Older age results in difficulties separating auditory and visual signals in time

Yu Man Chan
Department of Optometry and Vision Sciences, University of Melbourne, Parkville, Victoria, Australia

Michael J. Pianta
Department of Optometry and Vision Sciences, University of Melbourne, Parkville, Victoria, Australia

Allison M. McKendrick
Department of Optometry and Vision Sciences, University of Melbourne, Parkville, Victoria, Australia

Previous research provides conflicting evidence regarding whether older adults have altered tolerance to timing differences between auditory and visual events. We examine the potential impact of age-related unisensory decline on audiovisual synchrony perception. Fifteen younger (21–32 years) and 13 older (60–72 years) adults participated. To assess unisensory sensitivity, visual Gabor contrast detection thresholds and auditory masked tone pip detection thresholds were measured. Four multisensory conditions were then tested: suprathreshold and near-threshold stimuli (based on individual unisensory psychometric functions), each tested with a masked tone pip stimuli at 0.5 and 4 kHz sound frequencies. Two audiovisual pairs (one synchronous, the other asynchronous) were presented in a two-interval forced-choice procedure, with observers identifying the interval containing the asynchronous stimulus. Older adults required a larger physical asynchrony to perceive the stimuli as asynchronous, particularly for low frequency sounds. Our results demonstrate that the impact of age on audiovisual synchrony perception cannot be explained by decline in unisensory sensitivity alone.

Introduction

Multisensory perception is key to a natural sensory experience. Audiovisual synchrony tasks that measure the perception of timing difference between auditory and visual stimuli are one method that has been used to behaviorally study multisensory perception (Laurienti, Burdette, Maldjian, & Wallace, 2006; Love, Petrini, Cheng, & Pollick, 2013; Peiffer, Mozolic, Hugenschmidt, & Laurienti, 2007; Van Eijk, Kohlrausch, Juola, & van de Par, 2008; Vatakis, Navarra, Soto-Faraco, & Spence, 2008). Differences in the speed of sound and light in air, as well as differential speeds of neural transmission between the senses, make it unlikely that auditory and visual information arising from the same event will reach the cortex at the same time. Therefore, some tolerance to audiovisual asynchrony is required to produce a coherent percept of the world. However, to avoid incorrect pairings of auditory and visual information from separate events, it is also important to be able to correctly segregate stimuli in time. There has been minimal study of how this ability is altered during normal, healthy aging. A recent study has shown that middle-aged adults (50–60 years) have smaller thresholds for audio-lead speech stimuli than younger adults (20–30 years; Alm & Behne, 2013). Two other studies have explored whether audiovisual synchrony perception is altered by hearing loss, and these have included adults of a variety of ages (Baskent & Bazo, 2011; Hay-McCutcheon, Pisoni, & Hunt, 2009). Hay-McCutcheon et al. (2009) reported that older adults are less able to distinguish timing mismatch between auditory and visual components of complex speech stimuli, but Baskent and Bazo (2011) did not find a main effect of older age. Hay-McCutcheon and colleagues (2009) suggested that age differences in peripheral sensory processing of the higher frequency sounds in speech stimuli might be an important factor for the poorer ability to separate temporally offset audiovisual signals with age.

It is well known that age affects the detectability of both auditory and visual cues. Older people have reduced visual contrast sensitivity (Derefeldt, Lennerstrand, & Lundh, 1979; Owsey, Sekuler, & Siemsen, 1983) and reduced sound detection thresholds, partic-
ularly at higher sound frequencies (Gordon-Salant, 2005; Mazelová, Popelar, & Syka, 2003). Consequently, when presented with stimuli that are at a fixed (suprathreshold) physical stimulus level, the stimulus will be closer to detection threshold for older adults than for younger observers. Given that neural transmission is slower for near threshold stimuli than for suprathreshold stimuli in both the visual (Barlow, 1972; Maunsell & Newsome, 1987) and auditory (Kerkhof, 1978; Kerkhof & Uhlenbroek, 1981; Squires, Hillyard, & Lindsay, 1973) pathways, differences in detectability may lead to variability in the neural transmission rate of stimulus content. Given the complex nature of audiovisual speech stimuli, it is impossible to directly determine whether differences in unisensory processing can explain the poorer ability of older adults to segregate auditory and visual signals in time (Hay-McCutcheon et al., 2009).

An additional factor that has not been considered previously is the observation that older people often have more conservative response criteria than younger people (Ratcliff, Thapar, & McKoon, 2001). In typical audiovisual synchrony judgment experiments, participants are presented with a single audiovisual stimulus pair, and are asked to judge if the pair was perceived to be synchronous or not (Dixon & Spitz, 1980; Fujisaki & Nishida, 2005; Zampini, Guest, Shore, & Spence, 2005; Zampini, Shore, & Spence, 2005). Response criterion bias can confound the interpretation of such audiovisual synchrony judgment results (Baskent & Bazo, 2011; Yarrow, Jahn, Durant, & Arnold, 2011). Several studies have shown that it is difficult to eliminate the effects of criterion bias from audiovisual timing experiments of this nature (Spence, Shore, & Klein, 2001; Zampini, Shore et al., 2005). Zampini, Shore et al. (2005) showed that when observers directed their attention to vision rather than audition, synchrony judgment responses were biased towards the attended sensory modality—audiovisual asynchrony was perceived when the visual stimulus was presented before the auditory stimulus by a markedly smaller sound-lag. The contribution of response bias to previous studies that have used such methods to explore the effects of hearing loss (and, indirectly, aging) on audiovisual synchrony perception (Baskent & Bazo, 2011; Hay-McCutcheon et al., 2009) is unknown.

Uncertainty remains about whether the normal aging process impairs the ability to discriminate auditory and visual signals that are separated in time and, if so, whether performance differences can be explained by reduced sensitivity to the visual and/or auditory stimuli in older adults. Our study aimed to determine the impact of aging on audiovisual synchrony perception. Simple visual Gabor and auditory masked tone pip stimuli were used to enable correction for interindividual differences in visual contrast and auditory sensitivities. Two sound frequencies were used to compare between stimuli for which auditory sensitivities are, and are not, affected by presbycusis. Stimuli were scaled according to detectability for each individual observer and tested at near threshold and suprathreshold levels. Finally, a two-interval forced-choice (2IFC) paradigm was used to minimize any age- or observer-related differences in response criterion bias.

### Methods

#### Participants

Participants were recruited from the University of Melbourne community and the public through advertisements posted around the University and in community newspapers. Fifteen younger and 14 older adults met the inclusion criteria. One older participant had difficulty performing the task as instructed despite considerable training, and did not complete the experiments. Therefore the final sample size consisted of 15 younger (21–32 years, mean age of 25 years; five males) and 13 older adults (61–72 years, mean age of 66 years; seven males). There was insufficient data from previous audiovisual ageing studies to conduct a meaningful power analysis. Our required sample size was estimated based on previous behavioral aging studies in vision and in audiovisual research that have found a significant age difference with a sample size of 10 to 20 in each age group (Hay-McCutcheon et al., 2009: 12Y, 10O; DeLoss et al., 2013: 12Y, 12O). All participants had normal or corrected-to-normal vision of 6/7.5 or better, and normal hearing for their age. Normal hearing was defined as having audiometric thresholds less than 35 decibels hearing level (dB HL) at 4 kHz, and less than 25 dB HL at all other tested frequencies (0.25, 0.5, 1, and 2 kHz), according to the International Organization for Standardization (ISO) standard on hearing by age and sex (ISO 7029:2000 Acoustics). No cognitive assessment was performed. Informed consent was obtained from all participants as approved by the Human Research Ethics Committee of University of Melbourne and in accordance with the Declaration of Helsinki.

#### Equipment

The experiment was controlled by software written in MATLAB 7.6.0 (R2008a; Mathworks, Boston, MA) and run on a personal computer (Dell Precision T3500, Round Rock, TX). The visual stimulus was presented using a ViSaGe (Cambridge Research Systems, Cambridge, UK) to drive a cathode ray tube monitor (Sony...
Trinitron Multiscan G520, mean luminance: 100 cd/m², frame rate: 100 Hz, 800 × 600 pixels; Tokyo, Japan) that was gamma corrected on a weekly basis. Responses were collected using a CB6 response box. The ViSaGe also initiated sound presentation through a set of headphones (Sennheiser HD 205, Wedemark, Germany), by triggering a multifunction processor (Tucker-Davis Technologies [TDT] RX6, Alachua, FL) that drove a programmable attenuator (TDT PA5) and a headphone driver (TDT HB7). Physical synchrony between the audio and visual stimulus was verified prior to starting on the main experiment using an oscilloscope. Participants stabilized their head position by resting on a chin rest positioned 100 cm from the monitor.

Stimuli

The visual stimulus was a vertically striped Gabor of \(3 \text{ c}^/\text{o}\), with the standard deviation of the Gaussian envelope defined as the reciprocal of the spatial frequency (≈0.33°), that was presented for one monitor frame (frame rate: 100 Hz; Figure 1a). The auditory stimulus was a pure tone pip (10 ms duration; 2.5 ms onset and offset ramp) of 0.5 or 4 kHz presented binaurally through headphones over a pure tone mask (75 dB) of the same frequency.

Procedures

Measurement of detection thresholds

To equate for individual stimulus detectability, we first determined the auditory and visual psychometric functions. Masked thresholds for the auditory tone pip at 0.5 and 4 kHz were obtained using a 2IFC procedure with a method-of-constant stimuli (MOCS; seven sound intensity steps). As illustrated in Figure 1b, each interval was 1300 ms long, within which was a leading and trailing 200 ms period where the target tone pip was never presented. This was an attempt to minimize any masking effect from the onset and offset of the tone mask. There was a 500 ms
interstimulus interval before the second interval was shown. The target tone pip was randomly presented in either the first or second interval. Participants then indicated via a button press which interval contained the target. This was not designed to be a speeded detection task, so the participants were allowed to pace themselves at a comfortable speed, and breaks were given as required. The visual contrast psychometric function was measured in a similar manner (seven contrast steps), with participants indicating which interval contained the Gabor. To standardize conditions between observers, the fixation guidelines and auditory tone mask were present during all these measurements.

Each MOCS step was repeated 21 times (seven repeats per run, each run taking approximately 4 min). The proportion of correct responses was plotted against visual contrast (vision task) or sound intensity (auditory task). The data points were then fitted with a psychometric function using maximum likelihood estimation (Watson, 1979). The psychometric function was defined by

\[ f(t) = FP + (1 - FP - FN) \times \left( G(t, \mu, \delta) \right) \]

where \( G(t, \mu, \delta) \) was the cumulative Gaussian distribution with mean (\( \mu \)) and standard deviation (\( \delta \)) for stimulus value \( t \). FP and FN represented the proportions of false positive and false negative responses, respectively. False positive proportion was fixed at 0.5 assuming a 50% guess rate when the stimulus is well below detection threshold.

The mean and standard deviation extracted from each individual’s fitted psychometric function were then used to calculate the stimulus level for subsequent experiments. A suprathreshold stimulus level was defined as 10 SDs above the mean threshold, while a near threshold stimulus level was 3 SDs above threshold. This near threshold stimulus level was chosen as it represents a stimulus value that should be seen approximately 99.7% of the time. We confirmed this by testing each participant using a 2IFC procedure with 20 repeats of the near-threshold target compared to 20 trials without the stimulus. All participants were able to correctly identify the interval containing the stimulus 100% of the time.

Audiovisual asynchrony judgment

The sound-lead and sound-lag asynchrony detection threshold (i.e., the stimulus onset asynchrony for which the observer perceived the stimulus as asynchronous 79% of the time) was measured for a 0.5 and a 4 kHz masked tone pip and a Gabor, at suprathreshold and near-threshold stimulus levels, using a similar 2IFC method as described above. A 3-up-1-down staircase procedure was used with an initial step size of 80 ms, which was halved to 40 ms after the first reversal, and halved again to 20 ms after the second reversal. An asynchronous audiovisual stimulus was randomly presented in either the first or second interval, and a synchronous audiovisual stimulus was presented in the other interval. The participants were asked to indicate which interval contained the asynchronous stimulus (Figure 1b). Stimulus onset asynchrony (SOA) refers to the time interval between the onset of the tone pip and the Gabor: Physical synchrony is indicated by a zero SOA value, sound-lead is indicated by a negative SOA value, and sound-lag is specified by a positive SOA value. There were two interleaved staircases in each run: One staircase began with a clearly asynchronous sound-lead stimulus; another started with a clearly asynchronous sound-lag stimulus. A mean of the last two out of four reversals gave an estimate of the stimulus asynchrony that was detectable on approximately 79% of occasions (Figure 1c). Each run was repeated three times to give three estimates of both the sound-lead threshold and sound-lag threshold for each of the four test conditions (two sound frequencies and two stimulus threshold levels). The order of these 12 runs was randomized for each participant to avoid order effects. Prior to commencing the main experiment, practice runs (identical methods to the main experiment) were given to the participants until they were confident with the task and were providing reliable results. In the initial practice run, the researcher gave trial-by-trial verbal feedback to indicate which interval was physically synchronous. Markedly asynchronous stimuli were used for the initial training to verify understanding of the task.

Analysis

The width of an audiovisual synchrony window was defined as the difference between the mean sound-lead and sound-lag asynchrony detection thresholds. A repeated-measures analysis of variance (RM-ANOVA) was used to determine whether there were main effects or interactions between group factors of age, sound frequency, and threshold level (suprathreshold or near threshold) on the sound-lead asynchrony detection thresholds, sound-lag asynchrony detection thresholds, and width estimates.

Results

Detection thresholds

Consistent with previous literature (Derefeldt et al., 1979; Owsley et al., 1983), older adults had a
significantly elevated mean contrast detection threshold than the younger group, \( t(26) = 6.62, p < 0.001 \) (Figure 2a). There was no main effect of age on the auditory detection threshold, \( F(1, 26) = 0.03, p = 0.87 \) (Figure 2b, c). The group average auditory detection thresholds were significantly higher at 4 kHz than at 0.5 kHz (main effect of sound frequency: \( F[1, 26] = 398.88, p < 0.001 \)).

**Audiovisual synchrony window widths**

Figure 3 plots the mean synchrony window width for both 0.5 kHz (upper panel) and 4 kHz (lower panel) in the young and older groups. There was no main effect of age, \( F(1, 26) = 1.43, p = 0.24 \), nor sound frequency, \( F(1, 26) = 0.95, p = 0.34 \), on the width measures, but a main effect of stimulus detectability level on the widths, \( F(1, 26) = 29.51, p < 0.001 \). Average widths were higher for near threshold than suprathreshold stimulus level. There was a significant interaction between age and sound frequency, \( F(1, 26) = 4.90, p = 0.04 \), as illustrated by the wider average group window width in the older adults than in the young adults at 0.5 kHz, but relatively similar group means at 4 kHz. When the widths estimates at 0.5 kHz were analyzed independent from those at 4 kHz, there was a main effect of age, \( F(1, 26) = 4.37, p = 0.047 \).

**Audiovisual asynchrony detection thresholds**

To determine whether sound-lead and/or sound-lag contributed to the wider synchrony window in the older cohort at 0.5 kHz, the asynchrony detection thresholds for sound-lead and sound-lag stimuli were compared between groups, threshold levels, and sound frequencies. Figures 4 and 5 plot the average asynchrony detection threshold for the younger and older groups for suprathreshold and near threshold stimuli, respectively. The average asynchrony detection thresholds were significantly larger for near threshold (Figure 5) than suprathreshold stimuli (Figure 4; main effect of stimulus detectability level: \( F[1, 26] = 29.51, p < 0.001 \)). There was a significant interaction between sound frequency, threshold type, and age, \( F(1, 26) = 5.32, p = 0.03 \) and a significant interaction between sound frequency and age, \( F(1, 26) = 4.90, p = 0.04 \). As seen in Figures 4 and 5, the thresholds between the age groups were similar at 4 kHz. The older adults have a larger mean sound-lead asynchrony detection threshold and a larger mean sound-lag threshold only for low frequency sound stimulus at the suprathreshold stimulus level. When the asynchrony detection thresholds at 0.5 kHz was analyzed independently, there was a main effect of age, \( F(1, 26) = 4.37, p = 0.047 \). Exact values of the average window widths and the stimulus onset asynchrony range for 79% perceived asynchrony are...
presented in Table 1. No other main effects or interactions were significant.

Discussion

This study showed that, on average, older adults have a wider audiovisual synchrony window for low sound frequency stimuli. The wider audiovisual synchrony window is not an artefact of stimulus detectability differences in either vision or hearing, as our stimuli were scaled according to individual detection thresholds. We used a forced-choice paradigm to minimize the effects of age-related criterion differences in the judgment of synchrony.

Our older adult group was recruited from the community, was fit and active, and required to meet general screening of vision and hearing to ensure no significant age-related sensory organ damage. Out of the 13 older adults, there were five current or retired university staff. The rest were active older citizens from the community who were still involved in casual paid work and volunteer work. Our informal observation was that the older adult group showed increased motivation and interest in the research than the younger group. However, since we did not explicitly measure the cognitive performance of our participants, we cannot explicitly demonstrate whether our findings were influenced by differences in cognition or attention.

As expected for individuals of this age, hearing thresholds were less than 35 dB HL at 4 kHz, and less than 25 dB HL at all other tested frequencies (0.25, 0.5, 1, and 2 kHz; ISO 7029:2000 Acoustics). We also confirmed an age-related reduction in contrast sensitivity of around two-fold, consistent with previous aging studies in vision (Chan, Battista, & McKendrick, 2012; Derefeldt et al., 1979; Karas & McKendrick, 2009; Owsley et al., 1983).

Our measured audiovisual synchrony window widths for the suprathreshold stimulus condition were similar to those reported previously for young adults when a simple flash-noise burst stimulus was used (Van Eijk et al., 2008; Vatakis et al., 2008; Zampini, Guest, et al., 2005). As listed in Table 1, for suprathreshold stimuli,
the average window width measured in our younger group was 319 and 334 ms for 0.5 and 4 kHz sound frequencies, respectively. Using a simple LED flash and a 9-ms suprathreshold noise burst, the average window width reported by Zampini, Guest, et al. (2005) was 366 ms. However when stimuli were near threshold, we found significantly wider window widths of 484 and 528 ms at 0.5 and 4 kHz respectively (Table 1). The mechanisms underpinning this difference are not clear from our study but might reflect the relative slowing of neural latencies for low contrast stimuli (White & Jeffreys, 1982) or the relative reduction in signal to noise. It is worth noting that altered processing of near threshold multisensory stimuli is also demonstrated via electrophysiological and behavioral studies that have measured reaction times to audiovisual presentations. Such studies have demonstrated that reaction times to audiovisual signals are shorter than those in response to visual-only or auditory-only stimuli (Diederich & Colonius, 2004; Laurienti et al., 2006; Peiffer et al., 2007; Rach, Diedrich, & Colonius, 2011; Wallace, Meredith, & Stein, 1993), particularly when the stimuli are less salient: a phenomenon commonly referred to as inverse effectiveness. The relationship, if any, between speeded reaction times for low-contrast audiovisual stimuli and the width of the timing window for synchrony perception is unclear.

Another audiovisual perceptual phenomenon is the Colavita (1974) visual dominance effect, which refers to the dominance of vision over hearing when observers respond to an audiovisual stimulus. Under certain stimulus conditions, when visual and sound signals are presented bimodally, participants respond more often to the visual component and can report being unaware of the sound presentation. During our experiments, the experimenter (author YMC) regularly sought informal feedback from the participants, in order to maintain motivation on the task. None of the observers reported an absence of either the audio or visual stimulus during the experiments, so we consider it unlikely that the Colavita visual dominance effect influenced our outcomes.

Hay-McCutcheon et al. (2009) found a wider audiovisual (AV) synchrony window for older adults using speech stimuli, and suggested that this finding may arise due to age differences in stimulus detectability, especially of high spatial frequency visual information and high frequency auditory information contained within the complex speech stimulus. Our simple tone stimuli allowed us to test the effect of sound frequency on audiovisual synchrony perception in both age groups. We predicted that any aging effects would be greater for high than low sound frequency, on the assumption that the effects of presbycusis on localization of a sound in time might not be fully compensated for by scaling stimulus detectability. Our data demonstrated that sound frequency did not alter the average window width in the younger group. However, the older group had wider averaged window width at 500 Hz than 4kHz, opposite to our prediction.

Auditory temporal resolution is generally poorer for low than high frequency sounds. This is evidenced by gap detection threshold measurements in young adults showing a reduction in threshold with increasing sound frequency (Eddins, Hall, & Grose, 1992; Fitzgibbons, 1983; Formby & Muir, 1988; Shailer & Moore, 1985). By measuring gap detection thresholds in a group of younger and older adults, Snell (1997) showed that gap detection thresholds were uniformly elevated with age across a range of low to high sound spectral frequencies. Consequently, older observers may have poorer temporal representation of the sound onset for both sound frequencies tested in our study. This logic predicts a widening of the audiovisual synchrony window for both 0.5 and 4 kHz sound stimuli, rather than the frequency-specific age effect seen here. An explanation for the frequency-specific effect requires further experimental data. A neural processing difference between low and high frequency sounds is that sounds above 4 to 5 kHz are not phase-locked, whereas responses to low-frequency sounds demonstrate phase-locking (Woolf, Ryan, & Bone, 1981). Neural phase-locking is considered important for pitch discrimination and for the determination of interaural timing differences (Ochi, Yamasoba, & Furukawa, 2014), but whether it relates to our results and to audiovisual synchrony perception is unknown.

We scaled the auditory stimuli for detectability, but did not scale the mask (fixed at 75 dB for all participants). The mask was clearly suprathreshold for all observers, but due to differences in sound detection thresholds, was closer to threshold for older than younger adults, particularly at 4 kHz. If a lower mask intensity results in a narrowing of the audiovisual

---

Table 1. Audiovisual synchrony window width and the 79% perceived synchrony range in parentheses (sound-lead . . . sound-lag). Note: Values are indicated in milliseconds.

<table>
<thead>
<tr>
<th>Test condition</th>
<th>Younger</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suprathreshold 0.5 kHz</td>
<td>319 (−153…+166)</td>
<td>465 (−241…+224)</td>
</tr>
<tr>
<td>Suprathreshold 4 kHz</td>
<td>334 (−148…+186)</td>
<td>361 (−165…+195)</td>
</tr>
<tr>
<td>Near threshold 0.5 kHz</td>
<td>484 (−214…+270)</td>
<td>555 (−292…+263)</td>
</tr>
<tr>
<td>Near threshold 4 kHz</td>
<td>528 (−239…+290)</td>
<td>506 (−210…+297)</td>
</tr>
</tbody>
</table>
synchrony window (perhaps due to less neural noise), then it may be possible that any widening of the older adults’ window for the 4 kHz stimuli was opposed by the mask being relatively closer to threshold in the older group. To explore this possibility we repeated the experiments for four younger adults with a 50 dB mask (again with test stimuli scaled for detection). Window widths measured with a mask of 50 dB were not significantly different from widths measured with a 75 dB mask, (3) = 1.02, p > 0.05 (data not shown).

It is worth noting that different definitions of AV window width are used in different studies. Studies have commonly measured synchrony judgments across a range of temporal offsets, and then fitted a single Gaussian function through the data (Baskent & Bazo, 2011; Fujisaki & Nishida, 2005; Hay-McCutcheon et al., 2009; Zampini, Guest, et al., 2005). From the fitted function, window width was defined as the full-width-half-maximum (i.e., width at 50% asynchrony). However, this method assumes symmetry in the slope of the function at the ends of the synchrony window, and so may not reflect any asymmetry present. Therefore, using a modified synchrony judgment task, we adopted a method that independently assessed sensitivity at each end of the window using a 3-up-1-down adaptive staircase procedure. Two alternate forced-choice designs are generally considered to be less affected by observer criterion than yes-no procedures; however, it is worth noting that interval bias can exist in 2IFC experiments (Garcia-Perez & Alcala-Quintana, 2010; Yeshurun, Carrasco, & Maloney, 2008). This could mean that our finding was a result of a larger interval bias in the older group, instead of a genuine change in asynchrony discrimination sensitivity, resulting in a lower percentage correct in the 2IFC procedure.

Our measured magnitude of age-related widening of the audiovisual synchrony window at 0.5 kHz was 145.51 ms, which is larger than the magnitude of the age effect of 53.13 ms previously reported by Hay-McCutcheon et al. (2009). By estimating from Figure 2 in the paper by Hay-McCutcheon et al. (2009), we found that their data show an approximately twofold increase in the age-related widening of the window at 79% than at 50% perceived asynchrony, which explains the larger magnitude of age effect reported in our current study. For suprathreshold stimulus level at 0.5 kHz sound frequency, we found an earlier audio-lead threshold in the older (−241 ms) than the younger adults (−153 ms). Inconsistent with our findings, Alm and Behne (2013) reported a later audio-lead threshold in their older adults (−165 ms) than in their younger cohort (−216 ms). However, it is important to bear in mind that our study and previous reports have used significantly different methodologies; they used complex stimuli instead of simple stimuli and measured synchrony judgments instead of asynchrony discrimination thresholds.

A related, but different, methodology for investigating the ability to process the temporal properties of audiovisual events is the temporal order judgment (TOJ) task. TOJ differs from a synchrony judgment task by measuring the minimal temporal gap between the onset of a visual and an auditory stimulus for an observer to correctly determine which stimulus came first. While similar, TOJ tasks and synchrony tasks do not have the same requirements from observers (Love et al., 2013). In performing temporal order judgment tasks, observers have a presumption that the visual and auditory stimuli are not synchronous and that one modality should occur before the other (Van Eijk et al., 2008). A recent study reported that TOJ precision was the same between younger adults (18–29 years) and older adults (70–82 years; Fiacconi, Harvey, Sekuler, & Bennett, 2013). While there are a number of potential explanations for the difference in main finding between their study and ours, the comparison of results is consistent with the suggestion that TOJ and synchrony judgments involve at least partially different neural mechanisms, and that these mechanisms may be differentially susceptible to normal aging.

Traditional models of audiovisual synchrony detection proposed that signals are first processed independently in unisensory pathways and then audiovisual synchrony is determined by the relative arrival times of these signals at a central comparator (Rutschmann, 1966; Sternberg & Knoll, 1973; Ulrich, 1987). Differences in perceived audiovisual timing are considered to result from a change in the speed of neural transmission, causing a relative change in the unisensory signal arrival times at the central comparator. Older people (60–70 years) have a relatively larger magnitude of neural transmission delay than younger adults (20–30 years) in the auditory system (around 40 ms; Goodin, Squires, Henderson, & Starr, 1978) than in the visual system (between 10–25 ms; Celesia, Kaufman, & Cone, 1987; Dustman & Beck, 1969; Wright, Williams, Drasdo, & Harding, 1985). So the notion of a central comparator would predict an overall shift of the audiovisual asynchrony window toward sound-lead due to age-related changes in signal arrival times at the central comparator. As a result, the older people would perceive asynchrony when sound is presented much earlier than the visual signal, but be able to perceive asynchrony with a smaller magnitude of sound-lag. The overall widening of the audiovisual synchrony window observed in our study is not consistent with this idea.

More recent models include a temporal filter model (Burr, Silva, Cicchini, Banks, & Morrone, 2009) and a neural population code model (Roach, Heron, Whittaker, & McGraw, 2011). Because aging alters audiovisual processing, testing individuals of different ages...
with targeted experiments may provide a method to challenge or lend support to proposed models of how the human brain integrates and segregates auditory and visual events.

**Conclusions**

Older adults are less able to discriminate a timing difference between an auditory and visual stimulus than their younger counterparts, and these age-related differences are greater for low frequency, than high frequency sounds. Our data showed that scaling for stimulus detectability cannot fully account for the wider audiovisual synchrony window with age.

**Keywords:** audiovisual, aging, multisensory, perception, synchrony judgment

**Acknowledgments**

This research was funded by the Australian Research Council FT0990930 to author AMM.

Commercial relationships: none.
Corresponding author: Allison M. McKendrick.
Email: allisonm@unimelb.edu.au.
Address: Department of Optometry and Vision Sciences, University of Melbourne, Parkville, Victoria, Australia.

**References**


