Audiovisual Synchrony in Left-hemisphere Brain-lesioned Individuals with Aphasia

Haleh Farahbod[1], Corianne Rogalsky[2], Lynsey M. Keator[3], Julia Cai[2], Sara B. Pillay[4], Arianna N. LaCroix[5], Julius Fridriksson[3], Jeffrey R. Binder[4], Jonathan H. Venezia[6], Kourosh Saberi[1], and Gregory Hickok[1]

Abstract

■ We investigated the ability of 40 left-hemisphere brain-lesioned individuals with various diagnoses of aphasia to temporally synchronize the audio of a spoken word to its congruent video using a maximum-likelihood adaptive psychophysical procedure. We found a statistically significant effect of aphasia type, not explained by lesion volume, on measures of audiovisual (AV) synchrony. Brain-lesioned individuals with no symptoms of aphasia, and those with conduction aphasia performed on the synchrony task more similarly to age-matched neurotypical controls, whereas those with anomic aphasia performed the poorest. In addition, we examined the correlation between this ability and AV integration (fusion) and observed a significant correlation between measures of AV synchrony and fusion. An ROI analysis of stroke lesion maps showed that damage to the left posterior temporal regions adversely affected AV processing, although whole-brain univariate lesion-symptom mapping analyses did not yield any significant results. These findings contribute to a better understanding of the functional relationship between different AV processes in multimodal integration and their underlying cortical networks in the human brain. ■

INTRODUCTION

The perception of audiovisual (AV) synchrony has been extensively studied in psychophysics. Several studies have reported that, on average, the audio signal (e.g., speech) must be delayed relative to the visual (lip movements) by approximately 50 msec for the two to be perceived synchronous. Across studies, this delay has been reported to range from a few to over 100 msec depending on experimental paradigm and the nature, complexity, and ecological validity of stimuli (Venezia, Thurman, Matchin, George, & Hickok, 2016; van Eijk, Kohlrausch, Juola, & van de Par, 2008; Vatakis & Spence, 2006a, 2006b; Keetels & Vroomen, 2005; Spence, Baddeley, Zampini, James, & Shore, 2003; Zampini, Shore, & Spence, 2003; Jaśkowski, Jaroszyk, & Hojan-Jezierska, 1990; Dixon & Spitz, 1980; Dinnerstein & Zlotogura, 1968). Explanations for this phenomenon range from a need to account for the physical transmission time of sound (which travels approximately 1 ft per msec), to faster neural transduction of auditory signals from the sensory periphery to cortex relative to visual stimuli (10 vs. 50 msec), necessitating a visual "head start" (Vroomen & Keetels, 2010; van Eijk et al., 2008; Arrighi, Alais, & Burr, 2006; Pöppel, 1988; King & Palmer, 1985).

More recent studies have implicated critical contributions from higher-order cognitive factors, including attentional mechanisms, adaptation, object formation, predictive processes, and error reduction by the brain (Venezia et al., 2016; Vroomen & Keetels, 2010; Friston & Kiebel, 2009; van Eijk et al., 2008; Stekelenburg & Vroomen, 2007; Shore & Spence, 2005; Zampini, Shore, & Spence, 2005; Schneider & Bavelier, 2003; Shore, Spence, & Klein, 2001; Mattes & Ulrich, 1998; Stelmach & Herdman, 1991). Venezia and colleagues (2016), for example, using a paradigm that investigated the joint effects of AV synchrony and fusion (integration), showed that earlier arriving visual speech information may generate implicit predictive cues to both the content and timing of later-arriving auditory speech signals, not unlike how the visual trajectory of a bouncing ball (a classic paradigm in studies of AV synchrony) predicts the timing of the accompanying sound at the moment of impact. Other factors include stimulus intensity (faster processing of more intense stimuli; Jaśkowski & Verleger, 2000; Sanford, 1971) and individual differences (Stone et al., 2001; Mollon & Perkins, 1996). For reviews, see Zhou, Cheung, and Chan (2020), Venezia and colleagues (2016), and Vroomen and Keetels (2010).

Fewer studies have investigated the neural bases of AV synchrony for speech, but neuroimaging and stroke-induced lesion-symptom mapping studies have provided important information on the cortical mechanisms involved. The primary finding from these studies is that the superior temporal cortex is critical both to AV synchrony processing (Zhou et al., 2023; Marchant, Ruff, &

[1]University of California, Irvine, [2]Arizona State University, [3]University of South Carolina, [4]Medical College of Wisconsin, [5]Purdue University, [6]Loma Linda University

Journal of Cognitive Neuroscience 37:11, pp. 2190–2202
https://doi.org/10.1162/jocn_a_02316

Driver, 2012) and to multisensory integration (Hickok et al., 2018; Erickson, Heeg, Rauschecker, & Turkeltaub, 2014; Nath & Beauchamp, 2012) and, importantly, that not all AV processes are correlated with one another, either psychophysically or neurally (Zhou et al., 2020; Hickok et al., 2018; Erickson et al., 2014; Nath & Beauchamp, 2012; Stevenson, VanDerKlok, Pisoni, & James, 2011). In a large-scale study of 100 brain-lesioned individuals, our research group, in collaboration with six other institutions, conducted an extensive and detailed investigation of psychophysical and cortical AV processing and found that different AV processes engage functionally and anatomically different cortical regions (Hickok et al., 2018). We found no correlation between two behavioral measures of AV processing: AV advantage (benefits in intelligibility from adding visual to auditory speech) and AV integration (changes in speech-token identity resulting from AV fusion in a McGurk & MacDonald [1976] task). Furthermore, our study found that the posterior superior temporal regions of the brain are involved in AV integration, although not AV advantage. We concluded that: (1) not all AV integration tasks tap into the same generalized AV neural processes and (2) that different AV measures are detecting different combinations of abilities.

The current study expands on these findings from our group in two ways. First, in a new population of 40 left-hemisphere brain-lesioned individuals, we investigated the relationship between two other AV variables, synchrony and fusion (as contrasted to advantage vs. fusion). The results were informative and distinct from our earlier findings. Second, we examined how aphasia diagnosis resulting from stroke-induced brain lesions affects AV synchrony and fusion. We found that there was a statistically significant effect of aphasia type, not explained by lesion volume, on measures of AV synchrony. Brain-lesioned individuals with no symptoms of aphasia, and those with conduction aphasia performed more similarly to age-matched controls, whereas those with anomic aphasia performed the poorest (or most divergent from the control group). Furthermore, we found a statistically significant correlation between AV fusion rate and the ability to perform the AV synchrony task (unlike the orthogonal relation between AV advantage and fusion). Finally, consistent with our earlier work (Hickok et al., 2018), we found that damage to superior temporal areas adversely affects AV processing (especially fusion), with the additional novel finding that posterior and ventral temporal lesions may be disruptive to AV synchrony.

METHODS

Participants

Two categories of participants across four institutions took part in this experiment. The participant categories comprised 40 individuals with stroke-induced left-hemisphere brain lesions (mean age = 56.1 years, σ = 12.2 years, age range = 29–76 years) and 15 older adults with no neurological disease (mean age = 60.1 years, σ = 7.5 years, age range = 48–72 years). The brain-lesioned population was recruited by and participated in the experiment at Arizona State University (n = 12), the University of South Carolina (n = 20), and the Medical College of Wisconsin (n = 8). The older adult control group were recruited by and participated in the experiment at the University of California, Irvine. All lesion and control participants completed audiometric tests at their respective institutions. In general, these participants showed some characteristic hearing loss at higher frequencies above 2 kHz.[1] All participants had normal or corrected-to-normal vision. Participants signed written informed-consent forms approved at their respective institutions' institutional review board.

Stroke participants were included in the present study based on the following criteria: (i) a chronic focal (6 months or more postonset) lesion due to a stroke in the cerebrum, (ii) no significant anatomical abnormalities other than the signature lesion of their vascular event, (iii) an absence of a history of psychological or neurological disease other than stroke, (iv) native English speaker, (v) right-handed prestroke, and (vi) ability to follow task instructions. Stroke participants underwent 3 T or 1.5 T MRI scanning at their respective testing site, including T1 and T2 MRIs with 1-mm^3 resolution. Lesion mapping was performed using trained manual lesion demarcation and then enantiomorphic normalization as described in Rogalsky and colleagues (2022). Scans confirmed damage to the left cerebral hemisphere in all stroke participants. Figure 1 shows overlap maps (axial slices) of damaged areas in the stroke group participants, following transformation of the areas of lesion into standardized space using standard procedures (Rogalsky et al., 2022). The following aphasia types were identified via site-specific protocols that included the Western Aphasia Battery (WAB)–Revised, Boston Diagnostic Examination (BDAE), and clinical observations: Broca's (n = 13), conduction (n = 8), anomic (n = 9), Wernicke's (n = 1), global (n = 1), transcortical sensory (n = 2), transcortical motor (n = 1), and no aphasia (n = 5).

Stimuli and Procedures

AV Synchrony Task

Stimuli were prerecorded videos of a female native speaker of English who spoke one of eight monosyllabic words on each trial (doubt, give, knot, loan, pail, read, theme, and voice). The timing of the audio was adjusted to one of 201 delays (or advances) from 500 msec before the video onset to 500 msec after in steps of 5 msec, resulting in 1608 prerecorded videos (0 msec represents perfectly synchronized audio and video). These AV stimuli were then used in a 2-alternative forced-choice maximum-likelihood psychophysical procedure (Hautus, Macmillan, & Creelman, 2021; Saberi & Green, 1997;

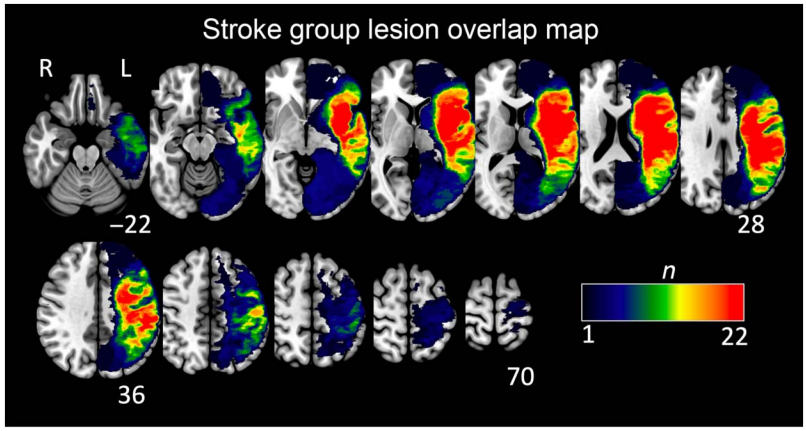

Figure 1. Representative axial slices (z = −22 to 70 Montreal Neurological Institute coordinates) depicting the areas of left hemisphere damage in the stroke group participants (n = 40). The areas of maximum overlap, n = 22, include left inferior frontal, superior temporal, and insular cortices.

Green, 1993) to adaptively track the delay at which the participant perceived synchronicity. On each of 64 trials of a run, the participant had to determine if the audio and video were synchronous or not (a yes/no task). The maximum-likelihood procedure tracked the sweetpoint (Green, 1993) of the psychometric function (i.e., a Gaussian probability density function) that best fit the cumulative response pattern of the participant. At the end of the run, the point of subjective synchrony (PSS) was determined from the psychometric function that optimally fit the participant's pattern of responses using maximum-likelihood rules (Saberi & Green, 1997; Green, 1993).[2] Each participant in the control group completed approximately 5 runs.[3] Each brain-lesioned participant completed a single run because of secondary fatigue and task difficulty. No statistically significant difference was observed between threshold estimates for the "first run" and averaged runs of the control group, $t(28) = 0.42$, $p = .68$, or first and last run of this group, $t(28) = 0.87$, $p = .39$, and therefore the averaged estimates were used in data analysis. This protocol is identical to that used in our earlier work with other auditory tasks on brain-lesioned individuals (Farahbod et al., 2023).

AV Fusion Task

The stimuli employed for the AV fusion task were adopted from an earlier study from our laboratory on the McGurk–MacDonald effect (Hickok et al., 2018). Details of the stimulus generation are fully described in that study and will be briefly outlined here. The AV fusion stimuli were of two types, matched (congruent) or mismatched (incongruent) audio–video. Videos were recorded of a male native speaker of English producing the following single syllables in isolation: /pa/ and /ka/. Separately, high-quality auditory recordings of these same syllables from the same talker were made in an anechoic chamber. The incongruent AV stimulus was created by aligning the onset of the consonant burst for /pa/ with that for the replaced audio of the natural /ka/ associated with the video of /ka/, resulting in the /pa/−/ka/ incongruent AV stimulus. These /pa/−/ka/ incongruent stimuli reliably generate a perception of /ta/ in control participants as reported by Hickok and colleagues (2018). To create the congruent AV /pa/ and /ka/ videos, we performed the same procedure but with congruent auditory and visual stimuli. The AV fusion task comprised 30 trials in which participants were asked to indicate which of three auditory stimuli was presented: /pa/, /ka/, or /ta/. Participants were instructed to pay close attention to both the face in the video and the auditory stimulus. The 30 trials consisted of 20 congruent trials (10 AV stimuli /pa/ and 10 AV stimuli /ka/) and 10 incongruent trials (auditory stimulus was /pa/, visual stimulus was /ka/). After each stimulus presentation, three printed response options were displayed horizontally across the computer screen, "Pa Ta Ka," with the serial positions of the three options presented in a fixed random order across trials for each participant. Participants were instructed to determine which of the three response options corresponded to the sound that they had heard on that trial. Fusion rate was defined as the proportion of incongruent trials on which the participant responded "Ta" (fusing the auditory /pa/ with the visual /ka/, resulting in the McGurk–MacDonald illusion of Ta). An auditory-only task was also employed in the same manner as the AV task, but no visual speech stimuli were presented. The auditory-only task consisted of 30 trials. Participants were given the same instructions as in the AV task, that is, to indicate which sound they heard. As in the AV task, three response options ("Pa, Ta, Ka") were presented on the screen. In 20 trials, the auditory stimulus /pa/ was presented, and in 10 trials, /ka/ was presented (see Hickok et al., 2018, for additional details).

Stimuli for both tasks were presented using a Dell Latitude E5450 computer. The audio portion was presented binaurally through digital-to-analog converters and Sennheiser headphones (HD360 Pro). All four research sites used identical computers, stimuli, and headphones calibrated and tested at The University of California, Irvine, and shipped to the other three research sites. The level of the audio signal was adjusted individually to what was reported by the participant to be a comfortable listening

level (Farahbod et al., 2023). Before data collection began, all participants were given instructions and listened to a few practice trials of each task until it was clear to the experimenter that they fully understood the task.

Lesion-symptom Mappings

To further investigate the neural correlates of AV synchrony, lesion-symptom mapping was conducted using the NiiStat toolbox for MATLAB (https://www.nitrc.org/projects/niistat) on the binary lesion maps in standard Montreal Neurological Institute space. Two univariate analyses fitting a general linear model were computed within each of the 94 left hemisphere ROIs in the Johns Hopkins University (JHU) atlas (Faria et al., 2012; Mori et al., 2008), as described in previous studies (e.g., Rogalsky et al., 2022; den Ouden et al., 2019). A general linear model was fit for (1) PSS (synchrony) values below 45 msec (i.e., lower than the control group's mean performance of 45 msec) and (2) PSS (synchrony) values above 45 msec, to identify regions in which the percentage of lesioned voxels was associated with PSS values below or above those of the control group. Overall lesion volume and data collection site were included as covariates. To correct for multiple comparisons ($p < .05$ controlled for family wise error), permutation thresholding was used with 4000 permutations.

RESULTS

Figure 2 shows results of the AV synchrony experiment for the two participant groups. Panel A shows PSS values individually for each participant arranged in ascending order. The abscissa shows participant number and the ordinate the PSS, with positive values representing the amount of audio delay relative to the video. While the PSS for the control participants is nearly always positive (audio must be delayed relative to the video to be perceived as synchronous with the video), the lesion group's results are symmetric about the "zero delay" point. Panel B shows mean PSS value for each participant group. For the control group, mean PSS was 44.6 msec, consistent with findings from prior studies (Vroomen & Keetels, 2010; van Eijk et al., 2008). Mean PSS for the lesion group was close to zero (−6.0 msec) with nearly equal distribution of PSS thresholds above and below zero (19 of 40 participants had negative PSS values). The mean PSS for the control group was statistically significantly higher than that for the lesion group, $t(53) = 2.07, p = .043$.[4]

Figure 3 shows PSS values individually for all 40 brain-lesioned participants plotted in ascending order and labeled by aphasia type, which is color coded as shown in the legend. A cursory examination suggests that brain-lesioned individuals with no aphasia (magenta symbols) and those with conduction aphasia (yellow) nearly always produced PSS values above zero, similar to the control group.

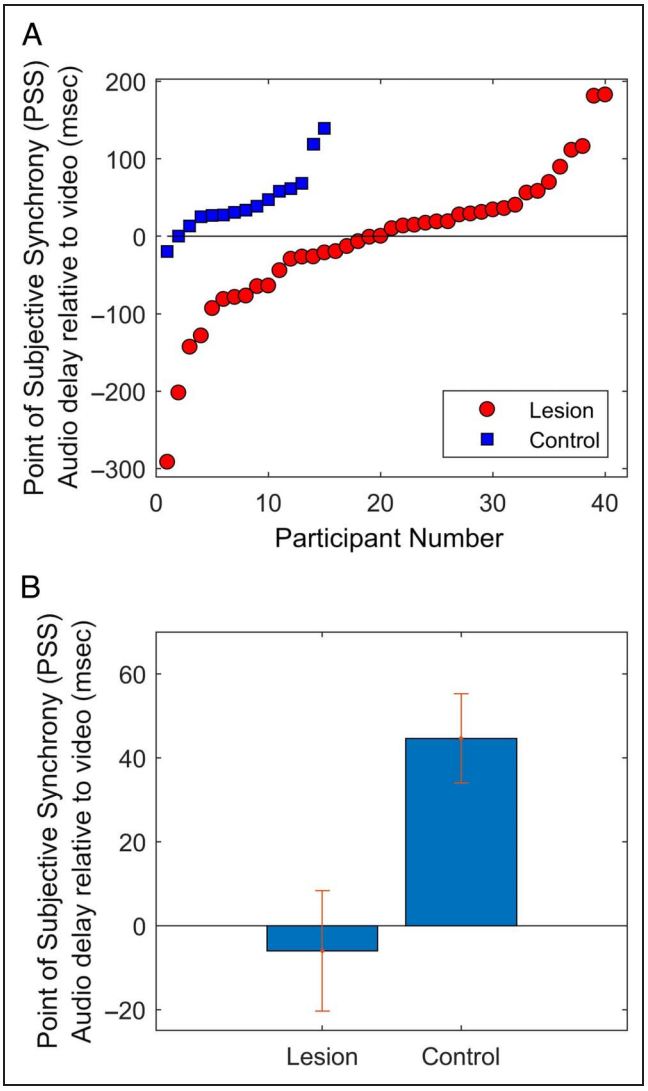

Figure 2. Results of the AV synchrony experiment. (A) PSS for individual participants. Positive values represent the amount of audio delay relative to video. (B) Averaged PSS for the two participant groups. Error bars are one standard error of the mean. $n = 40$ (lesion); $n = 15$ (control).

Figure 4 shows averaged PSS for four subgroups of brain-lesioned participants categorized by aphasia type: conduction, Broca's, anomic, and no aphasia. We excluded from this analysis aphasia types for which we did not have a sufficient number of participants (see the Methods section). The dashed horizontal line represents the average PSS for the control group. As noted earlier, the no-aphasia and conduction aphasia groups produced averaged positive PSS values, close to the control group (dashed line). The anomic and Broca's aphasia groups produced negative mean PSS values. A one-way between-groups ANOVA on the data of Figure 4A showed a significant effect of Aphasia Condition, $F(3, 31) = 3.55, p = .025$, $\eta^2 = .26$. Levene's test showed no violation of homogeneity of variance, $L(3, 31) = 1.589, p = .21$. Post hoc t tests (two-sided) showed a significant difference between the

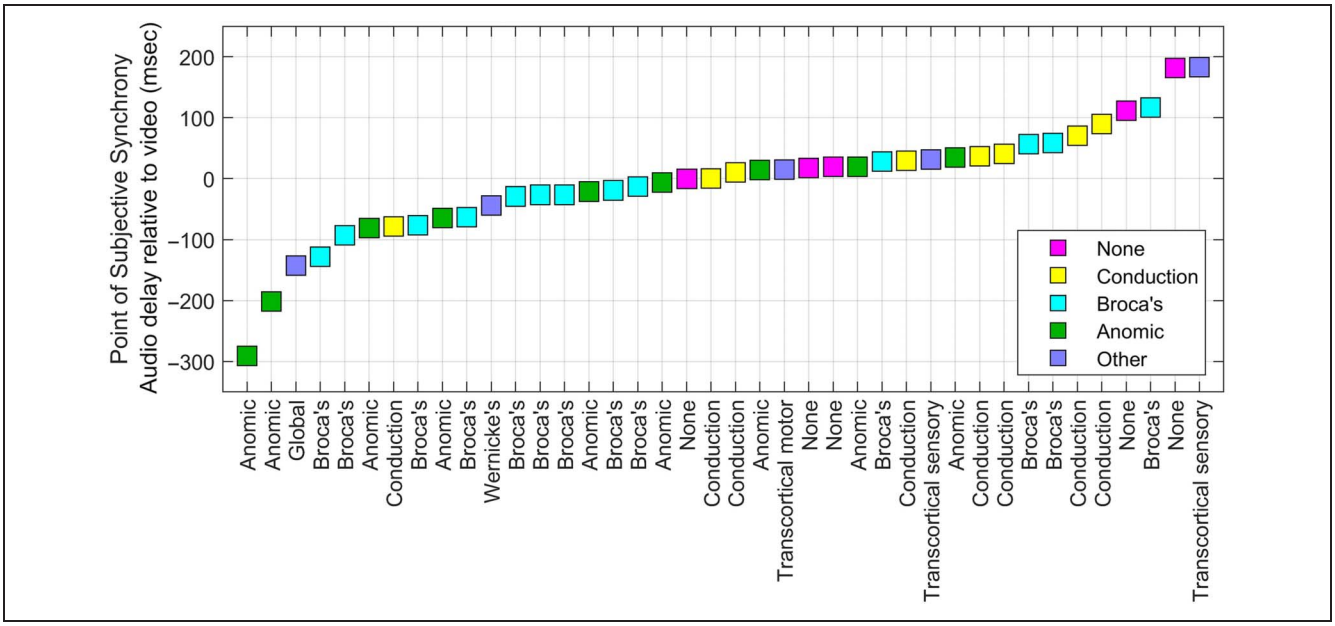

Figure 3. PSS values for 40 brain-lesioned participants with aphasia diagnosis color coded and shown along the abscissa.

no-aphasia and anomic group, $t(12) = 2.34$, $p = .037$; Cohen's $d = 1.31$, as well as between the conduction and anomic groups, $t(15) = 2.13$, $p = .05$; Cohen's $d = 1.03$. Given the counterintuitive finding of a negative mean PSS value for the anomic group (−66.4 msec), we conducted a one-sided post hoc t test of the hypothesis that this negative PSS is significantly below a referent point of zero. The result showed that the mean negative PSS for the anomic group is not significantly below zero, although it did approach significance, $t(8) = -1.79$, $p = .055$, Cohen's $d = -0.539$.

Inspection of the data shown in Figure 3 shows that two participants with anomic aphasia produced the lowest PSS values (at −202 and −291 msec). We considered whether these two points may be extreme, and if yes, how they may have affected our statistical analyses especially for the anomic participant group, which had a mean negative PSS (Figure 4A). On the basis of a 3IQR criterion (3 times the interquartile range), which identifies extreme points in a univariate population (IBM Corp., 2023; Hoaglin & Iglewicz, 1987), we found that neither point fell into this category. Note that the PSS value of −202 msec (the second largest negative PSS) is close in magnitude to PSS values of the two participants who produced the largest positive points (181 and 183 msec; Figure 3), and therefore, we do not consider this single point at −202 msec to be aberrant (i.e., all 3 points also fall well within 3 z-score units of the population mean). These findings were confirmed with a Dixon Q test for outliers (Rorabacher, 1991; Dixon, 1950), which verified that the largest negative PSS (−291 msec) was not a significant outlier at a 95% confidence level, $Q(n = 40) = 0.19$, $p > .05$. However, because this negative PSS value (−291 msec) is relatively large in magnitude, we repeated the ANOVA test with this data point excluded and still found a significant effect of Aphasia Type, $F(3, 30) = 3.00$, $p = .046$, $\eta^2 = .23$. A post hoc t test analysis showed that with this point excluded, there also remained a significant difference between the mean PSS values of brain-lesioned participants with no aphasia and the anomic group, $t(11) = 2.35$, $p = .038$ two sided, $p = .019$ one sided, Cohen's $d = 1.34$, and a near-significant difference between the conduction and anomic groups, $t(14) = 1.93$, $p = .074$ two-sided, significant with a one-tailed criterion $p = .037$, Cohen's $d = 0.965$. If we restrict our analysis only to the anomic subpopulation, then the large negative PSS values are even less of an issue as the lowest PSS of −291 msec is only about 2 z-units away from the anomic mean PSS (−66 msec). We should add that of the nine participants with an anomic diagnosis, three produced positive PSS values and six negative. Even with the PSS of −291 msec included in the analysis, the mean PSS for the anomic group is not significantly different than zero, but as noted earlier it approached significance ($p = .055$). Although this negative mean PSS for the anomic group is possibly statistical noise, its interesting counterintuitive direction and near significance, in addition to findings from several prior studies that some listeners consistently adjust the audio to negative PSS values (van Eijk et al., 2008), make this an interesting case for at least some consideration (we briefly revisit this point in the the Discussion section).

Figure 4B shows PSS as a function of brain lesion volume. We found no statistically significant correlation between lesion volume and the ability of participants to synchronize the audio and video of the stimuli, $r(38) = .17$; $p = .30$, nor did we find significant lesion volume effects as a function of aphasia type, $F(3, 31) = 2.09$, $p = .12$, $\eta^2 = .17$. In fact, the mean lesion volume for the conduction aphasia group was larger than that for the anomic

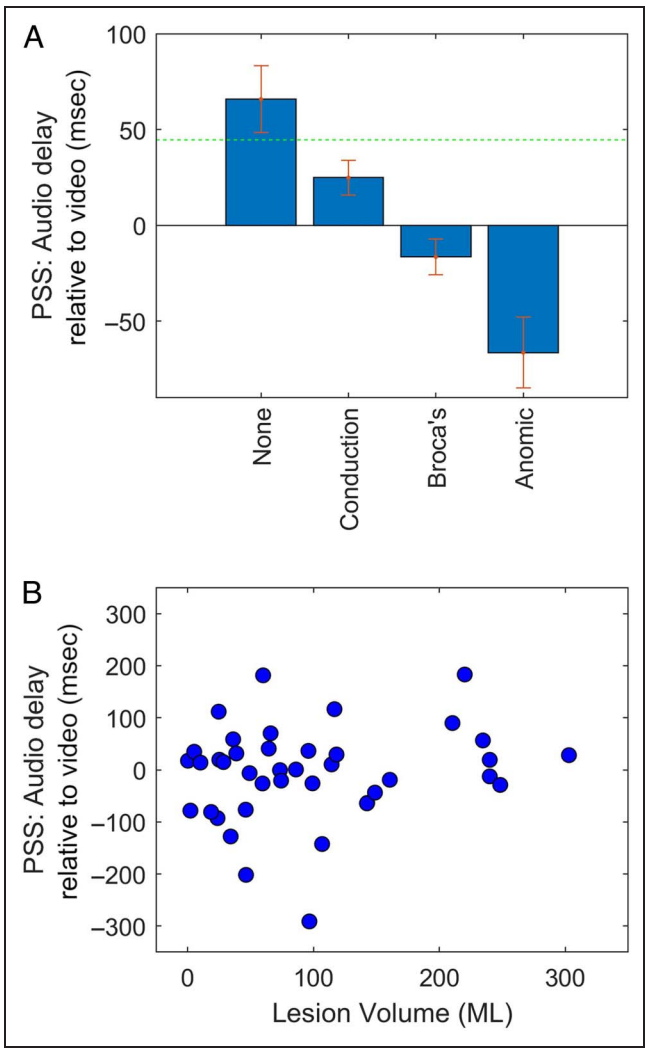

Figure 4. (A) Averaged PSS values for four subgroups categorized by aphasia diagnosis. $n = 35$. Error bars are one standard error. (B) PSS as a function of brain lesion volume for each individual participant. $n = 40$.

group (although not significantly: $t(11) = 0.36, p = .73$) in spite of the fact that the conduction group performed on the synchrony task more similarly to controls than other aphasia groups (anomic or Broca's; Figure 4A). There was also no statistically significant difference in lesion volume between the no-aphasia ("None" in Figure 4A) and the anomic group, $t(12) = 0.75, p = .47$), which are the two groups that produced the most divergent PSS values relative to each other.

Figure 5 shows results of the AV fusion experiment and its relation to AV synchrony. Because AV fusion requires processing of linguistic and phonemic content, and not just perceptual alignment (temporal matching) as in the synchrony task, we restricted data analysis for the AV fusion task to a subset of participants who were able to perform the *auditory only* syllable identification task with greater than 80% accuracy for both syllables (see methods: two auditory stimuli /pa/ and /ka/ and three response options that included /ta/). This procedure yielded a subgroup of 24 participants in the lesion group who, on average, performed the auditory-only syllable identification task with 95% accuracy (96% for /pa/ and 94% for /ka/).

Figure 5A shows AV fusion rate (i.e., percentage of "Ta" responses in the incongruent /pa/−/ka/ condition) plotted as a function of PSS (synchrony task) for each of the 24 lesion participants with 95% accuracy on the syllable identification task. The AV fusion rate represents the percentage of trials on which the McGurk–MacDonald illusion was perceived, with the participant reporting hearing a syllable that neither matched the auditory nor visual stimuli.

We found a significant correlation between AV synchrony and fusion, $r(22) = .59, p = .0013$.[5] This correlation cannot be explained by lesion volume as there was no significant correlation between lesion volume and fusion rate, $r(22) = .03$, *ns*. The higher the PSS values,

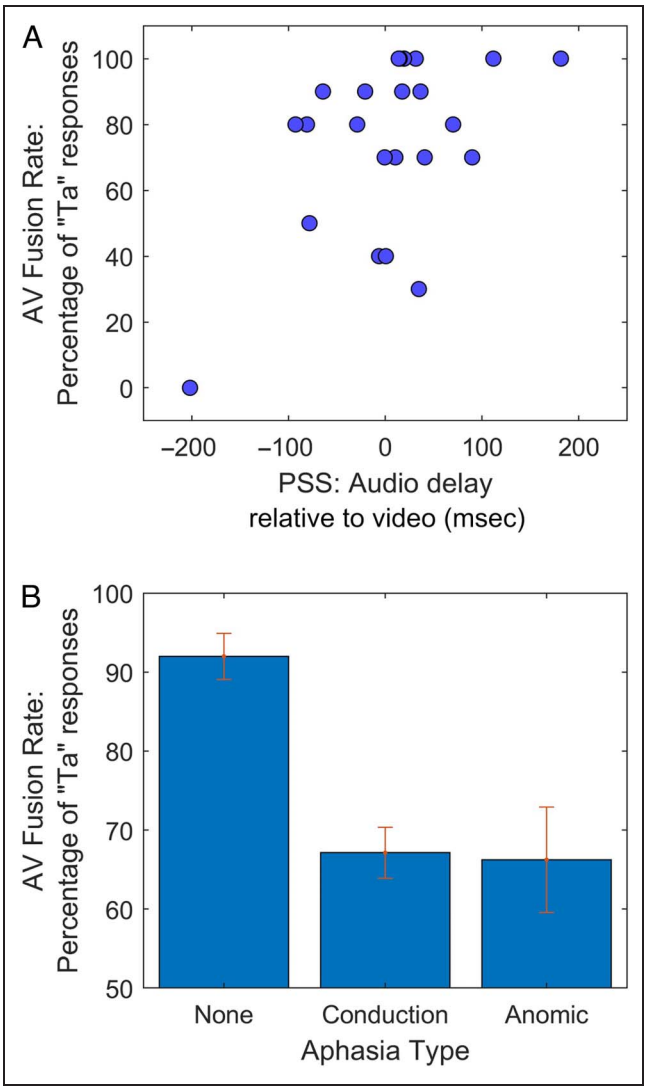

Figure 5. (A) AV fusion rate plotted as a function of AV synchrony (PSS) individually for each of 24 participants. (B) AV fusion rate as a function of aphasia diagnosis. $n = 20$.

the more likely it was for the same participant to display the McGurk–MacDonald illusion. Participants with PSS values that departed most from controls (negative PSS values in Figure 5A) were least likely to display the McGurk–MacDonald illusion, although they performed quite well in their ability to identify auditory syllables in isolation (no visual cue). In fact, the participant who produced the lowest PSS value (near −200 msec) in the synchrony task showed no McGurk–MacDonald illusion at all (0% Ta responses) even though this same participant performed perfectly (100% correct) in identifying auditory-alone syllables (/pa/ or /ka/).[6]

Figure 5B shows AV fusion rate as a function of aphasia diagnosis. We included only those categories of aphasia for which there were at least five participants. To determine if there was an effect of aphasia type on AV fusion, we first performed a Levene's test for the homogeneity of variance, the result of which was highly significant, $L(2, 17) = 7.04, p = .006$. This suggests that the variances of the fusion-rate measure across aphasia types are not homogeneous, and therefore, we used Welch's F test to determine if there was a significant difference in AV fusion rate across aphasia diagnoses. We found a significant effect of Aphasia Type (Welch's $F(2, 11.17) = 4.37, p = .04$), which as observed in Figure 5B was a result of the higher fusion rate of the "no-aphasia" brain-lesioned group. Furthermore, as before, this effect could not be attributed to differences in lesion volume across aphasia groups: There was no significant lesion volume difference between the no-aphasia and anomic groups, $t(11) = 0.87, p = .404$, nor between the no-aphasia and conduction aphasia groups, $t(10) = 0.26, p = .801$.

Several prior neuroimaging and lesion studies have implicated superior temporal areas, especially the posterior superior temporal gyrus (pSTG) as critical to AV processing (Krason et al., 2023; Zhou et al., 2023; Hickok et al., 2018; Erickson et al., 2014; Marchant et al., 2012; Nath & Beauchamp, 2012). We tested the hypothesis that cortical damage to pSTG adversely affects AV processing via a directional one-tailed t test comparing performance of individuals with any damage to the left pSTG to those without any pSTG damage. Left pSTG damage was defined as any lesion map including any left STG gray or white matter posterior to the posterior edge of Heschl's gyrus. The JHU atlas (i.e., the same atlas as used in the lesion-symptom mapping analyses; Faria et al., 2012; Mori et al., 2008) was used to define Heschl's gyrus and the STG (both gray and white matter). pSTG was defined as any portion of the STG (as delineated by the JHU atlas) that was posterior to the most posterior point of the JHU atlas' Heschl's gyrus region. We found a statistically significant effect on AV fusion at the full population level of 40 brain-lesioned participants, $t(38) = 2.12, p = .017$, and a near-significant effect on the subpopulation of 24 participants, $t(22) = 1.33, p = .079$. The lesion-symptom mapping analyses did not identify any regions in which damage was significantly associated with PSS values above or below 45 msec (i.e., control group PSS values), using appropriate multiple correction procedures. For exploratory purposes, the unthresholded z scores for each ROI were inspected. Damage to left posterior superior, middle, and inferior temporal gyri ROIs corresponded to the highest z scores for PSS (synchrony) values below 45 msec, with z scores ranging from 2.24 to 2.61. Damage to frontal and insula regions yielded z scores below 1. For PSS (synchrony) values above 45 msec, there were no regions in which the lesion-symptom mapping analyses yielded z scores > 1.

DISCUSSION

An important finding of the current study is that aphasia type selectively affects AV synchrony. Setting aside the no-aphasia brain-lesioned group who performed most similarly to the control group, the largest contrast in performance of aphasic populations was between those with conduction and anomic aphasia. Individuals with conduction aphasia adjusted the audio to be delayed by approximately 25 msec, on average, relative to the video for perceived synchrony. This value is close to (or at least in the same direction as) that of control participants (45 msec). Conversely, those with anomic aphasia performed most differently than controls, with an average adjustment of −66 msec, that is, with the video delayed relative to the audio (opposite to that of controls).

Anomic aphasia is characterized by difficulty in recalling words or finding the appropriate words to identify a person or object or numbers. Conduction aphasia, in contrast, is primarily a speech production impairment characterized by frequent phonemic speech errors and deficiency in verbatim repetition but with no difficulty in speech perception (Baldo, Klostermann, & Dronker, 2008; Goodglass, 1992; Damasio & Damasio, 1980; Benson et al., 1973). Although individuals with conduction aphasia performed more similarly to the control group, this is not to say that they were unaffected in AV tasks but rather that the adverse effect of lesions in this population is less pronounced. Figure 6 shows lesion overlap maps for the conduction and anomic aphasic participants. Most of the conduction aphasia lesions in our participants do not affect the ventral part of the pSTG, in the superior temporal sulcus, which is most often implicated in AV perception (Hickok et al., 2018). Anomic aphasia, however, is related to more significant lesions in the pSTS (sulcus) compared with conduction aphasia, which was associated with more superior lesions (pSTG but not STS). It is interesting that most participants with conduction aphasia have damage to pSTG and surrounding regions, yet still perform more similarly to the control group in the AV synchrony task. Perhaps the more posterior and ventral temporal lesions seen in the participants with anomic aphasia are more disruptive to AV synchrony.

It is unclear why the anomic group, on average, adjusted the video to lag the audio for perceived synchrony (why not simply a zero delay if there is a general systemic

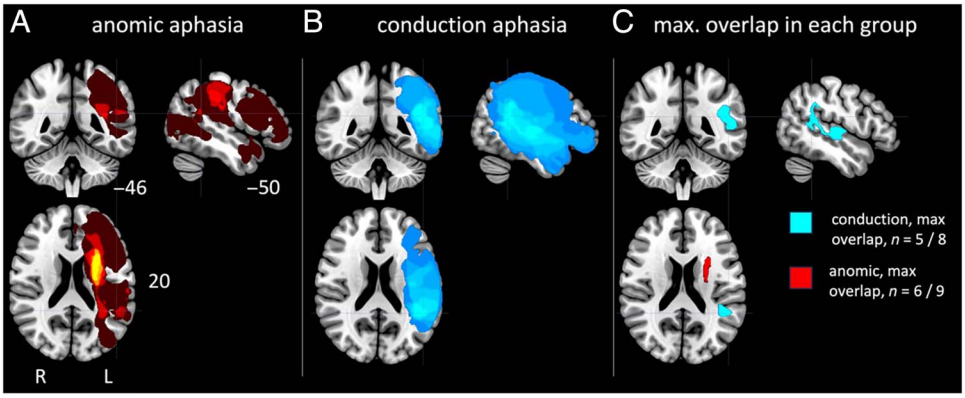

Figure 6. Representative orthogonal slices of the overlap of the lesions of individuals with (A) anomic aphasia range of $n = 1$ to $n = 6$ with warmer colors indicating greater overlap; (B) conduction aphasia with range of $n = 1$ to $n = 5$ with cooler colors indicating greater overlap, and (C) the areas of maximum overlap for each group.

failure?). Even though this effect (negative PSS) was not statistically significant when compared with a referent point of zero, it was near significant ($p = .055$), and therefore, some consideration of its potential origins may be warranted here. Perhaps, in the anomic group, disruption of the neural pathways that process the ascending auditory information may have delayed the arrival of the auditory signal to higher cortical centers that encode synchrony perception. Although this is speculative, a more circuitous distal route to higher centers caused by anomic-specific lesions may delay the arrival of the auditory neural signal and thus require a concomitant compensatory delay of the physical visual stimulus (to reproduce the original expected time difference). This would be the case if such cortical decision centers are calibrated to (and expecting) a fixed time difference between the arrival of auditory and visual information. Consistent with this interpretation, there is evidence from EEG and MEG studies that aphasia may in fact cause a delay in neural transmission of speech stimuli (Kielar, Shah-Basak, Deschamps, Jokel, & Meltzer, 2019; Giaquinto & Ranghi, 2009). We should also clarify that although, on average, most normal-hearing listeners adjust the audio to lag that of its corresponding video for perceived synchrony, several reports have shown that some listeners consistently adjust the audio to lead the video, that is, an ecologically invalid adjustment because an audio signal never naturally leads its corresponding visual component. The reason for this is unknown, but it is observed in different experimental designs including AV temporal-order judgments and synchrony decisions (van Eijk et al., 2008).

The few prior studies that have examined the relationship between aphasia type and AV processing have been individual case studies. Youse, Cienkowski, and Coelho (2004), for example, investigated AV fusion rates (McGurk–MacDonald illusion) in an individual with anomic aphasia and found that compared with an age-matched control, there was a significant decline in the illusion which they attributed to a perseverative response pattern. They suggested that this result may potentially be caused by the difficulty that aphasics have with nonsense syllable combinations (Duffy & Coelho, 2001). In another case study, Ramachandran, Rogers-Ramachandran, Altschuler, and Wisdom (1999) found a significant decline in performance of an individual with Broca's aphasia in an AV fusion task and concluded that Broca's area is critical to AV integration of speech. Andersen and Starrfelt (2015), however, in another case study of an individual with Broca's aphasia found, contrary to Ramachandran and colleagues, that their participant did in fact experience the McGurk–MacDonald illusion, indicating that an intact Broca's area is not required for AV integration of speech. Several other studies of AV processing by aphasics have reported group-level results for aphasia (in general) versus control groups but not selectively by aphasia type (Krason et al., 2023; Michaelis et al., 2020).

One caveat when interpreting our aphasia-type-specific results in the context of previous findings is that aphasia classification may vary based on the clinical assessment used to determine aphasia type (Crary, Wertz, & Deal, 1992; Swindell et al., 1984). For example, there is some evidence that the WAB and BDAE may have inconsistent agreement for aphasia type, including anomic aphasia potentially being underidentified by the BDAE compared with the WAB (Crary et al., 1992). Within the present study, our site-specific assessment procedures included (but were not limited to) either the WAB or BDAE. Each site has their own protocol for monitoring and maintaining the reliability of their diagnostic tests, employing teams of neurologists, speech-language pathologists, and experienced research assistants in classifying an aphasia subtype (with significant data-sharing and interaction across sites). Because all three sites contributed both anomic and conduction aphasics to this data set, if different sites on occasion classify a subject differently (adding noise to the data

and reducing statistical power), then our finding is still compelling in that we observed a statistically significant difference between the anomic and conduction groups in perception of AV synchrony in spite of the additional statistical noise. It is also notable that the lesion overlap maps for our anomic and conduction aphasia groups (Figure 6) align with a variety of previous studies localizing lesions associated with these aphasia subtypes and their defining impairments, further corroborating our diagnosis of the subtypes (Akinina et al., 2019; Yourganov, Smith, Fridriksson, & Rorden, 2015; Baldo, Arévalo, Patterson, & Dronkers, 2013; Dronkers & Baldo, 2009; Baldo et al., 2008). Nonetheless, because this potential variability in assessment protocol may be adding noise to our data, the aphasia type results should be viewed with this potential limitation in mind. Our main finding is, however, still empirically noteworthy and clinically relevant: There are two groups of aphasia participants who perceive AV synchrony in very different ways.

Another significant finding of the current study is the positive correlation between AV synchrony and fusion in the lesion population. Participants who performed most similar to controls in the AV synchrony task also showed the strongest AV fusion rate even when we controlled for lesion volume and for their ability to accurately perceive auditory-only syllables in isolation. This finding adds to results of Hickok and colleagues (2018) who found no correlation between AV fusion and AV advantage. They concluded that the absence of a correlation suggests that these two metrics do not necessarily measure the same underlying process. They further noted that AV advantage likely engages lip-reading as an orthogonal source of information to compensate for degraded speech (i.e., an additive effect instead of an interactive process). They observed that in the AV advantage task, unless speech perception is degraded to begin with, there is, by definition, no room for improvement by adding a video to the audio signal (i.e., a ceiling effect). This requirement does not exist for AV synchrony, which (in this sense) is a more proximate metric to AV fusion in measurement of perceptual integration. Further support for this view is provided by Michaelis and colleagues (2020) who demonstrated an inherent link between AV fusion and synchrony by desynchronizing the audio and video of an incongruent (McGurk–MacDonald) stimulus.[7] They found that if the audio lags the video, there is no significant reduction in AV fusion rate. However, if the audio leads the video, there is a significant decline in fusion (i.e., the illusion is much weaker). This temporal asymmetry may be explained by the fact that a delayed audio is expected for naturally occurring sounds (due to acoustic propagation time) but the reverse, the audio leading the visual, never occurs in nature. Their results suggest that temporally desynchronizing the audio and video is not a simple decoupling of orthogonal cues to fusion, but rather, its directional nature implies a robust integrative relationship between AV synchrony and fusion.

Concluding Remarks

Our findings have important implications for the cortical bases of multisensory integration and how different AV processes interact in encoding incoming stimuli in complex natural environments. Some AV processes may be interactive, as our findings on synchrony and fusion show, or additive, as the orthogonal cues reported for AV advantage and fusion (Hickok et al., 2018). Our findings additionally demonstrate that different AV processes engage functionally and anatomically different cortical regions, with superior temporal areas critical to AV fusion and the posterior and ventral temporal regions important in processing synchronous AV information. Finally, we would like to conclude by noting that beyond what it can reveal about the foundations of cortical organization in AV processing, a better understanding of how various metrics of AV integration are related has practical implications for clinical strategies in treatment of aphasia and other language dysfunctions (de Freitas, 2012; Prigatano, Roueche, & Fordyce, 1985; Geschwind, 1974). Treatment approaches to aphasia rely heavily on bimodal interventions that combine auditory and visual stimulus components (Duffy & Coelho, 2001). Insight into these AV interactions can aid speech pathologists in developing more appropriate treatment protocols, facilitate communication between aphasia patients and clinicians, and develop more effective therapeutic goals based on aphasia-specific diagnoses (Youse et al., 2004). The current findings may therefore additionally contribute to improvements in clinical interventions and rehabilitation strategies for treatment of language disorders.

Acknowledgments

We thank the Center for Hearing Research for assistance in administering clinical hearing tests. We also thank Alexandra Basilakos at the University of South Carolina for data collection, Adrijana Gombosev, the Institute for Clinical and Translational Science, and UCI Institute for Memory Impairments and Neurological Disorders for assistance with recruitment of older adult participants through the UCI Consent-to-Contact Registry. The Consent-to-Contact Registry is supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR001414, as well as by the UCI Alzheimer's Disease Research Center, through grant AG016573. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Corresponding author: Haleh Farahbod, Department of Cognitive Sciences, University of California, Irvine, 2240 SBSG, Irvine, CA 92617, e-mail: hfarahbo@uci.edu.

Data Availability Statement

Raw data and analysis programs associated with the current study are available on request.

Author Contributions

H. F., G. H., and K. S. designed the experiments and contributed to writing the manuscript with input from other authors. J.V. wrote the computer programs used in this study and contributed to the discussions. H. F., L. M. K., J. C., S. P., K. T., A. L., C. R., J. F., and J. B. contributed to data collection and discussions. H. F., C. R., G. H., and K. S. analyzed the results with input from other authors.

Funding Information

This work was supported by the Institute on Deafness and Other Communication Disorders (R01DC009659 and R01DC03681) and the Center for the Study of Aphasia Recovery at the University of South Carolina (National Institute on Deafness and Other Communication Disorders P50-DC014664).

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

Notes

1. This study is part of a larger project aimed at measuring the effects of auditory processing deficits on aphasia. Each of the four research sites had their own established protocol for measuring audiogram thresholds, but all included pure-tone thresholds measured for frequencies from 0.5 to 4 kHz. No significant differences were observed in overall hearing thresholds between the brain-lesioned and older adult control groups. The number of participants in the older adult control group (15) is consistent with previously published auditory psychophysical studies of older adults (Moore & Sęk, 2019 [n = 14]; Moore, Mariathasan, & Sęk, 2018, 2019 [n = 13]; Saberi, Farahbod, Turner, & Hickok, 2022 [n = 15]; Farahbod et al., 2023 [n = 15]). The number of participants in the brain-lesioned group (40) is also consistent with similar published studies (Michaelis et al., 2020 [n = 33]; Krason et al., 2023 [n = 36]).
2. There is a range of delays between the audio and video components at which a participant may respond "synchronous." The "synchrony" report probabilistically declines as the physical desynchrony between the audio and visual components increases, resulting in a psychometric function (fitted here with a Gaussian probability density function) whose expected value is operationally defined as the PSS.
3. Some individuals in the control group ran fewer and some more than five runs, averaging 4.9 runs across the 15 participants and ranging from four to seven runs.
4. The mean PSS value for the first run of the control group was also significantly higher than that of the lesion group (52.5 msec vs. −6 msec); $t(53) = 2.3$, $p = .03$ (two-tailed).
5. To calculate the Pearson correlation coefficient, the fusion-rate data were first z-transformed to account for compressive ceiling effects and nonlinearity of psychometric functions for data based on proportion values (Hautus et al., 2021; Gescheider, 1985; Green & Swets, 1966). By convention, we assumed a small inattention rate in setting a ceiling fusion rate of 99.5% and a floor of 0.5% (Hautus et al., 2021; Green, 1995; Saberi & Green, 1997). This avoids the practical problem of $z = \infty$ when estimating performance in small samples. Although our preference is to use this z-transform approach, the untransformed correlation coefficient was also significant, $r(22) = .51$, $p = .005$. Tests were one-tailed based on our hypothesis of an expected positive correlation between the two AV variables, fusion, and PSS.
6. Removing this one participant from analysis reduced the correlation coefficient from .59 to .37, although the correlation remained significant, $r(21) = .37$, $p < .05$. We should add that the data of the participant who had generated the largest negative PSS value of −291 msec had already been excluded from this correlational analysis because this individual had not met our inclusion criterion for the McGurk effect analysis (i.e., greater than 80% accuracy in the auditory-only syllable-identification task; see the Methods section).
7. Michaelis and colleagues (2020) did not measure the ability of brain-lesioned individuals to synchronize the audio and video of congruent stimuli and hence did not measure PSS for their participants.

REFERENCES

Akinina, Y., Dragoy, O., Ivanova, M. V., Iskra, E. V., Soloukhina, O. A., Petryshevsky, A. G., et al. (2019). Grey and white matter substrates of action naming. *Neuropsychologia*, *131*, 249–265. https://doi.org/10.1016/j.neuropsychologia.2019.05.015, PubMed: 31129278

Andersen, T. S., & Starrfelt, R. (2015). Audiovisual integration of speech in a patient with Broca's aphasia. *Frontiers in Psychology*, *6*, 435. https://doi.org/10.3389/fpsyg.2015.00435, PubMed: 25972819

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, *6*, 260–268. https://doi.org/10.1167/6.3.6, PubMed: 16643094

Baldo, J. V., Arévalo, A., Patterson, J. P., & Dronkers, N. F. (2013). Grey and white matter correlates of picture naming: Evidence from a voxel-based lesion analysis of the Boston Naming Test. *Cortex*, *49*, 658–667. https://doi.org/10.1016/j.cortex.2012.03.001, PubMed: 22482693

Baldo, J. V., Klostermann, E. C., & Dronker, N. F. (2008). It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace. *Brain and Language*, *105*, 134–140. https://doi.org/10.1016/j.bandl.2007.12.007, PubMed: 18243294

Benson, D. F., Sheremata, W. A., Bouchard, R., Segarra, J. M., Price, D., & Geschwind, N. (1973). Conduction aphasia: A clinicopathological study. *Archives of Neurology*, *28*, 339–349. https://doi.org/10.1001/archneur.1973.00490230075011, PubMed: 4696016

Crary, M. A., Wertz, R. T., & Deal, J. L. (1992). Classifying aphasias: Cluster analysis of Western Aphasia Battery and Boston Diagnostic Aphasia Examination results. *Aphasiology*, *6*, 29–36. https://doi.org/10.1080/02687039208248575

Damasio, H., & Damasio, A. R. (1980). The anatomical basis of conduction aphasia. *Brain*, *103*, 337–350. https://doi.org/10.1093/brain/103.2.337, PubMed: 7397481

de Freitas, G. R. (2012). Aphasia and other language disorders. *Frontiers of Neurology and Neuroscience*, *30*, 41–45. https://doi.org/10.1159/000333402, PubMed: 22377860

den Ouden, D.-B., Malyutina, S., Basilaskos, A., Bonilha, L., Gleichgerrcht, E., Yourganov, G., et al. (2019). Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Human Brain Mapping*, *40*, 2153–2173. https://doi.org/10.1002/hbm.24514, PubMed: 30666767

Dinnerstein, A. J., & Zlotogura, P. (1968). Intermodal perception of temporal order and motor skills: Effects of age. *Perceptual & Motor Skills*, *26*, 987–1000. https://doi.org/10.2466/pms.1968.26.3.987, PubMed: 5657762

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*, 719–721. https://doi.org/10.1068/p090719, PubMed: 7220244

Dixon, W. J. (1950). Analysis of extreme values. In *The annals of mathematical statistics* (Vol. 21, pp. 488–506). Institute of Mathematical Statistics. https://doi.org/10.1214/aoms/1177729747

Dronkers, N. F., & Baldo, J. V. (2009). Language: Aphasia. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (pp. 343–348). Oxford: Academic Press. https://doi.org/10.1016/B978-008045046-9.01876-3

Duffy, J. R., & Coelho, C. A. (2001). Schuell's stimulation approach to rehabilitation. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (pp. 745–808). Baltimore: Lippincott Williams & Wilkins.

Erickson, L. C., Heeg, E., Rauschecker, J. P., & Turkeltaub, P. E. (2014). An ALE meta-analysis on the audiovisual integration of speech signals. *Human Brain Mapping*, *35*, 5587–5605. https://doi.org/10.1002/hbm.22572, PubMed: 24996043

Farahbod, H., Rogalsky, C., Keator, L. M., Cai, J., Pillay, S. B., Turner, K., et al. (2023). Informational masking in aging and brain-lesioned individuals. *Journal of the Association for Research in Otolaryngology*, *24*, 67–79. https://doi.org/10.1007/s10162-022-00877-9, PubMed: 36471207

Faria, A. V., Joel, S. E., Zhang, Y., Oishi, K., van Zjil, P. C. M., Miller, M. I., et al. (2012). Atlas-based analysis of resting-state functional connectivity: Evaluation for reproducibility and multi-modal anatomy-function correlation studies. *Neuroimage*, *61*, 613–621. https://doi.org/10.1016/j.neuroimage.2012.03.078, PubMed: 22498656

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *364*, 1211–1221. https://doi.org/10.1098/rstb.2008.0300, PubMed: 19528002

Gescheider, G. A. (1985). *Psychophysics: Method, theory, and application* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Geschwind, N. (1974). Non-aphasic disorders of speech. In *Selected papers on language and the brain. Boston Studies in the Philosophy of Science* (Vol. 16, pp. 73–85). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-2093-0_6

Giaquinto, S., & Ranghi, F. (2009). Slowing of event-related potentials in primary progressive aphasia. A case report. *Scientific World Journal*, *9*, 633–638. https://doi.org/10.1100/tsw.2009.67, PubMed: 19618091

Goodglass, H. (1992). Diagnosis of conduction aphasia. In S. E. Kohn (Ed.), *Conduction aphasia* (pp. 39–49). Hillsdale, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203772188

Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes-no task. *Journal of the Acoustical Society of America*, *93*, 2096–2105. https://doi.org/10.1121/1.406696, PubMed: 8473622

Green, D. M. (1995). Maximum-likelihood procedures and the inattentive observer. *Journal of the Acoustical Society of America*, *97*, 3749–3760. https://doi.org/10.1121/1.412390, PubMed: 7790653

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hautus, M., Macmillan, N. A., & Creelman, C. D. (2021). *Detection theory: A user's guide*. 3rd Edition. Thames, Oxfordshire, England, UK: Routledge Publishing. https://doi.org/10.4324/9781003203636

Hickok, G., Rogalsky, C., Matching, W., Basilakos, A., Cai, J., Pillay, S., et al. (2018). Neural networks supporting audiovisual integration for speech: A large-scale lesion study. *Cortex*, *103*, 360–371. https://doi.org/10.1016/j.cortex.2018.03.030, PubMed: 29705718

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of American Statistical Association*, *82*, 1147–1149. https://doi.org/10.1080/01621459.1987.10478551

IBM Corp. (2023). *IBM SPSS Statistics for Windows*, Version 29.0.2.0. Armonk, NY: IBM Corp.

Jaśkowski, P., Jaroszyk, F., & Hojan-Jezierska, D. (1990). Temporal order judgments and reaction time for stimuli of different modalities. *Psychological Research*, *52*, 35–38. https://doi.org/10.1121/1.412390, PubMed: 2377723

Jaśkowski, P., & Verleger, R. (2000). Attentional bias toward low-intensity stimuli: An explanation for the intensity dissociation between reaction time and temporal order judgment? *Consciousness & Cognition*, *9*, 435–456. https://doi.org/10.1006/ccog.2000.0461, PubMed: 10993668

Keetels, M., & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research*, *167*, 635–640. https://doi.org/10.1007/s00221-005-0067-1, PubMed: 16175363

Kielar, A., Shah-Basak, P. P., Deschamps, T., Jokel, R., & Meltzer, J. A. (2019). Slowing is slowing: Delayed neural responses to words are linked to abnormally slow resting state activity in primary progressive aphasia. *Neuropsychologia*, *129*, 331–347. https://doi.org/10.1016/j.neuropsychologia.2019.04.007, PubMed: 31029594

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492–500. https://doi.org/10.1007/BF00236934, PubMed: 4076371

Krason, A., Vigliocco, G., Mailend, M.-L., Stoll, H., Varley, R., & Buxbaum, L. J. (2023). Benefit of visual speech information for word comprehension in post-stroke aphasia. *Cortex*, *165*, 86–100. https://doi.org/10.1016/j.cortex.2023.04.011, PubMed: 37271014

Marchant, J. L., Ruff, C. C., & Driver, J. (2012). Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human Brain Mapping*, *33*, 1212–1224. https://doi.org/10.1002/hbm.21278, PubMed: 21953980

Mattes, S., & Ulrich, R. (1998). Directed attention prolongs the perceived duration of a brief stimulus. *Perception & Psychophysics*, *60*, 1305–1317. https://doi.org/10.3758/BF03207993, PubMed: 9865072

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. https://doi.org/10.1038/264746a0, PubMed: 1012311

Michaelis, K., Erickson, L. C., Fama, M. E., Skipper-Kallal, L. M., Xing, S., Lacey, E. H., et al. (2020). Effects of age and left hemisphere lesions on audiovisual integration of speech. *Brain and Language*, *206*, 104812. https://doi.org/10.1016/j.bandl.2020.104812, PubMed: 32447050

Mollon, J. D., & Perkins, A. J. (1996). Errors of judgement at Greenwich in 1796. *Nature*, *380*, 101–102. https://doi.org/10.1038/380101a0, PubMed: 8600377

Moore, B. C. J., Mariathasan, S., & Sęk, A. P. (2018). Effects of age on the discrimination of amplitude and frequency modulation for 2- and 10-Hz rates. *Acta Acustica United with Acustica*, *104*, 778–782. https://doi.org/10.3813/AAA.919222

Moore, B. C. J., Mariathasan, S., & Sęk, A. P. (2019). Effects of age and hearing loss on the discrimination of amplitude and frequency modulation for 2- and 10-Hz rates. *Trends in Hearing*, *23*, 2331216519853963. https://doi.org/10.1177/2331216519853963, PubMed: 31250705

Moore, B. C. J., & Sęk, A. P. (2019). Discrimination of the phase of amplitude modulation applied to different carriers: Effects of modulation rate and modulation depth for young and older subjects. *Journal of the Acoustical Society of America*, *146*, 1696–1704. https://doi.org/10.1121/1.5126515, PubMed: 31590555

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., et al. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*, *40*, 570–582. https://doi.org/10.1016/j.neuroimage.2007.12.035, PubMed: 18255316

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage*, *59*, 781–787. https://doi.org/10.1016/j.neuroimage.2011.07.024, PubMed: 21787869

Pöppel, E. (1988). *Mindworks: Time and conscious experience* [Trans. of Grenzen des Bewußtseins by T. Artin]. Boston: Harcourt Brace Jovanovich.

Prigatano, G. P., Roueche, J. R., & Fordyce, D. J. (1985). Nonaphasic language disturbances after closed head injury. *Language Sciences*, *7*, 217–229. https://doi.org/10.1016/S0388-0001(85)80020-6

Ramachandran, V. S., Rogers-Ramachandran, D., Altschuler, E. L., & Wisdom, S. B. (1999). A test of Rizzolatti's theory of language evolution; McGurk effect and lip reading in Broca's aphasia? *Society for Neuroscience Abstracts*, *25*, 1636, Program No. 654.11.

Rogalsky, C., Basilakos, A., Rorden, C., Pillay, S., LaCroix, A. N., Keator, L., et al. (2022). The neuroanatomy of speech processing: A large-scale lesion study. *Journal of Cognitive Neuroscience*, *34*, 1355–1375. https://doi.org/10.1162/jocn_a_01876, PubMed: 35640102

Rorabacher, D. B. (1991). Statistical treatment for rejection of deviant values: Critical values of Dixon's 'Q' parameter and related subrange ratios at the 95% confidence level. *Analytical Chemistry*, *63*, 139–146. https://doi.org/10.1021/ac00002a010

Saberi, K., Farahbod, H., Turner, K., & Hickok, G. (2022). Detection of dynamic changes in interaural delay by older adults (L). *Journal of the Acoustical Society of America*, *152*, 2292–2295. https://doi.org/10.1121/10.0014833, PubMed: 36319225

Saberi, K., & Green, D. M. (1997). Evaluation of maximum-likelihood estimators in nonintensive auditory psychophysics. *Perception & Psychophysics*, *59*, 867–876. https://doi.org/10.3758/BF03205504, PubMed: 9270361

Sanford, A. J. (1971). Effects of changes in the intensity of white noise on simultaneity judgments and simple reaction time. *Quarterly Journal of Experimental Psychology*, *23*, 296–303. https://doi.org/10.1080/14640746908401824

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, *47*, 333–366. https://doi.org/10.1016/S0010-0285(03)00035-5, PubMed: 14642288

Shore, D. I., & Spence, C. (2005). Prior entry. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 89–95). Amsterdam: Elsevier, North-Holland. https://doi.org/10.1016/B978-012375731-9/50019-7

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*, 205–212. https://doi.org/10.1111/1467-9280.00337, PubMed: 11437302

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics*, *65*, 318–328. https://doi.org/10.3758/BF03194803, PubMed: 12713247

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience*, *19*, 1964–1973. https://doi.org/10.1162/jocn.2007.19.12.1964, PubMed: 17892381

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 539–550. https://doi.org/10.1037/0096-1523.17.2.539, PubMed: 1830091

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *Neuroimage*, *55*, 1339–1345. https://doi.org/10.1016/j.neuroimage.2010.12.063, PubMed: 21195198

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society B*, *268*, 31–38. https://doi.org/10.1098/rspb.2000.1326, PubMed: 12123295

Swindell, C. S., Holland, A. L., & Fromm, D. (1984). *Classification of aphasia: WAB type versus clinical impression*. University of Pittsburgh Clinical Aphasiology: Proceedings of the Conference 1984.

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 & Psychophysics*, *70*, 955–968. https://doi.org/10.3758/PP.70.6.955, PubMed: 18717383

Vatakis, A., & Spence, C. (2006a). Audiovisual synchrony perception for music, speech, and object actions. *Brain Research*, *1111*, 134–142. https://doi.org/10.1016/j.brainres.2006.05.078, PubMed: 16876772

Vatakis, A., & Spence, C. (2006b). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neuroscience Letters*, *393*, 40–44. https://doi.org/10.1016/j.neulet.2005.09.032, PubMed: 16213656

Venezia, J. V., Thurman, S. M., Matchin, W., George, S. E., & Hickok, G. (2016). Timing in audiovisual speech perception: A mini review and new psychophysical data. *Attention, Perception & Psychophysics*, *78*, 583–601. https://doi.org/10.3758/s13414-015-1026-y, PubMed: 26669309

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophysics*, *72*, 871–884. https://doi.org/10.3758/APP.72.4.871, PubMed: 20436185

Yourganov, G., Smith, K. G., Fridriksson, J., & Rorden, C. (2015). Predicting aphasia type from brain damage measured with structural MRI. *Cortex*, *73*, 203–215. https://doi.org/10.1016/j.cortex.2015.09.005, PubMed: 26465238

Youse, K. M., Cienkowski, K. M., & Coelho, C. A. (2004). Auditory–visual speech perception in an adult with aphasia. *Brain Injury*, *18*, 825–834. https://doi.org/10.1080/02699000410001671784, PubMed: 15204322

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research*, *152*, 198–210. https://doi.org/10.1007/s00221-003-1536-z, PubMed: 12879178

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience Letters*, *381*, 217–222. https://doi.org/10.1016/j.neulet.2005.01.085, PubMed: 15896473

Zhou, H.-Y., Cheung, E. F. C., & Chan, R. C. K. (2020). Audiovisual temporal integration: Cognitive processing, neural mechanisms, developmental trajectory and potential interventions. *Neuropsychologia*, *140*, 107396. https://doi.org/10.1016/j.neuropsychologia.2020.107396, PubMed: 32087206

Zhou, H.-Y., Zhang, Y.-J., Hu, H.-X., Yan, Y.-J., Wang, L.-L., Lui, S. S. Y., et al. (2023). Neural correlates of audiovisual speech synchrony perception and its relationship with autistic traits. *PsyCh Journal*, *12*, 514–523. https://doi.org/10.1002/pchj.624, PubMed: 36517928