.96	0.786
Age Group[Older] *Prediction Cue [What]	13.56	6.69	0.36 – 26.76	**0.044**
Age Group[Older] *Prediction Cue [When]	6.36	6.9	-7.25 – 19.98	0.358
Age Group[Older] *Prediction Cue [WW]	10.6	6.71	-2.65 – 23.84	0.116
Age Group[Older] *Modality[AV]	10.21	6.85	-3.29 – 23.72	0.138
Prediction Cue [What] * Modality[AV]	6.47	6.89	-7.12 – 20.06	0.349
Prediction Cue [When] * Modality[AV]	1.05	6.87	-12.49 – 14.60	0.878
Prediction Cue [WW] * Modality[AV]	1.42	6.79	-11.96 – 14.80	0.835
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	-19.74	9.58	-38.63 – -0.86	**0.041**
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	-9.53	9.72	-28.70 – 9.64	0.328
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	-12.29	9.56	-31.14 – 6.56	0.2
Random effects				
σ^2	173.44			
τ_{00}	227.84			
Marginal R^2 / Conditional R^2	0.113	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

15. Temporal and Phonetic Visual Predictive Cue analysis: N1 amplitude

Table 15.1 Estimated marginal means for N1 amplitude

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI LL**	**95% CI UL**	**MM**	**SE**	**df**	**95% CI LL**	**95% CI UL**	**MM**	**SE**	**df**	**95% CI LL**	**95% CI UL**
Control	**A**	-5.97	0.396	29	-6.78	-5.16	-6.75	0.643	29	-8.07	-5.44	-5.18	0.643	29	-6.49	-3.86
	V	-5.38	0.422	29	-6.24	-4.52	-5.8	0.675	29	-7.18	-4.42	-4.96	0.675	29	-6.34	-3.58
What	**A**	-5.22	0.396	29	-6.03	-4.41	-5.66	0.643	29	-6.98	-4.35	-4.78	0.643	29	-6.09	-3.46
	V	-5.03	0.422	29	-5.89	-4.17	-5.3	0.675	29	-6.68	-3.92	-4.76	0.675	29	-6.14	-3.38
When	**A**	-5.61	0.396	29	-6.42	-4.8	-6.03	0.643	29	-7.34	-4.71	-5.19	0.643	29	-6.5	-3.87
	V	-5.08	0.422	29	-5.95	-4.22	-5.32	0.675	29	-6.7	-3.94	-4.85	0.675	29	-6.23	-3.47
WW	**A**	-5.15	0.396	29	-5.96	-4.34	-5.56	0.643	29	-6.88	-4.25	-4.73	0.643	29	-6.04	-3.41
	V	-5.07	0.422	29	-5.93	-4.2	-5.78	0.675	29	-7.16	-4.4	-4.35	0.675	29	-5.73	-2.97

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 15.2 Type III Tests of Fixed Effects for N1 amplitude (marginal)

Effects and interactions	DF		F	p
(Intercept)	1	210	62.839	**<.0001**
Age group	1	29	2.417	0.131
Prediction cue	3	210	4.913	**0.003**
Modality	1	210	6.451	**0.012**
Hearing	1	29	3.251	0.082
Age group * Prediction cue	3	210	1.107	0.347
Age group * Modality	1	210	1.904	0.169
Prediction cue * Modality	3	210	2.163	0.094
Age group * Prediction cue * Modality	3	210	1.353	0.258

Note. DF = degrees of freedom. The r model was: N1_A ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~ Modality | Subject.

Table 15.3 Coefficients, standard errors and CI for the fixed and random effects for N1 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	-5.46	0.69	-6.82 – -4.10	**<0.001**
Age Group[Older]	1.57	1.01	-0.50 – 3.65	0.131
Prediction Cue [What]	1.09	0.34	0.41 – 1.77	**0.002**
Prediction Cue [When]	0.73	0.34	0.05 – 1.41	**0.036**
Prediction Cue [WW]	1.19	0.34	0.51 – 1.87	**0.001**
Modality[AV]	0.95	0.37	0.21 – 1.69	**0.012**
Hearing	-0.13	0.07	-0.27 – 0.02	0.082
Age Group[Older] *Prediction Cue [What]	-0.69	0.49	-1.64 – 0.27	0.16
Age Group[Older] *Prediction Cue [When]	-0.74	0.49	-1.70 – 0.22	0.131
Age Group[Older] *Prediction Cue [WW]	-0.74	0.49	-1.70 – 0.22	0.128
Age Group[Older] *Modality[AV]	-0.73	0.53	-1.77 – 0.31	0.169
Prediction Cue [What] * Modality[AV]	-0.58	0.49	-1.54 – 0.37	0.231
Prediction Cue [When] * Modality[AV]	-0.24	0.49	-1.20 – 0.71	0.616
Prediction Cue [WW] * Modality[AV]	-1.17	0.49	-2.13 – -0.21	**0.017**
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	0.38	0.69	-0.98 – 1.73	0.585
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	0.37	0.69	-0.99 – 1.72	0.596
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	1.32	0.69	-0.03 – 2.68	0.056
Random effects				
σ^2	0.95			
$\tau_{00\ SID}$	4.07			
$\tau_{11\ AV}$	0.34			
Marginal R^2 / Conditional R^2	0.097	0.84		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 15.4 Tukey-adjusted contrasts for N1 amplitude

Contrast	β	SE	df	T	P	d
A vs. AV	-0.344	0.16	210	-2.154	**0.0324**	-0.353
Control vs. What	-0.5469	0.172	210	-3.179	**0.0092**	-0.5619
Control vs. When	-0.328	0.172	210	-1.906	0.2285	-0.337
Control vs. WW	-0.5672	0.172	210	-3.297	**0.0063**	-0.5828
What vs. When	0.2189	0.172	210	1.272	0.5815	0.2249
What vs. WW	-0.0203	0.172	210	-0.118	0.9994	-0.0209
When vs. WW	-0.2392	0.172	210	-1.39	0.5067	-0.2458

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

16. Temporal and Phonetic Visual Predictive Cue analysis: N1 latency

Table 16.1. Estimated marginal means for N1 latency

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
					LL	**UL**				**LL**	**UL**				**LL**	**UL**
Control	A	118	2.15	29	114	123	118	3.25	29	111	125	119	3.25	29	112	125
	V	114	1.95	29	110	118	113	2.99	29	107	119	116	2.99	29	110	122
What	A	118	2.15	29	114	123	120	3.25	29	113	127	117	3.25	29	110	124
	V	119	1.95	29	115	123	119	2.99	29	113	125	119	2.99	29	113	125
When	A	117	2.15	29	113	122	121	3.25	29	115	128	113	3.25	29	107	120
	V	114	1.95	29	110	118	116	2.99	29	110	122	112	2.99	29	106	118
WW	A	117	2.15	29	112	121	120	3.25	29	113	126	114	3.25	29	107	120
	V	114	1.95	29	110	118	116	2.99	29	109	122	113	2.99	29	107	119

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 16.2 Type III Tests of Fixed Effects for N1 latency (marginal)

	DF		F	p
(Intercept)	1	210	1151.447	**<.0001**
Age group	1	29	0.011	0.917
Prediction cue	3	210	0.326	0.807
Modality	1	210	2.703	0.102
Hearing	1	29	2.058	0.162
Age group * Prediction cue	3	210	1.336	0.264
Age group * Modality	1	210	0.338	0.562
Prediction cue * Modality	3	210	0.414	0.743
Age group * Prediction cue * Modality	3	210	0.016	0.997

Note. DF = degrees of freedom. The r model was: N1_L ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~ Modality | Subject.

Table 16.3 Coefficients, standard errors and CI for the fixed and random effects for N1 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	114.34	3.37	107.70 – 120.99	**<0.001**
Age Group[Older]	0.52	4.87	-9.45 – 10.48	0.916
Prediction Cue [What]	1.81	3.23	-4.56 – 8.19	0.576
Prediction Cue [When]	3.19	3.23	-3.19 – 9.56	0.325
Prediction Cue [WW]	1.69	3.23	-4.69 – 8.06	0.602
Modality[AV]	-5.37	3.27	-11.82 – 1.07	0.102
Hearing	0.37	0.26	-0.16 – 0.90	0.162
Age Group[Older] *Prediction Cue [What]	-3.5	4.57	-12.52 – 5.52	0.445
Age Group[Older] *Prediction Cue [When]	-8.5	4.57	-17.52 – 0.52	0.065
Age Group[Older] *Prediction Cue [WW]	-6.69	4.57	-15.70 – 2.33	0.145
Age Group[Older] *Modality[AV]	2.69	4.62	-6.43 – 11.80	0.562
Prediction Cue [What] * Modality[AV]	4.37	4.57	-4.64 – 13.39	0.34
Prediction Cue [When] * Modality[AV]	-0.06	4.57	-9.08 – 8.95	0.989
Prediction Cue [WW] * Modality[AV]	1.19	4.57	-7.83 – 10.20	0.795
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	0.31	6.47	-12.44 – 13.06	0.962
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	1.13	6.47	-11.63 – 13.88	0.862
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	1.12	6.47	-11.63 – 13.88	0.862
Random effects				
σ^2	83.68			
τ_{00}	63.88			
$\tau_{11\ AV}$	3.66			
Marginal R^2 / Conditional R^2	0.073	0.43		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

17. Temporal and Phonetic Visual Predictive Cue analysis: P2 Amplitude

Table 17.1 Estimated marginal means for P2 amplitude

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
					LL	**UL**				**LL**	**UL**				**LL**	**UL**
Control	A	5.04	0.286	29	4.45	5.62	6.2	0.453	29	5.3	7.12	3.88	0.453	29	3	5
	V	4.03	0.277	29	3.46	4.59	4.59	0.441	29	3.7	5.49	3.47	0.441	29	3	4
What	A	4.45	0.286	29	3.87	5.04	5.23	0.453	29	4.3	6.16	3.67	0.453	29	3	5
	V	3.72	0.277	29	3.16	4.29	4.19	0.441	29	3.3	5.09	3.26	0.441	29	2	4
When	A	4.18	0.286	29	3.6	4.77	5.06	0.453	29	4.1	5.98	3.31	0.453	29	2	4
	V	3.91	0.277	29	3.34	4.47	5.06	0.441	29	4.2	5.96	2.76	0.441	29	2	4
WW	A	3.86	0.286	29	3.28	4.45	4.7	0.453	29	3.8	5.62	3.03	0.453	29	2	4
	V	3.46	0.277	29	2.9	4.03	4.22	0.441	29	3.3	5.12	2.71	0.441	29	2	4

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 17.2 Type III Tests of Fixed Effects for P2 Amplitude (marginal)

	DF		F	p
(Intercept)	1	210	168.46	**<.0001**
Age group	1	29	10.96	**0.003**
Prediction cue	3	210	10.30	**<.0001**
Modality	1	210	20.25	**<.0001**
Hearing	1	29	0.00	0.955
Age group * Prediction cue	3	210	1.45	0.231
Age group * Modality	1	210	5.64	**0.018**
Prediction cue * Modality	3	210	6.12	**0.001**
Age group * Prediction cue * Modality	3	210	3.47	**0.017**

Note. DF = degrees of freedom. The r model was: P2_A ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~ Modality | Subject.

Table 17.3 Coefficients, standard errors and CI for the fixed and random effects for P2 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	6.22	0.48	5.28 – 7.17	**<0.001**
Age Group[Older]	-2.32	0.70	-3.76 – -0.89	**0.002**
Prediction Cue [What]	-0.96	0.28	-1.52 – -0.41	**0.001**
Prediction Cue [When]	-1.14	0.28	-1.70 – -0.58	**<0.001**
Prediction Cue [WW]	-1.5	0.28	-2.06 – -0.95	**<0.001**
Modality[AV]	-1.61	0.36	-2.32 – -0.91	**<0.001**
Hearing	0	0.05	-0.10 – 0.09	0.955
Age Group[Older] *Prediction Cue [What]	0.76	0.40	-0.03 – 1.55	0.059
Age Group[Older] *Prediction Cue [When]	0.58	0.40	-0.21 – 1.36	0.152
Age Group[Older] *Prediction Cue [WW]	0.65	0.40	-0.13 – 1.44	0.103
Age Group[Older] *Modality[AV]	1.2	0.51	0.20 – 2.20	**0.018**
Prediction Cue [What] * Modality[AV]	0.56	0.40	-0.23 – 1.35	0.161
Prediction Cue [When] * Modality[AV]	1.61	0.40	0.83 – 2.40	**<0.001**
Prediction Cue [WW] * Modality[AV]	1.14	0.40	0.35 – 1.93	**0.005**
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	-0.57	0.57	-1.68 – 0.55	0.317
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	-1.76	0.57	-2.87 – -0.64	**0.002**
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	-1.05	0.57	-2.17 – 0.07	0.065
σ^2	0.64			
$\tau_{00\ SID}$	1.98			
$\tau_{11\ AV}$	0.77			
Marginal R^2 / Conditional R^2	0.275	0.817		
σ^2	0.64			

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 17.4 Tukey-adjusted contrasts for main effects and 2-way interactions for P2 amplitude

Contrast	β	SE	df	T	P	d
Younger vs. Older	1.64	0.622	29	2.643	**0.013**	2.050
A vs. AV	0.604	0.185	210	3.268	**0.001**	0.754
Control vs. What	0.4447	0.141	210	3.143	**0.010**	0.556
Control vs. When	0.4869	0.141	210	3.442	**0.004**	0.608
Control vs. WW	0.87	0.141	210	6.15	**<.0001**	1.087
What vs. When	0.0422	0.141	210	0.298	0.991	0.053
What vs. WW	0.4253	0.141	210	3.006	**0.016**	0.532
When vs. WW	0.3831	0.141	210	2.708	**0.037**	0.479
Younger A vs AV	0.783	0.261	210	2.998	**0.016**	0.979
Older A vs. AV	0.424	0.261	210	1.623	0.3681	0.53
Control: A vs. AV	1.0103	0.253	210	3.989	**0.0023**	1.2625
What: A vs. AV	0.7309	0.253	210	2.886	0.0806	0.9134
When: A vs. AV	0.2747	0.253	210	1.085	0.9594	0.3432
WW: A vs. AV	0.3984	0.253	210	1.573	0.7658	0.4979

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

Table 17.5 Tukey-adjusted contrasts for the 3-way interaction between Group, Prediction Cue and Modality for P2 amplitude

Contrast		Younger						Older					
Cue	Modality	**β**	**SE**	**df**	**T**	**P**	**d**	**β**	**SE**	**df**	**T**	**P**	**d**
Control	A vs. AV	1.612	0.358	210	4.5	**0.000**	2.014	0.409	0.358	210	1.141	0.947	0.511
What		1.049	0.358	210	2.928	0.072	1.311	0.413	0.358	210	1.153	0.944	0.516
When		-0.003	0.358	210	-0.007	1.000	-0.003	0.552	0.358	210	1.541	0.784	0.690
WW		0.475	0.358	210	1.326	0.888	0.594	0.322	0.358	210	0.899	0.986	0.402

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

18. Temporal and Phonetic Visual Predictive Cue analysis: P2 Latency

Table 18.1 Estimated marginal means for P2 latency

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
					LL	**UL**				**LL**	**UL**				**LL**	**UL**
Control	A	208	3.380	29	202	215	196	5.38	29	185	207	221	5.38	29	210	232
	V	203	3.940	29	195	211	193	6.1	29	181	206	212	6.1	29	200	225
What	A	209	3.380	29	202	216	199	5.38	29	188	210	218	5.38	29	207	229
	V	208	3.940	29	200	216	201	6.1	29	189	214	214	6.1	29	201	226
When	A	211	3.380	29	204	218	198	5.38	29	187	209	225	5.38	29	214	236
	V	198	3.94	29	190	206	193	6.1	29	181	205	203	6.1	29	191	216
WW	A	209	3.38	29	203	216	200	5.38	29	189	211	219	5.38	29	208	230
	V	206	3.94	29	198	214	196	6.1	29	183	208	217	6.1	29	205	230

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 18.2 Type III Tests of Fixed Effects for P2 Latency (marginal)

	DF		F	p
(Intercept)	1	210	1145.129	**<.0001**
Age group	1	29	8.439	**0.007**
Prediction cue	3	210	0.322	0.810
Modality	1	210	0.441	0.507
Hearing	1	29	0.271	0.607
Age group * Prediction cue	3	210	1.046	0.373
Age group * Modality	1	210	0.529	0.468
Prediction cue * Modality	3	210	0.511	0.675
Age group * Prediction cue * Modality	3	210	1.576	0.196

Note. DF = degrees of freedom. The r model was: P2_L ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~ Modality | Subject, weights = varIdent(form = ~ 1 | Modality).

Table 18.3 Coefficients, standard errors and CI for the fixed and random effects for P2 Latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	193.37	5.71	182.10 – 204.63	**<0.001**
Age Group[Older]	24.31	8.37	7.19 – 41.42	**0.007**
Prediction Cue [What]	2.75	4.04	-5.21 – 10.71	0.496
Prediction Cue [When]	1.31	4.04	-6.64 – 9.27	0.745
Prediction Cue [WW]	3.69	4.04	-4.27 – 11.64	0.362
Modality[AV]	-3.19	4.8	-12.65 – 6.27	0.507
Hearing	0.29	0.56	-0.85 – 1.43	0.607
Age Group[Older] *Prediction Cue [What]	-5.19	5.71	-16.44 – 6.06	0.364
Age Group[Older] *Prediction Cue [When]	3.06	5.71	-8.19 – 14.31	0.592
Age Group[Older] *Prediction Cue [WW]	-5.37	5.71	-16.63 – 5.88	0.347
Age Group[Older] *Modality[AV]	-4.94	6.79	-18.32 – 8.44	0.468
Prediction Cue [What] * Modality[AV]	5.56	6.46	-7.18 – 18.30	0.39
Prediction Cue [When] * Modality[AV]	-1.44	6.46	-14.18 – 11.30	0.824
Prediction Cue [WW] * Modality[AV]	-1.12	6.46	-13.86 – 11.61	0.862
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	-1.63	9.14	-19.64 – 16.39	0.859
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	-12.19	9.14	-30.20 – 5.83	0.184
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	7.5	9.14	-10.52 – 25.52	0.413
Random effects				
σ^2	130.3			
τ_{00}	234.61			
$\tau_{11\ AV}$	34.37			
Marginal R^2 / Conditional R^2	0.267	0.758		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 18.4 Tukey-adjusted contrasts for P2 Latency

Contrast	β	SE	df	T	P	d
Younger vs. Older	-19.200	7.68	29	-2.497	0.018	-1.680

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

19. Temporal and Phonetic Visual Predictive Cue analysis: N2 amplitude

Table 19.1 Estimated marginal means for N2 amplitude

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI LL**	**UL**	**MM**	**SE**	**df**	**95% CI LL**	**UL**	**MM**	**SE**	**df**	**95% CI LL**	**UL**
Control	A	-3.81	0.34	29	-4.50	-3.12	-4.45	0.531	29	-5.54	-3.37	-3.16	0.53	29	-4.25	-2.08
	V	-2.93	0.31	29	-3.57	-2.30	-3.16	0.495	29	-4.17	-2.15	-2.7	0.50	29	-3.72	-1.69
What	A	-3.05	0.34	29	-3.75	-2.36	-3.56	0.531	29	-4.64	-2.47	-2.55	0.53	29	-3.64	-1.46
	V	-2.34	0.31	29	-2.97	-1.70	-2.66	0.495	29	-3.67	-1.65	-2.01	0.50	29	-3.02	-1.00
When	A	-3.21	0.34	29	-3.90	-2.52	-3.27	0.531	29	-4.36	-2.19	-3.15	0.53	29	-4.24	-2.07
	V	-2.47	0.31	29	-3.10	-1.84	-2.85	0.495	29	-3.86	-1.83	-2.09	0.50	29	-3.11	-1.08
WW	A	-2.72	0.34	29	-3.41	-2.03	-3.14	0.531	29	-4.23	-2.06	-2.29	0.53	29	-3.38	-1.21
	V	-2.41	0.31	29	-3.04	-1.78	-2.81	0.495	29	-3.82	-1.79	-2.02	0.50	29	-3.03	-1.00

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 19.2 Type III Tests of Fixed Effects for N2 Amplitude (marginal)

	DF		F	p
(Intercept)	1	210	72.734	**<.0001**
Age group	1	29	2.476	0.127
Prediction cue	3	210	7.503	**0.0001**
Modality	1	210	9.170	**0.003**
Hearing	1	29	0.402	0.531
Age group * Prediction cue	3	210	2.674	**0.048**
Age group * Modality	1	210	1.906	0.169
Prediction cue * Modality	3	210	2.117	0.099
Age group * Prediction cue * Modality	3	210	2.029	0.111

Note. DF = degrees of freedom. The r model was: N2_A ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~ Modality | Subject.

Table 19.3 Coefficients, standard errors and CI for the fixed and random effects for N2 amplitude

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	-4.78	0.56	-5.89 – -3.68	**<0.001**
Age Group[Older]	1.29	0.82	-0.39 – 2.96	0.126
Prediction Cue [What]	0.89	0.3	0.29 – 1.49	**0.004**
Prediction Cue [When]	1.18	0.3	0.58 – 1.78	**<0.001**
Prediction Cue [WW]	1.31	0.3	0.71 – 1.91	**<0.001**
Modality[AV]	1.29	0.43	0.45 – 2.13	**0.003**
Hearing	0.03	0.05	-0.07 – 0.14	0.531
Age Group[Older] *Prediction Cue [What]	-0.28	0.43	-1.13 – 0.57	0.516
Age Group[Older] *Prediction Cue [When]	-1.16	0.43	-2.01 – -0.32	**0.007**
Age Group[Older] *Prediction Cue [WW]	-0.44	0.43	-1.28 – 0.41	0.312
Age Group[Older] *Modality[AV]	-0.83	0.6	-2.02 – 0.36	0.169
Prediction Cue [What] * Modality[AV]	-0.39	0.43	-1.24 – 0.45	0.361
Prediction Cue [When] * Modality[AV]	-0.86	0.43	-1.71 – -0.02	**0.046**
Prediction Cue [WW] * Modality[AV]	-0.95	0.43	-1.80 – -0.11	**0.028**
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	0.47	0.61	-0.73 – 1.67	0.438
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	1.46	0.61	0.26 – 2.66	**0.017**
Random effects				
σ^2	0.74			
τ_{00}	2.92			
τ_{11} A_V	1.43			
Marginal R^2 / Conditional R^2	0.138	0.810		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 19.4 Tukey-adjusted contrasts for N2 amplitude

Contrast	β	SE	df	T	P	d
A vs. AV	-0.661	0.237	210	-2.786	0.0058	-0.769
Control vs. What	-0.675	0.152	210	-4.44	**0.0001**	-0.785
Control vs. When	-0.529	0.152	210	-3.48	**0.003**	-0.615
Control vs. WW	-0.805	0.152	210	-5.30	**<.0001**	-0.936
What vs. When	0.146	0.152	210	0.96	0.772	0.170
What vs. WW	-0.130	0.152	210	-0.86	0.827	-0.152
When vs. WW	-0.276	0.152	210	-1.82	0.267	-0.322

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

Table 19.5 Tukey-adjusted contrasts for the 2-way interaction between Prediction Cue and Group for N2 amplitude

Contrast	Younger						Older					
	β	SE	df	T	P	d	β	SE	df	T	P	d
Control vs. What	-0.697	0.215	210	-3.241	**0.030**	-0.810	-0.653	0.215	210	-3.037	**0.054**	-0.759
Control vs. When	-0.746	0.215	210	-3.469	**0.014**	-0.867	-0.312	0.215	210	-1.450	0.833	-0.362
Control vs. WW	-0.831	0.215	210	-3.865	**0.004**	-0.966	-0.779	0.215	210	-3.626	**0.009**	-0.907
What vs. When	-0.049	0.215	210	-0.228	1.000	-0.057	0.341	0.215	210	1.588	0.757	0.397
What vs. WW	-0.134	0.215	210	-0.624	0.999	-0.156	-0.127	0.215	210	-0.589	0.999	-0.147
When vs. WW	-0.085	0.215	210	-0.395	1.000	-0.099	-0.468	0.215	210	-2.177	0.370	-0.544

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

20. Temporal and Phonetic Visual Predictive Cue analysis: N2 latency

Table 20.1 Estimated marginal means for N2 latency

Conditions		Everyone					Younger					Older				
Prediction cue	**Modality**	**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**		**MM**	**SE**	**df**	**95% CI**	
					LL	**UL**				**LL**	**UL**				**LL**	**UL**
Control	A	308	6.83	29	294	322	287	10.3	29	266	308	329	10.3	29	308	351
	V	298	6.83	29	284	312	283	10.3	29	262	304	313	10.3	29	292	334
What	A	289	6.83	29	275	303	282	10.3	29	261	303	297	10.3	29	275	318
	V	284	6.83	29	270	298	287	10.3	29	265	308	282	10.3	29	261	303
When	A	300	6.83	29	286	314	287	10.3	29	266	308	314	10.3	29	293	335
	V	284	6.83	29	270	298	278	10.3	29	256	299	291	10.3	29	270	312
WW	A	289	6.83	29	275	303	283	10.3	29	262	304	296	10.3	29	275	317
	V	269	6.83	29	255	283	273	10.3	29	252	294	266	10.3	29	244	287

Note. MM = estimated marginal means. SE = standard error of the MM; CI = confidence interval (LL = lower interval; UL = upper interval).

Table 20.2 Type III Tests of Fixed Effects for N2 latency (marginal)

	DF		F	p
(Intercept)	1	210	733.948	**<.0001**
Age group	1	29	7.476	**0.011**
Prediction cue	3	210	0.104	0.958
Modality	1	210	0.117	0.733
Hearing	1	29	0.246	0.624
Age group * Prediction cue	3	210	1.456	0.228
Age group * Modality	1	210	0.585	0.445
Prediction cue * Modality	3	210	0.355	0.786
Age group * Prediction cue * Modality	3	210	0.065	0.979

Note. DF = degrees of freedom. The r model was: N2_L ~ Age group * Prediction Prediction Cue * Modality + PTA, random = ~1 | Subject.

Table 20.3 Coefficients, standard errors and CI for the fixed and random effects for N2 latency

Predictors	β	SE	95% CI	p
Fixed effects				
(Intercept)	291.1	10.74	269.90 – 312.27	**<0.001**
Age Group[Older]	42.49	15.54	10.71 – 74.27	**0.011**
Prediction Cue [What]	-4.81	11.33	-27.14 – 17.52	0.671
Prediction Cue [When]	0.06	11.33	-22.27 – 22.39	0.996
Prediction Cue [WW]	-4	11.33	-26.33 – 18.33	0.724
Modality[AV]	-3.88	11.33	-26.21 – 18.46	0.733
Hearing	-0.41	0.84	-2.13 – 1.30	0.624
Age Group[Older] *Prediction Cue [What]	-27.94	16.02	-59.52 – 3.64	0.083
Age Group[Older] *Prediction Cue [When]	-15.63	16.02	-47.20 – 15.95	0.33
Age Group[Older] *Prediction Cue [WW]	-29.5	16.02	-61.08 – 2.08	0.067
Age Group[Older] *Modality[AV]	-12.25	16.02	-43.83 – 19.33	0.445
Prediction Cue [What] * Modality[AV]	8.38	16.02	-23.20 – 39.95	0.602
Prediction Cue [When] * Modality[AV]	-5.5	16.02	-37.08 – 26.08	0.732
Prediction Cue [WW] * Modality[AV]	-6.25	16.02	-37.83 – 25.33	0.697
(Age Group[Older] *Prediction Cue [What]) *Modality[AV]	-6.94	22.65	-51.60 – 37.72	0.76
(Age Group[Older] *Prediction Cue [When]) *Modality[AV]	-1	22.65	-45.66 – 43.66	0.965
(Age Group[Older] *Prediction Cue [WW]) * Modality[AV]	-8	22.65	-52.66 – 36.66	0.724
Random effects				
σ^2	1026.48			
τ_{00}	465.73			
Marginal R^2 / Conditional R^2	0.184	N/A		

Note. β = unstandardized regression coefficients. SE = standard error of the mean; CI = confidence interval; σ^2 = the random effect variance; τ00 = the random intercept variance (between-subject variance); τ_{11} = The random slope variance.

Table 20.4 Tukey-adjusted contrasts for N2 latency

Contrast	β	SE	df	T	P	d
Younger vs. Older	16.1	11.4	29	1.417	0.1672	0.503

Note. β = unstandardized regression coefficients. SE = standard error of the mean; d = Cohen d.

21. Mediation analyses with VE as the dependent variable

Table 21.1 Model #1				
X	Age group			
Y	VE overall			
M	P2A average (basic)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	-1.681	0.623	-2.700	**0.011**
b	0.019	0.048	0.403	0.690
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	-0.033	-0.229	0.170	0.698
ACME (Older)	0.421	0.011	1.020	**0.040**
ADE (Younger)	-0.341	-1.008	0.220	0.274
ADE (Older)	0.113	-0.269	0.470	0.496
Total Effect	0.080	-0.284	0.420	0.596
Prop. Mediated (Younger)	-0.031	-7.255	5.330	0.938
Prop. Mediated (Older)	1.439	-27.291	43.060	0.620
ACME (Average)	0.194	-0.011	0.510	0.084
ADE (Average)	-0.114	-0.561	0.280	0.608
Prop. Mediated (Average)	0.704	-11.696	18.790	0.656

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 21.2 Model #2				
X	Age group			
Y	VE overall			
M	P2L average (basic)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	19.782	7.732	2.558	**0.016**
b	0.005	0.005	0.869	0.393
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.094	-0.085	0.380	0.340
ACME (Older)	0.211	-0.083	0.650	0.190
ADE (Younger)	-0.116	-0.706	0.440	0.670
ADE (Older)	0.001	-0.364	0.360	0.990
Total Effect	0.095	-0.292	0.450	0.560
Prop. Mediated (Younger)	0.288	-5.860	6.310	0.660
Prop. Mediated (Older)	0.477	-14.300	15.390	0.700
ACME (Average)	0.153	-0.027	0.450	0.100
ADE (Average)	-0.058	-0.493	0.360	0.810
Prop. Mediated (Average)	0.383	-9.840	10.110	0.620

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 21.3 Model #3				
X	Age group			
Y	VE overall			
M	P2L average (integration)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	17.429	7.980	2.184	**0.037**
b	0.003	0.005	0.534	0.597
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.048	-0.169	0.320	0.650
ACME (Older)	0.192	-0.090	0.660	0.250
ADE (Younger)	-0.104	-0.694	0.430	0.700
ADE (Older)	0.040	-0.313	0.400	0.850
Total Effect	0.087	-0.278	0.410	0.610
Prop. Mediated (Younger)	0.134	-4.371	5.430	0.750
Prop. Mediated (Older)	0.278	-16.629	14.930	0.760
ACME (Average)	0.120	-0.036	0.390	0.220
ADE (Average)	-0.032	-0.469	0.360	0.870
Prop. Mediated (Average)	0.206	-9.243	10.380	0.710

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 21.4 Model #4				
X	Age group			
Y	VE overall			
M	N2A average (basic)			
COV	Hearing			
Results				
a	1.092	0.778	1.403	0.171
b	-0.001	0.048	-0.011	0.991
	β	**95%CI**		**p**
		LL	**UL**	
ACME (Younger)	-0.001	-0.170	0.140	0.990
ACME (Older)	0.153	-0.090	0.540	0.270
ADE (Younger)	-0.044	-0.452	0.320	0.890
ADE (Older)	0.110	-0.218	0.450	0.510
Total Effect	0.110	-0.220	0.430	0.480
Prop. Mediated (Younger)	0.015	-2.720	2.760	0.930
Prop. Mediated (Older)	0.570	-9.270	13.930	0.560
ACME (Average)	0.076	-0.052	0.290	0.320
ADE (Average)	0.033	-0.298	0.330	0.830
Prop. Mediated (Average)	0.292	-3.884	6.730	0.590

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 21.5. Model #5

X	Age group			
Y	VE overall			
M	N2L average (integration)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	0.750	11.128	0.067	0.947
b	0.000	0.006	0.006	0.995
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	-0.003	-0.195	0.160	0.990
ACME (Older)	-0.001	-0.075	0.070	0.970
ADE (Younger)	0.102	-0.214	0.440	0.540
ADE (Older)	0.104	-0.232	0.450	0.550
Total Effect	0.101	-0.215	0.430	0.530
Prop. Mediated (Younger)	0.003	-5.531	4.040	0.970
Prop. Mediated (Older)	0.002	-1.755	1.990	0.960
ACME (Average)	-0.002	-0.095	0.080	1.000
ADE (Average)	0.103	-0.215	0.430	0.490
Prop. Mediated (Average)	0.003	-2.763	2.720	0.960

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

22. Mediation analyses with PE as the dependent variable

Table 22.1 Model #1

X	Age group			
Y	PE overall			
M	P2A average (basic)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	-1.681	0.623	-2.700	**0.011**
b	-0.073	0.040	-1.827	0.079
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.132	-0.082	0.440	0.270
ACME (Older)	-0.129	-0.519	0.210	0.430
ADE (Younger)	0.005	-0.448	0.500	1.000
ADE (Older)	-0.256	-0.590	0.050	0.110
Total Effect	-0.124	-0.393	0.180	0.390
Prop. Mediated (Younger)	-0.407	-10.023	10.140	0.610
Prop. Mediated (Older)	0.513	-8.591	10.830	0.650
ACME (Average)	0.001	-0.205	0.250	0.960
ADE (Average)	-0.125	-0.455	0.230	0.430
Prop. Mediated (Average)	0.053	-5.147	6.640	0.900

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 22.2 Model #2

X	Age group			
Y	PE overall			
M	P2L average (basic)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	19.782	7.732	2.558	**0.016**
b	0.005	0.005	1.058	0.299
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.092	-0.060	0.320	0.280
ACME (Older)	0.009	-0.121	0.140	0.870
ADE (Younger)	-0.136	-0.369	0.090	0.270
ADE (Older)	-0.219	-0.513	0.050	0.110
Total Effect	-0.128	-0.374	0.130	0.310
Prop. Mediated (Younger)	-0.294	-10.681	11.660	0.540
Prop. Mediated (Older)	0.006	-4.701	3.840	0.970
ACME (Average)	0.050	-0.045	0.190	0.330
ADE (Average)	-0.178	-0.403	0.040	0.130
Prop. Mediated (Average)	-0.144	-6.721	6.210	0.600

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 22.3 Model #3

X	Age group			
Y	PE overall			
M	P2L average (integration)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	17.429	7.980	2.184	**0.037**
b	0.005	0.004	1.201	0.240
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.088	-0.027	0.280	0.190
ACME (Older)	0.021	-0.085	0.150	0.710
ADE (Younger)	-0.149	-0.421	0.110	0.260
ADE (Older)	-0.216	-0.510	0.060	0.120
Total Effect	-0.128	-0.380	0.120	0.320
Prop. Mediated (Younger)	-0.326	-9.985	6.510	0.470
Prop. Mediated (Older)	-0.050	-3.679	3.230	0.810
ACME (Average)	0.055	-0.024	0.180	0.200
ADE (Average)	-0.182	-0.441	0.070	0.180
Prop. Mediated (Average)	-0.188	-6.053	3.810	0.470

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 22.4 Model #4

X	Age group			
Y	PE overall			
M	N2A average (basic)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	1.092	0.778	1.403	0.171
b	0.065	0.041	1.575	0.127
	β	**95%CI**		**p**
		LL	**UL**	
ACME (Younger)	0.065	-0.097	0.330	0.450
ACME (Older)	-0.014	-0.107	0.070	0.740
ADE (Younger)	-0.132	-0.400	0.130	0.300
ADE (Older)	-0.211	-0.509	0.070	0.140
Total Effect	-0.145	-0.415	0.120	0.290
Prop. Mediated (Younger)	-0.144	-7.627	7.390	0.680
Prop. Mediated (Older)	0.046	-1.615	1.980	0.720
ACME (Average)	0.026	-0.066	0.150	0.610
ADE (Average)	-0.171	-0.425	0.070	0.170
Prop. Mediated (Average)	-0.049	-2.901	4.000	0.830

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

Table 22.5 Model #5				
X	Age group			
Y	PE overall			
M	N2L average (integration)			
COV	Hearing			
Results	**β**	**SE**	**t**	**p**
a	0.750	11.128	0.067	0.947
b	0.006	0.005	1.276	0.213
	β	**95% CI**		**p**
		LL	**UL**	
ACME (Younger)	0.008	-0.254	0.290	0.990
ACME (Older)	0.000	-0.183	0.170	0.990
ADE (Younger)	-0.146	-0.464	0.190	0.400
ADE (Older)	-0.154	-0.493	0.210	0.390
Total Effect	-0.146	-0.488	0.190	0.410
Prop. Mediated (Younger)	0.016	-4.590	5.790	0.950
Prop. Mediated (Older)	0.006	-3.730	4.320	0.950
ACME (Average)	0.004	-0.170	0.180	0.970
ADE (Average)	-0.150	-0.472	0.180	0.370
Prop. Mediated (Average)	0.011	-3.280	3.670	0.910

Note. **β** = estimate; SE = robust standard error of the estimate. **95% CI** = confidence interval of the estimate (LL = lower limit; UL = upper limit). ADE = Average direct effect. ACME = The average causal mediation effect (ACME) represents the expected difference in the potential outcome when the mediator took the value that would realize under the treatment condition as opposed to the control condition, while the treatment status itself is held constant.

CRediT author statement

Pascale Tremblay: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing- Original draft preparation. **Serge Pinto**: Conceptualization, Investigation, Writing—Reviewing and Editing. **Anahita Basirat**: Conceptualization, Writing—Reviewing and Editing. **Marc Sato**: Conceptualization, Methodology, Investigation, Resources, Project administration, Writing—Reviewing and Editing.

REVISED Manuscript (text with changes Marked) Click here to view linked References

Visual Prediction cues can facilitate behavioural and neural speech processing in young and older adults

Pascale Tremblay[1,2], Anahita Basirat[3], Serge Pinto[4], Marc Sato[4]

1 Département de Réadaptation, Faculté de Médecine, Université Laval, Quebec City, Canada

2 Cervo Brain Research Centre, Quebec City, Canada

3 Univ. Lille, CNRS, UMR 9193 – SCALab – Sciences Cognitives et Sciences Affectives, Lille, France

4 France Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

Corresponding author:

Pascale Tremblay, Ph. D.

Département de réadaptation, Université Laval

1050 avenue de la Médecine, Québec (QC)

CANADA G1V 0A6

Centre de recherche CERVO

Email: Pascale.Tremblay@fmed.ulaval.ca

Phone: +001 418 663–5000 ext. 8717

Fax: +001 418 656–5476

Abstract

The ability to process speech evolves over the course of the lifespan. Understanding speech at low acoustic intensity and in the presence of background noise becomes harder, and the ability for older adults to benefit from audiovisual speech also appears to decline. These difficulties can have important consequences on quality of life. Yet, a consensus on the cause of these difficulties is still lacking. The objective of this study was to examine the processing of speech in young and older adults under different modalities (i.e. auditory [A], visual [V], audiovisual [AV]) and in the presence of different visual prediction cues (i.e., no predictive cue (control), temporal predictive cue, phonetic predictive cue, and combined temporal and phonetic predictive cues.). We focused on recognition accuracy and four auditory evoked potential (AEP) components: P1-N1-P2 and N2. Thirty-four right-handed French-speaking adults were recruited, including 17 younger adults (28 ± 2 years; 20–42 years) and 17 older adults (67 ± 3.77 years; 60–73 years). Participants completed a forced-choice speech identification task. The main findings of the study are: (1) The faciliatory effect of visual information was reduced, but present, in older compared to younger adults, (2) visual predictive cues facilitated speech recognition in younger and older adults alike, (3) age differences in AEPs were localized to later components (P2 and N2), suggesting that aging predominantly affects higher-order cortical processes related to speech processing rather than lower-level *auditory* processes. (4) Specifically, AV facilitation on P2 amplitude was lower in older adults, there was a reduced effect of the temporal predictive cue on N2 amplitude for older compared to younger adults, and P2 and N2 latencies were longer for older adults, and finally (5) behavioural performance was associated with P2 amplitude in older adults. Our results indicate that aging affects speech processing at multiple levels, including audiovisual integration (P2) and auditory attentional processes (N2). These findings have important implications for understanding barriers to communication in older ages, as well as for the development of compensation strategies for those with speech processing difficulties.

Keywords: Event-related potentials, Speech perception, Aging, Audiovisual enhancement, Predictive coding, Auditory processing

1. Introduction

One of the most common complaints of older adults is difficulty in understanding speech at low acoustic intensity and in the presence of background noise (CHABA, 1988; Pichora-Fuller, 1997). These difficulties are associated with increased self-consciousness about one's communication competences, reduced self-confidence and disengagement from social activities, which can lead to isolation (Tobias, 1977). Yet, a unified theory of the etiology of these difficulties is still lacking and so are effective treatment options and prevention strategies (El-Assal & El-Gharib, 2019; Humes, Wilson, Barlow, & Garner, 2002).

Importantly, speech processing difficulty can arise during face-to-face conversations even when articulatory (visual) information from the speaker complements the acoustic speech signal. The influence of visual information on speech perception is well established. The most well-known example of the influence of visual cues on perception of speech sounds is the McGurk effect, in which, when presented with incongruent auditory and visual speech stimuli, most people report hearing a fusion of the two syllables presented (auditory and visual) (McGurk & MacDonald, 1976). In young adults, several studies have shown that audiovisual speech (AV) can be associated with a speech recognition gain compared to auditory (A) alone (e.g. Erber, 1969; Sumby & Pollack, 1954), a phenomenon that is often referred to as the *audiovisual speech advantage* or, more generally, *multimodal enhancement*. Importantly, in accordance with the principle of inverse effectiveness (i.e. multisensory enhancement is greatest when unimodal stimuli are least effective), AV speech recognition is more resistant to noise than A-only (Erber, 1969). These effects indicate that adding articulatory (visual) information helps reduce uncertainty, which facilitates auditory speech recognition, especially in noisy auditory conditions. A decline in the ability to extract and to use articulatory (visual) information to disambiguate speech could be contributing to the speech processing difficulties experienced by older adults. Several studies have shown that, compared to younger adults, middle-age and older adults exhibit *lower* performance enhancement for AV speech compared to A

speech (Tye-Murray, Sommers, Spehar, Myerson, & Hale, 2010; Yang & Ren, 2018). In contrast, other studies have found comparable (Avivi-Reich, Puka, & Schneider, 2018; Cienkowski & Carney, 2002; Sommers, Tye-Murray, & Spehar, 2005; Winneke & Phillips, 2011) or *enhanced* AV integration and binding abilities for older compared to younger adults (Ganesh, Berthommier, & Schwartz, 2017; Laurienti, Burdette, Maldjian, & Wallace, 2006; Sekiyama, Soshi, & Sakamoto, 2014). In an attempt to resolve these inconsistent findings, it has been suggested that there may be no overall decline in the capacity to benefit from audiovisual information in aging, but instead, that the conditions needed to benefit from it may change over the course of the lifespan, with older adults benefiting less when the (acoustic) signal-to-noise ratio is low (Jansen, Keebler, & Chaparro, 2018; Stevenson, Nelms, Baum, Zurkovsky, Barense, Newhouse, & Wallace, 2015) and when the visual information is degraded (Gordon & Allen, 2009). In sum, the ability for older adults to benefit from an audiovisual speech signal appears to evolve with age and to depend upon the clarity of the acoustic and/or visual sources.

While there have been several behavioural studies focusing on AV speech perception in aging, the neural mechanisms underlying AV speech perception performance and multimodal enhancement in aging, and, more generally, speech processing capacities, are still unclear. Neurophysiological studies have shown that, in young adults, prior knowledge of a speech input facilitate neural processing of deteriorated or missing speech (Cervantes Constantino & Simon, 2018), and that adding visual articulatory information to auditory speech modulates activity in primary auditory and associative regions of the temporal cortex including the superior temporal sulcus (Arnal, Morillon, Kell, & Giraud, 2009), affecting the P1-N1-P2 complex (e.g. Stekelenburg & Vroomen, 2007; van Wassenhove, Grant, & Poeppel, 2005). The P1-N1-P2 complex is a series of co-occurring auditory evoked potentials (AEPs), which indicate that a sound has reached the auditory cortex and that initial cortical acoustic-phonetic processing has begun. As such, the P1/N1/P2 complex indexes the capacity for speech sound processing and discrimination. It consists of positive and negative voltage deflections peaking around 50 ms (P1), 100 ms (N1), and 200 ms (P2) after stimulus onset. AEPs components such as the P1-N1-

P2 complex are typically identified by their polarity (positive or negative), latency (peak occurrence after stimulus onset, measured in milliseconds) and amplitude (in µV). Specifically, it is well established that, compared to unimodal auditory perception, adding visual articulatory information to auditory speech leads to an attenuated amplitude and earlier latency of the N1/P2 complex (e.g., Besle, Fort, Delpuech, & Giard, 2004; Klucharev, Mottonen, & Sams, 2003; Treille, Cordeboeuf, Vilain, & Sato, 2014a; Treille, Vilain, Kandel, & Sato, 2017; Treille, Vilain, & Sato, 2014b; Treille, Vilain, Schwartz, Hueber, & Sato, 2018; van Wassenhove et al., 2005). Hence, this complex is usually seen as a reliable marker of AV integration.

Winneke & Phillips were the firsts to investigate the behavioural benefit of and the neural processes (P1/N1/P2 complex) associated with AV perception of spoken words presented in babbling noise in young and older adults (Winneke & Phillips, 2011). Behaviourally, their results showed no difference in the ability for younger and older adults to benefit from audiovisual information. Yet, EEG results revealed an age-related difference in P1 amplitude, with an increased reduction from A to AV speech (i.e. multisensory integration) for older compared to younger adults. There was also an age difference in N1 latency with a more pronounced latency shift (i.e. facilitation) for older compared to younger adults when comparing A to AV speech. The authors concluded that older adults, compared with younger adults, "*are not better lip readers per se but rather are better "lip/speech integrators*." (p. 436), perhaps as a compensation for more laborious auditory processing in older adults. Though these results critically await replication, abnormal cortical response patterns have been shown during speech and speech in noise tasks in older adults in several studies (e.g. Brodbeck, Presacco, Anderson, & Simon, 2018; Presacco, Simon, & Anderson, 2016a, 2016b).

During audiovisual speech processing, perceptual experience is aided by prior crossmodal associations and integration mechanisms, which can help reduce sensory uncertainty. Predictive coding theories postulates that the brain actively predicts upcoming sensory input rather than simply registering it. According to this view, bottom-up sensory information is compared with top-down

predictions from higher levels to estimate prediction errors (Clark, 2013; Friston, 2005, 2010; Rao & Ballard, 1999). For speech, audiovisual integration operates through temporal expectations and speech-specific predictions. From a Bayesian perspective, perceptual experience derives from the processing and integration of multisensory (AV) inputs based on their predictability and joint probability (Massaro, 1998; van Wassenhove, 2013). Therefore, while visual (articulatory) information can facilitate speech recognition and modulate the N1/P2 complex, other types of predictive cues could also have a facilitatory impact on speech perception and AEPs and could potentially facilitate speech processing in older adults. Notably, we showed in a previous EEG study that visual temporal and phonetic predictions can attenuate the amplitude of the N1/P2 components during *auditory* speech perception, and that P2 (amplitude and latency), but not N1, is modulated (facilitated) by phonetic prediction during *audiovisual* speech perception (Pinto, Tremblay, Basirat, & Sato, 2019), demonstrating an enhanced sensitivity for P2 to phonetic cues, consistent with previous work (e.g. Baart, Stekelenburg, & Vroomen, 2014). The effect of age on the ability to process different kinds of visual predictive cues, however, remains unclear.

The first objective of the present study was to investigate the effect of age and audiovisual integration on speech perception performance (Objective 1). Specifically, we wanted to determine whether audiovisual integration (objective 1a: adding natural speech movements through video) and/or other visual prediction cues (objective 1b: temporal onset and/or visuo-orthographic (e.g. /pa/) cues) facilitate auditory speech syllable perception similarly for younger and older adults. The second objective of the study was to compare the neurophysiological response to speech in younger and older adults focusing on AEPs (P1-N1-P2 and N2). First, we compared audiovisual speech integration in younger and older adults using an additive model (objective 2a; AV vs. A+V). This analysis focused strictly on visual articulatory movements in order to confirm and extend previous EEG findings (Brodbeck et al., 2018; Presacco et al., 2016a, 2016b; Winneke & Phillips, 2011). Next, we tested whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural

processing of speech similarly for younger and older adults (Objective 2b; AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). Finally, the third objective of the study was to examine the relationship between speech perception performance and AEPs to shed new lights on brain aging and its impact on human behaviour.

To achieve these goals, the predictability of auditory syllables was manipulated experimentally by adding unnatural visual information indicative of their temporal unfolding (*when*) and phonetic content (*what*) (Pinto et al., 2019). The presentation modality (auditory (A), visual (V) and audiovisual (AV)) was manipulated to examine whether the visual predictive cues would be processed similarly when the speech signal was audio and audiovisual.

Based on prior studies, we hypothesized that older adults would show similar or enhanced audiovisual and predictive gain. Specifically, we expected a facilitation of AEPs (shorter latency and/or reduced amplitude) during AV speech compared to A speech, as well as during the processing of temporal and phonetic predictive cues. We expected those patterns to either be identical for the young and older adults, or to be enhanced in older adults, based on their lifetime of experience processing speech. Indeed, predictive coding for speech could be heightened with age and serve as a compensation strategy to overcome declining unisensory processing. Alternatively, predictive coding could, like other higher-order cognitive functions, decline with age and this decline could be associated with age-related speech processing difficulties. Given the well-established hearing loss and cognitive decline that occurs in aging, and the known association between these two factors (e.g.Humes, Busey, Craig, & Kewley-Port, 2013; Humes, Kidd, & Lentz, 2013; Lin, 2011; Lin, Ferrucci, Metter, An, Zonderman, & Resnick, 2011; Lin, Yaffe, Xia, Xue, Harris, Purchase-Helzner, Satterfield, Ayonayon, Ferrucci, Simonsick, & Health, 2013; Wayne & Johnsrude, 2015), we expected that the auditory N2, which indexes late cognitive processes, such as executive functions and attention (Fritz, Elhilali, David, & Shamma, 2007), would show either an increased amplitude, reflecting compensatory activity, or perhaps a decline, reflecting a disruption in processing.

2. Material and Methods

2.1. Participants

Thirty-four right-handed French-speaking adults participated in the study after giving informed consent. All participants reported normal or corrected-to-normal vision and were allowed to wear their glasses or lenses during the experiment. Participants reported no history of hearing, speaking, language, neurological and/or neuropsychological disorders. The cognitive functioning of all participants was evaluated using the Montreal Cognitive Assessment scale (MoCA) (Nasreddine, Chertkow, Phillips, Bergman, & Whitehead, 2003; Nasreddine, Phillips, Bedirian, Charbonneau, Whitehead, Collin, Cummings, & Chertkow, 2005). None of the participant had mild cognitive decline using the criteria by Larouche et al. (Larouche, Tremblay, Potvin, Laforest, Bergeron, Laforce, Monetta, Boucher, Tremblay, Belleville, Lorrain, Gagnon, Gosselin, Castellano, Cunnane, Macoir, & Hudon, 2016).

Participants were divided into a younger and an older group. The younger group included 17 adults (14 females), with a mean age of 28 ± 2 years (20–42 years) and an average of 15.6 ± 2.61 years of education (range: 11–20 years). The data from the younger group were published in Pinto et al. (2019). The older group included 17 adults (8 females), with a mean age of 67 ± 3.77 years (60–73 years) and an average of 15.25 ± 3.55 years of education (range: 9–20 years). Participants' characteristics are detailed in Table 1. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (World Medical, 2013) and participants were compensated for the time spent in the study.

Table 1. Participants' characteristics

	Young (N = 17; 13 W)		**Older** (N = 17; 8 W)	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**
Age	28.20 (6.87)	20–42	67 (3.88)	60–73
MoCA (/30)	29.07 (1.38)	25–30	26.38 (1.66)	23–29

Handedness	84.8 (.15)	55–100	95.9 (11.5)	58–100
Right ear PTA	4.77 (5.22)	-5 – 15	13.699 (6.41)	4.16 – 26.6
Left ear PTA	5.55 (4.3)	-3.33 – 13.33	14.79 (6.26)	3.33 – 25
Education	15.6 (2.61)	11 – 20	15.25 (3.55)	9 – 20

Table 1 Participants' characteristics. SD = standard deviation. W = women. PTA = Pure tone average.

2.2. Hearing Assessment

To ensure that participants had normal hearing, pure tone audiometry was performed using a clinical audiometer (Resonance R17A, MRS, Italy) for each ear separately, at the following frequencies: .5, 1, 2 kHz. For each participant, a standard pure tone average (PTA: average of thresholds at .5, 1 and 2 kHz) was computed for the left and right ear. The result of the hearing assessment is provided in Table 1. Figure 1 illustrates the average threshold at each frequency tested separately for each Age Group.

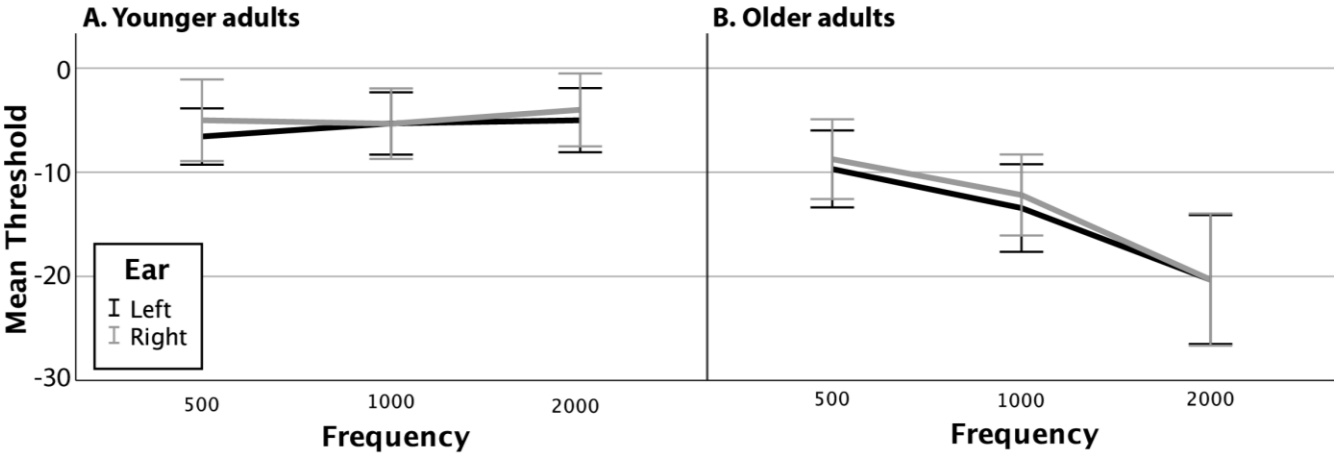

Figure 1 . Hearing thresholds. The line charts present an overview of participants' hearing thresholds, separately for each ear, and the younger (A) and the older adults. Each line in the figure represents the average threshold for a group of participants. Error bars represent the confidence interval of the means.

A linear mixed model (LMM) analysis was conducted using the r package version 4.0.3 for Mac (Team, 2019), with Age Group (younger, older) as the between-subject factor, Ear (Left, Right), and Frequency (500, 1000, 2000 Hz) as the within-subject factors. The random intercepts for participants were included in the model. The LMM results show main effects of Age Group ($p \leq .001$)

and Frequency ($p \leq .001$), as well as a 2-way interaction between Age Group and Frequency ($p \leq .001$). Tukey-corrected post hoc contrasts indicated that thresholds were higher in the older adults at 1000 Hz (p = .0022, d = -1.230) and 2000 Hz ($p \leq .001$, d = -2.588). The descriptive statistics and the detailed results of the LMM analyses as well as the pairwise contrasts are provided in Supplementary Material 1. Because of these (expected) Group differences, Hearing (average PTA) was included in all analyses as a covariate.

2.3. Speech Perception Assessment

All procedures were carried out in a sound-attenuated room at the LPL Lab in Aix-en-Provence. Participants sat in front of a computer monitor at a distance of approximately 50 cm. The acoustic stimuli were presented through loudspeakers located on each side of a computer monitor. Stimuli were presented using Presentation software (Neurobehavioral Systems, Albany, USA), which was also used to record participants' behavioural responses. During the task, participants were asked to complete a forced-choice speech identification task. On each trial, they identified one syllable by pressing one of three keys on a keyboard with their left hand. No feedback was provided. The stimuli were the syllables /pa/, /ta/, and/ka/. All stimuli were presented in quiet. The response key designation was counterbalanced. To dissociate sensory/perceptual from motor responses on the EEG recordings, each stimulus was followed, after 600 ms, by a brief auditory tone and question mark (?), which served as "GO" cues. The inter-trial interval was 3s. The experiment lasted approximately 45 min and was divided in four sessions of ~11 minutes with short breaks in between sessions.

The syllables were presented in three modalities (auditory [A], visual [V], audiovisual [AV]) and under 4 different cue conditions (control [no cue], *when*, *what*, *what-when*), which resulted in 12 experimental conditions: control (A, V, AV), when (A_{when}, V_{when}, AV_{when}), what (A_{what}, V_{what}, AV_{what}) and what-when ($A_{what\text{-}when}$, $V_{what\text{-}when}$, $AV_{what\text{-}when}$). The experiment consisted of 864 trials presented in a pseudo-randomized order, including 72 trials in each of the 12 experimental conditions.

A full description of the stimuli can be found in Pinto et al. (2019). Briefly, the auditory stimuli consisted of an acoustic syllable dubbed on a static image of a neutral mid-open mouth position of a speaker. The visual stimuli consisted of the visual speech movements displayed without any sound. The audiovisual stimuli started with an initial neutral mid-open mouth position followed by visual speech movements (30 frames, 1200 ms) before the acoustic consonantal burst and the syllable (5 frames, 200 ms). For all stimuli, the auditory signal intensity was normalized using a common maximal amplitude criterion. Importantly, the audiovisual stimuli were first created. The visual and auditory-only stimuli were created from the audiovisual stimuli, by removing the acoustic signal (visual stimuli) or by replacing the visual speech movements by a static face (auditory stimuli). Participants were informed that visual cues were always coherent with the auditory syllable.

In all conditions, "##" orthographic symbols and a static timeline were visually presented during the first 15 frames (0 to 600 ms). In the *when* conditions, a moving visual timeline indicative of the temporal consonantal onset of the acoustic syllable replaced the static timeline during the subsequent 15 ± 2 frames (600 ± 80 to 1200 ms). In the *what* conditions, a visuo-orthographic cue indicative of the syllable (/pa/, /ta/or/ka/) replaced the "##" symbols during the subsequent 15 ± 2 frames (600 ± 80 to 1200 ms). In the *what-when* conditions, both the visual timeline and visuo-orthographic cues were presented. Finally, in the control conditions, speech signals were presented only with the "##" orthographic symbols and the static timeline during 15 ± 2 frames (600 ± 80 to 1200 ms). The Prediction conditions are illustrated in Figure 2.

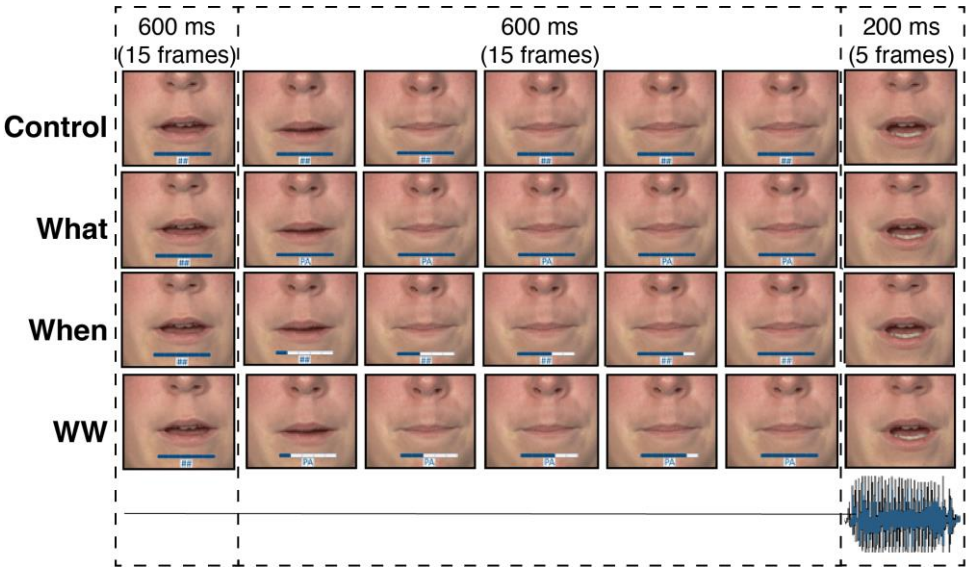

Figure 2 Prediction Cue conditions. WW = What+When condition. The sound wave at the bottom right represents the auditory speech signal.

2.4. EEG Recordings and Pre-Processing

EEG data were recorded continuously from 9 scalp electrodes (Electro-Cap International, INC, according to the international 10–20 system) using the Biosemi Active Two AD-box EEG system operating at a 512 Hz sampling rate. Because N1/P2 AEPs have maximal response over central sites on the scalp (Naatanen & Picton, 1987; Scherg & Von Cramon, 1986), EEG was only collected from fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) (Pinto et al., 2019). This minimal recording procedure has been used in several EEG studies on audiovisual speech integration and auditory evoked responses that showed classical audiovisual speech interactions on N1/P2 AEPs (Stekelenburg & Vroomen, 2007; Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; Vroomen & Stekelenburg, 2010).

Two additional electrodes were used as ground electrodes (Common Mode Sense [CMS] active and Driven Right Leg [DRL] passive electrodes). In addition, one external reference electrode was set at the top of the nose. Horizontal (HEOG) and vertical (VEOG) eye movements were recorded using electrodes positioned at the outer canthus of each eye, as well as above and below the right eye. Before the experiment, the impedance of all electrodes was adjusted to get low offset voltages and stable DC.

EEG data were processed using EEGLAB (Delorme & Makeig, 2004) running on Matlab (Mathworks, Natick, USA). EEG data were first re-referenced offline to the nose channel and band-pass filtered using a 2-way least square FIR filtering (3–30 Hz) to reduce slow drifts and high frequency noise (see Pinto et al. (2019). Next, the data was segmented into 500 ms epochs including a 100 ms prestimulus baseline (from −100 to −0 ms relative to the acoustic syllable onset). Epochs with an amplitude change exceeding ±100 uV at any channel (including HEOG and VEOG channels) were rejected (mean (±SD): 2% (±2%) trials). Responses from/pa/, /ta/and/ka/syllables were first averaged together in order to provide 72 trials per condition. For each participant and each condition (i.e., A, V, AV, A_{when}, V_{when}, AV_{when}, A_{what}, V_{what}, AV_{what}, $A_{what\text{-}when}$, $V_{what\text{-}when}$, $AV_{what\text{-}when}$), data were then averaged over the nine electrodes. Finally, the maximal amplitude and peak latency of the N1-P2-N2 complex were determined using a fixed temporal window for each component (N1: 70–150 ms; P2: 150–250 ms; N2: 180–325 ms). For P1, instead of extracting the maximal amplitude and peak latency, we computed the area under the curve for amplitude (10–100 ms), and the 50% area latency. The V condition was included in the study design to examine behaviour in this condition, as well as to compare the neural response to AV with the A+V signal (Objective 2a, see section 2.5.3).

2.5. Statistical analyses

2.5.1. Behavioural data

Behavioural data analyses were conducted to address the first objective of the project, to investigate the effect of age and audiovisual integration on speech perception performance. To achieve this goal, three complementary behavioural indexes of performance were computed to characterize speech perception performance: a classical percentage of correct responses, and two difference scores: a visual (VE) effect scores and predictive cue effect (PE) scores.

First, a classical measure of accuracy was computed and used to compare perception accuracy in young and older adults during speech perception. The percentage of correct responses was

determined for each participant and condition (note that RTs were not recorded due to the 600 ms delay between the stimuli and the "GO" cues for the manual responses). A linear mixed model (LMM) analysis was conducted using r version 4.0.3 for Mac (Team, 2019), using the nlme package, with Age Group (younger, older) as the between-subject factor, Modality (A, V, AV) and Predictive cue (Control, What, When, WW) as the within-subject factors, and hearing (PTA) as between-subject continuous fixed factor. The random intercepts for participants were also included in the model. Model selection included testing models, using likelihood ratio tests, with and without PTA as well as with different random effect structures (with or without slopes for either Modality or Predictive cue condition) and covariance structure for the residuals (homogeneous vs. heterogeneous variance across levels of Modality or Predictive Cue). The model with the best fit, assessed using likelihood ratio tests, was kept. When fit was similar across one model or more, the simplest model was kept. The same procedure was used for all LMM analyses (accuracy, VE, PE and all EEG analyses) and will not be repeated hereafter. For each analysis (accuracy, VE, PE and all EEG analyses), the final model is provided in the corresponding supplementary materials.

Next, visual effect (VE) scores were calculated to determine whether adding natural speech movements through video to auditory speech facilitates auditory speech perception through audiovisual integration similarly in young and older adults (objective 1a). These scores were derived from the percentage correct scores for each participant and each of the cue condition (Control, What, When, WW) with the following formula: VE = (AV—A) / (100-A). VE measures have been used in several AV studies as they circumvent the bias inherent to calculating the difference between AV and A, in which higher values of A necessarily lead to lower values of VE (Sommers et al., 2005). To examine Age Group differences in VE, a LMM analysis was conducted on the VE scores, with Age Group (younger, older) as the between-subject factor, Prediction Cue condition (Control, What, When, WW) as the within-subject factor, and hearing (PTA) as a between-subject continuous fixed factor. The random intercepts for participants were also included in the model. The model selection procedure was

the same as described for accuracy. The VE value for 8 participants (3 young and 5 elderly) that had perfect score in the auditory modality could not be included in the analysis because the equation resulted in an error when attempting to divide by zero. Two outliers were removed from the final analyses (one young and one older adult).

Finally, predictive cue effect (PE) scores were calculated to determine whether adding other temporal and/or phonetic visual cues to the speech signal would facilitate speech perception similarly in young and older adults (objective 1b). The PE scores were inspired by the VE scores and used to circumvent the bias inherent to calculating the difference between the score in Prediction and Control conditions, in which higher values of Prediction necessarily lead to lower values of PE. A PE score was calculated for each prediction cue (What, When, WW) and each modality. The What effect (WhatE) score was calculated from the percentage correct scores for each participant. WhatE = (What—Control)/(100-Control). The formula for the When effect (WhenE) (WhenE = (When—Control)/(100-Control)) and the What+When effect (WWE) scores were identical: WWE = (WWE—Control)/(100-Control). To examine Age Group differences in prediction effect, a LMM analysis was conducted on the PE scores, with Age Group (younger, older) as the between-subject factor, Modality (A, V, AV), and Predictive cue condition (What, When, WW) as the within-subject factors, and hearing (PTA) as the between-subject continuous covariate. The random intercepts for participants were also included in the model. The model selection procedure was the same as described for accuracy.

2.5.2. EEG Analyzes: Audiovisual Integration (additive model)

The first set of analyses focused on objective 2a of the study, to compare audiovisual speech integration in younger and older adults using an additive model (AV vs. A+V). To address this aim, we used an additive model (e.g. Pilling, 2009; Treille et al., 2014a; Treille et al., 2014b; van Wassenhove et al., 2005) in which the bimodal audiovisual EEG signal was compared to the sum of auditory and visual unimodal EEG signals (AV ≠ A + V). The resulting signal may contain task-related neural

activity common to all modalities unrelated to crossmodal integration, such as movement-related cortical potentials, characterized by a slow negative deflection on fronto-central sites starting around 1000 ms prior to the manual response (e.g. Kornhuber & Deecke, 1965; Libet, Gleason, Wright, & Pearl, 1983), as well as slow anticipatory potentials that precedes perceptual decisions and discriminative responses, characterized by a slow positive deflection on fronto-central sites (e.g. Teder-Salejarvi, McDonald, Di Russo, & Hillyard, 2002). In order to minimize the temporally contingent influence of movement-related cortical potentials on auditory evoked potentials, an 800 ms delay was introduced between the acoustic consonantal burst of all syllables and the "GO" cues. Moreover, a 3 Hz high-pass filter was applied on the EEG data to minimize the contribution of movement-related as well as slow anticipatory potentials (Teder-Salejarvi et al., 2002).

A series of LMM analysis was conducted for each dependent measure: P1 peak amplitude and peak latency, N1 peak amplitude and peak latency, P2 peak amplitude and peak latency, and N2 peak amplitude and peak latency. In all analyses, the within-subject (repeated) fixed factors were Modality (AV, A+V), and the between-subject factor was Age Group (younger, older). Hearing (PTA) was included as a between-subject continuous fixed factor (covariate). Predictive cues were not included in this analysis because here we meant to assess audiovisual integration. The analyses were conducted on the signal from the control (Control) condition). The random intercepts for participants were also included in the model. The model selection procedure was the same as described for accuracy.

2.5.3. EEG Analyzes: Temporal and Phonetic Visual Predictive Cues

This second set of EEG analyses focused on objective 2b of the study, to determine whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural processing of speech similarly for younger and older adults (Objective 2b; AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). To address this aim, a series of linear mixed model (LMM) analysis was run to examine age differences on the neural processing of

Predictive Cues under different modalities (A, AV). One analysis was conducted for each dependent measure: P1 peak amplitude and peak latency, N1 peak amplitude and peak latency, P2 peak amplitude and peak latency, and N2 peak amplitude and peak latency. In all analyses, the within-subject (repeated) fixed factors were Modality (A, AV) and Predictive Cue (Control, What, When, WW), the between-subject factor was Age Group (younger, older), and hearing (PTA) was included as between-subject continuous fixed factors (covariate). The random intercepts for participants were included in the model. The model selection procedure was the same as described for accuracy.

2.5.4. Electrophysiology-Behaviour Relationship

The third objective of the study was to examine the relationship between performance in young and older adults and AEPs. To achieve this goal, a series of simple mediation analyses were conducted using the r packages NLME and Mediation, a package to conduct causal mediation analyses (Tingley D, 2013). All analyses were run with the same seed (2021). In all analyses, Age Group was used as the categorical predictor variable (X), one AEP component was included as continuous mediators (M) in separate analyses, and hearing (PTA) was used as a continuous between-subject covariate. The dependent variables (Y) were those that showed Age Group differences or interactions. A quasi-Bayesian approximation based on normal approximation (Imai, Keele, & Tingley, 2010) was used to compute percentile confidence intervals of the mediation effects with 1000 Monte Carlo draws. Robust (heteroskedasticity-consistent) standard errors were computed for the quasi-Bayesian simulations for the mediation effects. Because each analysis aimed at relating one set of Tukey-corrected behavioural findings to relevant AEP to get at underlying mechanism, these analyses were not corrected for multiple comparisons.

3. Results

3.1. Behavioural Results

First, we report the results for accuracy, then we report the results for the VE and PE scores. A summary of all behavioural results is provided in Table 2.

3.1.1. Accuracy

This analysis focuses on the comparison of the facilitation effect of audiovisual integration (objective 1a) and other visual predictive cues (objective 1b) on speech perception in younger and older adults.

Results of the LMM analyses (marginal fixed effects) revealed a main effect of Modality ($F_{(2,330)}$ = 194.503, p <.0001). Tukey-corrected post hoc contrasts indicated that accuracy was higher in A compared to V (β = 13.78, SE = 0.797, p <.0001) and in AV compared to V (β = 3.651, SE = 0.206, p = <.0001). There was also an interaction between Age Group and Modality ($F_{(2,330)}$ = 6.215, p = .002). Tukey-corrected post hoc contrasts indicated that accuracy in the V condition was lower for older compared to younger adults (β = 8.350, SE =2.24, *p* = .001) (Figure 3A). Finally, the LMM analysis also revealed an interaction between Prediction Cue and Modality ($F_{(6,330)}$ = 39.422, p <.0001) (Figure 3B). Tukey-corrected post hoc contrasts indicated that, in the A condition, accuracy was higher in the WW condition compared to the Control condition (β = -2.99, SE = .996, p = .015). In the V condition, accuracy was higher in What compared to the Control condition (β = -23.9, SE = .996, p <.0001), in the When compared to the Control condition (β = -3.60, SE = .996, p = .002), and in WW compared to the Control condition (β = 23.9, SE = .996, p <.0001). Of all the cues, the When cue had the smallest impact on performance. Finally, in the AV condition, the cues did not affect accuracy. The descriptive statistics and the detailed results of the LMM analyses as well as the pairwise contrasts are provided in Supplementary Material 2, along with a figure showing accuracy in all experimental conditions separately.

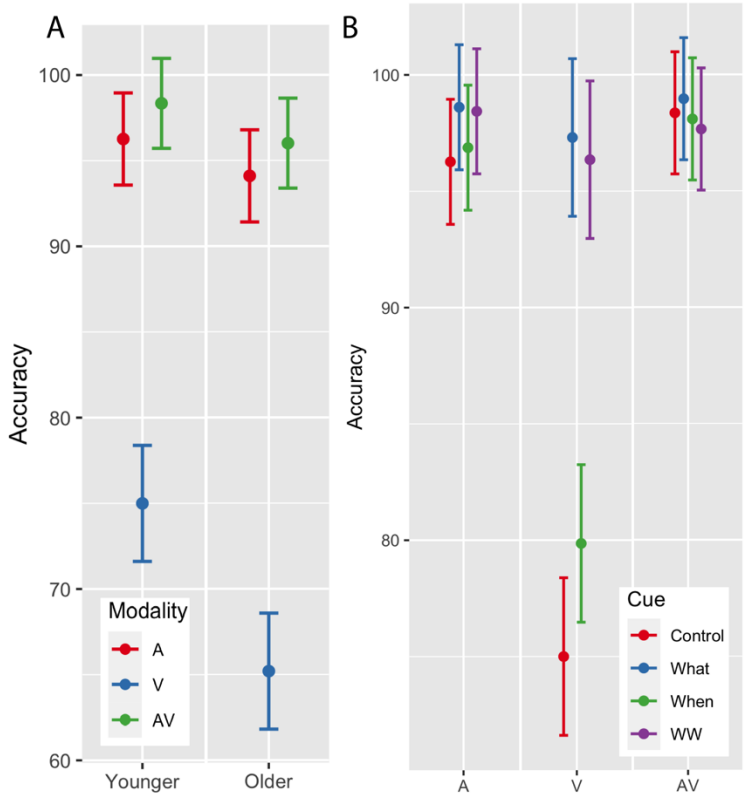

Figure 3 Accuracy results (marginal means). A. The plots display Accuracy as a function of Modality, separately for each group. B. The plots display Accuracy as a function of Modality, separately for each cue. Refer to the text for the list of the significant contrasts. The error bars represent the confidence intervals of the marginal means.

3.1.2. VE Scores

The analysis of the VE scores aimed to determine whether adding natural speech movements to auditory speech facilitates auditory speech perception in older adults through audiovisual integration (objective 1a). Results of the LMM analyses (marginal fixed effects) revealed a main effect of Age Group (Younger > Older) ($F_{(1,28)} = 8.469$, $p = .007$) (Figure 4A).

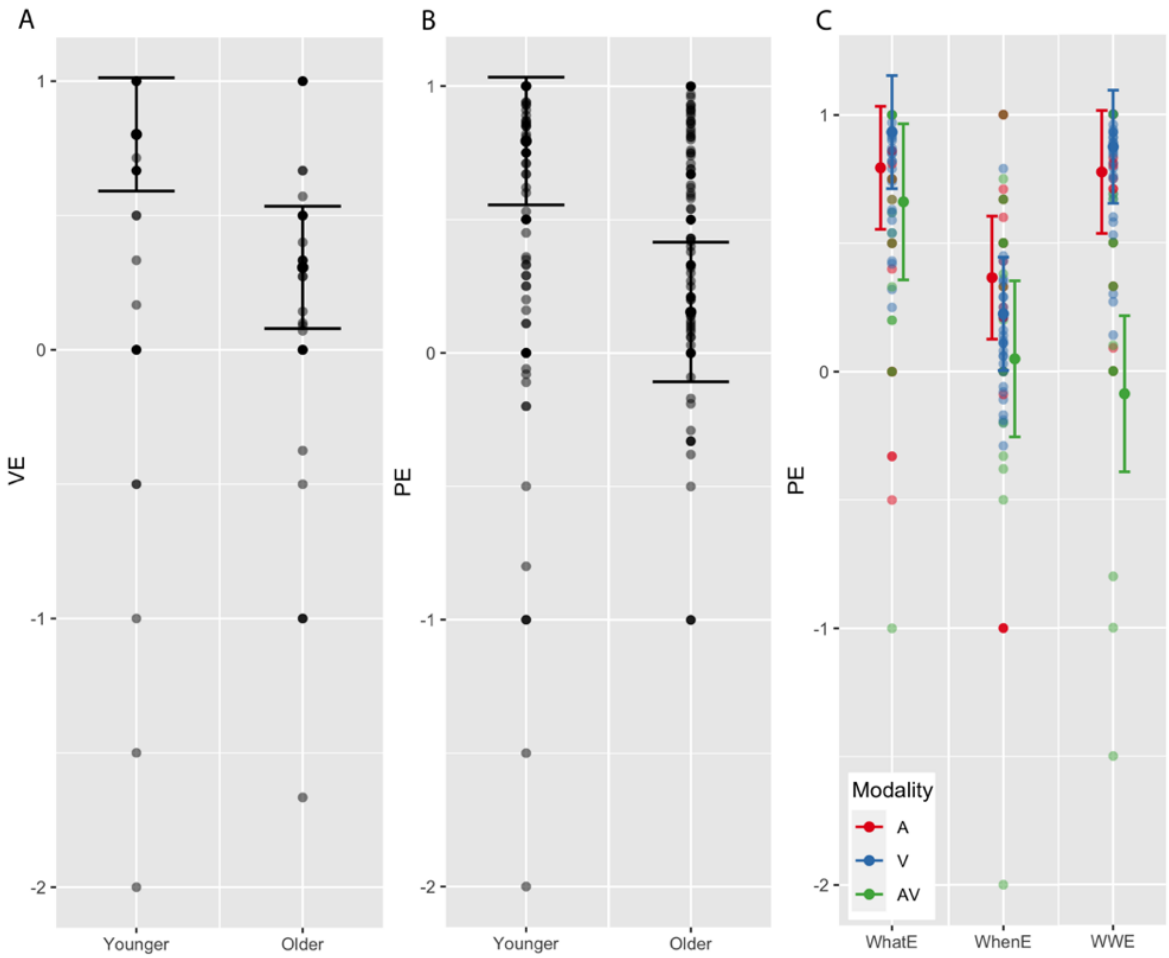

Figure 4 VE and PE results (marginal means). A. Age Group differences in VE. The groups are represented on the X-axis while VE is displayed in the y-axis. B. Age Group differences in PE. C. Cue by Modality interaction on PE. The Cue conditions are represented on the X-axis while PE is displayed in the y-axis. Modalities are colour coded. In all graphs, the error bars represent standard errors of the marginal means. Each dot is a participant.

3.1.3. PE Scores

The PE scored aimed to determine whether adding other visual Prediction Cues to the speech signal would facilitate speech perception similarly in young and older adults (objective 1b). The LMM results (marginal means) revealed a main effect of Age Group ($F_{(1,29)} = 11.822$, $p = .002$), with overall higher PE values in the younger adults (Figure 4B). The analysis also showed an interaction between Modality and Prediction Cue ($F_{(4,173)} = 3.646$, $p = .007$) (Figure 4C). This interaction revealed that PE scores did not differ as a function of Modality in the What and When conditions, but only in the WW condition. The condition with the most information (WW_{AV}) showed the lowest PE. Tukey-corrected

post hoc tests revealed that in the WW condition, the effect of adding a prediction cue on performance was stronger for V then AV (β = -0.647 SE = 0.122, p = .001) and for A compared to AV (β = 0.555, SE = 0.129, p = .001), while the effect was similar for the unisensory conditions A and V (β = -0.092, SE = 0.109, p = .995). The descriptive statistics and the detailed results of the LMM analyses and are provided in Supplementary Material 4.

Table 2. Type III Tests of Fixed Effects (marginal) for the behavioural analyses

Effects and interactions	DF		F	p
A. Accuracy				
Age Group	1	29	1.095	0.304
Prediction cue	3	330	1.348	0.2587
Modality	2	330	109.03	**<.0001**
Hearing	1	29	0.266	0.6101
Age Group * Prediction cue	3	330	0.373	0.7727
Age Group * Modality	2	330	6.215	**0.0022**
Prediction cue * Modality	6	330	39.422	**<.0001**
Age Group * Prediction cue * Modality	6	330	1.559	0.1583
B. VE				
Age Group	1	28	8.469	**0.007**
Prediction cue	3	47	1.872	0.147
Hearing	1	28	1.145	0.294
Age Group * Prediction cue	3	47	0.871	0.463
C. PE				
Age Group	1	29	11.822	**0.002**
Prediction cue	2	173	4.741	**0.010**
Modality	2	173	1.257	0.287
Hearing	1	29	0.540	0.468
Age Group * Prediction cue	2	173	2.366	0.097
Age Group * Modality	2	173	1.699	0.186
Prediction cue * Modality	4	173	3.646	**0.007**
Age Group * Prediction cue * Modality	4	173	1.042	0.387

3.2. EEG Analyzes: Audiovisual Integration (additive model)

The first set of EEG analyses addresses objective 2a, to compare audiovisual speech integration in younger and older adults using an additive model (AV vs. A+V). To achieve this goal, we compared the AV signal to the sum of the unisensory signals (A+V) in young and older adults. Figure 5 shows the average neurophysiological response for each Age Group and each modality (AV, A+V). The descriptive and inferential statistics are presented in Supplementary Materials 5-6 (P1), 7-8 (N1), 9-10 (P2) and 11-12 (N2). In line with previous EEG studies (e.g., Besle et al., 2004; Klucharev et al., 2003;

Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; van Wassenhove et al., 2005), the difference between AV and A+V was significant on the N1/P2/N2 complex in both younger (Figure 5A) and older adults (Figure 5B). As can be seen in the figure, overall, the amplitude of N1 and N2 was more negative for A+V than for AV, and the amplitude of P2 was more positive for A+V than AV. In terms of latency, the A+V signal had a longer N1 and P2 peak latency. As detailed in the following paragraphs and illustrated in Table 3, Age Group differences were found on P2 and N2 (latency).

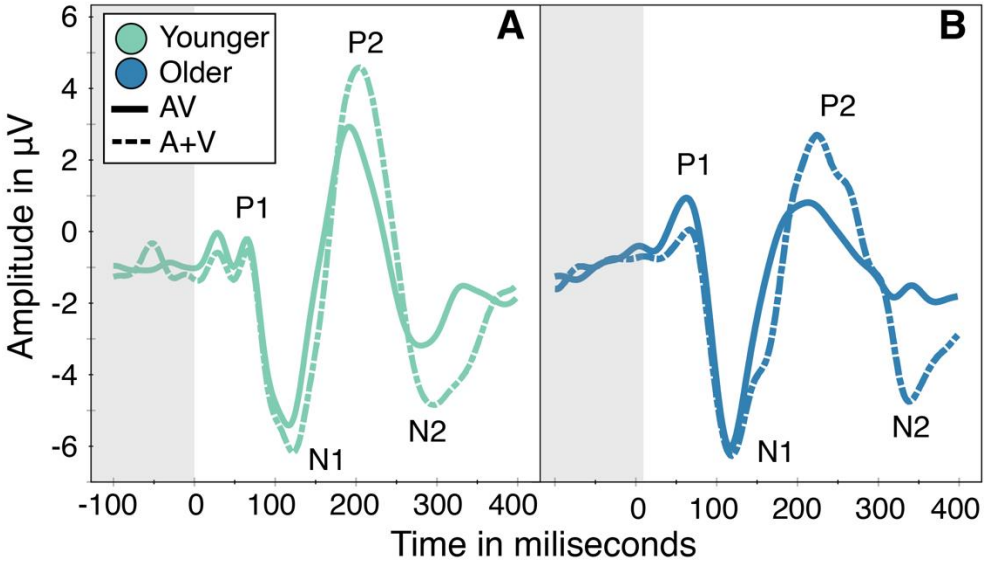

Figure 5. Group average electrophysiological responses for the younger adults (A) and older adults (B). In both panels, the response in the AV condition is represented as a plain line while the A+V is represented as a dotted line. In each plot, time is displayed in the X-axis while amplitude in µV is displayed in the y-axis. The 0 ms corresponds to the time of stimulus presentation. The shaded area represents the baseline period. The 4 components of interest are identified in each plot: P1, N1, P2 and N2.

The LMM analyses (marginal means) revealed no effect on P1. For **N1 amplitude**, there was a marginally significant effect of Modality ($F_{(1,30)} = 4.198$, $p = .049$), with a more negative signal amplitude for A+V compared to AV (Figure 6A). For **N1 latency**, there was also a significant effect of Modality ($F_{(1,30)} = 7.169$, $p = .012$), with a longer N1 latency for A+V compared to AV (Figure 6B). For **P2 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,30)} = 23.039$, $p <.001$), with higher P2 amplitude for A+V compared to AV (Figure 6C). For **P2 latency**, the LMM analyses revealed a main effect of Age Group ($F_{(1,30)} = 4.487$, $p = .043$), with P2 peaking later in the older

compared to the younger group (Figure 6D). There was also a main effect of Modality on P2 latency ($F_{(1,30)}$ = 4.983, p = .033), with a longer latency for A+V compared to AV (Figure 6E).

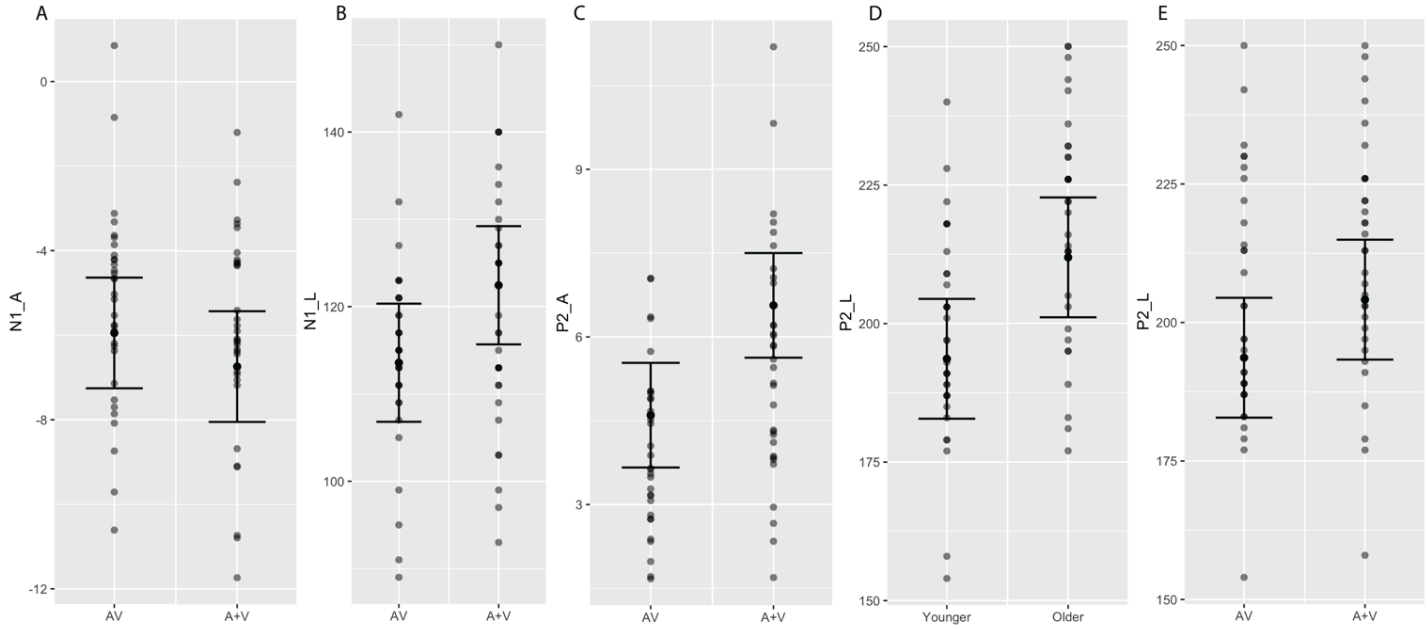

Figure 6 Results for N1 and P2 for the integration analyses (additive model). A. Modality difference in N1 amplitude. B. Modality difference in N1 latency (N1-L). C. Modality difference in P2 amplitude (P2_A). D. Age difference in P2 latency (P2_L). E. Modality difference in P2 latency (P2_L). In all figures, N1/P2 amplitude/latency is displayed in the y-axis. Each dot is a participant. The error bars represent standard errors of the marginal means.

For **N2 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,30)}$ = 22.341, p <.001), with a more negative N2 amplitude for A+V compared to AV (Figure 7A). For **N2 latency**, the LMM analyses revealed a main effect of Age Group ($F_{(1,29)}$ = 5.872, p = .022), with N2 peaking later in the older compared to the younger group (Figure 7B).

Table 3 summarizes all effects for this set of analyses.

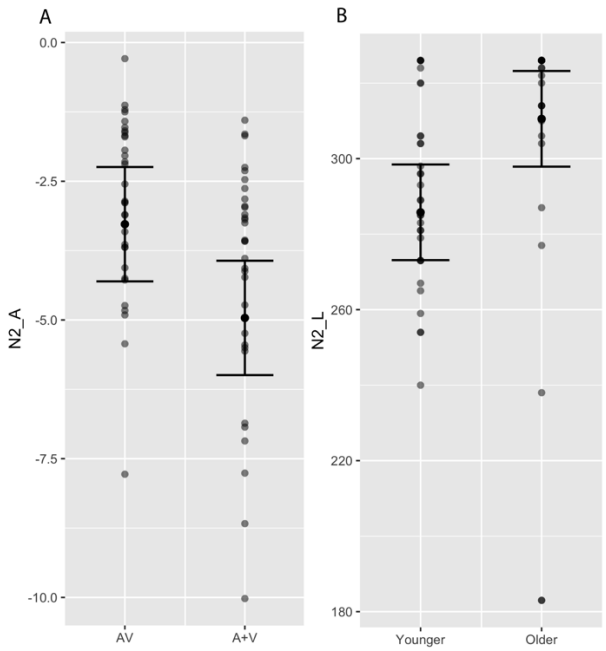

Figure 7 Results for N2 for the integration analyses (additive model). A. Modality difference in N2 amplitude. B. Age difference in N2 latency. D. Modality difference in P2 latency. N2 amplitude/latency is displayed in the y-axis. Each dot is a participant. In all graphs, the error bars represent standard errors of the marginal means.

Table 3. Type III Tests of Fixed Effects (marginal) for the EEG integration analyses (additive model)

Effects and interactions	DF		F	p
A. P1 amplitude				
Age Group	1	29	2.176	0.151
Modality	1	29	0.090	0.767
Hearing	1	29	0.030	0.865
Age Group * Modality	1	29	0.202	0.656
B. P1 latency				
Age Group	1	28	0.968	0.334
Modality	1	25	0.055	0.816
Hearing	1	28	0.005	0.943
Age Group * Modality	1	25	0.566	0.459
C. N1 amplitude				
Age Group	1	29	1.155	0.291
Modality	1	30	4.198	**0.049**
Hearing	1	29	4.605	**0.040**
Age Group * Modality	1	30	0.182	0.672
D. N1 latency				
Age Group	1	29	0.075	0.787
Modality	1	30	7.169	**0.012**
Hearing	1	29	2.555	0.121
Age Group * Modality	1	30	1.923	0.176
E. P2 amplitude				
Age Group	1	29	2.348	0.136
Modality	1	30	23.039	**<.0001**
Hearing	1	29	0.000	0.995
Age Group * Modality	1	30	2.147	0.153
F. P2 latency				
Age Group	1	29	4.487	**0.043**
Modality	1	30	4.983	**0.033**
Hearing	1	29	0.499	0.485
Age Group * Modality	1	30	0.248	0.622

		G. N2 amplitude		
Age Group	1	29	0.680	0.416
Modality	1	30	22.341	**0.0001**
Hearing	1	29	0.017	0.898
Age Group * Modality	1	30	1.897	0.179
		H. N2 latency		
Age Group	1	29	5.872	**0.022**
Modality	1	30	1.326	0.259
Hearing	1	29	0.080	0.780
Age Group * Modality	1	30	1.618	0.213

3.3. EEG Analyzes: Temporal and Phonetic Visual Predictive Cues

The second set of EEG analyses addresses objective 2b, to test whether adding articulatory movement and temporal and/or phonetic visual cues would facilitate neural processing of speech similarly for younger and older adults (AV vs. A, AVwhat vs. Awhat, AVwhen vs. Awhen and AVwhat-when vs. Awhat-when). To address this objective, first, we report the average response for each Age Group (younger, older) and each Predictive cue condition (Control, What, When, WW), separately for each modality (A, V, AV), in Figure 8. Though the visual condition was not analyzed, we present it in the figure for the sake of transparency and completeness. As can be seen in the Figure, the average response patterns were similar for the A and AV conditions, with reduced responses in the AV compared to the A condition, for both younger and older adults.

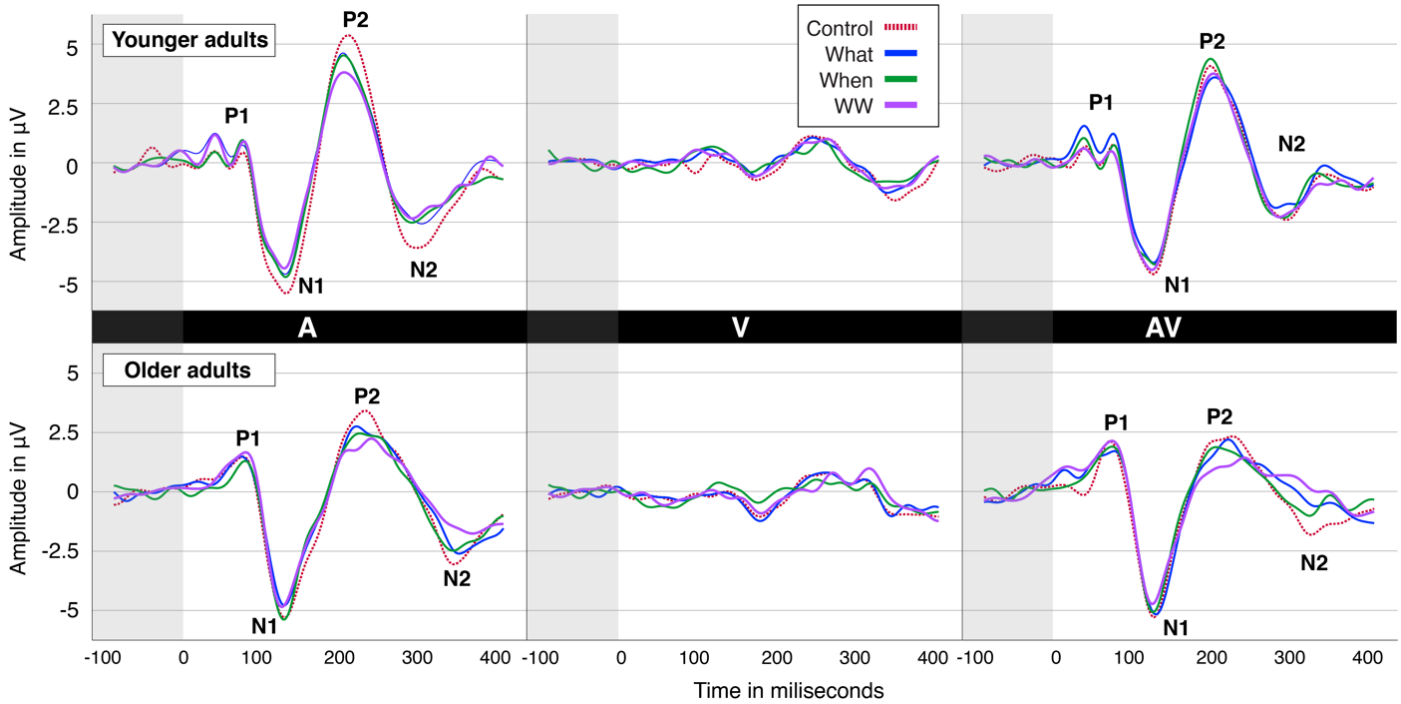

Figure 8 Group average electrophysiological responses for the younger adults (top panels) and older adults (bottom panels), separately for the A condition (left panels), V condition (central panels) and AV condition (right panel). In each plot, time is displayed in the X-axis while amplitude in µV is displayed in the y-axis. The 0 ms corresponds to the time of stimulus presentation. The shaded area represents the baseline period. The 4 components of interest are identified in each plot: P1, N1, P2 and N2.

For each component (P1, N1, P2 and N2), Q-Q plot and histograms were computed, which revealed that the residuals followed a normal or fairly normal distribution. The descriptive statistics for each dependent variable and the detailed results of the LMM analyses are provided in Supplementary Materials 13-14 (P1), 15-16 (N1), 17-18 (P2) and 19-20 (N2). As detailed in the following paragraphs and illustrated in Table 4, Age Group differences were found on P2 (both amplitude and latency) and N2 (latency).

For **P1 and N1 latency,** the marginal fixed effect tests revealed no main effects and no interaction. For **N1 amplitude**, the LMM analyses revealed a main effect of Modality ($F_{(1,210)} = 6.451$, $p = .012$). Tukey-corrected post hoc tests showed facilitation for AV compared to A ($\beta = -0.344$, SE = 0.26, p = .0324) on N1 amplitude (Figure 9A). There was also a main effect of Prediction Cue on N1

amplitude ($F_{(3,210)}$ = 4.913, p = .003). Tukey-corrected post hoc tests were conducted to decompose the effect of Cue on N1 amplitude. The contrasts showed facilitation for What compared to the control condition (β = -0.5469, SE = 0.172, p = .0092), and for WW compared to the Control condition (β = 0.344, SE = 0.172, p = .0063) (Figure 9B).

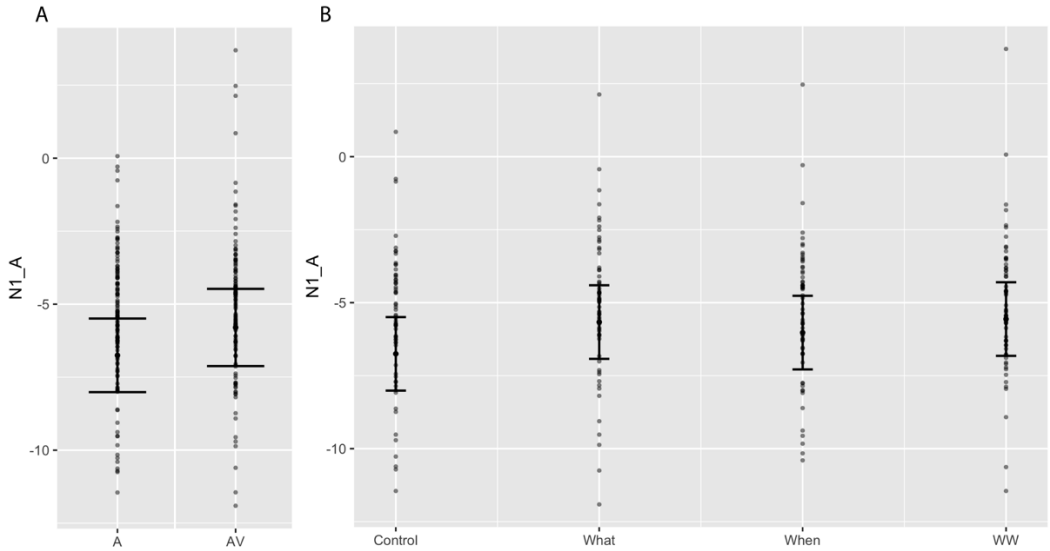

Figure 9 N1 amplitude (marginal means). A. Modality differences in N1 amplitude (µV; N1_A). Each modality is represented on the X-axis while amplitude is displayed in the y-axis. Each dot is a participant. B. Prediction Cue differences in N1 amplitude (µV; N1_A). The Cue conditions are represented on the X-axis while amplitude is displayed in the y-axis. Modalities are colour coded. Il all graphs, the error bars represent standard errors of the marginal means.

For **P2 amplitude**, the LMM analyses (marginal fixed effects) revealed a main effect of Age Group ($F_{(1,29)}$ = 10.958, p = .0003) (Younger > Older) as well as a main effect of Modality ($F_{(1,210)}$ = 20.25, p < .001) (A > AV), and Prediction Cue ($F_{(3,210)}$ = 10.30, p < .001). There was also a 2-way interaction between Age Group and Modality ($F_{(1,210)}$ = 5.642, p = .0184) and between Cue and Modality ($F_{(3,210)}$ = 6.12, p = .001) on P2 amplitude. Finally, the LMM analyses also revealed a 3-way interaction between Age Group, Modality and Prediction Cue ($F_{(3,210)}$ = 3.468, p = .0171) on P2 amplitude. Tukey-corrected post hoc tests were conducted to decompose the 3-way interaction on P2 amplitude. These tests showed P2 amplitude facilitation for AV compared to A in the Control condition

(β = 1.612, SE = 0.358, p = .0003), and a similar trend in the What condition (β = 1.049, SE = 0.358, p = 072), for younger adults, but no such P2 amplitude facilitation in older adults (Figure 10A). For **P2 latency**, we found a main effect of Age Group ($F_{(1,29)}$ = 8.439, *p* = .007), with longer P2 peak latency in the older group (Figure 10B).

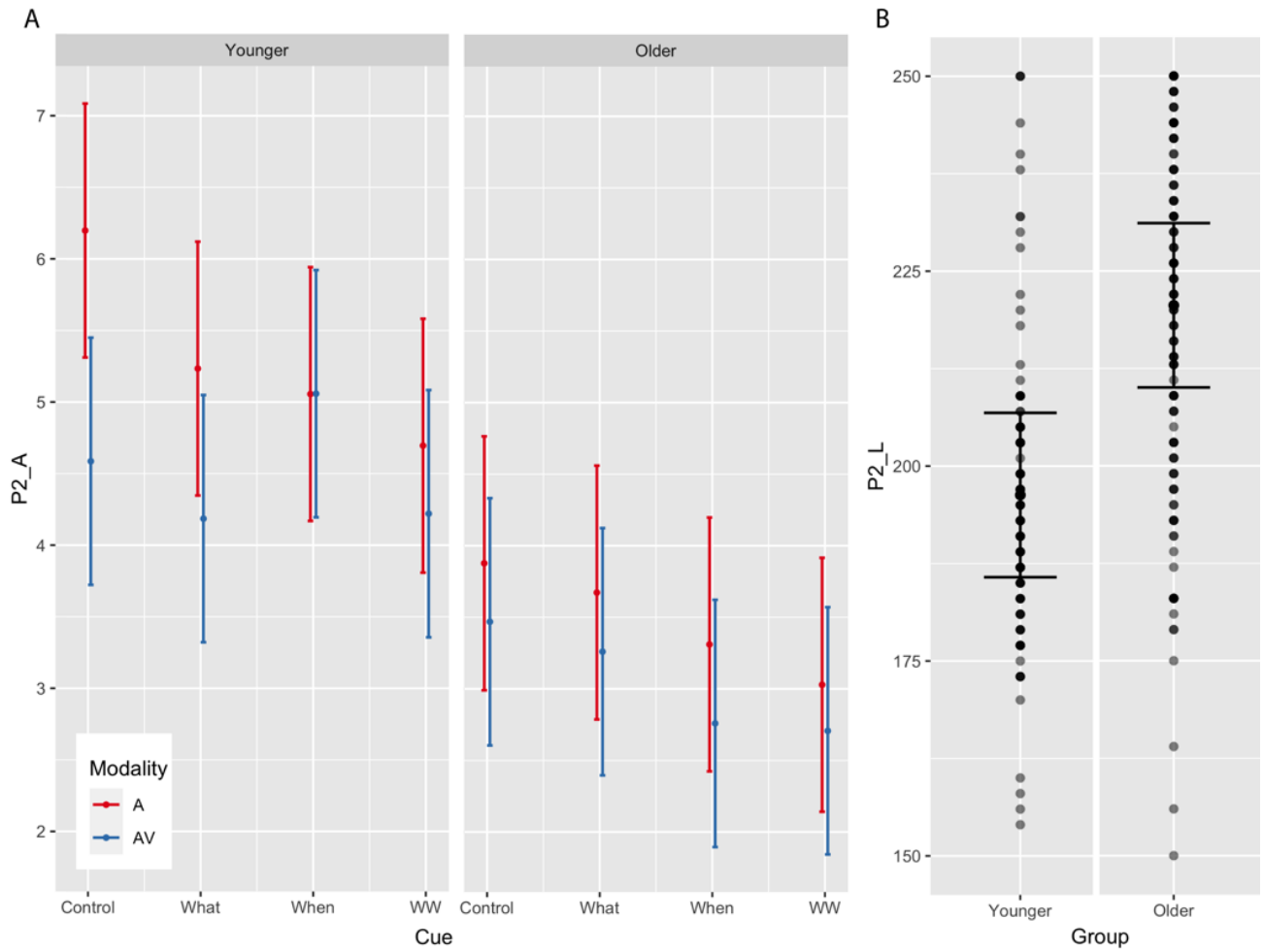

Figure 10. P2 results (marginal means). A. 3-way interaction between Age Group, Cue and Modality on P2 amplitude (P2_A). B. Group difference in P2 latency (P2_L). In all graphs, P2 amplitude/latency is displayed in the y-axis. The error bars represent standard errors of the marginal means. Each dot is a participant.

For **N2 amplitude**, the LMM analyses (marginal fixed effects) revealed a main effect of Prediction Cue ($F_{(3,210)}$ = 7.503, *p* = .0001) and a main effect of Modality ($F_{(1,210)}$ = 9.170, *p* = .003). Tukey-corrected post hoc tests were conducted to decompose the effect of Modality. These contrasts showed N2 amplitude facilitation for AV compared to A (β = -0.661, SE = 0.237, p = .0058) (Figure

11A). There was also as a marginally significant interaction between Age Group and Cue on N2 amplitude ($F_{(3,210)} = 2.6742$, $p = .048$) (Figure 11B). Tukey-corrected post hoc tests were conducted to decompose the interaction between Age Group and Cue on N2 amplitude. The tests revealed that, in younger adults, there was a difference in N2 amplitude between the Control and the What (β = -0.697, SE = 0.215, p = .03), When (β = -0.746, SE = 0.215, p = .014), and WW conditions (β = -0.831, SE = 0.215, p = .004). In the older adults, there was a difference in N2 amplitude between the Control and the What (β = -0.653, SE = 0.215, p = .054) and WW conditions (β = -0.779, SE = 0.215, p = .009), but not between the Control condition and the When condition (β = -0.312, SE = 0.215, p = .833). For **N2 latency**, the LMM analyses (marginal fixed effects) revealed a main effect of Age Group ($F_{(3,210)} = 4.1524$, $p = .007$), with longer latency for the older compared to the younger adults (Figure 11C).

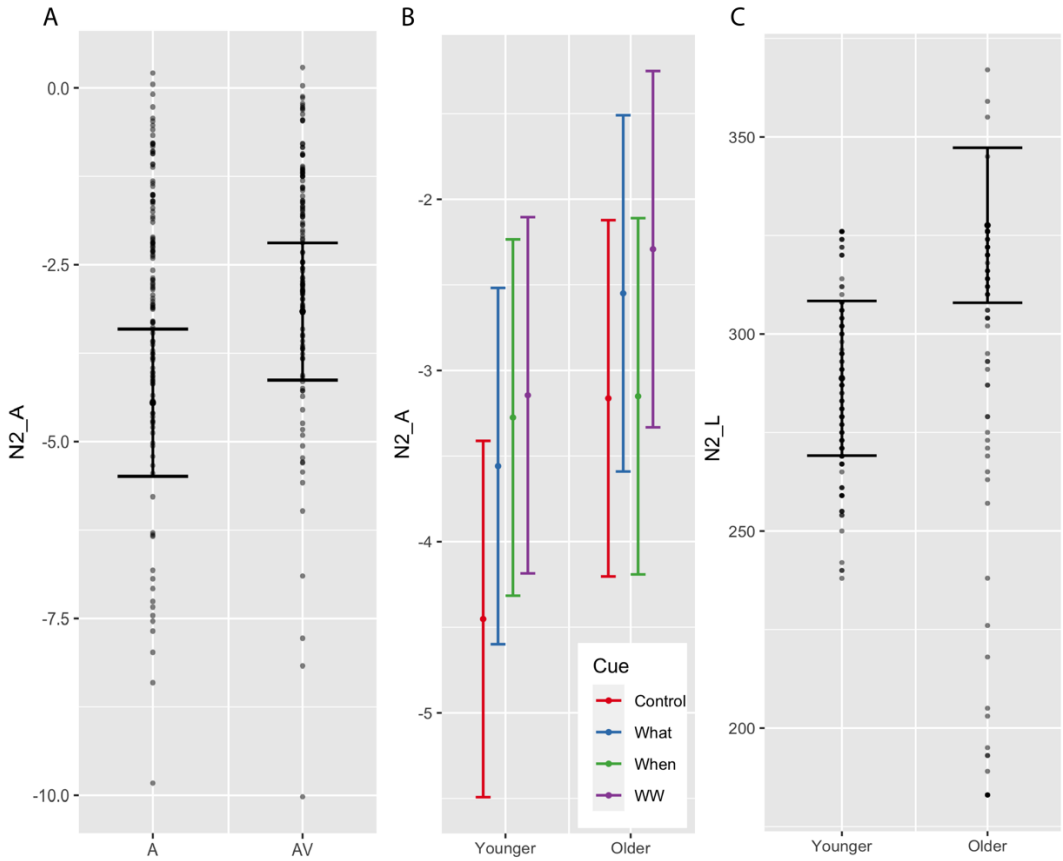

Figure 11. N2 results (marginal means). A. Main effect of Modality on N2 amplitude (N2_A). B. 2-way interaction between Age Group and Prediction Cue on N2 Amplitude (N2_A). C. Group difference in N2 latency (N2_L). In all graphs, the error bars represent standard errors of the marginal means. Each dot is a participant.

Table 4. Type III Tests of Fixed Effects (marginal) for the EEG analyses

Effects and interactions	DF		F	p
B. P1 amplitude				
Age Group	1	29	2.753	0.108
Prediction cue	3	210	0.703	0.551
Modality	1	210	0.412	0.522
Hearing	1	29	0.034	0.854
Age Group * Prediction cue	3	210	0.469	0.704
Age Group * Modality	1	210	0.086	0.770
Prediction cue * Modality	3	210	0.027	0.994
Age Group * Prediction cue * Modality	3	210	0.631	0.596
B. P1 latency				
Age Group	1	29	0.017	0.896
Prediction cue	3	198	1.579	0.196
Modality	1	198	0.150	0.699
Hearing	1	29	0.075	0.786
Age Group * Prediction cue	3	198	1.540	0.205
Age Group * Modality	1	198	2.223	0.138
Prediction cue * Modality	3	198	0.356	0.785
Age Group * Prediction cue * Modality	3	198	1.448	0.230
I. N1 amplitude				
Age Group	1	29	2.417	0.131
Prediction cue	3	210	4.913	**0.003**
Modality	1	210	6.451	**0.012**
Hearing	1	29	3.251	0.082
Age Group * Prediction cue	3	210	1.107	0.347
Age Group * Modality	1	210	1.904	0.169
Prediction cue * Modality	3	210	2.163	0.094
Age Group * Prediction cue * Modality	3	210	1.353	0.258
J. N1 latency				
Age Group	1	29	0.011	0.917
Prediction cue	3	210	0.326	0.807
Modality	1	210	2.703	0.102
Hearing	1	29	2.058	0.162
Age Group * Prediction cue	3	210	1.336	0.264
Age Group * Modality	1	210	0.338	0.562
Prediction cue * Modality	3	210	0.414	0.743
Age Group * Prediction cue * Modality	3	210	0.016	0.997
K. P2 amplitude				
Age Group	1	29	10.96	**0.003**
Prediction cue	3	210	10.30	**<.0001**
Modality	1	210	20.25	**<.0001**
Hearing	1	29	0.00	0.955
Age Group * Prediction cue	3	210	1.45	0.231
Age Group * Modality	1	210	5.64	**0.018**
Prediction cue * Modality	3	210	6.12	**0.001**
Age Group * Prediction cue * Modality	3	210	3.47	**0.017**
L. P2 latency				
Age Group	1	29	8.439	**0.007**
Prediction cue	3	210	0.322	0.810
Modality	1	210	0.441	0.507
Hearing	1	29	0.271	0.607
Age Group * Prediction cue	3	210	1.046	0.373
Age Group * Modality	1	210	0.529	0.468
Prediction cue * Modality	3	210	0.511	0.675
Age Group * Prediction cue * Modality	3	210	1.576	0.196
M. N2 amplitude				
Age Group	1	29	2.476	0.127
Prediction cue	3	210	7.503	**0.0001**
Modality	1	210	9.170	**0.003**
Hearing	1	29	0.402	0.531
Age Group * Prediction cue	3	210	2.674	**0.048**
Age Group * Modality	1	210	1.906	0.169
Prediction cue * Modality	3	210	2.117	0.099
Age Group * Prediction cue * Modality	3	210	2.029	0.111

N. N2 latency				
Age Group	1	29	7.476	**0.011**
Prediction cue	3	210	0.104	0.958
Modality	1	210	0.117	0.733
Hearing	1	29	0.246	0.624
Age Group * Prediction cue	3	210	1.456	0.228
Age Group * Modality	1	210	0.585	0.445
Prediction cue * Modality	3	210	0.355	0.786
Age Group * Prediction cue * Modality	3	210	0.065	0.979

3.4. Electrophysiology-Behaviour Relationship

This final set of analysis addresses the third objective of the study, to examine the relationship between speech perception performance and AEPs in young and older adults to shed new lights on brain aging and its impact on human behaviour. This analysis focused on overall VE and overall PE scores, and on AEPs (P2A, P2L, N2A and N2L) measures that showed Age Group effects or interactions in the statistical analyses detailed in the previous subsections.

For the analysis with overall **VE as dependent variable**, the predictor variables were the AEP components that showed an effect of Age Group or an interaction with Age Group: P2 amplitude average (basic analysis), P2 latency average (basic and integration analyses), N2 amplitude average (basic analysis) and N2 average latency (integration analysis). There were thus 5 analyses, which are detailed in **Supplementary Material 21**. The analyses revealed a significant mediation effect of P2 amplitude on overall VE for the older adults. As shown in Figure 12B and C, in older adults with higher overall VE, the amplitude of P2 was less positive.

For the analysis with overall **PE as dependent variable**, the predictor variables were the AEP components that showed an effect of Age Group or an interaction with Age Group: P2 amplitude average (basic analysis), P2 latency average (basic and integration analyses), N2 amplitude average (basic analysis) and N2 average latency (integration analysis). There were thus 5 analyses, which are detailed in **Supplementary Material 22**. There were no significant mediation effects of AEPs and no direct effect of AEPs on overall VE scores.

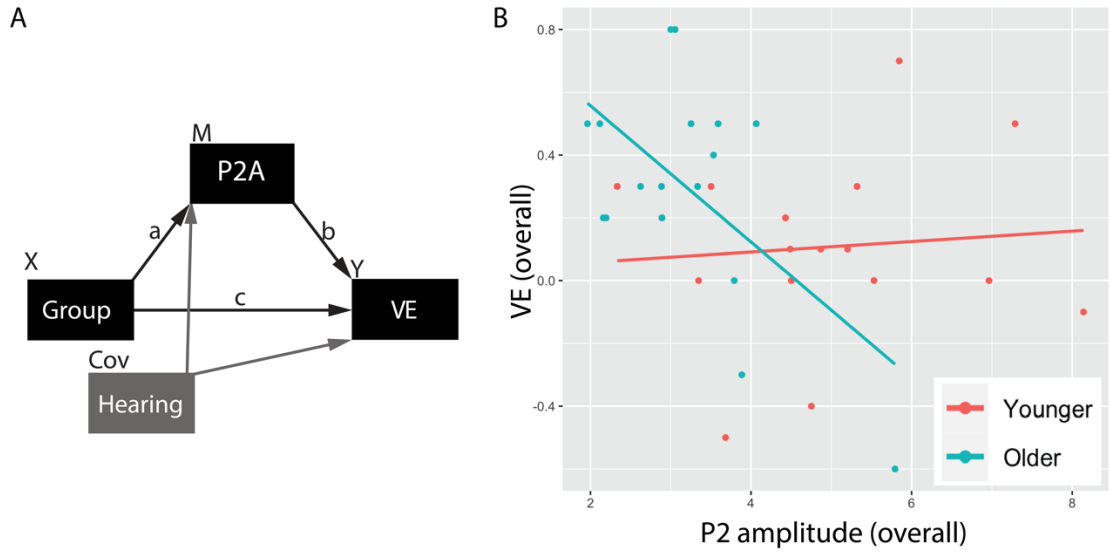

Figure 12 Significant mediation results illustrating the relationship between VE and P2 amplitude for younger and older adults. A. Mediation model for the significant mediation. X = the predictor variable; M = the mediator variable; Y = the dependent variable; Cov = covariate; a represents the effect of X on M; b represents the effect of M on Y. B. The scatterplot displays the mediating effect of Age Group on VE through P2 amplitude. In both the scatterplots, each dot represents a participant. The red dots represent the younger participants while the blue dots represent the older participants.

4. Discussion

The general objective of the present study was to fill a knowledge gap about the neural processing of auditory speech in aging under different levels of prediction. By measuring multiple early and late auditory evoked potentials (P1-N1-P2 and N2), we aimed to shed new lights on the locus of age-related differences in neural speech processing. To examine the effect of prediction, we compared the processing of auditory and audiovisual speech, and we manipulated prior knowledge on auditory syllables by presenting participants with visual information indicative of the temporal unfolding (*when*) or the phonetic content (*what*) of auditory syllables. Based on Winneke and Phillips (2011), we hypothesized that, compared to younger adults, older adults would show similar or enhanced audiovisual and predictive effects. Specifically, we expected an enhanced facilitation of AEPs (shorter latency and/or reduced amplitude) during AV compared to A speech, as well as during the processing of temporal and phonetic predictive cues, reflecting a lifetime of experience with speech, or perhaps a

compensation strategy to overcome possible hearing loss. Given the well-established cognitive decline that occurs in aging (e.g., Park, Lautenschlager, Hedden, Davidson, Smith, & Smith, 2002; Salthouse, 1996; Salthouse, 2009), we expected that the auditory N2, which indexes cognitive and executive processes (e.g. Czigler, Csibra, & Ambró, 1997; Falkenstein, Hoormann, & Hohnsbein, 1999; Folstein & Van Petten, 2008), would show either an increased amplitude, reflecting compensatory activity, or perhaps a decline, reflecting a disruption in processing.

The main findings of the study are as follows: the behavioural results show that (1) there was a VE effect in both groups (as shown in Figure 4 and detailed in Table 2B), but, contrary to our prediction, it was lower in older compared to younger adults, (2) predictive cues facilitated speech recognition in younger and older adults alike (Table 2C). The EEG results showed that (3) age differences in AEPs are localized to later components (P2 and N2, see Tables 3 and 4), suggesting that aging predominantly affects higher-order cortical processes related to speech processing rather than lower-level auditory processes. (4) Specifically, the latency of the P2 and N2 components were delayed in older adults (Tables 3 and 4). Moreover, there was reduced AV facilitation on P2 amplitude in older adults (Figure 10), while there was a reduced effect of the When cue on N2 amplitude for older compared to younger adults (Figure 11). Finally, the mediation analyzes on EEG-behaviour relationship revealed that (5) overall VE scores were associated with P2 amplitude (Figure 12).

4.1. Prediction Facilitation as a Function of Age

The main objective of the study was to examine whether adding visual cues (natural or unnatural) to the speech signal would facilitate speech perception similarly in younger and older adults, behaviourally (objective 1) and at the neural level (objective 2).

First, our behavioural results revealed that overall speech recognition accuracy in a straightforward syllable identification task performed in quiet was lower for older adults in the V condition, suggestive of poorer lipreading ability in older compared to younger adults. Here,

participants reported normal or corrected-to-normal vision, and they were allowed to wear their glasses or contact lenses during the experiment, although their visual acuity was not assessed. The low accuracy in the V modality could therefore reflect degraded sensory input or reduced visual processing efficiency. While the clinical relevance of reduced lipreading should be investigated in future work, this finding is consistent with previous studies showing that, compared to younger adults, middle-aged and/or older adults exhibit reduced lipreading abilities (Cienkowski & Carney, 2002; Sommers et al., 2005; Tye-Murray et al., 2010; Winneke & Phillips, 2011). Future studies need to determine if adequate correction for age-related visual impairment can restore lip-reading ability.

Despite a lower performance in the V condition, however, our study is the first to show that visual cues can improve speech processing accuracy in older adults. Accuracy in the visual modality was enhanced for the young and older adults when visual predictive cues were provided (especially the What and the WW cues). Moreover, the prediction effect (PE) scores did not show any effect of age groups. In the V modality, older adults improved maximally upon presentation of the dual (WW) cue, as did younger adults. This suggests that the visual cues were processed and that they were helpful to older adults. Given that visual speech information can be processed, even at the categorical level, by the human brain (O'Sullivan, Crosse, Di Liberto, & Lalor, 2016), the finding that prediction cues can improve lipreading in older adults has potential implications for the care and rehabilitation of older adults with communication difficulties, especially for those with significant hearing loss who rely more heavily on visual speech and audiovisual integration (Puschmann, Daeglau, Stropahl, Mirkovic, Rosemann, Thiel, & Debener, 2019). Identifying strategies to improve visual speech processing is of key importance for this population, but additional empirical evidence is needed to determine if visual cues such as the ones used in the present study have beneficial impacts on elderly populations with hearing impairments.

Though our results show that older adults are capable of integrating AV information to decipher speech, older participants in the present study exhibited poorer VE compared to younger ones, which

suggests that, though preserved, the ability to benefit from natural visual information is reduced in aging. Importantly, VE was unaffected by hearing. This result is in line with studies demonstrating that performance enhancement for AV speech compared to A speech is lower for older compared to younger adults (Tye-Murray et al., 2010; Yang & Ren, 2018), but it is at odds with other studies that did not find a decline in enhancement in aging (Avivi-Reich et al., 2018; Cienkowski & Carney, 2002; Ganesh et al., 2017; Sommers et al., 2005; Winneke & Phillips, 2011). Recent investigations have suggested an age-related change in the conditions needed for older adults to benefit from AV speech, rather than an absolute incapacity to benefit from an audiovisual signal, with older adults benefiting less when signal-to-noise ratio is low (Jansen et al., 2018; Stevenson et al., 2015) and when the visual information is degraded (Gordon & Allen, 2009). Crucially, in the present study, an age difference was observed even in quiet, using non-degraded stimuli. In addition to the context, it is possible that changes in AV integration are progressive. Indeed, in several previous studies, participants classified as "older" were, on average, younger than 60 (Jansen et al., 2018; Stevenson et al., 2015), which is young. In the present study, the older group was older (average of 67 years), and a group difference was observed in the absence of noise. Additionally, one cannot exclude that effects of dual-tasking and attention-sharing might be present in our experimental tasks, preventing older participants from maximally benefiting from the visual cues. Additional studies are needed to investigate the environmental conditions and participants characteristic (including visual acuity and cognitive capabilities) that affect audiovisual facilitation.

Together, these findings suggest that predictive coding capabilities, in general, may be slightly reduced in older adults. Predictive coding theories suggest that our perceptual experience is determined by a fine balance between internal predictions based on priors acquired over the course of our lifetimes and incoming sensory evidence (Clark, 2013; Friston, 2005, 2010; Rao & Ballard, 1999). Sensory evidence and priors are thought to be fused in a Bayesian way, leading to a prediction about the state of the world. Aging offers a unique opportunity to probe this notion, as the amount of phonetic

information accumulated, especially about talker variability, increases over the lifetime, as well as knowledge about communication contexts, while at the same time predictive coding mechanisms probability become more fine-tuned and efficient. Importantly, however, the precision of sensory processing degrades with age, in both the auditory and visual modalities. While the first factor (accumulated knowledge) strengthens the influence of predictions, the second (reduced sensory processing) lowers the influence of sensory evidence, especially unisensory signals. One possibility is that experience plays a stronger role than sensory decline, but our results suggest that impoverished sensory processing may be playing a stronger part in the interplay between these opposing forces. Additional studies are needed to compare predictive coding in younger and older adults with various levels of visual and auditory acuity to determine the threshold at which sensory processing becomes the dominant mechanisms driving predictive coding.

4.2. Age Differences in Electrophysiological Responses Are Localized to P2 and N2

A central objective of the present study was to shed new lights on the locus of age differences in the neural processing of speech sounds, which has major implications for understanding speech processing difficulties and to guide rehabilitation research and practice. This was achieved by investigating several ERP components: the P1-N1-P2 complex as well as the later N2 component.

Our results show that, controlling for hearing thresholds in the low frequency range (PTA), age differences in electrophysiological response to speech were predominantly located to P2 and N2. This suggests that speech processing difficulties may result from changes to higher-order cortical processes rather than lower-level auditory ones.

While several studies have reported age differences in the neural responses to speech in older adults, and more generally, the neural responses to sounds (e.g. Anderson & Karawani, 2020), uncertainty remains regarding the nature of these changes and how they affect speech perception performance. Previous AV speech studies investigating AEPs in aging have largely focused on the P1-

N1-P2 complex. The P1-N1-P2 complex represents the obligatory sensory-evoked response of the auditory system (Naatanen & Picton, 1987). The P1-N1-P2 complex can be used to indirectly assess the integrity of the central auditory system, up to the cortex. The earlier P1 primarily originates from the primary auditory cortex, and is thought to have subcortical contributions, including the reticular activating system (e.g., Erwin & Buchwald, 1987). The N1 component has been localized to the primary and secondary auditory associative regions (e.g., Eggermont & Ponton, 2002; Naatanen & Picton, 1987). It is related to stimulus detection and the encoding of auditory stimulus properties. Both components are obligatory auditory responses unaffected by attention. In early studies of syllable discrimination, Tremblay et al. reported evidence of neural disruptions in older adults in the form of delayed N1 latency, potentially reflecting age-related changes in neural synchrony (Tremblay, Piskosz, & Souza, 2002, 2003). A similar age-related delayed N1 latency was reported by Bidelman et al. during a vowel categorization task (Bidelman, Villafuerte, Moreno, & Alain, 2014). However, Soros and colleagues, using MEG, found no delays in P1 or N1m either during passive listening of rapid sequences of speech sounds, but instead stronger amplitude for older adults (Soros, Teismann, Manemann, & Lutkenhoner, 2009). Likewise, Roque et al. reported age differences in P1 during an auditory word identification task performed in quiet, with older adults exhibiting earlier P1 peak latency and larger amplitude compared to younger adults; but no difference on N1 was found (Roque, Karawani, Gordon-Salant, & Anderson, 2019). In the present study, the P1 and N1 showed no signs of age-related disruptions. This suggests that basic auditory processing during a simple syllable recognition in quiet task, at least in our sample, was largely unchanged with age.

In contrast to early components, the later auditory components—P2 and N2—showed evidence of age-related differences, with overall lower amplitude and longer latencies. Latency generally reflects the time point of the peak neurophysiological response peak relative to the eliciting stimulus. Latency is related to neural conduction time and site of excitation: the time it takes for the sound to travel through the peripheral auditory system to the place of excitation in the central nervous system (Alain &

Tremblay, 2007). Earlier peaks are therefore presumed to reflect the output of a process that is completed earlier in time. Several studies have shown an age-related increase in P2 latency (e.g., Billings, Penman, McMillan, & Ellis, 2015; Czigler, Csibra, & Csontos, 1992; Goodin, Squires, Henderson, & Starr, 1978; Iragui, Kutas, Mitchiner, & Hillyard, 1993; Tremblay, Billings, & Rohila, 2004; Tremblay et al., 2002, 2003). An increase in latency could reflect increased neural conduction time with normal aging, or more laborious neural processing which could be related to decline in the structure of the cerebral cortex, for example.

The auditory P2 is thought to reflect synchronous neural activation in the thalamic-cortical segment of the central nervous system, mainly originating from the supratemporal plane of the auditory cortex (e.g., Naatanen & Picton, 1987). Previous neurophysiological studies, as well as the present one, have shown that the N1/P2 complex occurs earlier and its amplitude is lower for AV compared to unimodal (A) speech processing (e.g., Besle et al., 2004; Klucharev et al., 2003; Treille et al., 2014a; Treille et al., 2017; Treille et al., 2014b; Treille et al., 2018; van Wassenhove et al., 2005). P2 is often considered as an index of AV integration. Here, while P2 latency showed an age-related delay, it showed no evidence of interaction between group and modality, either in the basic analysis or in the integration analyses, suggesting that AV integration is taking place in older adults in a manner that is similar to younger adults. However, in terms of amplitude, there was some evidence of reduced P2 facilitation in older adults compared to younger adults (i.e. in the Control and the What conditions). These findings suggest that P2's sensibility to prediction is slightly reduced in aging. Interestingly, the brain-behaviour analyses revealed a relationship between P2 amplitude difference and overall VE scores. Older adults who exhibited facilitation (overall lower P2 amplitude) also exhibited a higher overall VE score, suggesting that a lowering of P2 amplitude with age may be normal and even beneficial to AV speech processing, perhaps reflecting increased experience and reduced processing need. Together, these results suggest that normal age-related changes to P2 affect speech processing skills in older adults, with those exhibiting a more efficient audiovisual integration process, as reflected

by lower P2 amplitude, maintaining better AV speech skills. This notion is consistent with a prior study that found that increased N1-P2 signal was negatively correlated with speech classification performance in young and older adults (Bidelman et al., 2014). The finding of reduced amplitude and delayed P2 in older adults suggest that audiovisual integration may be less energy consuming in older adults. Interestingly, Anderson and Karawani (2020) have suggested that an imprecise stimulus representation associated with age could lead to delays in P2. Future studies with larger sample sizes and using more challenging speech tasks will help clarify the mechanism that underlies age-related changes in P2, and more generally, in audiovisual integration for speech. Critically, such analyses will need to take into account not just hearing but also visual capabilities.

In addition to age differences in the auditory P2, we also found significant differences in the auditory N2. Similar to P2, N2 was generally delayed in older adults, but it also showed a decline in facilitation effects compared to younger adults. The N2 is a broad, slow negativity that occurs in the 140–300 ms latency range after stimulus onset. It is thought to index attentional discrimination process and response monitoring. Previous studies have shown delayed N2 latency for older compared to younger adults (Dushanova & Christov, 2013; Goodin et al., 1978; Schiff, Valenti, Andrea, Lot, Bisiacchi, Gatta, & Amodio, 2008) and lower N2 response amplitudes in older compared to younger adults during a speech perception in noise task (Billings et al., 2015). In a series of studies, Czigler et al. showed that attention-related processes in the visual modality, as indexed by event-related potentials—have longer latency in the elderly (Czigler et al., 1997). This delay in the posterior N2 (selection negativity) was considered as evidence for age-related delay in stimulus evaluation. As a consequence of the slowing down of attentional processes, as indicated by the slowing of the auditory N2, the use of prediction in speech perception may become less efficient in older adults, however, we did not find significant relationship between N2 and behaviour.

5. Conclusion

Our results indicate that aging affect speech processing at multiple levels, including sensory (visual) processing, multimodal integration and auditory attentional processes. Despite impaired lipreading abilities, older adults were able to use visual prediction cues to help decipher speech, suggesting preserved AV integration capabilities. These findings have important implications for rehabilitation research and interventions by suggesting that improving communication in older ages may rest on the need to enhance sensory processing, but also information integration and attention processes.

6. Acknowledgements

PT holds a Career award from the Fonds de la Recherche en Santé du Quebec (FRQ-S, #35016). We thank all participants.

7. CRediT author statement

Pascale Tremblay: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing- Original draft preparation. **Serge Pinto**: Conceptualization, Investigation, Writing—Reviewing and Editing. **Anahita Basirat**: Conceptualization, Writing—Reviewing and Editing. **Marc Sato**: Conceptualization, Methodology, Investigation, Resources, Project administration, Writing—Reviewing and Editing.

8. References

Alain, C., & Tremblay, K. (2007). The role of event-related brain potentials in assessing central auditory processing. *Journal of the American Academy of Audiology, 18*(7), 573-589.

Anderson, S., & Karawani, H. (2020). Objective evidence of temporal processing deficits in older adults. *Hear Res, 397*, 108053. doi:10.1016/j.heares.2020.108053

Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A. L. (2009). Dual neural routing of visual facilitation in speech processing. *J Neurosci, 29*(43), 13445-13453. doi:10.1523/JNEUROSCI.3194-09.2009

Avivi-Reich, M., Puka, K., & Schneider, B. A. (2018). Do age and linguistic background alter the audiovisual advantage when listening to speech in the presence of energetic and informational masking? *Atten Percept Psychophys, 80*(1), 242-261. doi:10.3758/s13414-017-1423-5

Baart, M., Stekelenburg, J. J., & Vroomen, J. (2014). Electrophysiological evidence for speech-specific audiovisual integration. *Neuropsychologia, 53*, 115-121. doi:10.1016/j.neuropsychologia.2013.11.011

Besle, J., Fort, A., Delpuech, C., & Giard, M.-H. (2004). Bimodal speech: Early suppressive visual effects in human auditory cortex. *European Journal of Neuroscience, 20*, 2225-2234.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging, 35*(11), 2526-2540. doi:10.1016/j.neurobiolaging.2014.05.006

Billings, C. J., Penman, T. M., McMillan, G. P., & Ellis, E. M. (2015). Electrophysiology and Perception of Speech in Noise in Older Listeners: Effects of Hearing Impairment and Age. *Ear and hearing, 36*(6), 710-722. doi:10.1097/AUD.0000000000000191

Brodbeck, C., Presacco, A., Anderson, S., & Simon, J. Z. (2018). Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust United Acust, 104*(5), 774-777. doi:10.3813/AAA.919221

Cervantes Constantino, F., & Simon, J. Z. (2018). Restoration and Efficiency of the Neural Processing of Continuous Speech Are Promoted by Prior Knowledge. *Frontiers in systems neuroscience, 12*, 56. doi:10.3389/fnsys.2018.00056

CHABA. (1988). Speech understanding and aging. Working Group on Speech Understanding and Aging. Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council. *J Acoust Soc Am, 83*(3), 859-895.

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear Hear, 23*(5), 439-449. doi:10.1097/00003446-200210000-00006

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci, 36*(3), 181-204. doi:10.1017/S0140525X12000477

Czigler, I., Csibra, G., & Ambró, A. (1997). Age and Information Processing. *European Psychologist*, 247-257. doi:https://doi.org/10.1027/1016-9040.2.3.247

Czigler, I., Csibra, G., & Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biological psychology, 33*(2-3), 195-206. doi:10.1016/0301-0511(92)90031-o

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21. doi:10.1016/j.jneumeth.2003.10.009

Dushanova, J., & Christov, M. (2013). Auditory event-related brain potentials for an early discrimination between normal and pathological brain aging. *Neural Regen Res, 8*(15), 1390-1399. doi:10.3969/j.issn.1673-5374.2013.15.006

Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: from single units to evoked potentials. *Audiology & neuro-otology, 7*(2), 71-99. doi:10.1159/000057656

El-Assal, H. A., & El-Gharib, A. M. (2019). Encoding of speech in noise in adults using hearing aids: effect of noise reduction algorithm. *Hearing, Balance and Communication* doi:https://doi.org/10.1080/21695717.2019.1667687

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *J Speech Hear Res, 12*(2), 423-425. doi:10.1044/jshr.1202.423

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol (Amst), 101*(2-3), 267-291. doi:10.1016/s0001-6918(99)00008-6

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170. doi:10.1111/j.1469-8986.2007.00602.x

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 360*(1456), 815-836. doi:W5T4QMCP8T4K0UP8 [pii]

10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci, 11*(2), 127-138. doi:10.1038/nrn2787

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention--focusing the searchlight on sound. *Curr Opin Neurobiol, 17*(4), 437-455. doi:10.1016/j.conb.2007.07.011

Ganesh, A. C., Berthommier, F., & Schwartz, J.-L. (2017). Audiovisual Binding for Speech Perception in Noise and in Aging. *Laanguage learning, 68*, 193-220. doi:https://doi.org/10.1111/lang.12271

Goodin, D. S., Squires, K. C., Henderson, B. H., & Starr, A. (1978). Age-related variations in evoked potentials to auditory stimuli in normal human subjects. *Electroencephalography and clinical neurophysiology, 44*(4), 447-458. doi:10.1016/0013-4694(78)90029-9

Gordon, M. S., & Allen, S. (2009). Audiovisual speech in older and younger adults: integrating a distorted visual signal with speech in noise. *Exp Aging Res, 35*(2), 202-219. doi:10.1080/03610730902720398

Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten Percept Psychophys, 75*(3), 508-524. doi:10.3758/s13414-012-0406-9

Humes, L. E., Kidd, G. R., & Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Frontiers in systems neuroscience, 7*, 55. doi:10.3389/fnsys.2013.00055

Humes, L. E., Wilson, D. L., Barlow, N. N., & Garner, C. (2002). Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults. *J Speech Lang Hear Res, 45*(4), 772-782. doi:10.1044/1092-4388(2002/062)

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychol Methods, 15*(4), 309-334. doi:10.1037/a0020761

Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology, 30*(1), 10-22. doi:10.1111/j.1469-8986.1993.tb03200.x

Jansen, S. D., Keebler, J. R., & Chaparro, A. (2018). Shifts in Maximum Audiovisual Integration with Age. *Multisens Res, 31*(3-4), 191-212. doi:10.1163/22134808-00002599

Klucharev, V., Mottonen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Brain Res Cogn Brain Res, 18*(1), 65-75. doi:10.1016/j.cogbrainres.2003.09.004

Kornhuber, H. H., & Deecke, L. (1965). [Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials]. *Pflugers Arch Gesamte Physiol Menschen Tiere, 284*, 1-17.

Larouche, E., Tremblay, M. P., Potvin, O., Laforest, S., Bergeron, D., Laforce, R., . . . Hudon, C. (2016). Normative Data for the Montreal Cognitive Assessment in Middle-Aged and Elderly Quebec-French People. *Arch Clin Neuropsychol*. doi:10.1093/arclin/acw076

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol Aging, 27*(8), 1155-1163. doi:10.1016/j.neurobiolaging.2005.05.024

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain, 106 (Pt 3)*, 623-642. doi:10.1093/brain/106.3.623

Lin, F. R. (2011). Hearing loss and cognition among older adults in the United States. *J Gerontol A Biol Sci Med Sci, 66*(10), 1131-1136. doi:10.1093/gerona/glr115

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25*(6), 763-770. doi:10.1037/a0024238

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., . . . Health, A. B. C. S. G. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern Med, 173*(4), 293-299. doi:10.1001/jamainternmed.2013.1868

Massaro, D. W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. Cambridge, London: The MIT Press.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

Naatanen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425. doi:10.1111/j.1469-8986.1987.tb00311.x

Nasreddine, Z. S., Chertkow, H., Phillips, N., Bergman, H., & Whitehead, V. (2003). Sensitivity and Specificity of The Montreal Cognitive Assessment (MoCA) for Detection of Mild Cognitive Deficits. *Can J Neurol Sci 30*(30).

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc, 53*(4), 695-699. doi:10.1111/j.1532-5415.2005.53221.x

O'Sullivan, A. E., Crosse, M. J., Di Liberto, G. M., & Lalor, E. C. (2016). Visual Cortical Entrainment to Motion and Categorical Speech Features during Silent Lipreading. *Front Hum Neurosci, 10*, 679. doi:10.3389/fnhum.2016.00679

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and aging, 17*(2), 299-320.

Pichora-Fuller, M. K. (1997). Language comprehension in older listeners. *Journal of Speech-Language Pathology and Audiology, 21*(2), 125–142.

Pilling, M. (2009). Auditory event-related potentials (ERPs) in audiovisual speech perception. *J Speech Lang Hear Res, 52*(4), 1073-1081. doi:10.1044/1092-4388(2009/07-0276)

Pinto, S., Tremblay, P., Basirat, A., & Sato, M. (2019). The impact of when, what and how predictions on auditory speech perception. *Exp Brain Res*. doi:10.1007/s00221-019-05661-5

Presacco, A., Simon, J. Z., & Anderson, S. (2016a). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol, 116*(5), 2356-2367. doi:10.1152/jn.00373.2016

Presacco, A., Simon, J. Z., & Anderson, S. (2016b). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol, 116*(5), 2346-2355. doi:10.1152/jn.00372.2016

Puschmann, S., Daeglau, M., Stropahl, M., Mirkovic, B., Rosemann, S., Thiel, C. M., & Debener, S. (2019). Hearing-impaired listeners show increased audiovisual benefit when listening to speech in noise. *NeuroImage, 196*, 261-268. doi:10.1016/j.neuroimage.2019.04.017

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience, 2*(1), 79-87. doi:10.1038/4580

Roque, L., Karawani, H., Gordon-Salant, S., & Anderson, S. (2019). Effects of Age, Cognition, and Neural Encoding on the Perception of Temporal Speech Cues. *Front Neurosci, 13*, 749. doi:10.3389/fnins.2019.00749

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol Rev, 103*(3), 403-428.

Salthouse, T. A. (2009). Decomposing age correlations on neuropsychological and cognitive variables. *J Int Neuropsychol Soc, 15*(5), 650-661. doi:10.1017/S1355617709990385

Scherg, M., & Von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and clinical neurophysiology, 65*(5), 344-360. doi:10.1016/0168-5597(86)90014-6

Schiff, S., Valenti, P., Andrea, P., Lot, M., Bisiacchi, P., Gatta, A., & Amodio, P. (2008). The effect of aging on auditory components of event-related brain potentials. *Clin Neurophysiol, 119*(8), 1795-1802. doi:10.1016/j.clinph.2008.04.007

Sekiyama, K., Soshi, T., & 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

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*(3), 263-275.

Soros, P., Teismann, I. K., Manemann, E., & Lutkenhoner, B. (2009). Auditory temporal processing in healthy aging: a magnetoencephalographic study. *BMC Neurosci, 10*, 34. doi:10.1186/1471-2202-10-34

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci, 19*(12), 1964-1973. doi:10.1162/jocn.2007.19.12.1964

Stevenson, R. A., Nelms, C. E., Baum, S. H., Zurkovsky, L., Barense, M. D., Newhouse, P. A., & Wallace, M. T. (2015). Deficits in audiovisual speech perception in normal aging emerge at the level of whole-word recognition. *Neurobiol Aging, 36*(1), 283-291. doi:10.1016/j.neurobiolaging.2014.08.003

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(212-215).

Team, R. C. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Teder-Salejarvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res Cogn Brain Res, 14*(1), 106-114. doi:10.1016/s0926-6410(02)00065-4

Tingley D, Y. T., Hirose K, Keele L, Imai K. (2013). mediation: R Package for Causal Mediation Analysis. (Version R package version 4.4.2). Retrieved from http://CRAN.R-project.org/

Tobias, J. V. (1977). The need for preventive audiology. *Audiol. Hear. Educ. , 3*(2), 37-38.

Treille, A., Cordeboeuf, C., Vilain, C., & Sato, M. (2014a). Haptic and visual information speed up the neural processing of auditory speech in live dyadic interactions. *Neuropsychologia, 57*, 71-77. doi:10.1016/j.neuropsychologia.2014.02.004

Treille, A., Vilain, C., Kandel, S., & Sato, M. (2017). Electrophysiological evidence for a self-processing advantage during audiovisual speech integration. *Exp Brain Res, 235*(9), 2867-2876. doi:10.1007/s00221-017-5018-0

Treille, A., Vilain, C., & Sato, M. (2014b). The sound of your lips: electrophysiological cross-modal interactions during hand-to-face and face-to-face speech perception. *Front Psychol, 5*, 420. doi:10.3389/fpsyg.2014.00420

Treille, A., Vilain, C., Schwartz, J. L., Hueber, T., & Sato, M. (2018). Electrophysiological evidence for Audio-visuo-lingual speech integration. *Neuropsychologia, 109*, 126-133. doi:10.1016/j.neuropsychologia.2017.12.024

Tremblay, K. L., Billings, C., & Rohila, N. (2004). Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol, 15*(3), 226-237; quiz 264. doi:10.3766/jaaa.15.3.5

Tremblay, K. L., Piskosz, M., & Souza, P. (2002). Aging alters the neural representation of speech cues. *Neuroreport, 13*(15), 1865-1870.

Tremblay, K. L., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol, 114*(7), 1332-1343.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., & Hale, S. (2010). Aging, audiovisual integration, and the principle of inverse effectiveness. *Ear Hear, 31*(5), 636-644. doi:10.1097/AUD.0b013e3181ddf7ff

van Wassenhove, V. (2013). Speech through ears and eyes: interfacing the senses with the supramodal brain. *Front Psychol, 4*, 388. doi:10.3389/fpsyg.2013.00388

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, 102*(4), 1181-1186.

Vroomen, J., & Stekelenburg, J. J. (2010). Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J Cogn Neurosci, 22*(7), 1583-1596. doi:10.1162/jocn.2009.21308

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res Rev, 23*(Pt B), 154-166. doi:10.1016/j.arr.2015.06.002

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*(2), 427-438. doi:10.1037/a0021683

World Medical, A. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191-2194. doi:10.1001/jama.2013.281053

Yang, W., & Ren, Y. (2018). Attenuated audiovisual integration in middle-aged adults in a discrimination task. *Cogn Process, 19*(1), 41-45. doi:10.1007/s10339-017-0838-1