NEUROSCIENCE

RESEARCH ARTICLE

Neural Correlates of Enhanced Audiovisual Processing in the Bilingual Brain

Gavin M. Bidelman [a,b,c*] **and Shelley T. Heath** [a]

[a] *School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA*

[b] *Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA*

[c] *University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA*

Abstract—**Bilingualism is associated with enhancements in perceptual and cognitive processing necessary for juggling multiple languages. Recent psychophysical studies demonstrate bilinguals also show enhanced multi-sensory processing and more restricted temporal binding windows for integrating audiovisual information. Here, we probed the neural mechanisms of bilinguals' audiovisual benefits. We recorded neuroelectric responses in mono- and bi-lingual listeners to the double-flash paradigm in which auditory beeps concurrent with a single visual flash induces the perceptual illusion of multiple flashes. Relative to monolinguals, bilinguals showed less susceptibility to the illusion (fewer false perceptual reports) coupled with stronger and faster event-related potentials to audiovisual information. Source analyses of EEG data revealed monolinguals' increased propensity for erroneously perceiving audiovisual stimuli was attributed to increased activity in primary visual (V1) and auditory cortex (PAC), increases in multisensory association areas (BA 37), but reduced frontal activity (BA 10). Regional activations were associated with an opposite pattern of behaviors: whereas stronger V1 and PAC activity predicted slower behavioral responses, stronger frontal BA10 responses elicited faster judgments. Our results suggest bilinguals' higher precision in audiovisual perception reflects more veridical sensory coding of physical cues coupled with superior top-down gating of sensory information to suppress the generation of false percepts. Findings underscore that the plasticity afforded by speaking multiple languages shapes extra-linguistic brain regions and can enhance audiovisual brain processing in a domain-general manner.** © 2019 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: audiovisual integration, bilingualism, double-flash illusion, temporal binding, experience-dependent plasticity.

INTRODUCTION

Modern society places an ever-increasing demand for multilingual communication. It is now estimated that nearly half the world speaks multiple languages (Grosjean, 2010). Consequently, there is interest in understanding how different language experiences (bilingualism) sculpt brain function and potentially enhance different perceptual-cognitive skills (Bialystok et al., 2012; Ressel et al., 2012; Costa and Sebastian-Galles, 2014). In this regard, studies have suggested that bilingualism is associated with better executive functioning (Bialystok and DePape, 2009), inhibitory/cognitive control (Crinion et al., 2006; Bialystok and DePape, 2009; Blumenfeld and Marian, 2011; Bialystok et al., 2012; Krizman et al.,

2012; Moreno et al., 2014; Schroeder et al., 2016), speech processing (Bidelman et al., 2011; Krizman et al., 2012, 2016), working memory (Bidelman et al., 2013; Calvo et al., 2016), and neuroprotective effects against cognitive aging (Bialystok et al., 2007; Gold et al., 2013a,b). While manipulating speech sounds of multiple languages is arguably an auditory skill, language is a multimodal experience. Consequently, audiovisual (AV) interactions are highly relevant to speech perception. Indeed, the combination of visual and auditory cues enhances spoken word recognition (Sumby and Pollack, 1954; Vatikiotis-Bateson et al., 1998; Ross et al., 2007). Given its importance in shaping our perceptual world, understanding how (or if) different experiential factors or disorders can modulate AV processing is of interest to examine the extent to which these fundamental mechanisms of perception–action are malleable to the neuroplasticity of different human experiences (e.g., Foss-Feig et al., 2010; Bidelman, 2016).

Recent studies have indeed suggested that bilingualism may alter multisensory processing of

---

[*]Correspondence to: G. M. Bidelman, School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA.

E-mail address: gmbdlman@memphis.edu (G. M. Bidelman).
Abbreviations: AV, audiovisual; BF, Bayes factor; L2, second language; PAC, primary auditory cortices; PCA, principal component analysis; SOA, stimulus onset asynchrony.

speech stimuli (Burfin et al., 2014; Reetzke et al., 2016). During speech comprehension, integrating a speaker's visual cues with sound is larger when information from the auditory modality is unfamiliar, as in the case of listening to nonnative or accented speech (Banks et al., 2015). Thus, enhanced multisensory processing might help bilinguals improve their second language (L2) understanding by providing better integration of the auditory and visual elements of speech. Indeed, in a recent psychophysical report, we demonstrated that bilinguals show enhanced multisensory processing and temporal binding of AV cues even for *non-speech* stimuli (Bidelman and Heath, 2019). Using the "double-flash illusion" paradigm (Shams et al., 2000; Foss-Feig et al., 2010; Bidelman, 2016), we found that bilinguals' temporal binding windows for audiovisual perception were roughly 2× narrower than in monolinguals, indicating more precise integration of multisensory cues (Bidelman and Heath, 2019). Our findings implied that the plasticity afforded by speaking multiple languages enhances not only auditory processing (e.g., Krizman et al., 2012; Bidelman et al., 2013), but also how the bilingual brain combines information from multiple senses. Nevertheless, the purely behavioral nature of the study did not allow us to delineate the underlying neural mechanisms responsible for bilinguals' perceptual advantages in AV processing.

Here, we extend these results by recording neuroelectric brain activity in monolinguals and bilinguals during the "double-flash illusion" (Shams et al., 2000, 2002) in order to characterize the neural correlates of bilinguals' enhanced AV processing and perceptual binding of multisensory events (Bidelman and Heath, 2019). In this paradigm, the presentation of multiple auditory beeps concurrent with a *single* visual flash induces a perceptual illusion of multiple flashes. These non-speech stimuli show the degree to which auditory cues influence the visual input, mirroring other well-known audiovisual influences in the reverse (V → A) direction (McGurk and MacDonald, 1976). They are also ideal for studying domain-generality in AV processing since the stimuli are void of lexical-semantic meaning that might otherwise confound interpretation in a cross-linguistic study. We varied the onset asynchrony between auditory and visual events (leads and lags) to assess group differences in the "temporal window" for binding audiovisual perceptual objects. Source imaging of the EEG data assessed the relative contribution of different brain regions underlying the integration of multisensory information and generation of illusory AV percepts, including auditory, visual, associative, and frontal cortices. Our findings show that bilinguals' higher precision in AV perception, as indexed by the double-flash paradigm, reflects more veridical coding of physical stimulus cues coupled with superior top-down gating of sensory information to suppress the generation of false percepts.

EXPERIMENTAL PROCEDURES

Participants

The current sample represents a subset of individuals who participated in our previous psychophysical study on bilingualism and AV processing (Bidelman and Heath, 2019) for which neuroimaging data were also obtained. Twenty-three young adults participated in the experiment: 13 monolinguals (2 male; 11 female) and 10 bilinguals (4 male; 6 female). A language history questionnaire assessed linguistic background (Li et al., 2006; Bidelman et al., 2011). Monolinguals were native speakers of American English unfamiliar with a L2 of any kind. Bilingual participants were classified as late sequential, bi- or multi-linguals having received formal instruction in their L2, on average, for 21.4 ± 3.1 years. Average L2 onset age was 6.0 ± 3.9 years. All reported using their first language $58.8 \pm 37.3\%$ 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: 5.7 (1.1); writing: 5.6 (1.1); speaking: 5.5 (0.9); listening: 5.6 (0.9)]. Participants reported their primary language as French (2), Mandarin (2), Korean (1), Odia (1), Farsi (1), Spanish (2), and Tegula (1). Four bilinguals also reported speaking three or more languages. While all bilinguals spoke English as their L2, we specifically recruit bilinguals with a range of L1 backgrounds to increase the external validity of the study (Bidelman and Dexter, 2015; Bidelman and Heath, 2019).

The two groups were otherwise similar in age (Mono: 24.9 ± 3.5 yrs, Biling: 27.4 ± 3.3 yrs; $t_{19} = -1.67$, $p = 0.11$) and years of formal education (Mono: 18.1 ± 2.3 yrs, Biling: 18.6 ± 2.1 yrs; $t_{19} = -0.59$, $p = 0.56$). All showed normal hearing thresholds (i.e., < 25 dB HL; 500–8000 Hz, octave frequencies), normal or corrected-to-normal vision, were right-handed ($> 70\%$ laterality; Oldfield, 1971), and had no previous history of neuro-psychiatric illness. Musicianship is known to enhance audiovisual processing and perception of the double-flash illusion (Lee and Noppeney, 2011; Bidelman, 2016). Consequently, all participants were required to have < 3 years of lifetime musical training. 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 comprised the double-flash illusion (Shams et al., 2000, 2002; Foss-Feig et al., 2010; Bidelman, 2016; Bidelman and Heath, 2019). In this paradigm, the presentation of multiple auditory stimuli (beeps) concurrent with a *single* visual object (flash) induces an illusory perception of multiple flashes (Shams et al., 2000). Complete details of the psychometrics of the illusion in monolingual and bilingual individuals can be found in our companion paper (Bidelman and Heath, 2019). We parametrically varied the stimulus onset asynchrony (SOA) between the auditory and visual stimulus pairs to either promote or deny the illusory percept. The illusion (i.e., erroneously perceiving two flashes) is higher at shorter SOAs when beeps are in closer proximity to the flash and conversely, is less likely (i.e., individuals perceive a single flash) at longer SOAs when the auditory and visual objects are temporally separated.

On each trial, participants were asked to report the number of *visual flashes* they perceived (Fig. 1A). 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 ~4.5° visual angle. In illusory trials, a single flash (F) was accompanied by a pair of 80 dB SPL auditory beeps (B), 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. The SOA between beeps and the single flash varied from −300 to +300 ms (cf. Foss-Feig et al., 2010) (Fig. 1A). This allowed us to quantify the temporal spacing by which listeners bind auditory and visual cues (i.e., report the illusory percept). 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 illusory (1F/2B) trials, non-illusory (2F/2B) trials were run at SOAs of ±300, ±150, ±100, ±50, ±25 ms. A total of 30 trials were run for each of the positive/negative SOA conditions, spread across three blocks. Thus, there were a total of 300 illusory (1F/2B) and 300 non-illusory (2F/2B) SOA trials in aggregate. We interleaved illusory and non-illusory conditions to help to minimize response bias (Mishra et al., 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. 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*30 trials; spread over three blocks). Stimulus delivery and response data collection were controlled by E-prime® 2.0 (Psychological Software Tools, Inc.).

Electrophysiological recordings

Recording procedures followed our previous multichannel EEG studies (e.g., Bidelman and Howell, 2016; Bidelman, 2018). Participants were seated ~90 cm from the computer monitor within an IAC electro-acoustically shielded booth. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts). Stimulus delivery and responses data collection was controlled by E-prime® (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 via ER-2 insert earphones (Etymotic Research). Stimulus intensity was calibrated using a Larson-Davis SPL meter (Model LxT) and measured in a 2-cc coupler (IEC 60126).

Neuroelectric activity was recorded from 64 sintered Ag/AgCl electrodes at standard 10-10 scalp locations (Oostenveld and Praamstra, 2001). EEGs were digitized at a sampling rate of 500 Hz (SynAmps RT amplifiers; Compumedics Neuroscan) using an online passband of DC-250 Hz. Electrodes placed on the outer canthi of the eyes and the superior and inferior orbit monitored ocular activity. During online acquisition, electrodes were referenced to an additional sensor placed ~1 cm posterior to Cz. Data were re-referenced off-line to a common average reference for subsequent analyses. Contact impedances were maintained ≤5 kΩ.

Subsequent pre-processing was performed in Curry 7 (Compumedics Neuroscan) and custom routines coded in MATLAB. Ocular artifacts (saccades and blinks) were first corrected in the continuous EEG using a principal component analysis (PCA) (Picton et al., 2000). The PCA decomposition provided a set of independent components which best explained the topography of the blink/saccadic artifacts. The scalp projection of the first two PCA loadings was subtracted from the continuous EEG traces to nullify ocular contamination in the recorded data. Cleaned EEGs were bandpass filtered between 1.5 and 12 Hz,[1] epoched (1500 ms). Traces were then baseline-corrected (i.e., mean voltage within the epoch subtracted from the signal), and averaged to obtain ERPs for each SOA per condition and participant. ERPs were peak time-aligned such that $t = 0$ corresponded to the concurrent flash-beep response wavelet, which was present across all SOA conditions. This resulted in a final epoch window of −750 to 750 ms (see Fig. 2). The experimental protocol including behavioral and electrophysiological testing took ~2 h to complete.

Data analysis

Behavioral data. For each SOA per listener, we computed the mean percentage of trials for which two flashes were reported. For 1F/2B presentations (illusory trials), higher percentages indicate listeners erroneously perceived two flashes when only one was presented (i.e., the illusion); lower values reflect better performance. RTs were calculated as the median response time between the end of stimulus presentation and button press.

ERP data. Following our previous reports and for data reduction purposes, we collapsed and analyzed a subset of the sensor-level ERPs from the average potential in a central cluster of electrodes (C1, Cz, C2) (Chung and Bidelman, 2016; Bidelman and Yellamsetty, 2017) (Fig. 2, inset). This cluster was chosen given that both auditory and visual ERPs are readily recorded at central (vertex) scalp locations (Crottaz-Herbette and Menon, 2006) and visual inspection of the scalp topographies, which indicated a frontocentral distribution (Fig. 2). Amplitude and latency were taken as the peak negativity in a 100-ms search window surrounding the visual-evoked potential wavelet (i.e., complex at $t = 0$; see dotted lines, Fig. 2). Latency was normalized to the 1F/1B condition, which allowed us to assess the relative change in

---

[1] We found it necessary to use a restricted filter passband to reduce noise (e.g., low-frequency drift) that would preclude proper baseline definition of the ERPs and enable better visualization of the multiple visual, and auditory wavelets of the response. A lower, high frequency cutoff is also beneficial for more accurate source analyses (Alain et al., 2017; Bidelman and Howell, 2016).
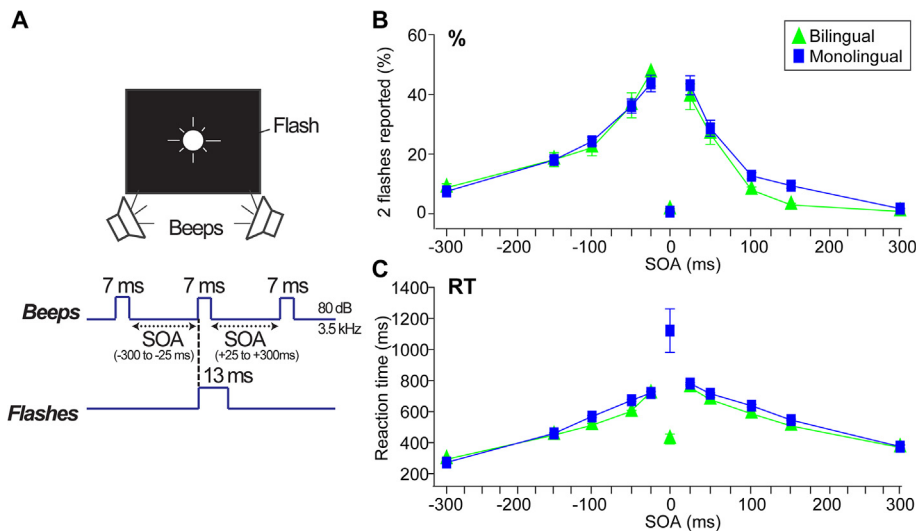
Fig. 1. Bilinguals show lower susceptibility for perceiving the double-flash illusion and more precise temporal binding of audiovisual information. (A) Double-flash illusion stimulus paradigm. Flashes were presented concurrent with auditory beeps delivered via headphones (*top*). Single trial time course (*bottom*). A single beep was always presented simultaneous with the onset of the flash. A second beep was then presented either before (+SOAs) or after (−SOAs) the first. While only a single flash is presented, listeners perceive two illusory flashes indicating that auditory cues modulate the visual percept. (B) Percentage of "2 flash" (i.e., illusion) reports to 1F/2B stimuli. Psychometric functions reveal the illusion is strongest for short SOAs and progressively weakens with increasing asynchrony between auditory and visual cues. Bilinguals show less susceptibility to the illusion. (C) Reaction times. Bilinguals are faster at responding to AV stimuli. Panel A, adapted from Bidelman and Heath (2019) with permission from Cambridge University Press. errorbars = ±1 s.e.m.

response (Δt) between the veridical (non-illusory) stimulus and conditions evoking the perceptual illusion. Δt was computed as a simple difference between the illusory and non-illusory conditions (i.e., 1F/2B − 1F/1B) for each SOA. While the scalp ERPs allowed us to assess the gross presence/absence of group differences in neural activity, the volume-conducted nature of sensor-space (i.e., electrode) responses did not allow us to separate the underlying sources that contribute to apparent functional differences between groups. Consequently, subsequent analyses were conducted in source space to directly assess the source generator characteristics underlying AV processing in monolinguals vs. bilinguals.

Distributed source imaging. We used Classical Low Resolution Electromagnetic Tomography Analysis Recursively Applied (CLARA) [BESA® (v7)] (Iordanov et al., 2014) to estimate the neuronal current density underlying the sensor data recorded at the scalp (e.g., Alain et al., 2017; Bidelman, 2018; Bidelman et al., 2018). CLARA models the inverse solution as a large collection of elementary dipoles distributed over nodes on a mesh of the cortical volume. The algorithm estimates the total variance of the scalp-recorded data and applies a smoothness constraint to ensure current changes minimally between adjacent brain regions (Picton et al., 1999; Michel et al., 2004). CLARA renders more focal source images by iteratively reducing the source space during repeated estimations. On each iteration (×3), a spatially smoothed LORETA solution (Pascual-Marqui et al., 2002) was recomputed and voxels below a 1% max amplitude threshold were removed. This provided a spatial weighting term for each voxel on the subsequent step. Three iterations were used with a voxel size of

7 mm in Talairach space and regularization (parameter accounting for noise) set at 0.01% singular value decomposition. Group-level statistical (*t*-stat) maps were computed using the 'ft_sourcestatistics' function in the MATLAB FieldTrip toolbox (Oostenveld et al., 2011) and threshold at $\alpha = 0.05$. Source activations were then interpolated and visualized by projecting them onto the MNI adult brain template (Fonov et al., 2009).

From each participants' functional images, we extracted the amplitude of source activations at the centroid of several regions of interest including prefrontal cortex (BA 10), bilateral primary auditory cortices (PAC; BA 41)—pooled across hemispheres—and primary visual cortex (V1; BA 17) (see Fig. 4). These seeds were selected given their hypothesized role in controlling (BA 10) the gating and sensory coding of auditory (PAC) and visual (V1) cues in our stimulus paradigm (cf. Knight et al., 1989). Source activations were computed at the response latency corresponding to the second auditory beep (see ◄ icons, Fig. 2), where the auditory stimulus creates the illusion of a secondary flash. To assess relations between region-specific neural activity and behavior, we regressed voxel activations with listeners' behavioral accuracy (%) and RT data (collapsed across stimuli). This allowed us to evaluate the degree to which specific brain regions predicted listeners' propensity and speed of perceiving double-flashes (i.e., AV binding).

Statistical analysis

Unless otherwise noted, dependent measures were analyzed using two-way mixed model ANOVAs (SAS 9.4, GLIMMIX) with fixed effects of group as the between-subject factor and SOA as the within-subject factor. Subjects nested within group served as a random effect. Multiple comparisons were adjusted using Tukey–Kramer corrections. Dependent variables were SQRT-transformed to satisfy homogeneity of variance assumptions for parametric statistics. Behavioral responses to non-illusory trials are reported elsewhere (Bidelman and Heath, 2019) and were not analyzed here. Robust regression (bisquare weighting) between brain and behavioral measures was performed using the 'fitlm' function in MATLAB® (The MathWorks, Inc.). This procedure minimizes potential influential outliers in the model by negatively weighting observations with more extreme residuals in the fit. The alpha level was set at $\alpha = 0.05$. Effect sizes are reported as η^2.

In addition to null hypothesis statistical testing (NHST), we performed Bayes factor (BF) analyses. Under NHST,

Bilinguals

Monolinguals
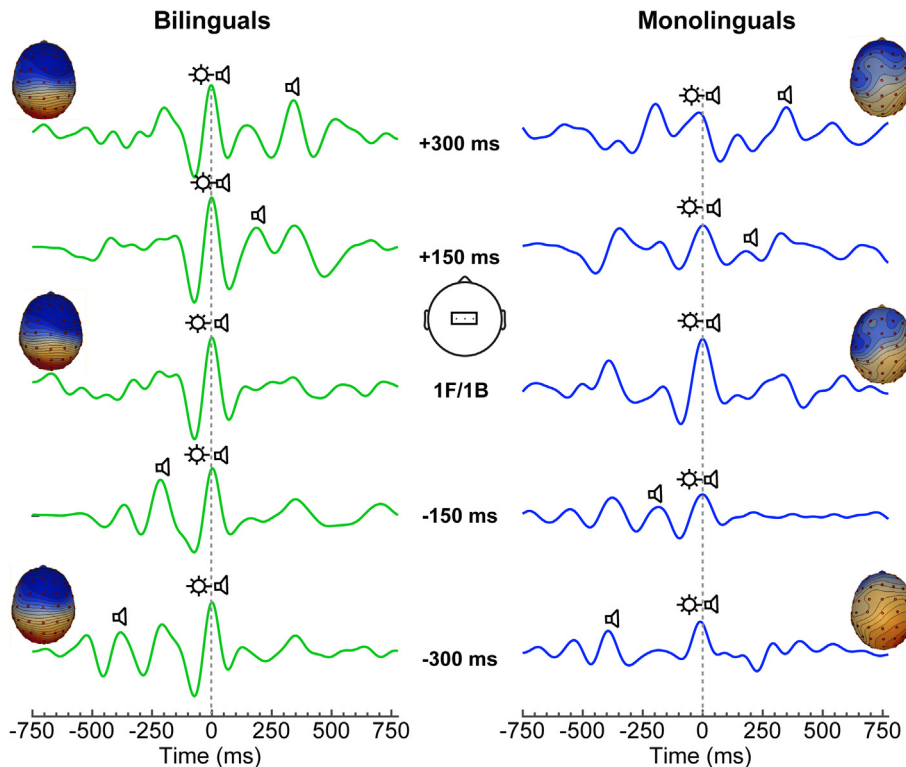
Fig. 2. ERPs to double-flash (illusory) stimuli. Traces are shown for a subset of SOAs averaged across a central cluster of electrodes (C1, Cz, C2) (*inset*). 0-ms SOA is the 1F/1B condition. Symbols above waveforms mark locations of flashes (☼) and sound beeps (◀). Strong responses to the visual token are apparent in the ERPs at ~0 ms, which are flanked by auditory responses that vary according to the SOA. Inset scalp topographies show the voltage distribution at the peak negativity in the 100-ms search window surrounding $t = 0$ (dotted line), where sensor-level data were quantified (see Fig. 3).

reporting two flashes indicate a greater strength or susceptibility to the illusion and poorer AV binding. Consistent with previous studies (Foss-Feig et al., 2010; Neufeld et al., 2012; Bidelman, 2016), both groups showed a similar pattern of responses where the illusion was stronger for short SOAs (± 25 ms), progressively weakened with increasing asynchrony, and was absent for the longest intervals (± 300 ms). Yet, bilinguals showed less susceptibility to the illusion overall, demonstrating lower incidence of perceiving two illusory flashes. These observations were confirmed with a two-way ANOVA (conducted on only illusory 1F/2B conditions), which revealed a significant group × SOA interaction [$F_{9, 189} = 4.41$, $p < 0.0001$, $\eta^2 = 0.18$]. Tukey–Kramer contrasts revealed bilinguals reported fewer illusory double flashes for the positive + 150 ms SOA ($p = 0.0019$; $BF_{H1} = 9.87$). A Bayes factor of 9.87 indicates that the alternative hypothesis (H1) is 9.87 times more likely than the null (H0), and thus represents strong evidence favoring a group difference. These findings suggest that bilinguals showed a lower propensity to perceive the double-flash illusion and more accurately parsed audiovisual cues—indicative of more precise AV binding windows.

Group reaction times across SOAs are shown in Fig. 1C. An ANOVA revealed a significant group × SOA interaction on behavioral RTs to illusory trials [$F_{9, 189} = 11.27$, $p < 0.0001$, $\eta^2 = 0.35$]. While bilinguals responded faster than monolinguals overall [group effect: $F_{1, 21} = 4.89$, $p = 0.038$, $\eta^2 = 0.19$], Tukey contrasts revealed this effect was apparent for all SOAs *except* the −150, ±25, and 300 ms SOAs (p-values 0.0001 to 0.047). Of note, bilinguals had faster responses than monolinguals for the + 150 ms SOA ($p = 0.0079$; $BF_{H1} = 5.06$), where they also showed fewer illusory reports (cf. Fig. 1B). Collapsing across SOAs, we also found that behavioral RTs were correlated with % scores [Pearson's-$r = 0.36$, $p = 0.0003$], such that slower responses corresponded to a higher propensity of illusory two-flash reports. Collectively, behavioral findings indicate that bilinguals were both more accurate and faster at judging the composition of audiovisual stimuli than their monolingual peers.

significant results cannot differentiate insensitive data and evidence in favor of the null hypothesis. BFs determine the degree to which experimental data support the null (H0) or alternative (H1) hypothesis, or instead represent ambiguous evidence (Rouder et al., 2009; Dienes, 2014). A BF_{H1} of 5, for example, is directly interpreted as the data being 5 times more likely in favor of H1 compared to H0. Bayes factor analysis is well suited to assess the likelihood of study replication (Anderson and Maxwell, 2016), which is particularly useful for smaller sample studies. For relevant group contrasts (e.g., *t*-tests, Fig. 4), we calculated BFs using default Cauchy priors (conservative scaling factor $r = 0.707$) (Rouder et al., 2009) as implemented in the Bayes Factor Calculator (http://pcl.missouri.edu/bayesfactor) with the Scaled-Information Bayes Factor option. Similarly, BFs for the linear regressions were computed following Liang et al. (2008). BFs ranging from 1-3 are considered anecdotal evidence, 3–10 moderate, 10–30 strong, 30–100 very strong, and > 100 extreme evidence for the respective hypothesis (Lee and Wagenmakers, 2013).

RESULTS

Behavioral data

The proportion of two-flash reports for each SOA and group is shown in Fig. 1B. Higher proportions in

ERP data

ERPs and scalp topographies for each group are shown for a subset of SOAs in Fig. 2. Latency and amplitude
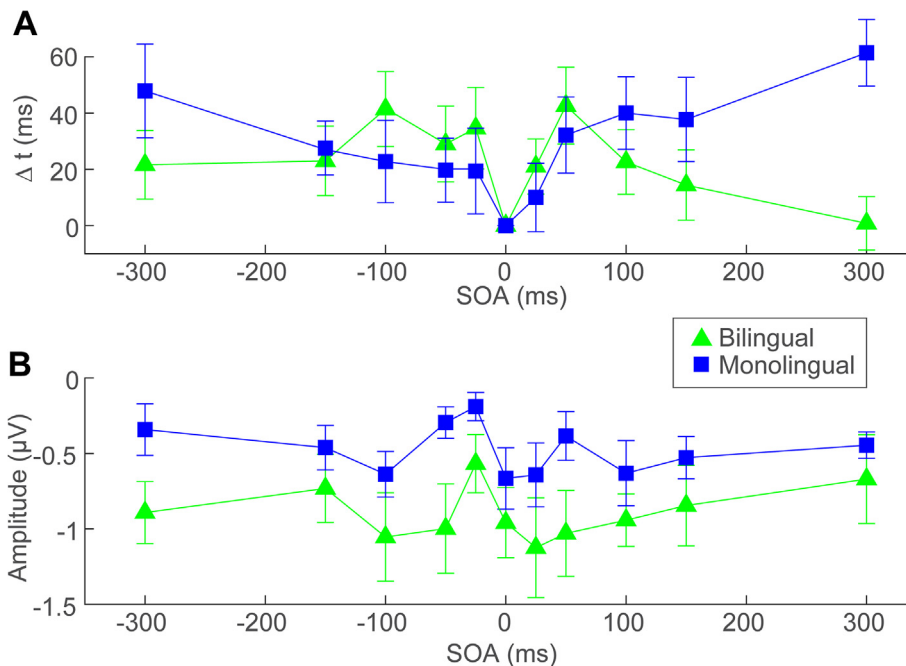
Fig. 3. Cross-linguistic differences in ERP latency (A) and amplitude (B) to audiovisual stimuli. Latencies reflect the shift (Δt) in response between illusory conditions and the non-illusory stimulus (e.g., 1F/2B − 1F/1B). This shows the effect of the illusion on response timing to the visual event. errorbars = ±1 s.e.m.

characteristics are shown in Fig. 3. ERPs were marked by a series of wavelets corresponding to responses to the visual token (∼0 ms) that were flanked by auditory responses appearing at the SOA spacing between visual and auditory events. An ANOVA (including all SOAs) revealed strong modulations in response latency and amplitude dependent on both group membership and SOA. We found a group × SOA interaction on ERP Δt latencies [$F_{10, 210} = 2.53$, $p = 0.0067$, $\eta^2 = 0.11$] (Fig. 3A). Inspection of the pattern of the data (Fig. 3A) suggested that responses to illusory stimuli were delayed relative to the non-illusory (1F/1B) condition. Post hoc comparisons revealed a group effect for the +300-ms SOA ($p = 0.0016$; $BF_{H1} = 10.79$). This Bayes factor suggests strong evidence for a group difference at positive SOAs. Analysis of amplitudes revealed that sensor-level (electrode) responses did not differ between groups [$F_{1, 21} = 2.87$, $p = 0.1049$, $\eta^2 = 0.12$] but was nevertheless modulated by stimulus SOA [$F_{10, 210} = 3.07$, $p = 0.0012$, $\eta^2 = 0.13$] (Fig. 3B). These ERP findings indicate a differential pattern of responses to AV stimuli in bilingual vs. monolingual listeners.

Distributed source-level imaging data

We performed distributed CLARA source imaging (Iordanov et al., 2014; Alain et al., 2017; Bidelman, 2018; Bidelman et al., 2018) of the scalp (electrode-level) data to parse region-specific activity that may contribute to group differences observed in the ERPs and help delineate mechanistic differences in AV processing between language groups. CLARA maps were computed at a latency corresponding to the second auditory beep

(◄, Fig. 2), where the auditory token creates the illusion of a second flash.

Functional CLARA maps are shown in Fig. 4 for select SOA conditions. Visual inspection suggested that monolinguals and bilinguals showed an opposite pattern of activation of the frontal cortex between the control (1F/1B) and illusory test conditions (cf. 1F/1B vs. 150 ms; Fig. 4A). For the +150-ms SOA which produced a relatively strong illusion and showed group effects behaviorally (i.e., Fig. 1), bilinguals showed stronger activity in frontal cortex (BA 10) [$t_{21} = 2.40$, $p = 0.026$; $BF_{H1} = 3.43$], indicative of increased top-down control and/or response inhibition regulating AV percepts (Fig. 4B). Visual cortex (V1) and BA 37 in the occipitotemporal junction are implicated in multisensory processing (e.g., Stock et al., 2017). However, we found no group differences in V1 or BA 37 responsiveness ($ps > 0.33$; BF favoring H0 null = 1.3–1.5). Nevertheless, monolinguals showed stronger source activation in bilateral PAC [$t_{21} = -2.26$, $p = 0.0034$; $BF_{H1} = 2.84$] during the illusion than bilinguals.

To assess the behavioral relevance of source responses, we conducted robust regression between the amplitude of region-specific activations and listeners' behavioral %-scores and RTs (Fig. 4C). Behavioral responses were averaged across SOAs to reduce the dimensionality of the data. %-scores were not predicted by V1, PAC, nor frontal source activity ($ps \gg 0.05$). In contrast, brain-behavior regressions revealed stronger source activity in frontal cortex (BA10) was associated with *faster* perceptual RTs [$t_{21} = -2.70$, $p = 0.01$; $R^2 = 0.27$; Bayes factor favoring H1 = 4.8], whereas stronger V1 activity was associated with *slower* behavioral decisions [$t_{21} = 2.32$, $p = 0.03$; $R^2 = 0.21$; $BF_{H1} = 2.5$]. Stronger bilateral PAC activity similarly predicted *slower* RTs [$t_{21} = 3.16$, $p = 0.0047$; $R^2 = 0.33$; $BF_{H1} = 10$]. BA 37 activations were not associated with behavioral RTs ($p = 0.63$; BF favoring H0 = 2.4).

DISCUSSION

By measuring electrophysiological responses in monolinguals and bilinguals to the double-flash illusion (Shams et al., 2000; Stevenson et al., 2014; Bidelman, 2016; Bidelman and Heath, 2019), our findings reveal cross-linguistic differences in the neural mechanisms underlying fundamental AV processing. In particular, we find that the perception of illusory double-flashes, a mar-
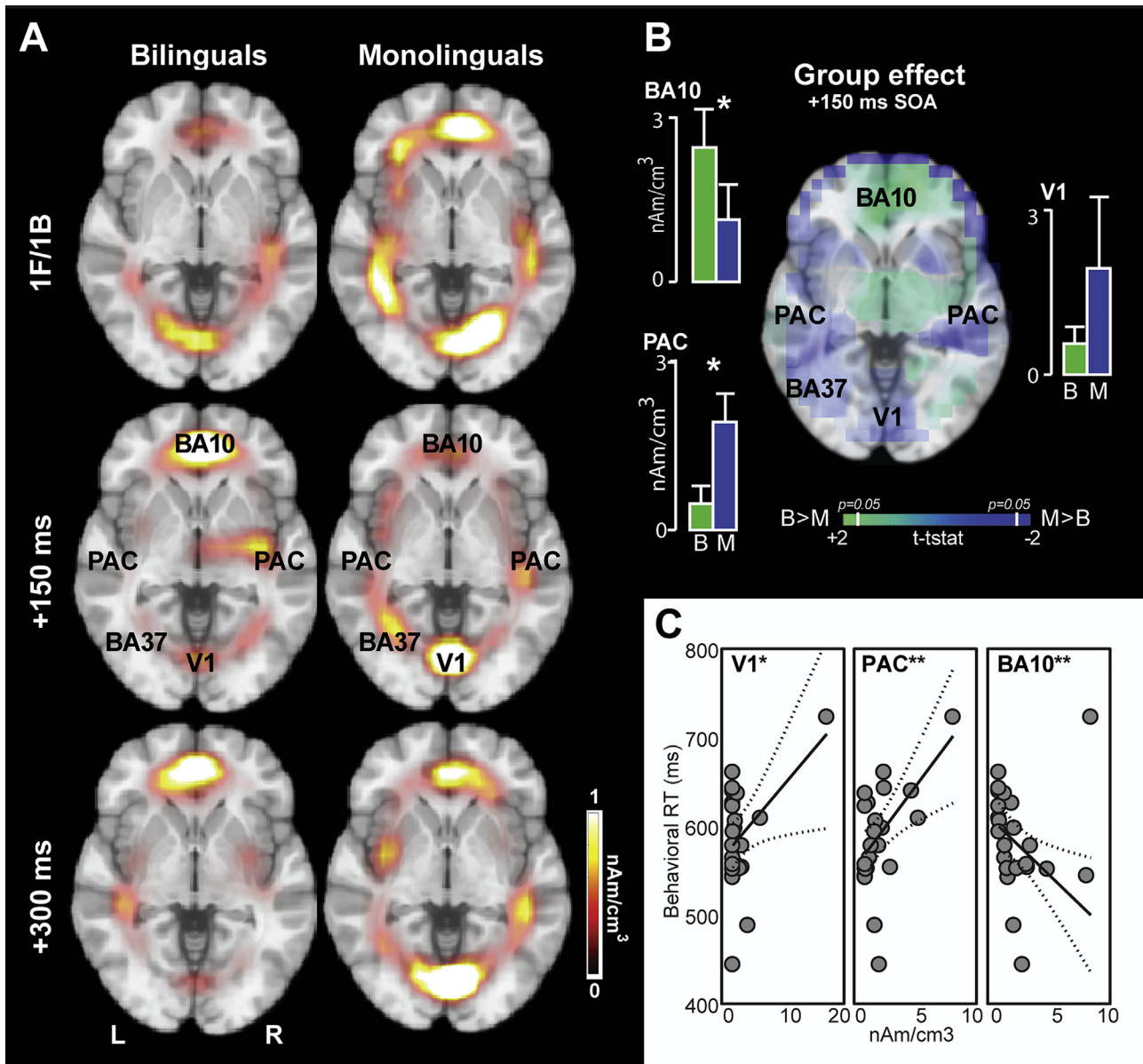
Fig. 4. Source activations reveal mechanistic differences in AV brain networks between monolinguals and bilinguals. (A) Grand average CLARA source maps by group for three SOA conditions. For illusory stimuli (+150 ms, +300 ms), maps were computed at a latency corresponding to the second auditory beep (i.e., ◀, Fig. 2), where the auditory event creates the illusion of a second flash. 1F/1B maps are computed at $t = 0$. Functional data are overlayed onto the MNI adult brain template (Fonov et al., 2009). Note the stronger frontal activation in bilinguals, particularly for stimuli producing strong AV illusions (e.g., 150 ms). (B) Group t-stat contrast maps ($p < 0.05$ masked, uncorrected) for the +150 ms condition, which produced a strong illusion and group difference behaviorally (i.e., Fig. 1). Barcharts = voxel cluster activations in select regions of interest. During the illusion, bilinguals show stronger activity in frontal cortex (BA 10) whereas monolinguals show stronger activation in bilateral PAC (V1 and multimodal BA 37 were *n.s.*). (C) Brain-behavior relations between behavioral RTs (collapsed across SOAs) and region-specific amplitudes. Stronger source responses in frontal cortex (BA 10) are associated with faster perceptual RTs, whereas stronger V1 and PAC activity is associated with slower judgments. BA 10, prefrontal cortex; PAC, primary auditory cortex (BA 41); V1, primary visual cortex (BA 17); BA 37, visual association area (occipitotemporal junction). *$p < 0.05$, **$p < 0.01$. errorbars = ±1 s.e.m.

ker of audiovisual binding (Shams et al., 2000; Foss-Feig et al., 2010; Wallace and Stevenson, 2014; Bidelman, 2016; Bidelman and Heath, 2019), is governed by an interplay between primary sensory cortices, associative regions, and frontal cortices that is further modulated by language expertise. Our data show bilinguals are not only faster and more accurate at processing concurrent audiovisual objects than monolinguals, but these perceptual benefits are driven by physiological differences in the brain networks supporting multisensory integration.

Behaviorally, we found bilinguals had more accurate and efficient behavioral responses, indicating enhanced parsing of audiovisual objects and less susceptibility to false percepts (i.e., AV illusions). These results extend recent work on bilingualism and multisensory integration for *speech* stimuli (e.g., Burfin et al., 2014; Reetzke et al., 2016) by demonstrating comparable perceptual enhancements for *non-speech* AV stimuli. We also extend our previous behavioral studies (Bidelman and Heath, 2019) by revealing the neurophysiological mechanisms

that drive these cross-language differences in AV perception. While the double-flash task is typically interpreted as a measure of multisensory integration (Mishra et al., 2007; Powers et al., 2009; Foss-Feig et al., 2010; Bidelman, 2016; Bidelman and Heath, 2019), it remains possible that group differences result from enhanced processing in unisensory brain regions rather than multisensory integration, *per se*. In this regard, our ERP source imaging approach was useful to disentangle the underlying neural mechanisms of bilinguals' AV processing benefits.

ERPs revealed a differential pattern of responses in bilingual vs. monolingual listeners. Relative to monolinguals, bilinguals showed less illusion-induced changes in ERP latency (Fig. 3A), paralleling the group effect observed in the behavioral RTs. Source analysis further suggested that the degree of illusory perception in the double-flash paradigm—a marker of audiovisual binding (Foss-Feig et al., 2010; Wallace and Stevenson, 2014; Bidelman, 2016; Bidelman and Heath, 2019)— was governed by a distributed neural network involving visual cortex and surrounding association areas (BA37), auditory cortex, and frontal brain regions. While overall ERP amplitude at the scalp did not differ between language groups (Fig. 3B), these underlying sources showed a differential activation depending on language experience. This implies that while mono- and bi-linguals were similar in overall neural responsivity, each group executed the task by differentially engaging the underlying fronto-temporal-occipital network (i.e., different task strategy). In particular, monolinguals showed increased activity in auditory cortex (PAC; BA 41) than their bilingual peers. Monolinguals also showed a tendency for increased visual activation, but this did not differ statistically from V1 activity of bilinguals. Still, both auditory and visual activity predicted RTs, indicating a link between activation in audiovisual sensory regions and behavioral responses to the double-flash illusion. Conceivably, this series of activations could provide a putative neural correlate of the illusory visual percept (e.g., Fig. 2), whereby perception of the double-flash illusion is created by stronger (erroneous) interplay between auditory and visual encoding, perhaps also by way of converging signals in multisensory areas (e.g., BA 37). BA 37 is a visual associative region implicated in audiovisual convergence (e.g., Büchel et al., 1998; Stock et al., 2017) and plays a key role in multimodal processing, reading acquisition (Pugh et al., 2001), and the alternation between external and internal (proprioceptive) information (Stock et al., 2017). While this region was active in both groups, we would have expected group differences in this multisensory area. Still, group differences in primary sensory and frontal brain areas alone account for the increased propensity of illusions reported among monolinguals in the current study.

Contrasting monolinguals' increased neural activity in sensory regions, bilinguals showed stronger responses in higher order frontal cortices (BA 10). Brain-behavior relations further suggested that these increased frontal responses were associated with improved AV perception in the form of faster decisions in the behavioral task (Fig. 4C). The opposite pattern was observed in sensory cortices where stronger PAC and V1 responses were linked with slower behavioral judgments. Given that slower RTs were associated with increased reports of illusory percepts (worse performance), a parsimonious account of our data is that more exaggerated sensory representations (whether veridical or induced) lead to more erroneous AV percepts in the double-flash paradigm (i.e., more frequent ''two flash'' reports for 1F/2B stimuli).

Nevertheless, we find that these false sensory representations can be counteracted by higher level processes located in the frontal lobes. We posit that frontal systems act to gate or control incoming sensory information and regulate the generation of false percepts. In the case of monolinguals, stronger auditory coding appears to induce false visual activity which fails to be suppressed by top-down modulation as in bilinguals. Support for this notion stems from previous neuroimaging studies. For example, in their source analysis of ERPs to flash-beep stimuli, Meylan and Murray (2007) reported an attenuation within low-level visual cortices dependent on a second auditory stimulus. Interestingly, auditory-induced attenuation of visual processing occurred within an integration window of ∼160 ms, consistent with the SOA where we observe maximal group differences in the AV illusion. Other neuroimaging studies investigating flash-beep stimuli show engagement of both unisensory (auditory, visual) and polysensory brain areas (Mishra et al., 2007, 2008). Taken as a whole, both behavioral and neurophysiological data indicate that bilinguals' brain responses reflect more veridical sensory coding of the physical cues coupled with superior top-down gating of AV information that more effectively suppresses the generation of false AV percepts.

Our data are broadly consistent with the notion that experience-dependent plasticity of intensive language experience improves the integration of information from multiple sensory systems (audition and vision). Results also extend previous studies documenting similar experience-dependent plasticity in AV processing among other highly skilled individuals (e.g., trained musicians: Bidelman, 2016). Bilinguals' increased frontal control revealed by our source analyses, also converges with reports of enhanced selective attention, inhibition, and executive control in these listeners (Bialystok et al., 2003; Bialystok et al., 2007; Bialystok, 2009; Bialystok and DePape, 2009; Krizman et al., 2014; Schroeder et al., 2016). Distributing attention across sensory modalities can enhance performance in complex audiovisual tasks (Mishra and Gazzaley, 2012). This may account for the improved performance we find in our bilingual cohort. That said, diffusion tensor imaging (DTI) and fMRI studies have shown increased structural and functional connectivity in frontal-occipital fasciculus pathways in both bilingual adults and children (Luk et al., 2011; Mohades et al., 2012). Our functional data may reflect similar increases in long-range neural communication. Still, while the group differences and correlational effects observed here (e.g., $\eta^2 > 0.20$; $r > 0.5$) are considered intermediate to large effects (Cohen, 1988), were adequately powered (e.g., PAC correlation: $r = 0.57$;

α = 0.05, power 85%, two-tailed) (G*Power; Faul et al., 2007), and provide moderate to strong evidence favoring the alternative hypotheses (Bayes factors = 3–10), we acknowledge the limitation of our smaller sample size. Additional studies on a larger population of monolinguals and bilinguals are needed to replicate and confirm the present findings. Regardless, our data implicate frontal brain systems (e.g., BA 10) as a putative mechanism underlying bilinguals' improved AV perceptual-cognitive control and an avenue to explore in future studies.

In summary, our results show that compared to monolinguals, bilinguals show enhanced behavioral and neurophysiological parsing of audiovisual stimuli. While improved AV perception is itself advantageous, the broader implications of bilinguals' more precise and domain general AV processing remains somewhat speculative. Visual speech movements are known to augment second language perception by way of multisensory integration (Navarra and Soto-Faraco, 2007). Nonnative listeners also show difficulty parsing the speech of their L2 in noisy listening environments (Rogers et al., 2006; Bidelman and Dexter, 2015). Yet, it is well known that combining visual and auditory cues enhances spoken word recognition in adverse listening environments (Sumby and Pollack, 1954; Vatikiotis-Bateson et al., 1998; Ross et al., 2007; Bidelman et al., under review). Presumably, bilinguals might compensate for deficits in both normal and degraded L2 speech perception if they are better able to combine and integrate the auditory and visual modalities as observed here. Future studies are needed to test these possibilities and the putative impact of bilinguals' superior AV processing across a wider range of perceptual and cognitive tasks.

ACKNOWLEDGMENTS

REFERENCES

Alain C, Arsenault JS, Garami L, Bidelman GM, Snyder JS (2017) Neural correlates of speech segregation based on formant frequencies of adjacent vowels. Sci Rep 7:1–11.

Anderson SF, Maxwell SE (2016) There's more than one way to conduct a replication study: beyond statistical significance. Psychol Methods 21:1–12.

Banks B, Gowen E, Munro KJ, Adank P (2015) Audiovisual cues benefit recognition of accented speech in noise but not perceptual adaptation. Front Hum Neurosci 9:422.

Bialystok E (2009) Bilingualism: the good, the bad, and the indifferent. Bilingualism: Language and Cognition 12:3–11.

Bialystok E, Craik FI, Freedman M (2007) Bilingualism as a protection against the onset of symptoms of dementia. Neuropsychologia 45:459–464.

Bialystok E, Craik FI, Luk G (2012) Bilingualism: consequences for mind and brain. Trends Cogn Sci 16:240–250.

Bialystok E, DePape AM (2009) Musical expertise, bilingualism, and executive functioning. J Exp Psychol Hum Percept Perform 35:565–574.

Bialystok E, Majumder S, Martin MM (2003) Developing phonological awareness: is there a bilingual advantage? Appl Psycholinguist 24:27–44.

Bidelman GM (2016) Musicians have enhanced audiovisual multisensory binding: experience-dependent effects in the double-flash illusion. Exp Brain Res 234:3037–3047.

Bidelman GM (2018) Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. Neuroimage 175:56–69.

Bidelman GM, Brown B, Mankel K, Price CN (under review), Psychobiological responses reveal audiovisual noise differentially challenges speech recognition.

Bidelman GM, Davis MK, Pridgen MH (2018) Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. Hear Res 367:149–160.

Bidelman GM, 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 Lang 143:32–41.

Bidelman GM, Gandour JT, Krishnan A (2011) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci 23:425–434.

Bidelman GM, Heath S (2019) Enhanced temporal binding of audiovisual information in the bilingual brain. Bilingualism: Language and Cognition:1–11. https://doi.org/10.1017/S1366728918000408.

Bidelman GM, Howell M (2016) Functional changes in inter- and intra-hemispheric auditory cortical processing underlying degraded speech perception. Neuroimage 124:581–590.

Bidelman GM, 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 e60676.

Bidelman GM, Yellamsetty A (2017) Noise and pitch interact during the cortical segregation of concurrent speech. Hear Res 351:34–44.

Blumenfeld HK, Marian V (2011) Bilingualism influences inhibitory control in auditory comprehension. Cognition 118:245–257.

Büchel C, Price C, Friston K (1998) A multimodal language region in the ventral visual pathway. Nature 394:274.

Burfin S, Pascalis O, Ruiz Tada E, Costa A, Savariaux C, Kandel S (2014) Bilingualism affects audiovisual phoneme identification. Front Psychol 5:1179.

Calvo N, Ibáñez A, García AM (2016) The impact of bilingualism on working memory: a null effect on the whole may not be so on the parts. Front Psychol 7.

Chung W-L, Bidelman GM (2016) Cortical encoding and neurophysiological tracking of intensity and pitch cues signaling English stress patterns in native and nonnative speakers. Brain Lang 155–156:49–57.

Cohen J (1988) Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates.

Costa A, Sebastian-Galles N (2014) How does the bilingual experience sculpt the brain? Nat Rev Neurosci 15:336–345.

Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, et al. (2006) Language control in the bilingual brain. Science 312:1537–1540.

Crottaz-Herbette S, Menon V (2006) Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. J Cogn Neurosci 18:766–780.

Dienes Z (2014) Using Bayes to get the most out of non-significant results. Front Psychol 5:781.

Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Meth 39:175–191.

Fonov VS, Evans AC, McKinstry RC, Almli CR, Collins DL (2009) Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage 47:S102.

Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT (2010) An extended multisensory temporal binding window in autism spectrum disorders. Exp Brain Res 203:381–389.

Gold BT, Johnson NF, Powell DK (2013a) Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. Neuropsychologia 51:2841–2846.

Gold BT, Kim C, Johnson NF, Kryscio RJ, Smith CD (2013b) Lifelong bilingualism maintains neural efficiency for cognitive control in aging. J Neurosci 33:387–396.

Grosjean F (2010) Bilingual: life and reality. Harvard University Press.

Iordanov T, Hoechstetter K, Berg P, Paul-Jordanov I, Scherg M (2014) CLARA: classical LORETA analysis recursively applied. OHBM.

Knight RT, Scabini D, Woods DL (1989) Prefrontal cortex gating of auditory transmission in humans. Brain Res 504:338–342.

Krizman J, Bradlow AR, Lam SS-Y, Kraus N (2016) How bilinguals listen in noise: linguistic and non-linguistic factors. Bilingualism: Language and Cognition:1–10.

Krizman J, Marian V, Shook A, Skoe E, Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. PNAS 109:7877–7881.

Krizman J, Skoe E, Marian V, Kraus N (2014) Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. Brain Lang 128:34–40.

Lee H, Noppeney U (2011) Long-term music training tunes how the brain temporally binds signals from multiple senses. PNAS 108: E1441–E1450.

Lee MD, Wagenmakers EJ (2013) Bayesian cognitive modeling: a practical course. Cambridge: Cambridge University Press.

Li P, Sepanski S, Zhao X (2006) Language history questionnaire: a web-based interface for bilingual research. Behav Res Methods 38:202–210.

Liang F, Paulo R, Molina G, Clyde MA, Berger JO (2008) Mixtures of g priors for bayesian variable selection. J Am Stat Assoc 103:410–423.

Luk G, Bialystok E, Craik FI, Grady CL (2011) Lifelong bilingualism maintains white matter integrity in older adults. J Neurosci 31:16808–16813.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264:746–748.

Meylan RV, Murray MM (2007) Auditory–visual multisensory interactions attenuate subsequent visual responses in humans. Neuroimage 35:244–254.

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, Grave de Peralta R (2004) EEG source imaging. Clin Neurophysiol 115:2195–2222.

Mishra J, Gazzaley A (2012) Attention distributed across sensory modalities enhances perceptual performance. J Neurosci 32:12294–12302.

Mishra J, Martinez A, Hillyard SA (2008) Cortical processes underlying sound-induced flash fusion. Brain Res 1242:102–115.

Mishra J, Martinez A, Sejnowski TJ, Hillyard SA (2007) Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. J Neurosci 27:4120–4131.

Mohades SG, Struys E, Van Schuerbeek P, Mondt K, Van De Craen P, Luypaert R (2012) DTI reveals structural differences in white matter tracts between bilingual and monolingual children. Brain Res 1435:72–80.

Moreno S, Wodniecka Z, Tays W, Alain C, Bialystok E (2014) Inhibitory control in bilinguals and musicians: event related potential (erp) evidence for experience-specific effects. PLoS ONE 9 e94169.

Navarra J, Soto-Faraco S (2007) Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. Psychol Res 71:4–12.

Neufeld J, Sinke C, Zedler M, Emrich HM, Szycik GR (2012) Reduced audio–visual integration in synaesthetes indicated by the double-flash illusion. Brain Res 1473:78–86.

Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9:97–113.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. Comput Intell Neurosci 2011:1–9.

Oostenveld R, Praamstra P (2001) The five percent electrode system for high-resolution EEG and ERP measurements. Clin Neurophysiol 112:713–719.

Pascual-Marqui RD, Esslen M, Kochi K, Lehmann D (2002) Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. Methods Find Exp Clin Pharmacol 24(Suppl C):91–95.

Picton TW, Alain C, Woods DL, John MS, Scherg M, Valdes-Sosa P, Bosch-Bayard J, Trujillo NJ (1999) Intracerebral sources of human auditory-evoked potentials. Audiol Neurootol 4:64–79.

Picton TW, van Roon P, Armilio ML, Berg P, Ille N, Scherg M (2000) The correction of ocular artifacts: a topographic perspective. Clin Neurophysiol 111:53–65.

Powers AR, Hillock AR, Wallace MT (2009) Perceptual training narrows the temporal window of multisensory binding. J Neurosci 29:12265–12274.

Pugh KR, Mencl WE, Jenner AR, Katz L, Frost SJ, Lee JR, Shaywitz SE, Shaywitz BA (2001) Neurobiological studies of reading and reading disability. J Commun Disord 34:479–492.

Reetzke R, Lam BPW, 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 e0168048.

Ressel V, Pallier C, Ventura-Campos N, Diaz B, Roessler A, Avila C, Sebastian-Galles N (2012) An effect of bilingualism on the auditory cortex. J Neurosci 32:16597–16601.

Rogers CL, Lister JJ, Febo DM, Besing JM, Abrams HB (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. Appl Psycholinguist 27:465–485.

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb Cortex 17:1147–1153.

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. Psychon Bull Rev 16:225–237.

Schroeder SR, Marian V, Shook A, Bartolotti J (2016) Bilingualism and musicianship enhance cognitive control. Neural Plas 2016:4058620.

Shams L, Kamitani Y, Shimojo S (2000) What you see is what you hear. Nature 408:788.

Shams L, Kamitani Y, Shimojo S (2002) Visual illusion induced by sound. Cognitive Brain Res 14:147–152.

Stevenson RA, Siemann JK, Schneider BC, Eberly HE, Woynaroski TG, Camarata SM, Wallace MT (2014) Multisensory temporal integration in autism spectrum disorders. J Neurosci 34:691–697.

Stock AK, Gohil K, Huster RJ, Beste C (2017) On the effects of multimodal information integration in multitasking. Sci Rep 7:4927.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26:212–215.

Vatikiotis-Bateson E, Eigsti IM, Yano S, Munhall KG (1998) Eye movement of perceivers during audiovisual speech perception. Percept Psychophys 60:926–940.

Wallace MT, Stevenson RA (2014) The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. Neuropsychologia 64C:105–123.