

# Enhanced temporal binding of audiovisual information in the bilingual brain\*

GAVIN M. BIDELMAN

School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA

SHELLEY T. HEATH

School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

(Received: October 20, 2016; final revision received: February 07, 2018; accepted: March 04, 2018)

*We asked whether bilinguals' benefits reach beyond the auditory modality to benefit multisensory processing. We measured audiovisual integration of auditory and visual cues in monolinguals and bilinguals via the double-flash illusion where the presentation of multiple auditory stimuli concurrent with a single visual flash induces an illusory perception of multiple flashes. We varied stimulus onset asynchrony (SOA) between auditory and visual cues to measure the "temporal binding window" where listeners fuse a single percept. Bilinguals showed faster responses and were less susceptible to the double-flash illusion than monolinguals. Moreover, monolinguals showed poorer sensitivity in AV processing compared to bilinguals. The width of bilinguals' AV temporal integration window was narrower than monolinguals' for both leading and lagging SOAs (Biling.: -65–112 ms; Mono.: -193–112 ms). Our results suggest the plasticity afforded by speaking multiple languages enhances multisensory integration and audiovisual binding in the bilingual brain.*

Keywords: audiovisual integration, bilingualism, experience-dependent plasticity, multisensory facilitation, temporal binding window

## Introduction

The perceptual world consists of a rich combination of multisensory experiences. Audiovisual (AV) interactions are especially apparent in speech perception. For instance, combining auditory and visual cues acts to enhance speech recognition (Sumby & Pollack, 1954), particularly in noisy environments (Erber, 1975; Ross, Saint-Amour, Leavitt, Javitt & Foxe, 2007; Sumby & Pollack, 1954; Vatikiotis-Bateson, Eigsti & Munhall, 1998). Given its importance in shaping the perceptual world, understanding how (or if) different experiential factors or disorders can modulate AV processing is of interest in order to examine the extent to which this fundamental process is subject to neuroplastic effects.

In this vein, several studies have shown that multisensory processing is impaired in certain neurodevelopmental disorders including autism, dyslexia, and specific language impairments (Foss-Feig, Kwakye, Cascio, Burnette, Kadivar, Stone & Wallace, 2010; Kaganovich, Schumaker, Leonard, Gustafson & Macias,

2014; Wallace & Stevenson, 2014). The consensus of these studies is that certain disorders may extend the brain's "temporal window" for integrating sensory cues, producing an aberrant binding of multisensory features and deficits in creating a single unified percept. While temporal binding might be prolonged in disordered populations, a provocative question that arises from these studies is whether AV binding might be enhanced by certain human experiences. Indeed, while somewhat controversial (Rosenthal, Shimojo & Shams, 2009), there is some evidence that the AV temporal binding window can be shortened with acute perceptual learning (Powers, Hillock & Wallace, 2009). Recent studies also demonstrate that one form of human experience – musical training – can improve the brain's ability to combine auditory and visual cues for speech (Lee & Noppeney, 2014; Musacchia, Sams, Skoe & Kraus, 2007) and non-speech (Bidelman, 2016) stimuli. Here, we asked if another salient human experience, namely second language expertise, similarly bolsters AV processing.

Several lines of evidence support the notion that bilingualism might tune multisensory processing and the temporal binding of AV information. Second language (L2) acquisition requires the assimilation of novel auditory cues that are not present in a bilingual's first language (Kuhl, Ramirez, Bosseler, Lin & Imada,

\* The authors thank Haley Sanders for assistance in data collection and Kelsey Mankel for comments on earlier versions of this manuscript.

Supplementary material can be found online at <https://doi.org/10.1017/S1366728918000408>

Address for correspondence:

Gavin M. Bidelman, PhD, School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN, 38152  
[g.bidelman@memphis.edu](mailto:g.bidelman@memphis.edu)

2014; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). Because of the more unfamiliar auditory input of their L2, bilinguals might place a heavier reliance on vision to aid in spoken word recognition. Under certain circumstances, visual cues alone can contain adequate information for speakers to differentiate between languages (Ronquest, Levi & Pisoni, 2010; Soto-Faraco, Navarra, Weikum, Vouloumanos, Sebastian-Gallés & Werker, 2007). However, the potential improvement in speech comprehension from the integration of a speaker's visual cues with sound tends to be larger when information from the auditory modality is unfamiliar, as in the case of listening to nonnative or accented speech (Banks, Gowen, Munro & Adank, 2015). Under this hypothesis, bilinguals might improve their L2 understanding by better integrating the auditory and visual elements of speech.

Recent behavioral studies have in fact shown differences between monolingual and bilingual listeners' ability to exploit audiovisual cues in phoneme recognition tasks (Burfin, Pascalis, Ruiz Tada, Costa, Savariaux & Kandel, 2014). In early life, infant bilinguals also gaze longer at the face and mouth of a caregiver to parse L1/L2 (Pons, Bosch & Lewkowicz, 2015). There are also suggestions that bilingualism improves cognitive control including selective attention and executive function (Bialystok, 2009; Krizman, Skoe, Marian & Kraus, 2014; Schroeder, Marian, Shook & Bartolotti, 2016). Collectively, previous studies imply that in order to effectively juggle the speech from multiple languages, bilingualism might facilitate multisensory processing and improve the control of audiovisual information.

In the present study, we adopted the "double-flash illusion" paradigm (Shams, Kamitani & Shimojo, 2000; Shams, Kamitani & Shimojo, 2002) to determine if bilinguals show enhanced audiovisual processing and temporal binding of multisensory cues. In this paradigm, the presentation of multiple auditory stimuli (beeps) concurrent with a SINGLE visual object (flash) induces an illusory perception of multiple flashes. These nonspeech stimuli have no relation to familiar speech stimuli and are thus ideal for studying audiovisual processing in the absence of lexical-semantic meaning that might otherwise confound interpretation in a cross-linguistic study. By parametrically varying the onset asynchrony between auditory and visual events (leads and lags) we quantified group differences in the "temporal window" for binding audiovisual perceptual objects in monolingual and bilingual individuals. We hypothesized that bilinguals would show both faster and more accurate processing of concurrent audiovisual cues than their monolingual peers. Our predictions were based on recent evidence from our lab demonstrating that other intensive multimodal experiences (i.e., musicianship) can enhance the temporal binding of audiovisual cues as indexed by the double-

flash illusion (Bidelman, 2016). Our findings show that bilinguals have a more refined multisensory temporal binding window for integrating the auditory and visual senses than monolinguals.

## Methods

### Participants

Twenty-six young adults participated in the experiment: 13 monolinguals (2 male; 11 female) and 13 bilinguals (7 male; 6 female). A language history questionnaire assessed linguistic background (Bidelman, Gandour & Krishnan, 2011; Li, Sepanski & Zhao, 2006). Monolinguals were native speakers of American English unfamiliar with a L2 of any kind. Bilingual participants were classified as late sequential, unimodal multilinguals having received formal instruction in their L2, on average, for  $21.9 \pm 3.01$  years. Average L2 onset age was  $5.8 \pm 3.6$  years. All reported using their first language  $58 \pm 35\%$  of their daily use. Self-reported language aptitude indicated that all were fluent in L2 reading, writing, speaking, and listening proficiency [1(very poor)–7(native-like) Likert scale; reading:  $5.69(0.95)$ ; writing:  $5.53(0.96)$ ; speaking:  $5.46(0.88)$ ; listening:  $5.62(0.87)$ ]. Participants reported their primary language as Bengali (2), French (2), Mandarin (2), Korean (1), Odia (1), Farsi (1), Spanish (2), Teluga (1), and Portuguese (1). Five bilinguals also reported speaking three or more languages. We specifically recruited bilinguals with diverse language backgrounds to increase external validity/generalizability of our study.

The two groups were otherwise similar in age (Mono:  $24.5 \pm 3.4$  yrs, Biling:  $27.7 \pm 3.6$  yrs) and years of formal education (Mono:  $17.9 \pm 2.1$  yrs, Biling:  $18.7 \pm 1.9$  yrs). All showed normal audiometric sensitivity (i.e., pure tone thresholds  $< 25$  dB HL at octave frequencies between 500–8000 Hz), normal or corrected-to-normal vision, were right-handed, and had no previous history of neuropsychiatric illnesses. Musicianship is known to enhance audiovisual binding (Bidelman, 2016; Lee & Noppeney, 2011). Consequently, all participants were required to have minimal ( $< 3$  years) musical training at any point in the lifetime. All were paid for their time and gave informed consent in compliance with a protocol approved by the Institutional Review Board at the University of Memphis.

### Stimuli

Stimuli were constructed to replicate the sound-induced double-flash illusion (Bidelman, 2016; Foss-Feig et al., 2010; Shams et al., 2000; Shams et al., 2002). In this paradigm, the presentation was of multiple auditory stimuli (beeps) concurrent with a SINGLE visual

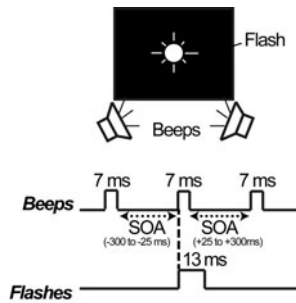


Figure 1. Task schematic for double-flash illusion. Flashes (13.33 ms white disks) were presented on the computer screen concurrent with auditory beeps (7 ms, 3.5 kHz tone) delivered via headphones (*top*). Single trial time course (*bottom*). A single beep was always presented simultaneous with the onset of the flash. A second beep was then presented either before (negative SOAs) or after (positive SOAs) the first. SOAs ranged from  $\pm 300$  ms relative to the single flash. Despite seeing only a single flash, listeners report perceiving two visual flashes indicating that auditory cues modulate the visual percept. The strength of this double-flash illusion varies with the proximity of the second beep (i.e., SOA). Adapted from Bidelman (2016) with permission from Springer-Verlag.

object (flash), that induces an illusory perception of multiple flashes (Shams et al., 2000) (for examples, see: <https://shamslab.psych.ucla.edu/demos/>). Full details of the psychometrics of the illusion with parametric changes in stimulus properties (e.g., number of beeps *re.* flashes, spatial proximity of the visual and auditory cues) can be found in previous psychophysical reports (Innes-Brown & Crewther, 2009; Shams et al., 2000; Shams et al., 2002). Most notably, stimulus onset asynchrony (SOA) between the auditory and visual stimulus pairing can be parametrically varied to either promote or deny the illusory percept. The illusion (i.e., erroneously perceiving two flashes) is higher at shorter SOAs, i.e., when beeps are in closer proximity to the flash. The illusion is less likely (i.e., individuals perceive only a single flash) at long SOAs when the auditory and visual objects are well separated in time. A schematic of the stimulus time course is shown in Figure 1.

On each trial, participants reported the number of flashes they perceived. Each trial was initiated with a fixation cross on the screen. The visual stimulus was a brief (13.33 ms; a single screen refresh) uniform white disk displayed on the center of the screen on a black background, subtending  $\sim 4.5^\circ$  visual angle. In illusory trials, a single flash was accompanied by a pair of auditory beeps, whereas non-illusory trials actually contained two flashes and two beeps. The auditory stimulus consisted of a 3.5 kHz pure tone of 7 ms duration including 3 ms of onset/offset ramping (Shams et al., 2002). In illusory (single flash) trials, two beeps were presented with varying

SOA relative to the single flash. We parametrically varied the SOA between beeps and the single flash from  $-300$  and  $+300$  ms (cf. Foss-Feig et al., 2010) (see Fig. 1). This allowed us to quantify the temporal spacing by which listeners bind auditory and visual cues (i.e., report the illusory percept) and compare the temporal window for audiovisual integration between groups. The onset of one beep always coincided with the onset of the single flash. However, the second beep was either delayed ( $+300$ ,  $+150$ ,  $+100$ ,  $+50$ ,  $+25$  ms) or advanced ( $-300$ ,  $-150$ ,  $-100$ ,  $-50$ ,  $-25$  ms) relative to flash offset. In addition to these illusory (1F/2B) trials, non-illusory (2F/2B) trials were run at SOAs of:  $\pm 300$ ,  $\pm 150$ ,  $\pm 100$ ,  $\pm 50$ ,  $\pm 25$  ms. A total of 30 trials were run for each of the positive/negative SOA conditions, spread across three blocks. Thus in aggregate, there was a total of 300 illusory (1F/2B) and 300 non-illusory (2F/2B) SOA trials. We interleaved illusory and non-illusory conditions to help to minimize response bias effects in the flash-beep task (Mishra, Martinez, Sejnowski & Hillyard, 2007). In addition, 30 trials containing only a single flash and one beep (i.e., 1F/1B) were intermixed with the SOA trials. 1F/1B trials were included as control catch trials and were dispersed randomly throughout the task. Non-illusory trials allowed us to estimate participants' response bias as these trials do not evoke a perceptual illusion and are clearly perceived as having one (1F/1B) or two (2F/2B) flashes, respectively. Illusory (1F/2B) and non-illusory (2F/2B or 1F/1B) conditions were interleaved and trial order was randomized throughout each block. In total, participants performed 630 trials of the task ( $=21$  stimuli $\times 30$  trials).

### Procedure

Listeners were seated in a double-walled sound attenuating chamber (Industrial Acoustics, Inc.)  $\sim 90$  cm from a computer monitor. Stimulus delivery and responses data collection was controlled by E-prime<sup>®</sup> (Psychological Software Tools, Inc.). Visual stimuli were presented as white flashes on a black background via computer monitor (Samsung SyncMaster S24B350HL; nominal 75 Hz refresh rate). Auditory stimuli were presented binaurally using high-fidelity circumaural headphones (Sennheiser HD 280 Pro) at a comfortable level (80 dB SPL). On each trial of the task, listeners indicated via button press whether they perceived "1" or "2" flashes. Participants were aware that trials would also contain auditory stimuli but were instructed to make their response based solely on their perception of the visual stimulus. They were encouraged to respond as accurately and quickly as possible. Both response accuracy and reaction time (RT) were recorded for each stimulus condition. Participants were provided a break after each of the three blocks to avoid fatigue.

## Data analysis

### Behavioral data (% , $d'$ -prime, and RT)

For each SOA per subject, we first computed the mean percentage of trials for which two flashes were reported. For 1F/2B presentations (illusory trials), higher percentages indicate that listeners erroneously perceived two flashes when only one was presented (i.e., the illusion). However, our main dependent measures of behavioral performance were based on signal detection theory (Macmillan & Creelman, 2005), which allowed us to account for listeners' sensitivity and bias in the double-flash task. Signal detection also incorporates both a listeners' sensitivity (hits) and false alarms in perceptual identification and thus is more nuanced than raw %-scores. Behavioral sensitivity ( $d'$ ) was computed using hit ( $H$ ) and false alarm ( $FA$ ) rates for each SOA (i.e.,  $d' = z(H) - z(FA)$ , where  $z(\cdot)$  represents the  $z$ -transform). Bias was computed as  $c = -0.5[z(H) + z(FA)]$ . In the present study, hits were defined as 2F/2B (non-illusory) trials where the listener correctly responded "2 flashes", whereas false alarms were considered 1F/2B (illusory) trials where the listener erroneously reported "2 flashes". Tracing the presence of the double flash illusion across SOAs allowed us to examine the temporal characteristics of multisensory integration and the audiovisual synchrony needed to bind auditory and visual cues. RTs were also computed per condition for each participant, calculated as the median response time between the end of stimulus presentation and execution of the response button press.

Unless otherwise noted, the main dependent measures ( $d'$ , RTs) were analyzed using two-way mixed model ANOVAs with fixed effects of group as the between-subjects factor and SOA as the within-subjects factor. Subjects were modeled as a random effect. Following this omnibus analysis, post hoc multiple comparisons were employed; pairwise contrasts were adjusted using Tukey-Kramer corrections to control Type I error inflation. Unless otherwise noted, the alpha level was set at  $\alpha = 0.05$  for all statistical tests.

### Temporal window quantification

We measured the width of each participant's temporal window to characterize the extent required to accurately perceive the double-flash illusion. Using a  $d' = 1$  (~70% correct performance) as a criterion level of performance (Macmillan & Creelman, 2005, p.9), we quantified the breadth of each person's sensitivity function (see Fig. 2A) as the temporal width where the skirts of each listener's behavioral function exceeded a  $d' = 1$ . This was achieved by spline interpolating ( $N = 1000$  points) each listener's function to provide a more fine-grained step size for measurement. We then repeated this procedure for both the negative (left side) and positive (right side) SOAs of the psychometric function, allowing us to quantify

the width of each portion of the curve and examine possible asymmetries in the temporal window for leading vs. lagging AV stimuli. This procedure was repeated per listener, allowing for a direct comparison between the widths of the temporal binding windows between groups.

## Results

### Behavioral data ( $d'$ -prime)

Sensitivity ( $d'$ ) and response bias for the double-flash task is shown at each SOA in Figures 2A and B, respectively. Results reported in the form of raw proportion of two-flash responses (cf. % correct) is shown in Figure S1 (see Supplementary Materials). Higher  $d'$  is indicative of greater success in AV perception and better sensitivity in differentiating illusory and non-illusory stimuli – i.e., correctly reporting "2F/2B" on actual two flash trials (high hit rate) and avoiding erroneously reporting "2 flashes" for 1F/2B trials (low false alarm rate). Consistent with previous reports (Foss-Feig et al., 2010; Neufeld, Sinke, Zedler, Emrich & Szyck, 2012), both groups showed a similar pattern of responses where the illusion was strong for short SOAs ( $\pm 25$  ms), progressively weakened with increasing asynchrony, and was absent for the longest intervals outside  $\pm 150$ -200 ms (e.g., Fig. 1, Supplementary Materials). Yet, differences in double-flash perception emerged between groups when considering signal detection metrics. A two-way ANOVA conducted on  $d'$  scores revealed a significant group x SOA interaction [ $F_{9, 216} = 7.19, p < 0.0001$ ]. Follow up Tukey-Kramer contrasts revealed higher  $d'$  in bilinguals at SOAs of  $-300$ ,  $-150$ , and  $+300$  ms. These findings reveal that bilinguals better parsed audiovisual cues across several SOA conditions.

### Bias and asymmetry of the psychometric functions

Differences between bilinguals and monolinguals could result from group-specific response biases, e.g., if monolinguals had a higher tendency to report "two flashes." To rule out this possibility, we analyzed bias via signal detection metrics. In the context of the current task, bias values differing from zero indicate a tendency to respond either "2 flashes" (negative bias) or "1-flash" (positive bias) (Stanislaw & Todorov, 1999). Across conditions, we found that response bias was minimal between groups (Fig. 2B). The small positive bias suggests that if anything, listeners tended to more often report "1-flash" across stimuli. Furthermore, while there was a group x SOA interaction in bias ( $F_{9, 216} = 8.99, p < 0.0001$ ), this effect was driven by bilinguals having higher bias at positive SOAs ( $+100$ ,  $+150$ ,  $+300$  ms) where the illusion is generally weakest and group effects in sensitivity ( $d'$ -prime) were not observed (see Fig. 2A).

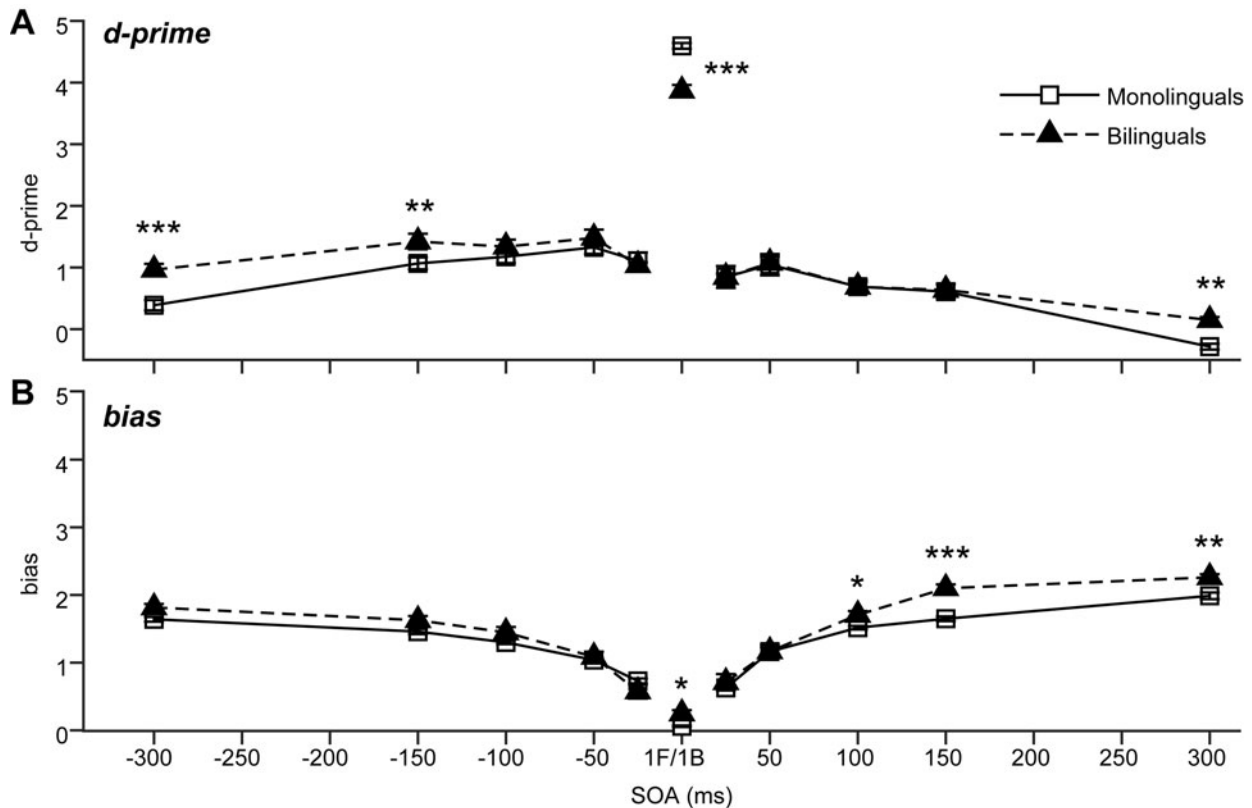


Figure 2. Group differences in perceiving the double-flash illusion. (A)  $d'$  sensitivity scores for correctly reporting “2 flashes” in 2F/2B trials adjusted for false alarms (i.e., “2 flashes” erroneously reported in 1F/2B trials). For the corresponding data expressed as %-accuracy, see Fig. S1 (Supplementary Materials) (B) Response bias. Bilinguals show higher sensitivity in AV processing, particularly at negative SOAs. errorbars =  $\pm 1$  s.e.m.; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Together, signal detection results indicate that bilinguals were more sensitive in correctly identifying veridical (non-illusory) trials and showed less susceptibility to illusory trials (i.e., they better parsed AV events). Moreover, the low bias coupled with the opposite pattern of group effects observed in  $d'$  scores suggests that results are not driven by listeners' inherent decision process or tendency toward a certain response, per se, but rather their sensitivity for audiovisual processing and adjudicating true from illusory flash-beep percepts.

#### Group differences in the temporal binding window

Figure 3A shows the group comparison of the duration of temporal binding window for monolinguals and bilinguals (cf. Bidelman, 2016; Foss-Feig et al., 2010). Results show that the width of monolingual's temporal window is wider than that of bilinguals overall (Biling.: [−65 – 112] ms, Mono.: [−192 – 112] ms;  $t_{24} = 2.72$ ,  $p = 0.0118$ ). This was attributable to bilinguals having shorter windows for negative SOA conditions ( $t_{24} = 3.18$ ,  $p = 0.0041$ ). Thus, bilinguals showed more precise multisensory processing (in terms of  $d'$ ) for lagging AV

stimuli, suggesting an asymmetry in audiovisual binding. Lastly, musical training was not correlated with temporal window durations in neither monolinguals [ $r = -0.07$ ,  $p = 0.81$ ] nor bilinguals [ $r = -0.09$ ,  $p = 0.77$ ]. However, this might be expected given that all participants had minimal (< 3 years) musical training.

We observed an asymmetry in the psychometric  $d'$  functions audiovisual stimuli (see Fig. 2A and 3A). To further quantify this asymmetry, we measured skewness of the psychometric functions computed as the third central statistical moment of the  $d'$  curves shown in Fig. 2A. Positive values denote asymmetry of the psychometric function with skewness tilted rightward and thus more susceptibility to the illusion (i.e., lower  $d'$ ) for positive (lagging) SOAs, whereas negative values reflect less susceptibility (higher  $d'$ ) for lagging SOAs. Psychometric skewness by group is shown in Fig. 3B. Bilinguals showed larger positive skew than monolinguals [ $z = -2.31$ ,  $p = 0.021$ ; Wilcoxon rank sum test (used given heterogeneity in variance)]. Larger positive skew in monolinguals' identification indicates they performed more poorly in the double-flash illusion particularly for audio lagging stimuli – and, conversely, that bilinguals performed better

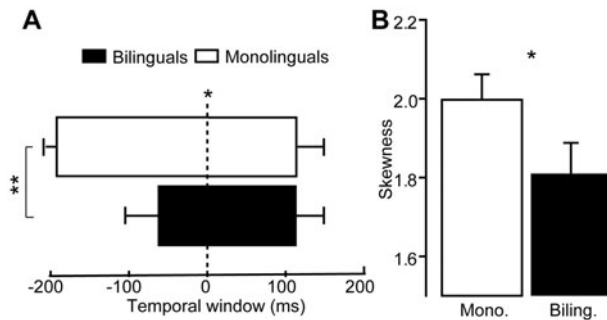


Figure 3. Temporal window duration and skewness of the psychometric functions for monolinguals and bilinguals. (A) Temporal binding window duration computed as the width (SOAs) at which each listener’s psychometric function (i.e., Fig. 2A) exceeded the criterion of  $d' = 1$ . Windows are shorter in bilinguals overall indicating more precise multisensory processing of AV stimuli. However, group differences are generally stronger in the negative SOA direction. (B) Skewness of the psychometric function, measured as the third statistical moment of the  $d'$  curves. Non-zero values denote asymmetry in psychometric function. Monolinguals’ psychometric functions are more positively skewed than bilinguals’, indicating poorer sensitivity in audio lagging conditions (i.e., positive SOAs). errorbars =  $\pm 1$  s.e.m.; \* $p < 0.05$ , \*\* $p < 0.01$ .

at negative, leading SOAs. These results corroborate the asymmetry observed in temporal binding windows between positive vs. negative SOAs (i.e., Fig. 3A).

**Reaction times (RTs)**

Group reaction times across SOAs are shown in Figure 4 for (A) illusory and (B) non-illusory trials. An omnibus 3-way ANOVA revealed a significant SOA x trial type x group interaction [ $F_{9, 480} = 13.02, p < 0.0001$ ]. To parse this three-way interaction, separate 2-way ANOVAs (group x SOA) were conducted on RTs split by illusory and non-illusory trials. This analysis revealed a significant group x SOA interaction on behavioral RTs to illusory trials [ $F_{9, 216} = 15.14, p < 0.0001$ ]. Follow-up contrasts revealed that bilinguals were faster at making their response than monolinguals for the majority of SOAs (all but  $-300, -150, -25,$  and  $+300$  ms). A similar pattern of results was found for non-illusory trials (Fig. 4B) [group x SOA interaction:  $F_{9, 216} = 8.38, p < 0.0001$ ], where bilinguals showed faster behavioral responses across all but the  $-50$  and  $\pm 25$  ms SOAs<sup>1</sup>. Collectively, RT findings

<sup>1</sup> Rather large group differences (bilinguals  $\ll$  monolinguals) were observed in RTs to the 1F/1B (control) trials. These trials were less frequent than the 1F/2B and 2F/2B trials and were the only to feature a single flash and single beep. Speculatively, it is possible that monolinguals found this condition more distracting or were waiting for an additional stimulus event (i.e., were more uncertain)

indicate that bilingual participants were not only more accurate at processing concurrent multisensory cues than monolinguals but were faster at judging the composition of audiovisual stimuli.

**Discussion**

We measured multisensory integration in monolinguals and bilinguals via the double flash illusion (Bidelman, 2016; Shams et al., 2000), a task requiring the perceptual binding of temporally offset auditory and visual cues. Collectively, our results indicate that bilinguals are (i) faster and more accurate at processing concurrent audiovisual objects than their monolingual peers and (ii) show more refined (narrower) temporal windows for multisensory integration and audiovisual binding. These findings reveal that experience-dependent plasticity of intensive language experience improves the integration of information from multiple sensory systems (audition and vision). Accordingly, our data also suggest that bilinguals may not have the same time-accuracy tradeoff in AV perception as monolinguals, since they achieve higher accuracy (sensitivity) without the expense of slower speeds (cf. Figs. 2 and 4). These data extend our previous studies showing similar experience-dependent plasticity in AV processing (Bidelman, 2016) and time-accuracy benefits (Bidelman, Hutka & Moreno, 2013) in trained musicians.

**Domain-general benefits of bilinguals’ plasticity**

The present data reveal that the benefits of bilingualism seem to extend beyond simple auditory processing and enhance multisensory integration. They further extend recent work on bilingualism and multisensory integration for SPEECH stimuli (e.g., Burfin et al., 2014; Reetzke, Lam, Xie, Sheng & Chandrasekaran, 2016) by demonstrating comparable enhancements to NON-SPEECH audiovisual stimuli. Here, we show that bilinguals experience a shorter temporal window for AV integration, have enhanced multimodal processing, and more efficient/accurate representations for perceptual audiovisual objects<sup>2</sup>. Accordingly, our data provide

than bilinguals. Why bilinguals did not experience this same lapse is unclear but could relate to the higher executive control noted in the bilingual literature (Bialystok et al., 2003; Bialystok et al., 2007; Bialystok, 2009, Bialystok and DePape, 2009, Krizman et al., 2014, Schroeder et al., 2016).

<sup>2</sup> In the present study, we interpret a narrower temporal binding window as an *enhancement* in AV processing. However, an argument in the opposite direction could be made such that having a wider binding window might be beneficial as it would allow for the integration of AV stimuli farther apart in time. That a *narrower* binding window in bilinguals represents *enhanced* AV processing is evident based on the nature of the double-flash task and previous studies. First, the

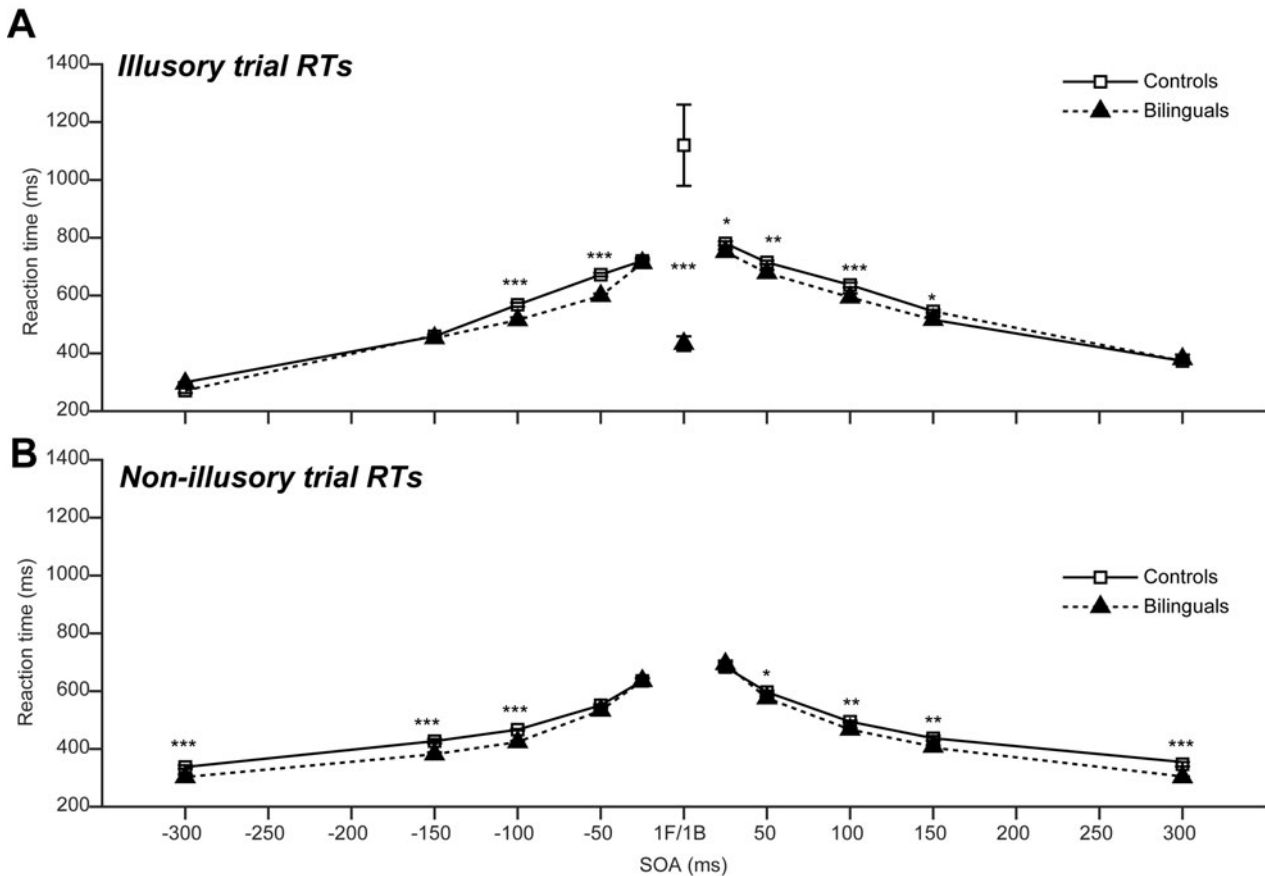


Figure 4. Reaction times by group. Across the board for both illusory (A) and non-illusory (B) trials, bilinguals show faster decisions than monolinguals when judging audiovisual stimuli. Bilinguals are not only more sensitive in processing concurrent audiovisual cues (e.g., Fig. 2) with a more precise temporal binding window (Fig. 3) but on average, also respond faster than monolinguals. errorbars =  $\pm 1$  s.e.m.; group difference ( $RT_{\text{biling}} < RT_{\text{mono}}$ ): \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

evidence that intense auditory experience afforded by speaking two languages hones AV processing and the multisensory binding window in an experience-dependent manner (cf. Ressel, Pallier, Ventura-Campos, Diaz, Roessler, Avila & Sebastian-Gallés, 2012). While our bilingual cohort included a variety of L1 backgrounds, our data cannot speak to how/if different native languages affect audiovisual temporal integration differentially. For example, bilinguals could be more accurate in temporal binding because their native languages entail audiovisual

integration on shorter timescales (i.e., temporal binding windows). Future studies are needed to determine if AV processing and temporal binding vary in a language-dependent manner.

Nevertheless, it is possible that the more refined audiovisual processing seen here in bilinguals might instead result from an augmentation of more general cognitive mechanisms. Bilinguals, for example, are known to have improved selective attention, inhibitory control, and executive functioning (Bialystok, 2009; Bialystok, Craik & Freedman, 2007; Bialystok & DePape, 2009; Bialystok, Majumder & Martin, 2003; Krizman et al., 2014; Schroeder et al., 2016). Distributing attention across the sensory modalities enhances performance in complex audiovisual tasks (Mishra & Gazzaley, 2012). Therefore, if bilingualism increases and/or enables one to deploy attentional resources more effectively (e.g., Krizman et al., 2014) – possibly across modalities – this could account for the cross-modal enhancements observed here. Future work is needed to tease apart these perceptual and cognitive accounts.

task itself requires individuals to adjudicate an audiovisual illusion; wider windows represent more false reports across a wider range of SOAs and thus a poorer perception of the physical characteristics of the AV stimuli. Second, studies using the double-flash and similar paradigms show that certain disorders (e.g., autism, language learning impairments) widen the AV temporal binding window (Foss-Feig et al, 2010; Kaganovich et al., 2014; Wallace & Stevenson, 2014) and produce a perceptual deficit rather than enhancement.

The double-flash illusion requires a behavioral decision on the visual stimulus that must be informed by the perception of a concurrent auditory event. As such, it is often considered a measure of multisensory integration (Foss-Feig et al., 2010; Mishra et al., 2007; Powers et al., 2009). Nevertheless, the better behavioral performance of bilinguals in the double-flash effect could result from enhanced unisensory or temporal processing (i.e., resolving multiple events) rather than multisensory integration, per se. We are unaware of data to suggest enhanced TEMPORAL resolution in bilinguals. Moreover, if this were the case, we might have expected more pervasive group differences across the board. Instead, we found an interaction in the behavioral pattern (e.g., Fig. 2A). Moreover, while neuroimaging studies of the double-flash illusion have shown engagement both unisensory (auditory, visual) and polysensory brain areas (Mishra, Martinez & Hillyard, 2008; Mishra et al., 2007), it is the latter (i.e., cross-modal interactions) which drive the illusory percept. Future neuroimaging studies could be used to evaluate the relative contribution of unisensory/multi-sensory brain mechanisms and the role of temporal processing in bilinguals' shorter temporal windows.

What might be the broader implications of bilinguals' enhanced AV processing? In addition to domain-general benefits in multisensory perception, one implication of bilingual's improved AV binding might be to facilitate speech perception for their L2, particularly in adverse listening conditions. Indeed, bilinguals show much poorer speech-in-noise comprehension when listening to their L2 (i.e., nonnative speech) (Bidelman & Dexter, 2015; Hervais-Adelman, Pefkou & Golestani, 2014; Rogers, Lister, Febo, Besing & Abrams, 2006; Tabri, Smith, Chacra & Pring, 2010; von Hapsburg, Champlin & Shetty, 2004; Zhang, Stuart & Swink, 2011). Speech-in-noise perception is improved with the inclusion of visual information from the speaker (Erber, 1975; Ross et al., 2007; Vatikiotis-Bateson et al., 1998) as in cases of lip-reading (i.e., "hearing lips": Bernstein, Auer Jr & Takayanagi, 2004; Navarra & Soto-Faraco, 2007). Visual speech movements are also known to augment second language perception by way of multisensory integration (Navarra & Soto-Faraco, 2007). Presumably, bilinguals could compensate for their normal deficits in degraded L2 speech listening (e.g., Bidelman & Dexter, 2015; Krizman, Bradlow, Lam & Kraus, 2016; Rogers et al., 2006) if they are better able to combine and integrate auditory and visual modalities.

#### ***Putative biological mechanisms of the double-flash illusion***

From a biological perspective, neurophysiological studies have shed light on how visual and auditory cues

interact within the various sensory systems. Visual evoked potentials to the double-flash stimuli used here show modulations in neural responses dependent on the perception of the illusion (Shams, Kamitani, Thompson & Shimojo, 2001). Interestingly, brain potentials for illusory flashes (1F/2B) are qualitatively similar to those elicited by an actual physical flash (Shams et al., 2001). These findings suggest that activity in visual cortex is not only modulated by the auditory input but that the pattern of neural activity is remarkably similar when one perceives an illusory visual object as when it actually occurs in the environment. That is, endogenously generated brain activity (representing the illusion) seems to closely parallel neural representations observed during exogenous stimulus coding.

Cross-modal interactions within sensory brain regions have also been observed in human neuromagnetic brain responses to auditory and visual stimuli (Raij, Ahveninen, Lin, Witzel, Jääskeläinen, Letham, Israeli, Sahyoun, Vasios, Stufflebeam, Hämäläinen & Belliveau, 2010). These studies reveal that while cross-sensory (auditory→visual) activity generally manifests later (~10–20 ms) than sensory-specific (auditory→auditory) activations, there is a stark asymmetry in the arrival of information between Heschl's gyrus and the Calcarine fissure. Auditory information is combined in visual cortex roughly 45 ms faster than the reverse direction of travel (i.e., visual→auditory) (Raij et al., 2010). Thus, auditory information seems to dominate when the two senses are integrated. An asymmetry in the flow and dominance of auditory→visual information may account for illusory percepts observed in our double-flash paradigm, where individuals perceive multiple flashes due to the presence of an "overriding" auditory cue.

Conceivably, bilingualism might change this brain organization and enhance functional connectivity between sensory systems that are highly engaged by speech-language processing (i.e., audition, vision, motor). In monolingual nonmusicians, prior studies have indicated that the likelihood of perceiving the double flash illusion is highly correlated with white matter connectivity between occipito-parietal regions, the putative ventral/dorsal streams comprising the "what/where" pathways (Kaposvari, Csete, Bogner, Csibri, Toth, Szabo, Vecsei, Sary & Kincses, 2015). This suggests that parallel visual channels play an important role in audiovisual interactions and the temporal binding of disparate cues as required by double-flash percepts (Shams et al., 2000; Shams et al., 2002). It is possible that bilinguals might show more refined temporal binding of auditory and visual events as we observe behaviorally due to increased functional connectivity between the auditory and visual systems or temporoparietal regions known to integrate disparate audiovisual information (Erickson, Zielinski, Zielinski,



Liu, Turkeltaub, Leaver & Rauschecker, 2014; Man, Kaplan, Damasio & Meyer, 2012). Additionally, recent EEG evidence suggests that alpha (~10 Hz) oscillations are a crucial factor in determining the susceptibility to the illusion and the size of the temporal binding window; individuals whose intrinsic alpha frequency is lower than average have longer (enlarged) temporal binding windows, whereas those having higher alpha frequency show more refined (shorter) windows (Cecere, Rees & Romei, 2015). Future neuroimaging experiments are warranted to test these possibilities and identify the neural mechanisms underlying bilingual's AV binding seen here and previously in other expert listeners (e.g., musicians; Bidelman, 2016).

### *Asymmetries in audiovisual processing*

Detailed comparison of each group's psychometric responses revealed that bilinguals did not show improved AV across the board. Rather, their enhanced temporal binding was restricted to certain (mainly negative) SOAs (see Fig. 2A and 3A). This perceptual asymmetry was corroborated via measures of psychometric skewness, which showed that monolinguals had more positively skewed behavioral responses than bilinguals, and were thus more susceptible to the double-flash illusion for audio lagging stimuli. The mechanisms underlying this perceptual asymmetry are unclear but could relate to the well-known psychophysical asymmetries observed in audiovisual perception. For instance, several studies have shown a differential sensitivity in detecting audiovisual mismatches for leading compared to lagging AV events (Cecere, Gross & Thut, 2016; van Eijk, Kohlrausch, Juola & van de Par, 2008; Wojtczak, Beim, Micheyl & Oxenham, 2012; Younkin & Corriveau, 2008). Interestingly, telecommunication broadcast standards exploit these perceptual asymmetries and allow for nearly twice the temporal offset for a delayed (compared to advanced) audio channel relative to the video signal (ITU, 1998; ATSC, 2003). Perceptual asymmetries in audiovisual lags vs. leads may reflect physical properties of electromagnetic wave propagation. Light travels faster than sound and thus implies a causal relation in the expected timing between modalities. As such, human observers naturally expect the arrival of visual information prior to auditory events. From a biological standpoint, recent studies suggest different integration mechanisms may underpin audio-first vs. visual-first binding (Cecere et al., 2016). Moreover, positive SOAs are also thought to be more critical for other forms of audiovisual processing (Cecere et al., 2016). Thus, both physical and physiological explanations may account for perceptual asymmetries observed in AV asynchrony studies and may underlie the differential pattern (i.e., skew) in AV responses observed between language groups and why we find they are

restricted to positive SOA conditions. Future studies are needed to fully explore the perceptual asymmetries in AV processing and how they are modulated by auditory training and/or language experience.

### **Supplementary material**

To view supplementary material for this article, please visit <https://doi.org/10.1017/S1366728918000408>

### **References**

- International Telecommunications Union (ITU). (1998). Relative timing of sound and vision for broadcasting (pp. 1–5). Technical Report, Geneva, Switzerland.
- ATSC. (2003). Relative Timing of Sound and Vision for Broadcast Operations. In *Implementation Subcommittee Finding: Doc. IS-191, 26 June, 2003*.
- Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Audiovisual cues benefit recognition of accented speech in noise but not perceptual adaptation. *Frontiers in Human Neuroscience, 9*, 422. doi:10.3389/fnhum.2015.00422
- Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*(1–4), 5–18. doi:<https://doi.org/10.1016/j.specom.2004.10.011>
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11.
- Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459–464. doi:10.1016/j.neuropsychologia.2006.10.009
- Bialystok, E., & DePape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 565–574. doi:10.1037/a0012735
- Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24*(01), 27–44. doi:10.1017/S014271640300002X
- Bidelman, G. M. (2016). Musicians have enhanced audiovisual multisensory binding: Experience-dependent effects in the double-flash illusion. *Experimental Brain Research, 234*(10), 3037–3047.
- Bidelman, G. M., & Dexter, L. (2015). Bilinguals at the “cocktail party”: Dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-noise recognition deficits. *Brain and Language, 143*, 32–41.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*(2), 425–434. doi:10.1162/jocn.2009.21362
- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch:

- Evidence for bidirectionality between the domains of language and music. *PLoS One*, 8(4), e60676. doi:[10.1371/journal.pone.0060676](https://doi.org/10.1371/journal.pone.0060676)
- Burfin, S., Pascalis, O., Ruiz Tada, E., Costa, A., Savariaux, C., & Kandel, S. (2014). Bilingualism affects audiovisual phoneme identification. *Frontiers in Psychology*, 5, 1179. doi:[10.3389/fpsyg.2014.01179](https://doi.org/10.3389/fpsyg.2014.01179)
- Cecere, R., Gross, J., & Thut, G. (2016). Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *European Journal of Neuroscience*, 43(12), 1561–1568. doi:[10.1111/ejn.13242](https://doi.org/10.1111/ejn.13242)
- Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology*, 25(2), 231–235. doi:<https://doi.org/10.1016/j.cub.2014.11.034>
- Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders*, 40(4), 481–492.
- Erickson, L. C., Zielinski, B. A., Zielinski, J. E. V., Liu, Turkeltaub, P. E., Leaver, A. M., & Rauschecker, J. P. (2014). Distinct cortical locations for integration of audiovisual speech and the McGurk effect. *Frontiers in Psychology*, 5(534). doi:[10.3389/fpsyg.2014.00534](https://doi.org/10.3389/fpsyg.2014.00534)
- Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203(2), 381–389.
- Hervais-Adelman, A., Pefkou, M., & Golestani, N. (2014). Bilingual speech-in-noise: Neural bases of semantic context use in the native language. *Brain and Language*, 132, 1–6.
- Innes-Brown, H., & Crewther, D. (2009). The impact of spatial incongruence on an auditory-visual illusion. *PLoS One*, 4(7), e6450. doi:[10.1371/journal.pone.0006450](https://doi.org/10.1371/journal.pone.0006450)
- Kaganovich, N., Schumaker, J., Leonard, L. B., Gustafson, D., & Macias, D. (2014). Children with a history of SLI show reduced sensitivity to audiovisual temporal asynchrony: An ERP study. *Journal of Speech, Language, and Hearing Research*, 57(4), 1480–1502. doi:[10.1044/2014\\_JSLHR-L-13-0192](https://doi.org/10.1044/2014_JSLHR-L-13-0192)
- Kaposvari, P., Csete, G., Bognar, A., Csibri, P., Toth, E., Szabo, N., Vecsei, L., Sary, G., & Kincses, Z. T. (2015). Audio-visual integration through the parallel visual pathways. *Brain Research*, 1624, 71–77. doi:[10.1016/j.brainres.2015.06.036](https://doi.org/10.1016/j.brainres.2015.06.036)
- Krizman, J., Bradlow, A. R., Lam, S. S.-Y., & Kraus, N. (2016). How bilinguals listen in noise: linguistic and non-linguistic factors. *Bilingualism: Language and Cognition*, 1–10. doi:[10.1017/S1366728916000444](https://doi.org/10.1017/S1366728916000444)
- Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. *Brain and Language*, 128(1), 34–40. doi:[10.1016/j.bandl.2013.11.006](https://doi.org/10.1016/j.bandl.2013.11.006)
- Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J., -F. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, doi:[10.1073/pnas.1410963111](https://doi.org/10.1073/pnas.1410963111). doi:[10.1073/pnas.1410963111](https://doi.org/10.1073/pnas.1410963111)
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606–608.
- Lee, H. L., & Noppeney, U. (2011). Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), E1441–1450. doi:[10.1073/pnas.1115267108](https://doi.org/10.1073/pnas.1115267108)
- Lee, H. L., & Noppeney, U. (2014). Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech and music. *Frontiers in Psychology*, 5(868), 1–9. doi:[10.3389/fpsyg.2014.00868](https://doi.org/10.3389/fpsyg.2014.00868)
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavioral Research Methods*, 38(2), 202–210.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.
- Man, K., Kaplan, J. T., Damasio, A., & Meyer, K. (2012). Sight and sound converge to form modality-invariant representations in temporoparietal cortex. *Journal of Neuroscience*, 32(47), 16629–16636. doi:[10.1523/jneurosci.2342-12.2012](https://doi.org/10.1523/jneurosci.2342-12.2012)
- Mishra, J., & Gazzaley, A. (2012). Attention distributed across sensory modalities enhances perceptual performance. *Journal of Neuroscience*, 32(35), 12294–12302. doi:<https://doi.org/10.1523/JNEUROSCI.0867-12.2012>
- Mishra, J., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Research*, 1242, 102–115. doi:[10.1016/j.brainres.2008.05.023](https://doi.org/10.1016/j.brainres.2008.05.023)
- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience*, 27, 4120–4131.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40), 15894–15898. doi:[10.1073/pnas.0701498104](https://doi.org/10.1073/pnas.0701498104)
- Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological Research*, 71(1), 4–12. doi:[10.1007/s00426-005-0031-5](https://doi.org/10.1007/s00426-005-0031-5)
- Neufeld, J., Sinke, C., Zedler, M., Emrich, H. M., & Szycik, G. R. (2012). Reduced audio-visual integration in synaesthetes indicated by the double-flash illusion. *Brain Research*, 1473, 78–86. doi:<https://doi.org/10.1016/j.brainres.2012.07.011>
- Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychol Sci*, 26(4), 490–498. doi:[10.1177/0956797614568320](https://doi.org/10.1177/0956797614568320)

- Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *Journal of Neuroscience*, *29*, 12265–12274.
- Raij, T., Ahveninen, J., Lin, F.-H., Witzel, T., Jääskeläinen, I. P., Letham, B., Israeli, E., Sahyoun, C., Vasios, C., Stufflebeam, S., Hämäläinen, M., & Belliveau, J. W. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, *31*(10), 1772–1782. doi:10.1111/j.1460-9568.2010.07213.x
- Reetzke, R., Lam, B. P. W., Xie, Z., Sheng, L., & Chandrasekaran, B. (2016). Effect of simultaneous bilingualism on speech intelligibility across different masker types, modalities, and signal-to-noise ratios in school-age children. *PLoS One*, *11*(12), e0168048. doi:10.1371/journal.pone.0168048
- Ressel, V., Pallier, C., Ventura-Campos, N., Diaz, B., Roessler, A., Avila, C., & Sebastian-Gallés, N. (2012). An effect of bilingualism on the auditory cortex. *Journal of Neuroscience*, *32*(47), 16597–16601. doi:10.1523/JNEUROSCI.1996-12.2012
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(03), 465–485. doi:10.1017/S014271640606036X
- Ronquest, R. E., Levi, S. V., & Pisoni, D. B. (2010). Language identification from visual-only speech signals. *Attention, perception & psychophysics*, *72*(6), 1601–1613. doi:10.3758/APP.72.6.1601
- Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, *21*(3-4), 185–192. doi:10.1007/s10548-009-0090-9
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*(5), 1147–1153. doi:10.1093/cercor/bh1024
- Schroeder, S. R., Marian, V., Shook, A., & Bartolotti, J. (2016). Bilingualism and Musicianship Enhance Cognitive Control. *Neural Plasticity*, *2016*, 4058620. doi:10.1155/2016/4058620
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(14), 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*, 147–152.
- Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *Neuroreport*, *12*(17), 3849–3852.
- Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastian-Gallés, N., & Werker, J. F. (2007). Discriminating languages by speech-reading. *Perception and Psychophysics*, *69*(2), 218–231.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. doi:10.3758/BF03207704
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212–215.
- Tabri, D., Smith, K. M., Chacra, A., & Pring, T. (2010). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders*, 1–12.
- van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception and Psychophysics*, *70*(6), 955–968. doi:10.3758/pp.70.6.955
- Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception and Psychophysics*, *60*, 926–940.
- von Hapsburg, D., Champlin, C. A., & Shetty, S. R. (2004). Reception thresholds for sentences in bilingual (spanish/english) and monolingual (english) listeners. *Journal of the American Academy of Audiology*, *15*, 88–98.
- Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64C*, 105–123. doi:10.1016/j.neuropsychologia.2014.08.005
- Wojtczak, M., Beim, J. A., Micheyl, C., & Oxenham, A. J. (2012). Perception of across-frequency asynchrony and the role of cochlear delays. *Journal of the Acoustical Society of America*, *131*(1), 363–377. doi:10.1121/1.3665995
- Younkin, A. C., & Corriveau, P. J. (2008). Determining the Amount of Audio-Video Synchronization Errors Perceptible to the Average End-User. *IEEE Transactions on Broadcasting*, *54*(3), 623–627. doi:10.1109/TBC.2008.2002102
- Zhang, J., Stuart, A., & Swink, S. (2011). Word recognition by English monolingual and Mandarin-english bilingual speakers in continuous and interrupted noise. *Canadian Journal of Speech-Language Pathology and Audiology*, *35*(4), 322–331.