

Published in final edited form as:

Exp Brain Res. 2012 May ; 219(1): 121–137. doi:10.1007/s00221-012-3072-1.

Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance

Ryan A. Stevenson¹, Juliane Krueger Fister^{1,2}, Zachary P. Barnett^{1,3}, Aaron R. Nidiffer¹, and Mark T. Wallace^{1,3,4,5,6}

¹Department of Hearing and Speech Sciences, Vanderbilt University Medical Center

²Neuroscience Graduate Program, Vanderbilt University Medical Center

³Vanderbilt Kennedy Center

⁴Vanderbilt Brain Institute

⁵Department of Psychology, Vanderbilt University

⁶Department of Psychiatry, Vanderbilt University

Abstract

In natural environments, human sensory systems work in a coordinated and integrated manner to perceive and respond to external events. Previous research has shown that the spatial and temporal relationships of sensory signals are paramount in determining how information is integrated across sensory modalities, but in ecologically plausible settings, these factors are not independent. In the current study we provide a novel exploration of the impact on behavioral performance for systematic manipulations of the spatial location and temporal synchrony of a visual-auditory stimulus pair. Simple auditory and visual stimuli were presented across a range of spatial locations and stimulus onset asynchronies (SOAs), and participants performed both a spatial localization and simultaneity judgment task. Response times in localizing paired visual-auditory stimuli were slower in the periphery and at larger SOAs, but most importantly, an interaction was found between the two factors, in which the effect of SOA was greater in peripheral as opposed to central locations. Simultaneity judgments also revealed a novel interaction between space and time: individuals were more likely to judge stimuli as synchronous occurring in the periphery at large SOAs. The results of this study provide novel insights into (a) how the speed of spatial localization of an audiovisual stimulus is affected by location and temporal coincidence and the interaction between these two factors, and (b) how the location of a multisensory stimulus impacts judgments concerning the temporal relationship of the paired stimuli. These findings provide strong evidence for a complex interdependency between spatial location and temporal structure in determining the ultimate behavioral and perceptual outcome associated with a paired multisensory (i.e., visual-auditory) stimulus.

Keywords

Audiovisual; inverse effectiveness; response time; race model; multisensory

Introduction

The different human sensory systems are continually challenged with stimulation from the external environment. These systems are each sensitive to a particular type and range of

energies, use distinct methods of transduction, and have dedicated sensory pathways into primary cortical domains. Because of their unique attributes, the different sensory systems have often been considered and studied in isolation. However, under natural circumstances, these systems must work in a coordinated manner in order for us to properly detect, interpret, and guide interactions with events in the external environment, and in order to form a coherent perceptual Gestalt of our world. To achieve this goal, information from the different sensory systems is centrally integrated to produce behavioral outcomes that can be more sensitive (Lovelace et al. 2003; Stein and Wallace 1996), accurate (Nelson et al. 1998; Wilkinson et al. 1996), and faster (Diederich and Colonius 2004; Hershenson 1962) than the responses seen with individual sensory inputs.

Over the past few decades, the study of the neural, behavioral and perceptual correlates of multisensory integration has grown dramatically. One of the most consistent set of findings in these studies is the profound effect that stimulus-level factors can have on integrative processes and outcomes. The best characterized of these factors are the temporal and spatial relationships of the stimuli to one another, as well as their relative effectiveness in evoking a response [*time* (Dixon and Spitz 1980; Meredith et al. 1987; Miller and D'Esposito 2005; Senkowski et al. 2007; Stevenson et al. 2010), *space* (Meredith and Stein 1986a, b; Wallace et al. 1992), and *effectiveness* (Meredith and Stein 1986b; Stevenson et al. 2007; Stevenson and James 2009)]. Collectively, the role of these factors has been encapsulated into the core *principles of multisensory integration*, in which the more spatially or temporally coincident and the less effective the sensory inputs that are combined (e.g., visual and auditory), the greater the multisensory enhancement when they are combined, with enhancement typically defined as increased performance relative to unisensory conditions. The applicability of these principles as general guidelines for integration has been demonstrated in both animal and human studies, as well as across a range of levels of analysis spanning single unit recordings, event-related potentials (ERPs), functional magnetic resonance imaging (fMRI), and behavior (for recent examples, see Wallace and Murray 2011). In the current study we will focus on the effects of space and time, with emphasis on elucidating the interaction between these two stimulus-level factors.

There are two particular areas where we can gain insight into the possible effects of spatial location on human behavior. The first is from single-cell recordings in the brainstem and cortex of cats and non-human primates. In this work, single-unit responses have been measured while varying the location of stimuli within the neuron's multisensory receptive fields. It was found that for those regions within the receptive fields in which unisensory (i.e., visual alone, auditory alone) responses were low, multisensory combinations produced the largest multisensory enhancement (Carriere et al. 2008; Krueger et al. 2009; Royal et al. 2009). This result has been interpreted in the context of the principle of inverse effectiveness, which holds that unisensory inputs that only elicit a minimal response when presented in isolation show maximal multisensory enhancement when presented in concert (Meredith and Stein 1986b). In extending this logic to human performance, one would predict that the largest multisensory enhancements should be observed at the locations that show the weakest unisensory responses to both auditory- and visual-only presentations. This hypothesis is in line with a second research finding from which we can gain insight. In a behavioral study in which cats were trained to localize a visual cue, their accuracy was lower in peripheral locations, and they showed greater multisensory enhancements at these locations with the addition of a spatially congruent auditory cue (Stein et al. 1988). The first aim of the current experiment is to examine what effect spatial location has on audiovisual integration by measuring the accuracy and response times to spatially-coincident audiovisual targets. Through comparisons of responses to multisensory (i.e., paired audiovisual) stimuli to those observed with unisensory stimuli, and in the case of RTs to a race model that is

derived from those unisensory responses, we set out to measure how multisensory integration changes across parametrically varied spatial locations.

In addition to being modulated by the spatial properties of the combined stimuli, audiovisual integration is also dependent upon the temporal relationships of these stimuli. As a general rule, the more temporally synchronous the paired stimuli are to one another, the greater the multisensory enhancement. This is a well-established finding, having been shown in single-unit recordings (Meredith et al. 1987; Meredith et al. 1992; Royal et al. 2009), ERPs (Schall et al. 2009; Senkowski et al. 2007; Talsma et al. 2009), fMRI (Macaluso et al. 2004; Miller and D'Esposito 2005; Stevenson et al. 2010; Stevenson et al. 2011), and behavioral studies (Conrey and Pisoni 2006; Conrey and Pisoni 2004; Dixon and Spitz 1980; Foss-Feig et al. 2010; Hillock et al. 2011; Keetels and Vroomen 2005; Powers et al. 2009; van Atteveldt et al. 2007; van Wassenhove et al. 2007; Wallace et al. 2004; Zampini et al. 2005a). In addition, this temporal "principle" has been shown across stimulus types ranging from the very simple to the complex (i.e., speech - Conrey and Pisoni 2006; Hirsh and Sherrick 1961; Keetels and Vroomen 2005; Miller and D'Esposito 2005; Stevenson et al. 2010; van Atteveldt et al. 2007; van Wassenhove et al. 2007; Vatakis and Spence 2006; Zampini et al. 2005a; Zampini et al. 2003). Multisensory enhancements are not only seen when the stimuli are synchronously presented, but also with a range of asynchronies that have come to define the multisensory temporal binding window (TBW). Within the TBW, two sensory inputs are highly likely to be perceptually bound or fused into a single Gestalt percept (for review, see Vroomen and Keetels 2010). The construct of the TBW has strong ethological validity, given that auditory and visual sensory inputs arising from the same source are typically not precisely temporally aligned due to differences in the speed at which light and sound travel (Pöppel et al. 1990), as well as differences in sensory transduction and processing time (Corey and Hudspeth 1979; King and Palmer 1985; Lamb and Pugh 1992; Lennie 1981). As such, in order to perceptually bind two sensory inputs that originate from a single external event, the multisensory systems must allow for a limited range of temporal discrepancy. Indeed, when presented with audiovisual stimuli at varying distances, an individual's point of subjective simultaneity varies; the further away the stimulus is from the subject, the longer the auditory lag needed for reporting simultaneity (Lewald and Guski 2004; Sugita and Suzuki 2003).

In any ecologically plausible setting, variations in space and time (and effectiveness) do not happen in isolation (Carriere et al. 2008; Krueger et al. 2009; Royal et al. 2009). External events occur across the full range of sensory space, and variations in stimulus distance cause asynchronies in the time at which the auditory and visual signals reach the peripheral sensory receptors and the primary sensory cortices (Pöppel et al. 1990). Despite this, while the impact of space and time on multisensory function are well established, only a few studies to date have addressed possible interactions between these two stimulus dimensions (Lewald and Guski 2003; Macaluso et al. 2004; Wallace et al. 2004). This is somewhat surprising, since it must be acknowledged that these factors are complexly intertwined in any naturalistic multisensory circumstance. Take the example of a ball bouncing away from you. As the ball moves in space, the relative timing and strength of the visual and auditory cues coming from it changes in a systematic manner. In many respects, investigating the way with which processing of these different stimulus factors interact is a next logical step in the study of multisensory integration, given that the ultimate goal is to better understand how dynamic and ethologically relevant cues are combined and integrated in order to drive behavior and shape perception.

In the current study we provide a novel exploration of the behavioral impact of the relationship between spatial location and temporal synchrony. Using parametric variations in the spatial location and temporal synchronies of audiovisual stimuli, we explored

interactions between these two factors. Two specific hypotheses were at the core of these experiments. First, we predict main effects of spatial location and temporal synchrony such that more peripheral stimulus pairs would be less effective in eliciting accurate localization, and thus would be associated with stronger multisensory enhancement as predicted by the principle of inverse effectiveness. Second, we predict a positive interaction between spatial location and temporal synchrony such that changes in multisensory enhancements observed when varying each factor independently are compounded when both factors are varied concurrently. We measured the effects of these factors on accuracy and response time (RT) using two tasks, a spatial localization task and a simultaneity judgment task. The results reveal main effects of stimulus location and temporal structure, and more importantly, a strong interaction between these processing of these factors.

Methods

Participants

Participants were fifty-one Vanderbilt undergraduate students (21 male, mean age = 18.9 yrs, $sd = 1$ yr, age range = 18–21 yrs) and were compensated with class credit. All recruitment and experimental procedures were approved by the Vanderbilt University Institutional Review Board.

Stimuli

Auditory and visual stimuli were presented using E-Prime version 2.0.8.79 (Psychology Software Tools Inc, Sharpsburg, PA; PST). Visual stimuli were presented on two Samsung Sync Master 2233RZ monitors running at a 100 Hz refresh rate arranged so that each monitor crossed the circumference of circle centered on the participants nasium at a distance of 45.7 cm at 0°, 30°, 60°, and 90° azimuth, with all presentations in the right visual field (see Figure 1A). All visual stimuli were white circles measuring 7 mm in diameter, or approximately 1° of visual angle. Visual stimulus durations were 10 ms, with timing confirmed using a Hameg 507 oscilloscope with a photovoltaic cell. Visual stimuli were presented at 215 cd/m² over a black background of 0.28 cd/m², measured with a Minolta Chroma Meter CS-100. Visual stimuli were presented at each spatial location for a total of four visual-only conditions (Figure 1C).

Auditory stimuli were presented via four separate speakers mounted on the top of the two monitors at 0°, 30°, 60°, and 90° azimuths angled towards the participant, matching the visual presentations. Speakers were mounted 20 mm, or approximately 2.5° of visual angle above their respective visual presentation (see Figure 1C). Speakers were constructed of all black materials, and were not visible in the darkened whisper room. Auditory stimuli consisted of a frozen white-noise burst generated at 44100 Hz with the Matlab *rand* function with a 5 ms rise/fall cosine gate (see Figure 1B). Auditory stimulus duration was held constant at 10 ms, with timing confirmed using a Hameg 507 oscilloscope. Auditory stimuli were presented at 64 dB SPL with a background noise at 41 dB SPL, measured with a Larson Davis sound level meter, Model 814. Auditory stimuli were presented at each spatial location for a total of 4 auditory-only conditions.

Audiovisual (AV) conditions consisted of pairs of the auditory and visual stimuli described above. Presentations were always spatially coincident. The temporal offset of the auditory and visual stimulus pairs were parametrically varied with SOAs of 0, 50, 100, and 200 milliseconds, with timing confirmed using a Hameg 507 oscilloscope. In all cases, visual onset preceded auditory onset. Visual-leading stimuli were chosen as they reflect the natural order in which sensory signals emanating from a single external event arrive at the retina

and cochlea (i.e., light travels much faster than sound). AV conditions were presented at each spatial location (4), and each SOA (4), for a total of 16 AV conditions.

Additionally, a blank, no stimulus condition was also included in which no auditory or visual stimulus was presented while all aspects of the trial remained consistent. In total 25 unique conditions were presented and analyzed, four visual-only, four auditory-only, sixteen AV, and one blank. This experiment was conducted as a part of a larger study in which additional conditions were concurrently run, but will not be discussed in this report. These additional conditions consisted of trials identical to those described below at a lower level of stimulus effectiveness (a lower visual luminance and a lower auditory intensity).

Procedure

Participants were seated inside an unlit WhisperRoom™ (Model SE 2000; Whisper Room Inc, Morristown, TN) with their forehead placed against a Headspot (University of Houston College of Optometry, Houston, TX) forehead rest locked in place, with a chinrest and chair height adjusted individually to the forehead rest. Participants were asked to fixate towards a fixation cross at all times, and were monitored by close circuit infrared cameras throughout the experiment to ensure compliance.

Each trial began with a fixation screen for 1 s with a fixation cross in visual center field at the 0° azimuth visual presentation point. Fixation was followed by blank screen with a randomly jittered time between 500 and 1000 ms, with stimulus presentation immediately after. Following stimulus presentation, the first of two response screens appeared, associated with the spatial location task. The response screen consisted of the prompt, “Where was it?” below the fixation cross, with the response options 1-4 displayed on screen at each spatial location coincident with stimulus presentation, and a fifth option (5) for no stimulus detected. Participants completed the spatial location task by responding via a five-button PST serial response box were 1 = 0°, 2 = 30°, 3 = 60°, 4 = 90°, and 5 = no stimulus, with specific instructions to respond “as quickly and as accurately as possible.” Following the participant’s response, the second of two response screens appeared which was associated with the simultaneity judgment task. The response screen consisted of the prompt “Was it synchronous?” below a fixation cross, with the response options listed below. Again, participants completed the simultaneity judgment task by responding via a five-button PST serial response box were 1 = synchronous, 2 = asynchronous, 3 = visual only, 4 = audio only, and 5=no stimulus (in the case that the participant only detected a stimulus in one modality, they were asked to indicate that with a visual-only or auditory-only response). Following the participant’s simultaneity judgment response, the fixation cross appeared, and the subsequent trial began. For a visual depiction of a trial, see Figure 1D.

Participants completed four sessions, each lasting approximately 20 minutes. Each session consisted of five stimulus presentations with each of the 25 conditions in a randomized order, for a total of 125 trials per session. Across the four sessions then, participants completed 20 trials per condition, for a total of 500 trials per participant. Participants were given breaks in between sessions as needed. Including breaks, total experiment time was approximately 90 minutes.

Analysis

Data from participants who did not detect synchronous audiovisual stimuli presented at the 0° spatial offset at a rate of 80% or greater were excluded from spatial localization analysis ($n = 5$), and data from participants who did not report synchrony for foveal, 0° azimuth target location, 0 ms SOA presentations at a rate of 50% were excluded from perceived synchrony analysis ($n = 5$). One further participant was excluded from all analysis for

responding “1” in all trials for both tasks, leaving a total of 40 participants included in all analyses.

Response accuracies were measured for the spatial location task for all conditions. Comparisons between unisensory and multisensory presentations at each location were compared, as well as between spatial location and SOA. Rates of perceived synchrony were also measured for all responses in which the participant perceived an audiovisual stimulus. As such, raw rates of perceived synchrony for each condition were calculated as:

$$\frac{\# \text{trials perceived synchronous}}{\# \text{trials perceived synchronous} + \# \text{trials perceived asynchronous}}.$$

The use of this calculation actively excludes trials in which the participant did not perceive either the auditory or visual stimulus (note that the denominator does not include trials in which participants responded “audio only”, “visual only”, or “no stimulus”). Rates of perceived synchrony were compared across spatial location and SOA. To measure and graphically represent changes in the rate of perceived synchrony accounting for possible differences in perception of synchrony across stimulus location (i.e. an interaction effect), normalized rates of perceived synchrony were calculated by normalizing data at each stimulus location to the 0ms SA condition at the respective location. Responses to the simultaneity judgment task were also used to assess the detection rates of each stimulus condition, where responses of “visual-only” and “auditory-only” reflected accurate detection of the stimulus in the visual-only and auditory-only conditions, respectively, and a response of “synchronous” or “asynchronous” reflected accurate detection of the stimulus in the audiovisual conditions.

RTs were calculated only for the spatial localization task given that the simultaneity judgment response followed the localization response, and was analyzed using multiple methods. First, mean RTs for correct trials associated with each condition were calculated and compared. Mean RTs were calculated across trials of each condition for each individual. Individual mean RTs were then averaged into a group mean RT for each condition. Mean RTs were compared across sensory modality, spatial location, and SOA. Additionally, cumulative distribution functions (CDFs) were calculated and compared for correct trials of each condition, again averaging within each participant, and then across participants. Finally, multisensory enhancement was measured by creating a race model based on individual’s unisensory auditory- and visual-only CDFs. Such race models account for statistical facilitation that may occur with multiple, redundant targets, and that do not necessarily imply the interaction between two processing streams. Violations of the race model imply enhancements beyond mere statistical facilitation that reflect interactions between processing streams, in this case, the integration of auditory and visual information. Race models (Raab 1962) were calculated for each participant and condition as well using the equation:

$$\widehat{p}(AV_t) = p(V_t) + p(A_{t-SOA}) - p(A_{t-SOA} \cap V_t),$$

where t is time in ms from 0-2000 ms post stimulus onset. For asynchronous trials in which $SOA \neq 0$, SOA was accounted for only in the auditory component due to the fact that auditory presentation always followed the visual presentation. Race models and the respective individual’s AV CDFs were compared at each spatial location and SOA. Violations of the race model where:

$$p(AV_t) > \widehat{p}(AV_t),$$

define interactions indicative of multisensory enhancement, or reaction times that are faster than predicted by the independent unisensory trials. The magnitude of violations, or multisensory enhancement, was then compared across location and SOA.

Results

Effects of spatial location on response accuracy for unisensory and synchronous audiovisual stimuli

For participant's accuracies in correctly identifying the location of a target stimulus (Figure 2, Table 1), effects of stimulus location were found for both unisensory auditory and visual presentations as well as for synchronous audiovisual presentations. A repeated-measures ANOVA was conducted across the three stimulus modalities (i.e., auditory, visual, synchronous audiovisual) and four locations (i.e., 0°, 30°, 60°, 90°), revealing a main effect of modality ($p < 1.00 \times 10^{-10}$, $F_{(2,39)} = 108.19$) and location ($p < 1.00 \times 10^{-10}$, $F_{(3,39)} = 77.82$), as well as an interaction between the two ($p < 1.00 \times 10^{-10}$, $F_{(6,39)} = 66.35$). Given these significant results, follow up t-tests were then conducted, with pair-wise statistics for accuracy with each modality are shown in Table 2. For visual stimuli, a pattern of progressive decreases in response accuracy as target location moved from central (i.e., 0°) to peripheral (i.e., 90°) locations was observed (Figure 2, blue). In comparison, response accuracies to auditory targets significantly decreased as stimulus location was moved from 0° to 60°, but showed an increase in accuracies from 60° to 90° (Figure 2, red), in agreement with previous findings (Nordlund 1962). Accuracy to synchronous AV targets significantly decreased as stimulus location was moved peripherally from 0° to 60°, but showed no difference between 60° and 90° (Figure 2, purple).

A significant multisensory-mediated enhancement in response accuracy is evident in the pattern of AV responses at intermediate locations (i.e., 30°, 60°), when both the unisensory visual and auditory responses were less accurate. Relative to their peak performance on unisensory trials, subjects' accuracy on multisensory trials was significantly greater in both the 30° ($p < 0.002$, $t = 2.90$) and 60° ($p < 0.008$, $t = 2.40$) conditions. In contrast, when one (or both) of the sensory channels was highly accurate, multisensory performance enhancements were weak or non-existent, as seen in the absence of effects at 0° ($t = 0.32$) and 90° ($t = 0.41$).

Effects of spatial location on response accuracy for asynchronous stimuli

In addition to the pattern of responses to synchronously presented AV targets, the effect of relative stimulus timing, specifically SOA, on target localization was also examined (Figure 3, Table 1). For this analysis, a repeated-measures ANOVA was conducted across the four levels of SOA and the four tested locations, revealing a main effect of location ($p < 2.39 \times 10^{-8}$, $F_{(3,39)} = 15.05$), but no effect of SOA ($F_{(3,39)} = 0.24$) and no interaction between the two ($F_{(3,39)} = 1.30$). Within this significant main effect, responses across spatial location were compared at each SOA. As for the synchronous AV conditions, response accuracies for asynchronously presented targets declined as stimulus location was moved peripherally from 0° to 60°, but showed no difference in accuracy between 60° to 90°. The overall pattern of results are summarized in Table 2 and Figure 3, and reveal near perfect performance for the 0° location and a significant decline in performance as the stimuli move toward the periphery, regardless of SOA.

Effects of spatial location on mean response times

The patterns of RTs to this same spatial task were analyzed using multiple methods including mean RTs, cumulative distribution functions (CDFs), and comparisons with race model predictions (see methods section for details).

Mean RTs to unisensory visual and auditory targets showed patterns similar to their respective accuracy results (Figure 4, Table 1). A repeated measures, ANOVA was conducted across the three stimulus modalities and four locations, revealing a main effect of modality ($p < 5.89 \times 10^{-9}$, $F_{(2,39)} = 24.40$) and location ($p < 1.00 \times 10^{-10}$, $F_{(3,39)} = 99.68$), as well as an interaction between the two ($p < 1.00 \times 10^{-10}$, $F_{(6,39)} = 24.66$). Given these significant results, follow up t-tests were then calculated. The more peripheral a visual target was, the slower the RTs (Figure 4A, blue), an effect that was significant at every location. With auditory stimuli, responses at more peripheral locations were significantly slower in all cases with the exception of the comparison between 60 and 90°, which was not significantly different (Figure 4A, red). For detailed statistics on each pair-wise comparison, see Table 2.

To assess the impact of spatial location and SOA on mean RTs for multisensory AV targets (Figure 4A, purple) a repeated-measures ANOVA was run across the four levels of SOA and spatial location. Mean AV RTs showed a significant main effect of SOA ($p < 1.00 \times 10^{-10}$, $F_{(3,39)} = 45.67$), such that the larger the asynchrony, the slower the mean RT. The main effect of spatial location was also significant ($p < 1.00 \times 10^{-10}$, $F_{(3,39)} = 92.96$), with the more peripheral stimulus presentations resulting in slower mean RT. Finally, the two-way interaction between SOA and spatial location was significant ($p < 0.002$, $F_{(9,39)} = 3.15$). Given the significant main effects of spatial location and SOA found in the ANOVA, pair-wise comparisons were made across spatial location and SOA. Stronger effects of location were observed with greater levels of asynchrony, and stronger effects of SOA were seen for more peripheral locations. For detailed statistics on pair-wise comparisons of mean RTs for the spatial localization task, see Table 2.

Effects of spatial location on response time CDFs

In order to provide a more nuanced view into the changes in RTs as a function of spatial location and SOA, cumulative distribution functions (CDFs) were generated for both the unisensory and multisensory conditions (Figure 4B-C). CDFs were compared across spatial location using step-wise Kolmogorov-Smirnov (KS) tests. Both unisensory and multisensory CDFs showed slower responses with peripheral presentations, again with the exception of the 90° auditory-only presentation. Detailed statistical results for step-wise KS tests can be found in Table 3.

To examine multisensory interactions on a subject-by-subject basis, each individual's multisensory CDFs were compared with the CDF predicted by the race model and which is based on the probability summation of the unisensory CDFs (see methods section for details on the computation of race models). Across conditions, an average of 35 out of 40 participants (range = 30 to 39 depending upon condition, see Table 4 for details) showed race model violations, indicative of multisensory enhancement. Average multisensory and race model CDFs were calculated and compared, with significant violations associated with more central and synchronous presentations (Figure 5A - for statistical details see Table 4, see Figure S1 for plots of all 16 conditions). These enhancements in group-averaged CDFs are best illustrated by the difference plots shown in Figure 5B in which positive values represent violations of the race model (see Figure S2 for comparisons of all 16 conditions). Note here that the only significant positive values are for the 0° location at an SOA of 0 ms.

Changes in multisensory enhancement were calculated as main effects across both spatial location (Figure 5C) and SOA (Figure 5D). To quantify these changes, differences between

an individual's multisensory CDFs and their respective race model CDFs were binned into 100 ms blocks. The effect of spatial location was measured by comparing these bins within subjects across spatial locations using paired-sample t-tests (Figure 5C). Significant effects were seen at all SOAs. The same analysis was performed with variations in SOA at each location, revealing that target presentations with shorter SOAs showed significantly more multisensory integration (as judged by race model violations - Figure 5D). This effect was stronger at central when compared with peripheral locations. Importantly, the two-way interaction between SOA and spatial location was quantified through a difference-of-differences measure, again structured in 100 ms bins (Figure 5E). This analysis reveals significant location by SOA interactions between 300 and 1000 ms. This interaction highlights that the decreases in multisensory enhancement seen with longer SOAs for central locations increases for more peripheral locations. Likewise, the decrease in multisensory enhancement in the periphery relative to central locations with synchronous presentations is compounded at increased levels of asynchrony.

In summary, mean RT measures across subjects in response to multisensory targets revealed main effects of target location and asynchrony (i.e., SOA), as well as an interaction between the two. CDF analyses revealed benefits relative to race model predictions, with more central and synchronously-presented targets showing the greatest multisensory enhancements. In addition, the interaction between these factors illustrates the reciprocal interdependency between space and time in influencing a subject's judgment of the location of a target stimulus.

Rates of simultaneity judgment

Following participants' judgments concerning the perceived location of an audiovisual stimulus pair, they were then asked to indicate if the stimuli were synchronous (i.e., simultaneous) or asynchronous. Rates of perceived synchrony were calculated for each individual for each condition (Figure 6A), and the data were subjected to a repeated-measures ANOVA. Not surprisingly, a significant main effect of SOA was found ($p < 1.00 \times 10^{-10}$, $F_{(3,39)} = 52.09$), with rates of perceived synchrony decreasing as SOA increased (Figure 6A). In addition, a significant main effect of spatial location was found ($p < 2.64 \times 10^{-4}$, $F_{(3,39)} = 6.87$), with higher rates of perceived synchrony being seen for peripheral when compared with central presentations. A significant two-way interaction between SOA and spatial location was also found ($p < 3.69 \times 10^{-4}$, $F_{(9,39)} = 3.69$). To visualize this interaction in the same manner as the raw curves, normalized rates of perceived synchrony were calculated for responses to stimulus presentations at each location (Figure 6B). To explore this interaction statistically, changes from the rate of perceived synchrony from synchronous presentations (0 ms SOA) were calculated for each individual and averaged for each condition (Figure 6C). Note that as target SOA increased, the influence of spatial location on synchrony judgments was amplified. Most notably, a highly asynchronous (i.e., 200 ms) target presented at 90° was significantly more likely to be judged as synchronous when compared with a 200 ms SOA target at 0°. Within the significant main effect of location, pair-wise comparisons were made and reported in Table 5.

In summary, these synchrony judgments reveal that target presentations in the periphery are more likely to be judged as synchronous when compared with central targets, and that this difference increases as the temporal offset between the auditory and visual stimulus increases.

Detection rates

The number of trials in which participants accurately detected the modality of stimulus presentation was assessed in order to ensure that results at the peripheral locations were not

reflecting stimulus judgments where detection rates were at chance levels. No detection rates were found to approach chance levels (Figure 7). Also, the finding of near-perfect modality detection rates at central locations with audiovisual stimuli presentations suggest that there was little, if any, Colavita effect, where the participants would misperceive the audiovisual presentation as visual-only (Colavita 1974).

Discussion

While it has been well established that the behavioral and perceptual enhancements seen under multisensory circumstances are modulated by stimulus-level factors such as space, time, and effectiveness, the current study adds to the growing set of observations illustrating that these factors are complexly interwoven with one another. In this study we provide novel insights into: (a) how the localization of spatially-concordant multisensory (i.e., auditory-visual) targets is dependent upon stimulus location, (b) how this localization is also dependent on the temporal structure of the paired stimuli, (c) how space and time interact in producing the pattern of localization responses, and (d) how judgments of perceived synchrony are similarly dependent upon these factors.

Of primary interest in the current study were the interactions between space and time in generating the final behavioral and perceptual responses. The current study is the first to detail such interactions, which are seen for both the localization of multisensory targets and judgments concerning the synchrony of these targets. For the localization responses, RTs slowed for more peripheral targets and for more asynchronous presentations. This increase in RTs was compounded when targets were both in the periphery and highly asynchronous.

Spatial Localization – Effects of Spatial Manipulations

The first task the participants performed was to localize a visual, auditory or combined visual-auditory (i.e., multisensory) target that varied in spatial location. While much research has been devoted to detailing changes in multisensory integration based on the spatial relationship of paired auditory and visual stimuli, little psychophysical work has looked at how stimulus location influences multisensory integration. The results under visual-alone and auditory-alone conditions confirmed the results of prior studies of unisensory localization (for respective reviews, see Blauert 1997; Matin 1986), in which accuracy for visual localization decreases parametrically from central to peripheral locations, whereas accuracy in localizing auditory targets is generally poorest for intermediate locations (Nordlund 1962). Under multisensory conditions (and where the auditory and visual stimuli were always presented in a spatially-concordant manner at the same location), accuracy was better at all target locations, but significant enhancements were seen at only the two intermediate locations (i.e., 30° and 60°). This finding is significant in the context that these two locations showed the poorest overall performance under unisensory conditions, a result congruent with the principle of inverse effectiveness, and which has been demonstrated at levels of analysis ranging from single-unit neurophysiological recordings to fMRI to behavior (Hairston et al. 2003; Meredith and Stein 1986b; Stevenson et al. 2009). The unique piece in the current study is that changes in effectiveness are usually accomplished via changes in stimulus salience, such as sound intensity and visual contrast. This is the first study in humans, to our knowledge, that has shown that changes in stimulus location result in changes in unisensory effectiveness that ultimately impact multisensory enhancement, reinforcing the interconnectedness of the spatial and inverse effectiveness principles. This finding is distinct from the typical finding of inverse effectiveness where stimulus effectiveness is modulated through stimulus salience, and suggests that inverse effectiveness may drive patterns of multisensory enhancement seen with other stimulus manipulations such as stimulus asynchrony or task difficulty (Hecht et al. 2008; Kim et al. In Press).

RTs during the spatial localization task showed slower responses (i.e. less effective responses) in the periphery for visual presentations, and similar to response accuracies, showed the slowest RTs at the 60° location. In such RT analyses, multisensory enhancements are best determined using the race model (Colonius and Diederich 2006; Miller 1982; Raab 1962), which provides a prediction as to the enhancement that would be seen if the visual and auditory processes were independent. In the current study, multisensory enhancement of RTs was greatest at central locations and decreased with more peripheral presentations, presentations that elicited slower responses on average. In contrast to the findings with localization accuracy, this result runs contrary to the principle of inverse effectiveness. Indeed, although inverse effectiveness has been shown in numerous studies, there are a small number of cases which failed to show support for this principle (Diederich and Colonius 2008; Kim and James 2010; Ross et al. 2007a; Ross et al. 2007b). This pattern of results strongly argues for differences in the underlying neural operations that support improvements in accuracy versus those supporting speeded RTs. Such a result is interesting in the context of prior work that has shown inverse effectiveness in both the accuracy and RT domains, but has never examined the relationship between these factors within the same task design. One possibility is that the current paradigm weights accuracy over speed of response, allowing enhancements in accuracy to follow inverse effectiveness but to be masked for RTs. Future work will strive to disentangle these seemingly discrepant findings.

Spatial Localization – Effects of Temporal Manipulations

In addition to parametric variations in target location, the temporal relationships of the paired visual-auditory stimuli were also varied. While the accuracy of spatial localization was greatly impacted by the location of the stimuli, it was not significantly impacted by the temporal synchrony of the presentations. However, individuals' RTs did show main effects of SOA and spatial location, with faster responses in more central locations and more synchronous presentations. Once again, these RT data showed an interaction between spatial location and SOA, with the location effects being amplified at longer SOAs. These findings reinforce that not only do these two stimulus factors modulate audiovisual integration, but that these effects strongly interact with one another. While the independent effect of temporal synchrony on multisensory interactions has been previously demonstrated (Dixon and Spitz 1980), these data are the first to show a novel interaction between SOA and target location in human psychophysical performance.

Simultaneity Judgment

The second task, which took place following the localization judgment, involved participants reporting if they perceived the audiovisual presentation as simultaneous. The effect of SOA on simultaneity judgments has been characterized in a number of prior studies (Dixon and Spitz 1980; Foss-Feig et al. 2010; Hillock et al. 2011; Miller and D'Esposito 2005; Powers et al. 2009; Stevenson et al. 2010; van Atteveldt et al. 2007; van Wassenhove et al. 2007; Vroomen and Keetels 2010; Zampini et al. 2005a). In general, this work has illustrated that the greater the SOA, the lower the probability that the stimulus pair will be perceived as synchronous. In the current study, a previously unknown interaction between spatial location and the perception of synchrony is reported. A simple synopsis of these results is that audiovisual stimuli presented at longer SOAs are more likely to be perceived as synchronous when they are in the periphery. It should also be noted that a second possibility is that the visual stimuli were detected at a slower rate in the periphery (and thus inducing the change in RT across locations with the spatial localization task seen in Figure 4, blue bars). While these current data were collected using a spatial localization task and thus cannot be used to test this hypothesis, this possibility should be explored in future experiments.

These data, in line with previous work, show that the probability of perceived simultaneity peaks when the auditory stimulus slightly lags the visual stimulus and gradually decreases as SOA increases, creating a temporal window within which auditory and visual inputs are likely to be perceptually bound (Dixon and Spitz 1980; Meredith et al. 1987; Pöppel et al. 1990; Roach et al. 2011; Seitz et al. 2006; Zampini et al. 2005a; Zampini et al. 2003, 2005b). These results show a novel finding that this TBW becomes significantly wider when stimuli are in the periphery. These results have very interesting implications for the multisensory TBW, in suggesting that this window may be wider for stimuli located in the peripheral field. Perceptually, this effect would result in a greater rate of multisensory binding for peripheral stimuli in ecological environments. Such a hypothesis would predict that measures of perceptual binding or fusion, such as the McGurk effect (McGurk and MacDonald 1976) or the sound-induced flash illusion (Shams et al. 2000) would occur more frequently in the periphery even with temporally disparate presentations. Neural support for this hypothesis has been seen as well, with greater multisensory enhancement observed at the edge of a neurons receptive field (Carriere et al. 2008; Royal et al. 2009).

Conclusions

The ability of stimulus factors such as space, time and effectiveness to modulate multisensory integration has been well established in prior work. However, only recently has attention been brought to bear on the interactions between these factors. Given that the field of multisensory integration is founded on the concept that individual sensory systems do not function in isolation, it is a logical next step to begin to acknowledge that factors that modulate multisensory integration likewise do not occur in isolation, and to explore the interactions between these factors. Here, we provide novel evidence showing that the spatial location and temporal structure of paired audiovisual stimuli strongly interact in order to influence behavioral judgments and perceptual processes. In particular, our findings that (a) rates of perceived simultaneity are dependent upon target location and (b) the level of temporal synchrony interacts with effects of target location on the ability of an individual to spatially localize an external event provided clear evidence of these interactions. That these interactions are seen using simple, low-level stimuli suggests that these interactions are likely present at all stimulus levels, and that these effects should be addressed in future research. Indeed, such interactions may become increasingly evident (and powerful) with the use of more complex and naturalistic stimuli, since these stimuli more faithfully reflect the statistical, semantic and contextual realities of real world events.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was funded in part through a grant from NIDCD awarded to Mark Wallace and Stephen Camarata, NIH # R34 DC010927, as well as an NIDCD grant awarded to Ryan Stevenson, NIH 1F32 DC011993. We would also like to acknowledge the support of the Vanderbilt Kennedy Center and the Vanderbilt Brain Institute.

References

- Blauert, J. Spatial hearing: the psychophysics of human sound localization. MIT Press; 1997.
- Carriere BN, Royal DW, Wallace MT. Spatial heterogeneity of cortical receptive fields and its impact on multisensory interactions. *Journal of neurophysiology*. 2008; 99(5):2357–2368.01386.200710.1152/jn.01386.2007 [PubMed: 18287544]
- Colavita F. Human sensory dominance. *Attention, Perception, & Psychophysics*. 1974; 16(2):409–412.10.3758/bf03203962

- Colonus H, Diederich A. The race model inequality: interpreting a geometric measure of the amount of violation. *Psychological review*. 2006; 113(1):148–154.2006-01885-00710.1037/0033-295X.113.1.148 [PubMed: 16478305]
- Conrey B, Pisoni DB. Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *The Journal of the Acoustical Society of America*. 2006; 119(6):4065–4073. [PubMed: 16838548]
- Conrey, BL.; Pisoni, DB. *Research on Spoken Language Processing*. Vol. 26. Indiana University; Bloomington: 2004. *Detection of Auditory-Visual Asynchrony in Speech and Nonspeech Signals*.
- Corey DP, Hudspeth AJ. Response latency of vertebrate hair cells. *Biophys J*. 1979; 26(3):499–506.S0006-3495(79)85267-4.10.1016/S0006-3495(79)85267-4 [PubMed: 318064]
- Diederich A, Colonius H. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & psychophysics*. 2004; 66(8):1388–1404. [PubMed: 15813202]
- Diederich A, Colonius H. When a high-intensity “distractor” is better than a low-intensity one: modeling the effect of an auditory or tactile nontarget stimulus on visual saccadic reaction time. *Brain research*. 2008; 1242:219–230.S0006-8993(08)01319-X.10.1016/j.brainres.2008.05.081 [PubMed: 18573240]
- Dixon NF, Spitz L. The detection of auditory visual desynchrony. *Perception*. 1980; 9(6):719–721. [PubMed: 7220244]
- Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT. An extended multisensory temporal binding window in autism spectrum disorders. *Experimental brain research Experimentelle Hirnforschung*. 2010; 203(2):381–389.10.1007/s00221-010-2240-4
- Hairston WD, Laurienti PJ, Mishra G, Burdette JH, Wallace MT. Multisensory enhancement of localization under conditions of induced myopia. *Experimental brain research Experimentelle Hirnforschung*. 2003; 152(3):404–408.
- Hecht D, Reiner M, Karni A. Multisensory enhancement: gains in choice and in simple response times. *Experimental brain research Experimentelle Hirnforschung*. 2008; 189(2):133–143.10.1007/s00221-008-1410-0
- Hershenson M. Reaction time as a measure of intersensory facilitation. *J Exp Psychol*. 1962; 63:289–293. [PubMed: 13906889]
- Hillock AR, Powers AR, Wallace MT. Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia*. 2011; 49(3):461–467.S0028-3932(10)00521-X.10.1016/j.neuropsychologia.2010.11.041 [PubMed: 21134385]
- Hirsh IJ, Sherrick CE Jr. Perceived order in different sense modalities. *J Exp Psychol*. 1961; 62:423–432. [PubMed: 13907740]
- Keetels M, Vroomen J. The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental brain research Experimentelle Hirnforschung*. 2005; 167(4):635–640.10.1007/s00221-005-0067-1
- Kim S, James TW. Enhanced effectiveness in visuo-haptic object-selective brain regions with increasing stimulus salience. *Hum Brain Mapp*. 2010; 31(5):678–693.10.1002/hbm.20897 [PubMed: 19830683]
- Kim S, Stevenson RA, James TW. Visuo-haptic neuronal convergence demonstrated with an inversely effective pattern of BOLD activation. *Journal of cognitive neuroscience*. In Press. 10.1162/jocn_a_00176
- King AJ, Palmer AR. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental brain research Experimentelle Hirnforschung*. 1985; 60(3):492–500.
- Krueger J, Royal DW, Fister MC, Wallace MT. Spatial receptive field organization of multisensory neurons and its impact on multisensory interactions. *Hearing research*. 2009; 258(1-2):47–54.S0378-5955(09)00197-X.10.1016/j.heares.2009.08.003 [PubMed: 19698773]
- Lamb TD, Pugh EN Jr. A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *The Journal of physiology*. 1992; 449:719–758. [PubMed: 1326052]
- Lennie P. The physiological basis of variations in visual latency. *Vision research*. 1981; 21(6):815–824.0042-6989(81)90180-2 [PubMed: 7314459]

- Lewald J, Guski R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Brain Res Cogn Brain Res*. 2003; 16(3):468–478.S0926641003000740 [PubMed: 12706226]
- Lewald J, Guski R. Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neuroscience letters*. 2004; 357(2):119–122.10.1016/j.neulet.2003.12.045S0304394003014691 [PubMed: 15036589]
- Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res*. 2003; 17(2):447–453.S0926641003001605 [PubMed: 12880914]
- Macaluso E, George N, Dolan R, Spence C, Driver J. Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage*. 2004; 21(2):725–732. [PubMed: 14980575]
- Martin, L. Handbook of perception and human performance. Vol. 1. Wiley-Interscience; New York: 1986. Visual localization and eye movements; p. 20.21-20.45.
- McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature*. 1976; 264(5588):746–748. [PubMed: 1012311]
- Meredith MA, Nemitz JW, Stein BE. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci*. 1987; 7(10):3215–3229. [PubMed: 3668625]
- Meredith MA, Stein BE. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain research*. 1986a; 365(2):350–354. [PubMed: 3947999]
- Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of neurophysiology*. 1986b; 56(3):640–662. [PubMed: 3537225]
- Meredith MA, Wallace MT, Stein BE. Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: multisensory properties of the tecto-reticulo-spinal projection. *Experimental brain research Experimentelle Hirnforschung*. 1992; 88(1):181–186.
- Miller J. Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*. 1982; 14(2):247–279. [PubMed: 7083803]
- Miller LM, D'Esposito M. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J Neurosci*. 2005; 25(25):5884–5893. [PubMed: 15976077]
- Nelson WT, Hettinger LJ, Cunningham JA, Brickman BJ, Haas MW, McKinley RL. Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Hum Factors*. 1998; 40(3):452–460. [PubMed: 9849103]
- Nordlund B. Physical factors in angular localization. *Acta Otolaryngol*. 1962; 54:75–93. [PubMed: 14480310]
- Pöppel E, Schill K, von Steinbüchel N. Sensory integration within temporally neutral systems states: A hypothesis. *Naturwissenschaften*. 1990; 77:89–91. [PubMed: 2314478]
- Powers AR 3rd, Hillock AR, Wallace MT. Perceptual training narrows the temporal window of multisensory binding. *J Neurosci*. 2009; 29(39):12265–12274.29/39/1226510.1523/JNEUROSCI.3501-09.2009 [PubMed: 19793985]
- Raab DH. Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*. 1962; 24:574–590. [PubMed: 14489538]
- Roach NW, Heron J, Whitaker D, McGraw PV. Asynchrony adaption reveals neural population code for audio-visual timing. *Proceedings of the Royal Society*. 2011; 278:9.
- Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb Cortex*. 2007a; 17(5):1147–1153.bhl02410.1093/cercor/bhl024 [PubMed: 16785256]
- Ross LA, Saint-Amour D, Leavitt VM, Molholm S, Javitt DC, Foxe JJ. Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia research*. 2007b; 97(1-3):173–183.S0920-9964(07)00365-910.1016/j.schres.2007.08.008 [PubMed: 17928202]
- Royal DW, Carriere BN, Wallace MT. Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Experimental brain research Experimentelle Hirnforschung*. 2009; 198(2-3):127–136.10.1007/s00221-009-1772-y

- Schall S, Quigley C, Onat S, Konig P. Visual stimulus locking of EEG is modulated by temporal congruency of auditory stimuli. *Experimental brain research Experimentelle Hirnforschung*. 2009; 198(2-3):137–151.10.1007/s00221-009-1867-5
- Seitz AR, JEN Sr, Holloway SR, Watanabe T. Perceptual Learning of Motion Leads to Faster Flicker Perception. *PLoS ONE*. 2006; 1(1)10.1371/journal.pone.0000028
- Senkowski D, Talsma D, Grigutsch M, Herrmann CS, Woldorff MG. Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*. 2007; 45(3):561–571. [PubMed: 16542688]
- Shams L, Kamitani Y, Shimojo S. Illusions. What you see is what you hear. *Nature*. 2000; 408(6814): 788. [PubMed: 11130706]
- Stein BE, Huneycutt WS, Meredith MA. Neurons and behavior: the same rules of multisensory integration apply. *Brain research*. 1988; 448(2):355–358. [PubMed: 3378157]
- Stein BE, Wallace MT. Comparisons of cross-modality integration in midbrain and cortex. *Progress in brain research*. 1996; 112:289–299. [PubMed: 8979836]
- Stevenson RA, Altieri NA, Kim S, Pisoni DB, James TW. Neural processing of asynchronous audiovisual speech perception. *NeuroImage*. 2010; 49(4):3308–3318.S1053-8119(09)01288-910.1016/j.neuroimage.2009.12.001 [PubMed: 20004723]
- Stevenson RA, Geoghegan ML, James TW. Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Experimental Brain Research*. 2007; 179(1):85–95.
- Stevenson RA, James TW. Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage*. 2009; 44(3):1210–1223.S1053-8119(08)00992-010.1016/j.neuroimage.2008.09.034 [PubMed: 18973818]
- Stevenson RA, Kim S, James TW. An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Experimental brain research Experimentelle Hirnforschung*. 2009; 198(2-3): 183–194.10.1007/s00221-009-1783-8
- Stevenson RA, VanDerKlok RM, Pisoni DB, James TW. Discrete neural substrates underlie complementary audiovisual speech integration processes. *NeuroImage*. 2011; 55(3):1339–1345.S1053-8119(10)01664-210.1016/j.neuroimage.2010.12.063 [PubMed: 21195198]
- Sugita Y, Suzuki Y. Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*. 2003; 421(6926):911.10.1038/421911a 421911a [PubMed: 12606990]
- Talsma D, Senkowski D, Woldorff MG. Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. *Experimental brain research Experimentelle Hirnforschung*. 2009; 198(2-3):313–328.10.1007/s00221-009-1858-6
- van Atteveldt NM, Formisano E, Blomert L, Goebel R. The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex*. 2007; 17(4):962–974.bhl00710.1093/cercor/bhl007 [PubMed: 16751298]
- van Wassenhove V, Grant KW, Poeppel D. Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*. 2007; 45(3):598–607.S0028-3932(06)00011-X10.1016/j.neuropsychologia.2006.01.001 [PubMed: 16530232]
- Vatakis A, Spence C. Audiovisual synchrony perception for music, speech, and object actions. *Brain research*. 2006; 1111(1):134–142.S0006-8993(06)01590-310.1016/j.brainres.2006.05.078 [PubMed: 16876772]
- Vroomen J, Keetels M. Perception of intersensory synchrony: a tutorial review. *Atten Percept Psychophys*. 2010; 72(4):871–884.72/4/87110.3758/APP.72.4.871 [PubMed: 20436185]
- Wallace, MH.; Murray, MM., editors. *Frontiers in the Neural Basis of Multisensory Processes*. Taylor & Francis; London: 2011.
- Wallace MT, Meredith MA, Stein BE. Integration of multiple sensory modalities in cat cortex. *Experimental brain research Experimentelle Hirnforschung*. 1992; 91(3):484–488.
- Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. Unifying multisensory signals across time and space. *Experimental brain research Experimentelle Hirnforschung*. 2004; 158(2):252–258.10.1007/s00221-004-1899-9

- Wilkinson LK, Meredith MA, Stein BE. The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental brain research Experimentelle Hirnforschung*. 1996; 112(1):1–10.
- Zampini M, Guest S, Shore DI, Spence C. Audio-visual simultaneity judgments. *Perception & psychophysics*. 2005a; 67(3):531–544. [PubMed: 16119399]
- Zampini M, Shore DI, Spence C. Audiovisual temporal order judgments. *Experimental brain research Experimentelle Hirnforschung*. 2003; 152(2):198–210.10.1007/s00221-003-1536-z
- Zampini M, Shore DI, Spence C. Audiovisual prior entry. *Neuroscience letters*. 2005b; 381(3):217–222.S0304-3940(05)00157-610.1016/j.neulet.2005.01.085 [PubMed: 15896473]

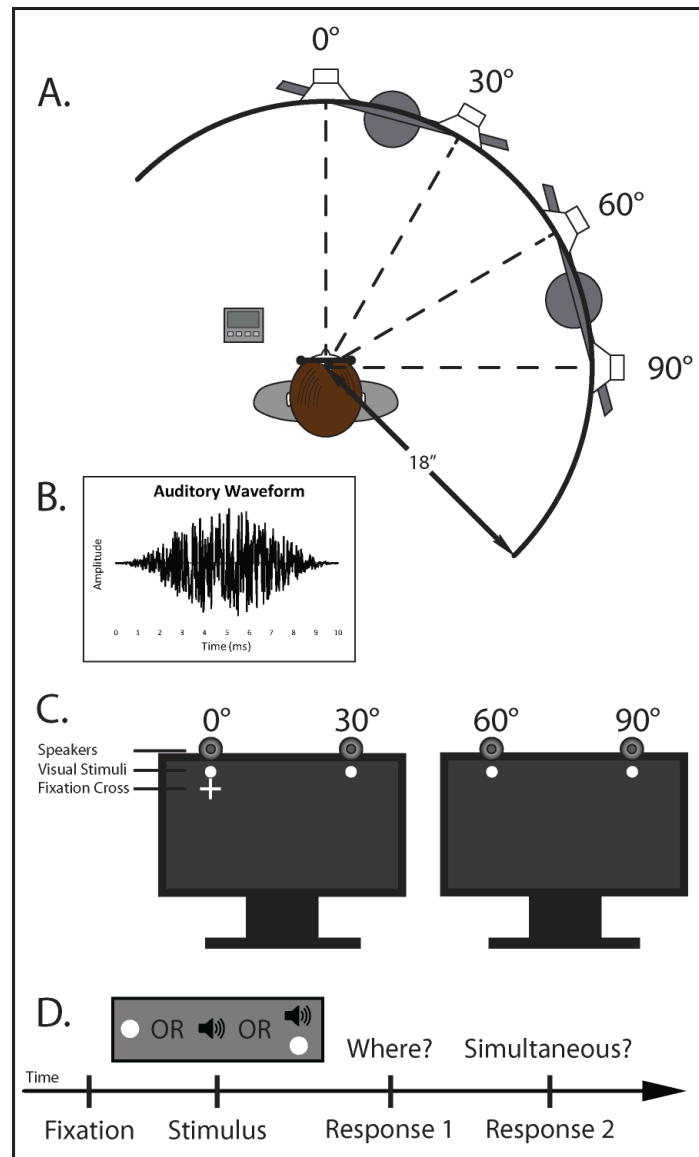


Figure 1. Stimulus Materials

Auditory and stimuli presentation at multiple angles and a constant distance from the participant was achieved via a pair of monitors with mounted speakers.

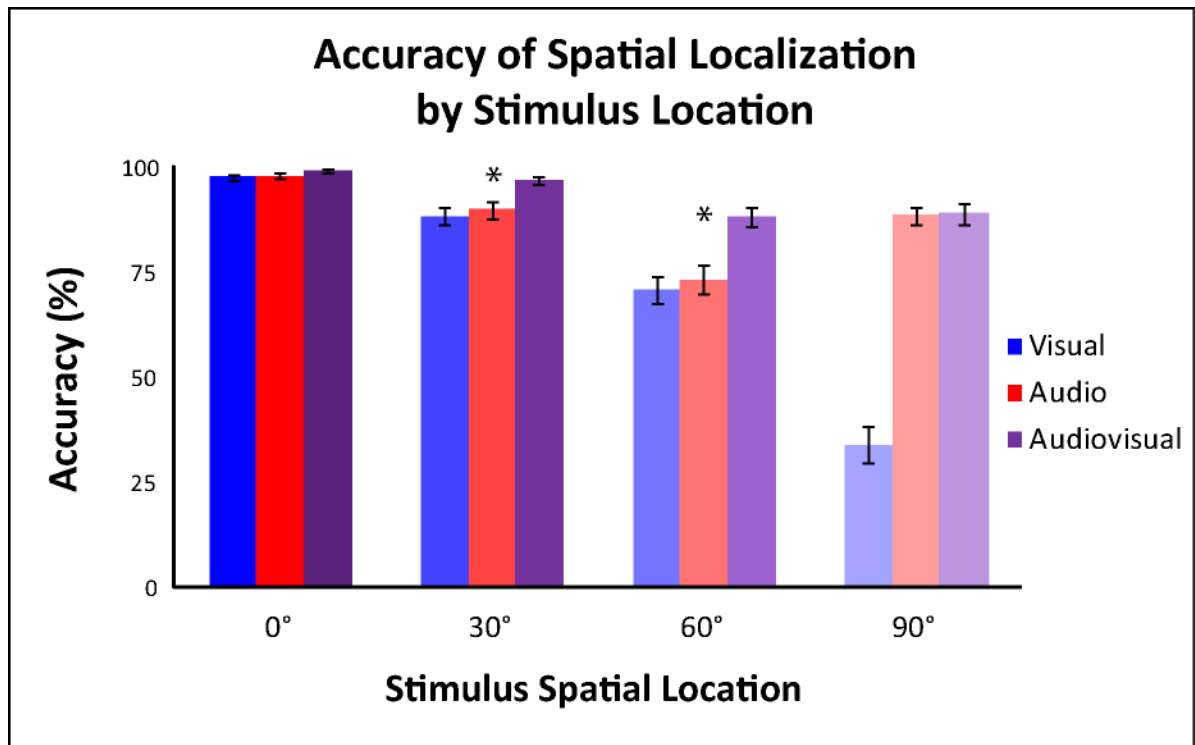


Figure 2. Visual, auditory and combined visual-auditory accuracy data for the spatial localization task

Accuracies associated with visual-only, auditory-only, and the multisensory conditions show a pattern of performance decline in the periphery. Multisensory enhancement was only significant for the 30° and 60° locations, where unisensory accuracies are weaker.

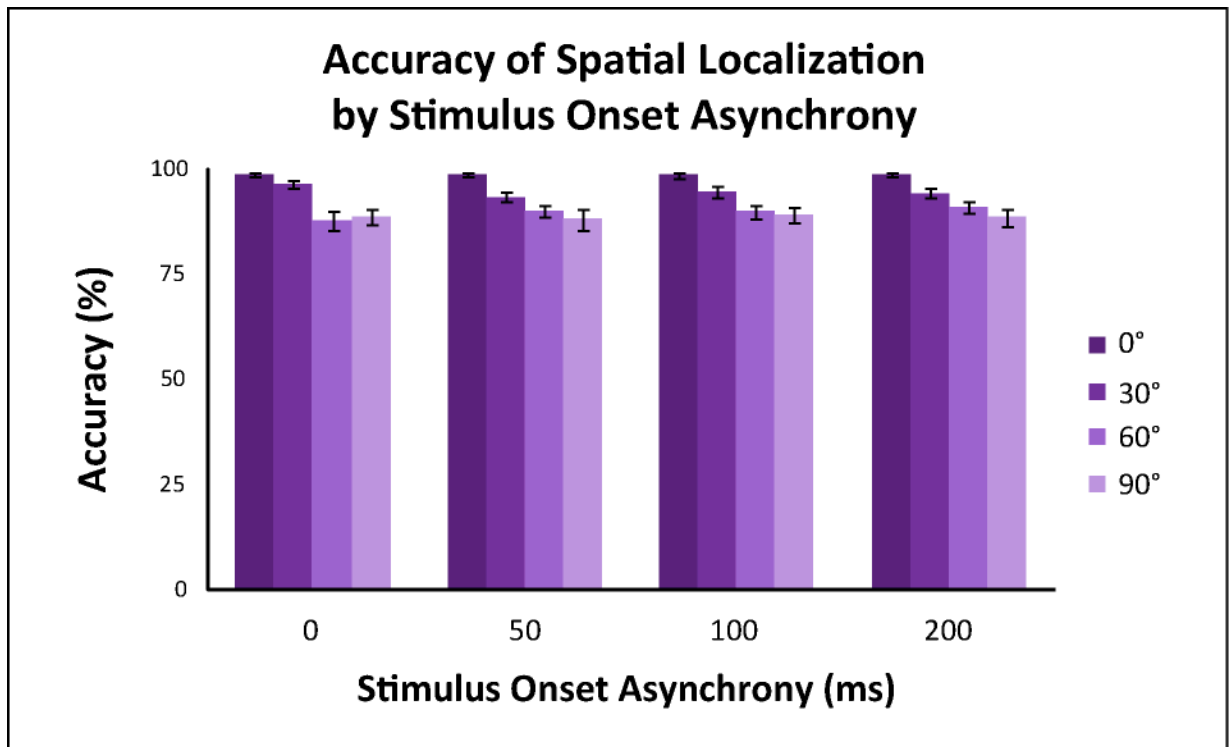


Figure 3. Combined visual-auditory accuracy data for the spatial localization task as a function of stimulus onset asynchrony

Accuracies associated with multisensory presentations at each stimulus onset asynchrony showed performance declines in the periphery.

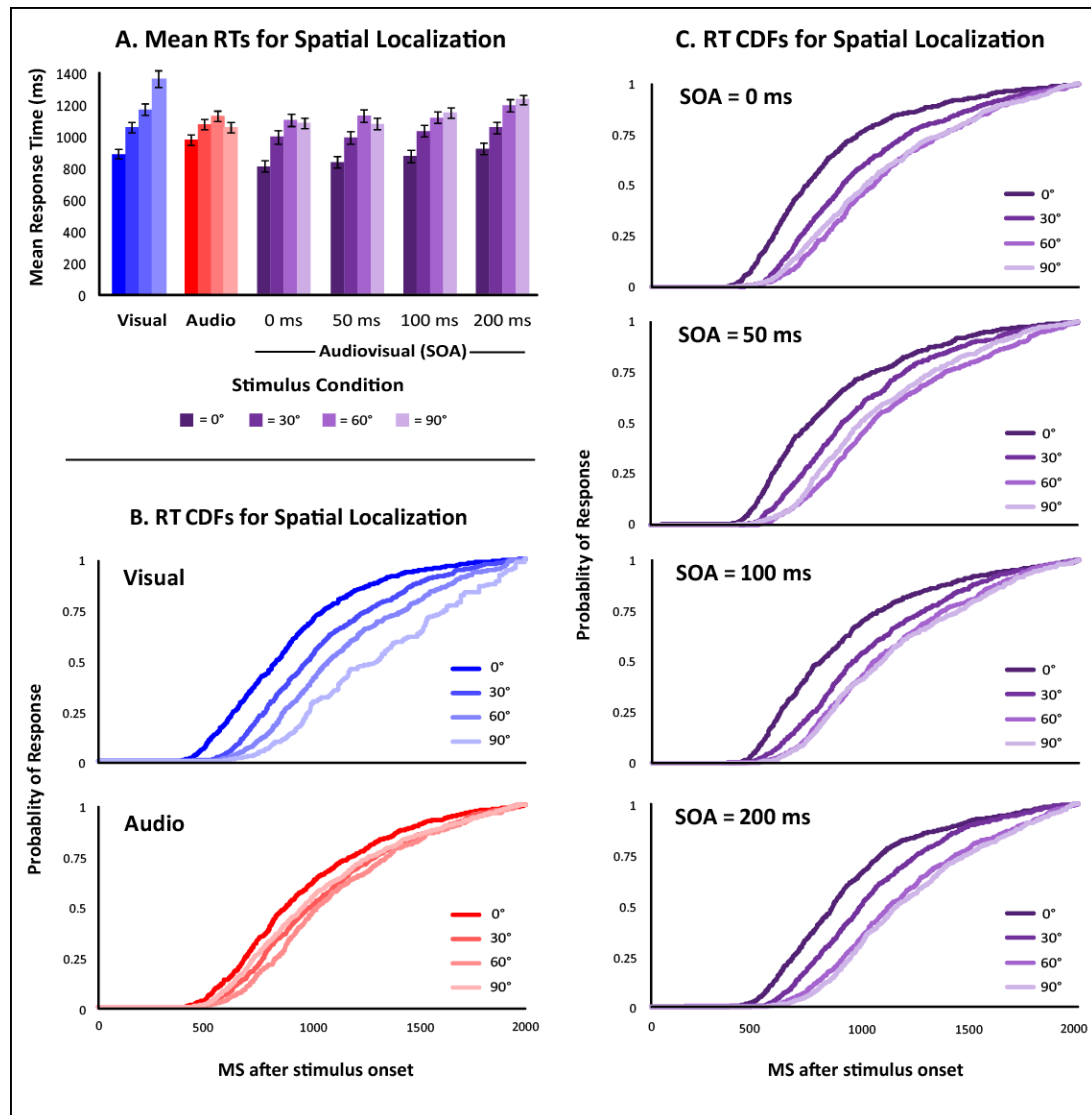


Figure 4. Response times for the localization task as a function of SOA and spatial location
 A. Mean response times for visual only (blue), auditory only (red), and multisensory (purple) presentations, with multisensory presentations clustered by stimulus onset asynchrony. Cumulative distribution functions for unisensory presentations are shown in Panel B, and for Multisensory presentations in Panel C grouped by SOA.

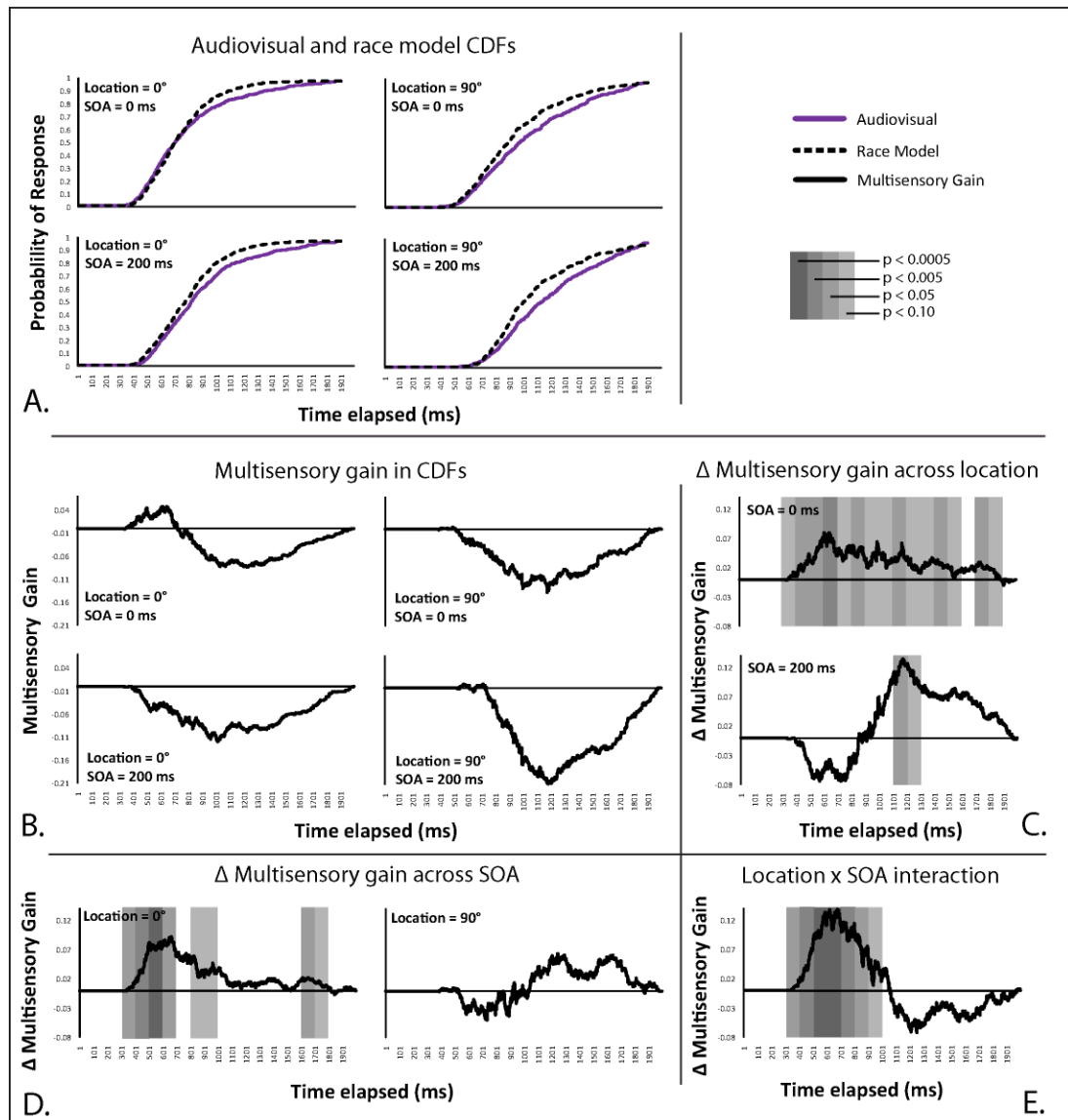


Figure 5. CDF analysis of interactions between location and SOA for the localization task
 A. Multisensory CDFs (purple) are compared to race model predictions (dashed), and are only shown for four representative conditions. B. Differences between multisensory CDFs and their respective race models are shown for these four conditions. C. and D. Changes in multisensory enhancement are shown across stimulus location (C, the differences from left to right in Panel B) and across SOA (D, the differences from top to bottom in Panel B). E. Plot of the interaction between stimulus location and SOA.

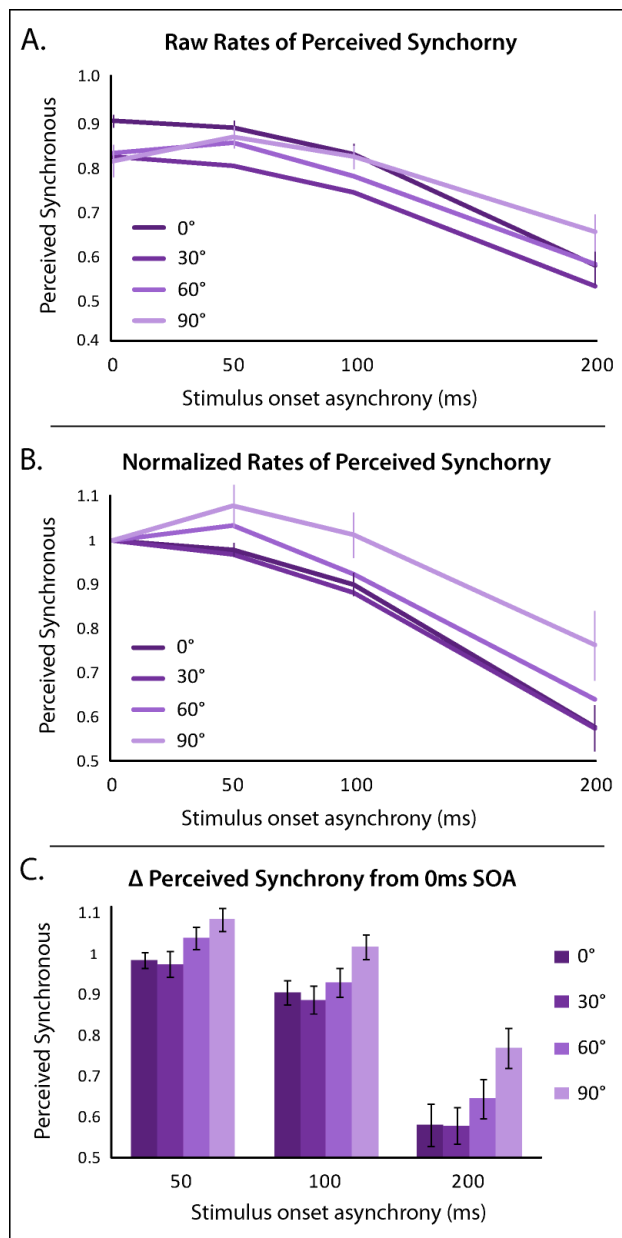


Figure 6. Perceived synchrony responses

A. Raw rates of perceived synchrony for each of the four target locations are plotted as a function of SOA. B. Normalized rates of perceived synchrony for each of the four target locations are plotted as a function of SOA. The normalized values are relative to the 0 ms SOA (synchronous presentation). C. shows the interaction effect where the change in perceived synchrony is greatest at peripheral locations.

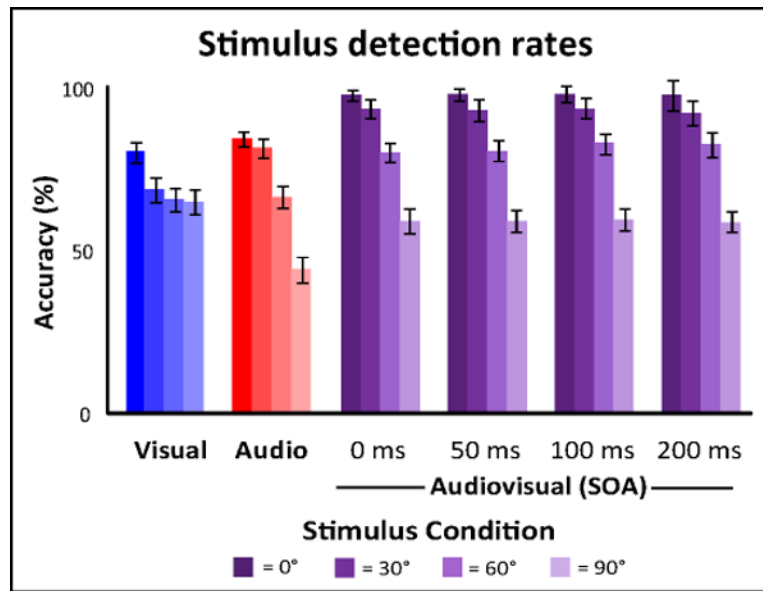


Figure 7. Stimulus detection rates

Responses with the synchrony judgment were used to assess the detection rates for each stimulus condition.

Table 1

Mean accuracy and response time values

Modality	0°		30°		60°		90°	
	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.
Visual-only	97.1	0.7	87.7	1.9	69.9	3.2	33.1	4.2
Audio-only	97.3	0.7	89.3	2.0	72.5	3.5	87.8	2.2
AV - 0 ms SOA	98.7	0.5	96.2	0.8	87.5	2.2	88.4	2.5
AV - 50 ms SOA	98.4	0.5	93.3	1.5	89.8	2.0	87.8	2.3
AV - 100 ms SOA	98.4	0.7	94.4	1.5	89.6	1.9	88.9	2.2
AV - 200 ms SOA	98.4	0.5	94.2	1.0	90.8	1.7	88.4	1.9

Modality	0°		30°		60°		90°	
	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.
Visual-only	870	30	1037	35	1148	35	1341	54
Audio-only	959	33	1057	34	1109	34	1036	34
AV - 0 ms SOA	799	35	982	39	1091	37	1072	32
AV - 50 ms SOA	824	37	977	38	1116	37	1065	36
AV - 100 ms SOA	863	39	1020	37	1105	38	1134	33
AV - 200 ms SOA	910	35	1041	35	1178	38	1216	30

Table 2

Within sensory modality comparison of spatial effect on localization task

Accuracy of spatial localization by location							
Modality	Location	30°		60°		90°	
		<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value
Visual	0°	1.70e ⁻⁶	5.42	3.65e ⁻¹¹	8.43	5.22e ⁻¹⁹	14.07
	30°			1.54e ⁻⁸	6.74	4.39e ⁻¹⁷	12.57
	60°					2.74e ⁻¹³	9.84
Audio	0°	2.82e ⁻⁵	4.61	3.72e ⁻⁹	7.13	1.05e ⁻⁵	4.90
	30°			1.34e ⁻⁴	4.13	n.s.	0.47
	60°					*2.14e ⁻⁶	*5.36
Audiovisual SOA = 0 ms	0°	5.47e ⁻³	2.90	5.84e ⁻⁶	5.07	9.42e ⁻⁵	4.25
	30°			3.32e ⁻⁴	3.85	2.51e ⁻³	3.18
	60°					n.s.	0.49
Audiovisual SOA = 50 ms	0°	6.51e ⁻⁴	3.64	7.48e ⁻⁵	4.32	4.48e ⁻⁵	4.47
	30°			0.12	1.59	0.03	2.19
	60°					n.s.	0.84
Audiovisual SOA = 100 ms	0°	6.70e ⁻³	2.83	2.21e ⁻⁵	4.68	7.74e ⁻⁵	4.30
	30°			0.01	2.52	0.01	2.66
	60°					n.s.	0.14
Audiovisual SOA = 200 ms	0°	2.32e ⁻⁴	3.97	3.12e ⁻⁵	4.58	1.75e ⁻⁶	5.41
	30°			0.09	1.70	8.30e ⁻³	2.75
	60°					n.s.	1.38
Mean response time of spatial localization by location							
Modality	Location	30°		60°		90°	
		<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value
Audio	0°	3.00e ⁻⁶	5.35	9.05e ⁻⁷	5.70	3.77e ⁻³	3.06

Accuracy of spatial localization by location							
Modality	Location	30°		60°		90°	
		<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value
Visual	30°			0.02	2.39	n.s.	0.77
	60°					8.45e ⁻³	2.76
	0°	6.22e ⁻⁸	6.50	1.80e ⁻¹⁴	11.20	2.86e ⁻¹¹	8.81
	30°			3.73e ⁻⁴	3.86	3.40e ⁻⁷	6.00
	60°					1.00e ⁻⁴	4.28
	0°	3.61e ⁻¹¹	8.74	2.71e ⁻¹⁵	11.86	9.64e ⁻¹³	9.88
	30°			8.26e ⁻⁶	5.05	3.30e ⁻⁴	3.68
	60°					n.s.	0.90
	0°	8.88e ⁻⁹	7.08	6.78e ⁻¹⁵	11.54	1.10e ⁻¹⁰	8.39
	30°			1.91e ⁻⁶	5.49	3.74e ⁻³	3.06
	60°					0.06	1.95
	0°	8.33e ⁻⁸	6.41	6.43e ⁻¹⁴	10.77	1.97e ⁻¹²	9.65
Audiovisual SOA = 50 ms	30°			3.06e ⁻⁵	4.65	4.77e ⁻⁵	4.51
	60°					n.s.	1.21
	0°	6.47e ⁻⁸	6.49	2.14e ⁻¹⁰	8.19	3.53e ⁻¹⁴	10.97
	30°			2.64e ⁻⁶	5.39	1.28e ⁻⁸	6.98
	60°					n.s.	1.50
	0°						
Mean response time of spatial localization by SOA							
Modality	SOA (ms)	50 ms		100 ms		200 ms	
		<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value
Audiovisual 0°	0	0.06	1.97	5.76e ⁻⁴	3.70	3.77e ⁻⁷	5.94
	50			0.03	2.25	4.07e ⁻⁵	4.55
	100					0.02	2.65
Audiovisual 30°	0	n.s.	0.30	0.02	2.46	1.31e ⁻³	3.43

\$watermark-text

\$watermark-text

\$watermark-text

Accuracy of spatial localization by location							
Modality	Location	30°		60°		90°	
		p-value	t-value	p-value	t-value	p-value	t-value
Audiovisual 60°	50			9.06e ⁻³	2.73	1.74e ⁻³	3.33
	100					n.s.	1.28
	0		n.s.	1.49	0.94	1.61e ⁻³	4.12
	50				n.s.	4.10e ⁻³	3.02
	100					9.91e ⁻⁴	3.52
Audiovisual 90°	0		n.s.	0.35	2.17e ⁻³	3.25	1.53e ⁻⁹
	50				2.78e ⁻³	3.17	3.11e ⁻⁸
	100					3.95e ⁻⁷	5.93

Table 3

Kolmogorov-Smirnov tests with spatial location

Modality	Contrast	<i>p</i>	<i>ks</i>
Audio	0 > 30°	1.97 e ⁻⁰⁵	0.1179
	30 > 60°	2.08 e ⁻⁰⁴	0.1162
	*60 < 90°	3.55 e ⁻⁰⁵	0.1294
Visual	0 > 30°	3.35 e ⁻¹⁵	0.2022
	30 > 60°	1.04 e ⁻⁰⁶	0.1442
	60 > 90°	9.10 e ⁻⁰⁴	0.1409
Audiovisual SOA 0 ms	0 > 30°	2.31 e ⁻²²	0.2415
	30 > 60°	5.69 e ⁻¹²	0.1845
	60 > 90°	n.s.	0.0346
Audiovisual SOA 50 ms	0 > 30°	9.57 e ⁻²⁰	0.2270
	30 > 60°	4.17 e ⁻¹¹	0.1757
	60 > 90°	n.s.	0.0099
Audiovisual SOA 100 ms	0 > 30°	5.31 e ⁻²¹	0.2353
	30 > 60°	1.97 e ⁻⁰⁶	0.1310
	60 > 90°	n.s.	0.0431
Audiovisual SOA 200 ms	0 > 30°	1.23 e ⁻¹⁵	0.2026
	30 > 60°	1.50 e ⁻¹⁰	0.1721
	60 > 90°	n.s.	0.0527
Audiovisual Location 0°	0 > 50 ms	0.062	0.0559
	50 > 100 ms	0.015	0.0689
	100 > 200ms	3.05 e ⁻⁰⁴	0.0966
Audiovisual Location 30°	0 > 50 ms	n.s.	0.0339
	50 > 100 ms	0.02	0.0691
	100 > 200ms	4.70 e ⁻⁰³	0.0811
Audiovisual Location 60°	0 > 50 ms	n.s.	0.0288
	50 > 100 ms	n.s.	0.0308
	100 > 200ms	6.77 e ⁻⁰⁵	0.1154
Audiovisual Location 90°	0 > 50 ms	n.s.	0.0511
	50 > 100 ms	0.012	0.0792
	100 > 200ms	1.73 e ⁻⁰⁷	0.1502

Table 4

Kolmogorov-Smirnov tests comparing AV CDFs to Race Models

SOA	Location	<i>p</i>	<i>Ks</i>	# individuals	<i>p</i>
0 ms	0°	7.96e-5	0.0735	36	9.28e-8
	30°	n.s.	0.0335	34	4.18e-6
	60°	n.s.	0.0111	32	9.11e-5
	90°	n.s.	0.0034	34	4.18e-6
50 ms	0°	1.44e-5	0.0796	35	6.91e-7
	30°	0.002	0.0617	38	4.67e-9
	60°	n.s.	0.0209	30	1.11e-3
	90°	n.s.	0.0188	35	6.91e-7
100 ms	0°	0.003	0.0576	35	6.91e-7
	30°	0.03	0.0466	34	4.18e-6
	60°	n.s.	0.0402	35	6.91e-7
	90°	n.s.	0.0115	35	6.91e-7
200 ms	0°	n.s.	0.0008	32	9.11e-5
	30°	n.s.	0.0119	38	4.67e-9
	60°	n.s.	0.0046	31	3.40e-4
	90°	n.s.	0.0151	35	6.91e-7

Table 5
Within sensory modality comparison of spatial effect on perceived-synchrony task

Modality	Location	Δ Rate of perceived synchrony					
		30°		60°		90°	
		<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value	<i>p</i> -value	<i>t</i> -value
Audiovisual SOA = 50 ms	0°						
	30°	n.s.	0.21	0.04	2.12	0.05	2.00
	60°			0.03	2.34	0.04	2.15
Audiovisual SOA = 100 ms	0°						
	30°	n.s.	0.21	n.s.	0.70	0.06	1.92
	60°			n.s.	0.21	n.s.	1.42
Audiovisual SOA = 200 ms	0°						
	30°	n.s.	0.05	0.06	1.91	0.02	2.46
	60°			0.03	2.25	0.007	2.86
						0.04	2.19

Table 6

Kolmogorov-Smirnov tests with simultaneity judgment task

Modality	Contrast	<i>p</i>	<i>ks</i>
Audiovisual SOA 0 ms	0 > 30°	1.49 e ⁻¹⁰	0.1803
	0 > 60°	7.33 e ⁻¹⁷	0.2398
	0 > 90°	4.39 e ⁻¹⁸	0.2723
Audiovisual SOA 50 ms	0 > 30°	4.01 e ⁻⁰⁶	0.1354
	0 > 60°	7.27 e ⁻¹⁴	0.2145
	0 > 90°	4.30 e ⁻¹⁸	0.2676
Audiovisual SOA 100 ms	0 > 30°	4.11 e ⁻⁰⁸	0.1656
	0 > 60°	4.23 e ⁻¹¹	0.1985
	0 > 90°	9.94 e ⁻¹⁶	0.2587
Audiovisual SOA 200 ms	0 > 30°	5.80 e ⁻⁰³	0.1139
	0 > 60°	1.07 e ⁻⁰⁵	0.1679
	0 > 90°	2.48 e ⁻¹⁰	0.2460