Mental Imagery Changes Multisensory Perception

Christopher C. Berger¹,* and H. Henrik Ehrsson¹
¹Department of Neuroscience, Karolinska Institutet, 17177 Stockholm, Sweden

Summary

Multisensory interactions are the norm in perception, and an abundance of research on the interaction and integration of the senses has demonstrated the importance of combining sensory information from different modalities on our perception of the external world [1–9]. However, although research on mental imagery has revealed a great deal of functional and neuroanatomical overlap between imagery and perception, this line of research has primarily focused on similarities within a particular modality [10–16] and has yet to address whether imagery is capable of leading to multisensory integration. Here, we devised novel versions of classic multisensory paradigms to systematically examine whether imagery is capable of integrating with perceptual stimuli to induce multisensory illusions. We found that imagining an auditory stimulus at the moment two moving objects met promoted an illusory bounce percept, as in the classic cross-bounce illusion; an imagined visual stimulus led to the translocation of sound toward the imagined stimulus, as in the classic ventriloquist illusion; and auditory imagery of speech stimuli led to a promotion of an illusory speech percept in a modified version of the McGurk illusion. Our findings provide support for perceptually based theories of imagery and suggest that neuronal signals produced by imagined stimuli can integrate with signals generated by real stimuli of a different sensory modality to create robust multisensory percepts. These findings advance our understanding of the relationship between imagery and perception and provide new opportunities for investigating how the brain distinguishes between endogenous and exogenous sensory events.

Results

Can what one imagines hearing change what one sees? Can what one imagines seeing change what one hears? In ordinary perception, multisensory integration is typical, and research has found that the integration of sensory information across sensory modalities can improve the detection and discrimination of events in our environment [1, 8, 17, 18] or lead to distortions, as in the case of multisensory illusions [3, 6, 9]. Multisensory illusions, such as the cross-bounce [3], ventriloquism [6], and McGurk [9] illusions, are classic examples of how sensory information in one modality can change what one perceives in another. For instance, in the cross-bounce illusion [3], the presentation of a sound close to the moment when two objects collide promotes the illusory perception that the objects collide. In the ventriloquism illusion [6], variation of the spatial relationship between audiovisual stimuli causes a translocation of the auditory stimulus toward the visual stimulus. In the McGurk illusion [9], an auditory stimulus of one phoneme (e.g., “ba”) paired with a visual stimulus of someone’s lip movements articulating a competing phoneme (e.g., “ga”), leads to a fused illusory auditory percept (e.g., “da”). Here, in a novel approach to investigating imagery and making use of these three classic multisensory illusions, we examined whether mental images are capable of leading to multisensory integration.

Auditory Imagery and the Cross-bounce Illusion

In experiment 1A, we closely followed the methodology of Sekuler et al.’s [3] original demonstration of the cross-bounce illusion. In each trial, the participants (n = 22) focused on a fixation cross while two blue disks appeared from the top-right and top-left corners of the screen, moved 165.1 mm/s at a 45° angle toward the opposite corner of the screen, crossed in the middle at a fixation cross, and disappeared off the screen in their respective corners. The participants imagined hearing a sound (150 ms “clink”) 500 ms before the disks coincided, at the moment of coincidence, 500 ms after they coincided, or not at all (all experiments were approved by the Regional Ethical Review Board of Stockholm—see the Supplemental Experimental Procedures for details; see Figure S1 available online for schematic overview of the experimental procedure). Given the known effects of tactile stimulation on the cross-bounce illusion [5], there was also an imagery condition involving motor and tactile imagery (i.e., an imagined finger tap) to investigate the extent or limitation of imagery on the perception of bounce (see the Supplemental Experimental Procedures and Supplemental Results for details).

The timing of the imagined sound significantly altered the perceived motion of the circles. Specifically, imagination of the sound at the moment the circles coincided led to a significant promotion of the perception that the circles bounced compared to the view-only condition, whereas imagination of the sound before or after coincidence did not (see Figure 1A).

To ensure that the promotion of the bounce percept was not due to nonspecific effects of imagery at the moment the circles coincided, we performed a control experiment (experiment 1B; n = 12) in which participants imagined a sound at the moment the circles coincided (as before), imagined a control motor stimulus (i.e., a finger lift) at the moment the circles coincided, or passively viewed the circles. A subsequent version of the experiment, in which the participants actually heard a sound or moved their finger at the moment of coincidence, was also conducted for comparison.

Auditory imagery of a sound at the moment of coincidence significantly promoted the perception of the bouncing percept compared to the passive viewing and motor imagery control conditions. These results were consistent with the results of the subsequent perceptual version of the experiment (see Figure 1B). Taken together, these results suggest that auditory imagery is capable of leading to multisensory integration.

Visual Imagery and Ventriloquism

In experiments 2A and 2B, we sought to determine whether the perceptual effects outlined above were specific to the
cross-bounce illusion or reflected a more general principle of imagery-perception multisensory interactions. Toward this end, we conducted two separate ventriloquism experiments. In experiment 2A (n = 21), we used an adapted version of the classic ventriloquism illusion [6, 7, 19]. The participants imagined a white circle appearing in one of four locations on a wall in front of them in a darkened room while maintaining fixation (0°). Auditory stimuli were presented at the same time and location as the imagined visual stimuli, alone in the same locations, or at disparities of 15° or 30° from the imagined visual stimuli (see the Supplemental Experimental Procedures for details). To determine the extent of the translocation of auditory stimuli toward the imagined stimuli, we calculated the percentages of the bias toward the imagined visual stimulus (%VB). To assess the extent to which auditory localization precision was enhanced when the participants imagined a visual stimulus in the same location, we calculated a multisensory enhancement index (MEI; see the Supplemental Experimental Procedures for equations). After the imagery version of this experiment, a version of the experiment using real visual stimuli was conducted, and the %VBs and MEIs were calculated for comparison with the imagery version of the experiment (see Figure S2 for the mean localization errors of all conditions).

Significant biases in the perceived sound location toward imagined visual stimuli were observed for disparities of 15° and 30°. A stronger %VB was observed for disparities of 30° compared to 15°. Correspondingly, in the perceptual version of the experiment, significant visual biases were observed for audiovisual disparities of 15° and 30°, with a stronger %VB observed for audiovisual disparities of 30° compared to 15° (see Figure 2A). Furthermore, a multisensory enhancement of auditory perception (i.e., increased sound localization accuracy) was observed when the participants imagined a visual stimulus in the same location as an auditory stimulus. Similarly, a significant multisensory enhancement of auditory perception was observed when a real visual stimulus was presented in the same location as an auditory stimulus in the perceptual version of the experiment (see Figure 2B).

In experiment 2B, we made use of a psychophysical staircase procedure. The strength of this method is that it eliminates any possible influence of voluntary postperceptual decisions on the ventriloquism effect [20]. Closely following the methodology of Bertelson and Aschersleben [20], we presented auditory stimuli to the participants (n = 18) from two randomly selected staircases that began at extreme left or right positions (48°) and gradually converged as the participants made dichotomic judgments of whether the sound came from the left or right of fixation (0°). That is, for the left staircase, the location of the sound moved one step to the right (toward fixation) after the participant indicated that the sound came from the left, and one step to the left (away from fixation) after the participant indicated that the sound came from the right. The opposite pattern was followed for the right staircase. This procedure continued until the participant made eight response reversals (i.e., responses that were different than the previous response) on each staircase. In this paradigm, the presentation of a visual stimulus at fixation has been found to lead to earlier uncertainty—in the form of response reversals—about the location of sounds presented at locations further from fixation compared to when no visual stimulus is presented at fixation [20]. This earlier uncertainty is due to the translocation of the auditory stimuli toward the visual stimuli presented at fixation. Thus, to test whether an imagined visual stimulus could lead to the same translocation of auditory stimuli, we had the participants imagine seeing a white circle appear at fixation or simply maintain fixation in two separate conditions that we counterbalanced across participants. All analyses were conducted on the first eight response reversals in each condition (see the Supplemental Experimental Procedures for further details).

We found that the staircases in the imagine-circle condition converged more slowly and began further away compared to the no-circle condition (Figure 2C), suggesting that the auditory stimuli were drawn toward the imagined visual stimuli as in experiment 2A. Moreover, the average distance between the staircases was greater for the imagine-circle condition than in the no-circle condition (Figure 2D). Taken together, the results from experiments 2A and 2B suggest that visual imagery is capable of leading to audiovisual integration, providing further support for the hypothesis...
that mental imagery is capable of integrating with perceptual stimuli of a different sensory modality.

Auditory Imagery in a Modified McGurk Illusion

In experiment 3, we made use of a modified version of the McGurk illusion to determine whether the imagery-induced multisensory effects from the previous experiments could be extended to speech perception. Toward this end, we tested whether auditory imagery of one phoneme could integrate with visual speech stimuli of another phoneme to promote an illusory speech percept. We supplanted the auditory stimuli from the classic McGurk illusion paradigm—in which an illusory fused “da” percept is heard when a conflicting auditory “ba” is dubbed over a visual stimulus of a person saying “ga,” but not when dubbed with a nonconflicting auditory stimulus, such as “ka” [9]—with the participants’ auditory imagery. Thus, the participants (n = 23) imagined either “ba” or “ka” while viewing videos of a person saying “ga.” The imagine-“ka” condition served to determine whether any effect was specific to a conflicting auditory stimulus and not a nonspecific effect of imagery per se. A baseline condition in which the participants passively viewed the videos was also included to assess the directionality of any observed differences between imagery conditions. After each video, the participants indicated whether they perceived the person in the video to be saying “ga” (i.e., nonillusory percept) or “da” (i.e., illusory percept). In this way, because the participants could not report what they heard (because the visual stimuli were silent) and instead reported what they perceived the person in the video to say, the effects in this modified McGurk paradigm may reflect the effects of a “reverse McGurk effect” [21, 22] in which auditory stimuli change one’s perception of visual stimuli. For the analysis, the participants were split into perceivers and nonperceivers based on a postexperiment perceptual test for the illusion (using real auditory stimuli of “ba” dubbed over visual stimuli of a person mouthing “ga”) that made use of the classic
McGurk paradigm [9]. Thus, in the perceptual test for the illusion, participants freely reported what they heard the person in the video say. A participant was considered to be a perceiver if he or she verbally reported hearing “da” at least one of the three times the video was played.

Auditory imagery of “ba” led to an increased perception of the “da” percept for perceivers compared to nonperceivers, whereas auditory imagery of “ka” did not lead to any difference in the perception of “da” between perceivers and nonperceivers (see Figure 3). Moreover, in perceivers, auditory imagery of “ba” led to a significant increase in the perception of “da” compared to auditory imagery of “ka,” whereas no such difference was observed in nonperceivers (see the Supplemental Results for additional analyses).

These results suggest that auditory imagery of a competing phoneme while viewing an ambiguous speech percept promotes an illusory speech percept, an effect that is specific to the type of auditory stimulus imagined and depends on whether one perceives the perceptual version of the illusion. These findings provide evidence in favor of the hypothesis that auditory imagery is capable of integrating with visual speech stimuli to promote an illusory speech percept.

Discussion

In three separate paradigms, using different forms of imagery, we found consistent evidence that imagery is capable of leading to perceptual illusions indicative of multisensory integration. In experiment 1A, we found that auditory imagery is capable of leading to a promotion of the illusory perception that two moving objects bounce off one another when imagined at the moment of coincidence. Moreover, in experiment 1B, we found that imagination of a control stimulus did not promote the bouncing percept, suggesting that this effect is not merely due to nonspecific effects of imagery at the moment of coincidence. In experiments 2A and 2B, we found that imagination of visual stimuli caused a translocation of spatially disparate auditory stimuli toward the imagined visual stimuli, and in experiment 2A we found that spatially congruent imagined visual and real auditory stimuli led to an enhancement of auditory localization. In experiment 3, we tested whether the findings from the first two experiments could be extended to complex speech stimuli in a modified version of the McGurk illusion, and we found that the auditory imagery of a competing phoneme while viewing lip movements of a different phoneme led to an increase in an illusory speech percept for McGurk illusion perceivers, an effect that was absent in nonperceivers. As in experiments 1A and 1B, the results from experiment 3 suggest that the multisensory effects of imagery are specific to the types of perceptual stimuli that lead to multisensory integration. Together, these findings suggest that imagery is capable of leading to multisensory integration and that imagery-induced multisensory illusions are restricted to the same temporal, spatial, and stimulus-specific characteristics as the perceptual versions of the illusions. To the best of our knowledge, these results provide the first direct behavioral evidence of imagery-induced multisensory illusions.

In light of previous research on the similarities between imagery and perception [13, 18, 23–26], the present findings suggest that the same neural mechanisms involved in multisensory integration of real stimuli are involved in integrating imagined stimuli with real stimuli. Research in neuroscience has linked the neuronal basis for multisensory interactions to specific areas in the frontal, parietal, and temporal association cortices, as well as to subcortical structures such as the superior colliculus and putamen [1, 27, 28]. These multisensory areas are anatomical zones of convergence for visual, tactile, and auditory signals and contain neurons that individually integrate multisensory signals [1]. Moreover, neuroimaging experiments have previously linked the cross-bounce [29], ventriloquism [7], and McGurk [30] illusions to activity in these multisensory areas, including the superior colliculus, posterior parietal cortex, insula, thalamus, and superior temporal sulcus [7, 29, 30]. However, although research on visual [10, 14, 15, 31], motor [16, 32], tactile [33], and auditory [26, 34] imagery has found that the processing of real and imagined stimuli share similar neural mechanisms, this line of research has yet to directly explore whether the same neuronal mechanisms involved in multisensory integration of perceptual stimuli can be activated by imagined stimuli. Evidence in favor of the influence of imagery on multisensory perception has come from neuroimaging experiments on haptic shape perception, which have found overlapping activation [24] and network connectivity [25] of brain areas involved in imagery of objects and haptic shape perception of familiar, but not unfamiliar, objects. This finding suggests that imagery may be functionally involved in certain multisensory percepts. However, our findings provide the first testable paradigms with which to directly investigate whether the neuronal signals generated by imagined stimuli are capable of integrating with those from perceptual stimuli of a different sensory modality.

Our results also corroborate the interpretation of a previous study that found indirect behavioral evidence of a possible crossmodal interaction between real and imagined stimuli [35]. Mast et al. [35] found that participants experienced a shift in the orientation of the visual horizon in the same direction as an imagined visual stimulus of rotating dots. The authors concluded that this finding is consistent with the literature.
on visual-vestibular multisensory interactions, which has demonstrated that such interactions lead to illusory vestibular motion. Because both the imagined stimulus and the reported perception are in the same modality, it is difficult to determine whether these results are an effect of an interaction between the imagined stimuli and the vestibular system or merely an effect of imagery per se. In our experiments, however, we directly addressed whether imagery in one modality directly affects perception in a different modality in three different classic multisensory illusions, and we found consistent evidence for multisensory integration. Furthermore, whereas previous studies have found that imagined stimuli can interact with perceptual stimuli within the same modality [23, 35, 36] or lead to top-down effects on perception [15, 37], our results demonstrate that imagery can interact with a perceptual stimulus from a different sensory modality to have a direct functional impact on perception.

Together, our results provide strong support for perceptually based theories of imagery [10, 12] and represent an unparalleled example of how imagination can change perception [11, 13, 38]. This unique approach to investigating the functional impact of imagery on multisensory perception broaches an exciting new paradigm to investigate how the brain distinguishes between internally and externally generated sensory signals [39]. Future research may determine the neural substrates of imagery-induced multisensory perception as well as the neural mechanisms behind one’s ability to distinguish between endogenous and exogenous sensory events.

Supplemental Information

Supplemental Information includes Supplemental Results, Supplemental Experimental Procedures, and two figures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.06.012.

Acknowledgments

We thank Ezequiel Morsella for valuable discussions about the McGurk illusion. This study was funded by the European Research Council, the Swedish Foundation for Strategic Research, the James S. McDonnell Foundation, the Swedish Research Council, and Söderbergska Stiftelsen.

Received: January 15, 2013
Revised: April 30, 2013
Accepted: June 5, 2013
Published: June 27, 2013

References

Supplemental Information

Mental Imagery Changes
Multisensory Perception

Christopher C. Berger and H. Henrik Ehrsson

Inventory of Supplemental Information

1. Supplemental Figures
   a. Figure S1: Schematic Overview of Experimental Procedures
   b. Figure S2: Average Localization Errors in Experiment 2A

2. Supplemental Results
   a. Experiment 1A: Additional Analysis (Finger Tap Condition)
   b. Experiment 1B: Imagery-Perception Correlation Analysis
   c. Experiment 2A: Imagery-Perception Correlation Analysis
   d. Experiment 2B: Additional Analysis (Binomial Test)
   e. Experiment 3: Additional Analyses

3. Supplemental Experimental Procedures
   a. Participants
   b. Experiment 1A Materials and Procedures: Imagery Cross-Bounce
   c. Experiment 1B Materials and Procedures: Imagery Cross-Bounce Control
   d. Experiment 2A Materials and Procedures: Imagery Ventriloquism Experiment
   e. Experiment 2B Materials and Procedures: Psychophysical Staircase Experiment
   f. Experiment 3 Materials and Procedures: Imagery McGurk Experiment

4. Supplemental References
Figure S1. Schematic Overview of Experimental Procedures, Related to Figures 1–3 and the Results
(A) The time-course for one example trial in experiment 1. (B) The time-course for one example trial in experiment 2A. (C) The time-course for one example trial in experiment 3.
Figure S2. Average Localization Errors in Experiment 2A, Related to Figure 2

(A) The average (absolute) localization error for the 40° and 60° locations (averaged across hemifields) when a sound was presented alone (Auditory Only), when a sound was presented in the same location as an imagined visual stimulus (AV Same), and when the auditory stimuli were presented in locations 15° (AV 15°) and 30° (AV 30°) away from the imagined visual stimuli (averaged across hemifields). (B) The average (absolute) localization error for the 40° and 60° locations (averaged across hemifields) when a sound was presented alone (Auditory Only), when a sound was presented in the same location as a visual stimulus (AV Same), and when the auditory stimuli were presented in locations 15° (AV 15°) and 30° (AV 30°) away from the visual stimuli. Error bars represent ±SEM.
Supplemental Results

Experiment 1A: Finger-Tap Condition Results
We found that imagining a finger tap at the moment of coincidence led to a significant increase in the perception of bounce ($M = 0.53$, $SD = 0.29$) as compared to the view-only condition [$t(21) = 3.38$, $p = .003$]. No significant difference was observed between the imagine finger tap and imagine sound at conditions [$t(21) = .98$, $p = .34$]. This finding is consistent with previous work on the motion bounce illusion, demonstrating an increased perception of the bouncing percept in the cross-bounce illusion if a tactile stimulus is presented close to the moment when the two objects meet [1]. Moreover, this finding stands in contrast to the results of experiment 1B in which imagining a finger movement that did not contain a tactile component but had the same temporal profile (i.e., a finger lift), did not lead to an increase in the perception of bounce compared to the view only condition (see Figure 1B). Together, these findings demonstrate that the effects of imagined stimuli on motion perception are specific to the effects of perceptual stimuli on motion perception.

Experiment 1B: Imagery-Perception Correlation
There was a very strong positive correlation across participants [$r(10) = .830$, $p = .001$] between the proportion of perceived bounce when a sound was imagined at the moment of coincidence and the proportion of perceived bounce when the participants heard the real sound in experiment 1B. This finding provides convergent evidence of the similarities between perceptual-perceptual and imagery-perceptual multisensory integration.

Experiment 2A: Imagery-Perception Correlation
There was a strong positive correlation across participants [$r(20) = .613$, $p = .003$] between the participants’ MEIs in the imagery and perceptual versions of experiment 2A. There was also a strong positive correlation across participants [$r(20) = .477$, $p = .029$] between participants’ VB (averaged across disparities of 15° and 30°) in the imagery and perceptual versions of experiment 2A. These findings provide convergent evidence of the similarities between perceptual-perceptual and imagery-perceptual multisensory integration.

Experiment 2B: Additional Analysis (Binomial Test)
An overall greater distance between the left and right staircases for the imagine circle condition was observed in 14 of the 18 participants (78%). A binomial (sign) test revealed that this proportion was significantly greater than chance ($p = .031$, two-tailed).

Experiment 3: Supplemental Results
In the main analysis, there was no significant main effect of condition (imagine ‘ba,’ imagine ‘ka,’ view only) [$F(2, 21) = .547$, $p = n.s.$]. There was also no main effect of perceiver status (perceivers vs. non-perceivers) [$F(2, 21) = 1.21$, $p = n.s.$].

For the non-perceivers, there was no significant effect of condition (imagine ‘ba,’ imagine ‘ka,’ view only) [$F(2, 9) = 1.69$, $p = n.s.$]. For perceivers, however, there was a significant effect of condition [$F(2, 10) = 3.89$, $p = .03$]. Moreover, planned comparisons revealed a significant increase in perceivers’ proportion of perceived ‘da’ for the imagine ‘ba’ condition compared to the imagine ‘ka’ control condition [$t(11) = 2.46$, $p = .032$]. Together, these findings demonstrate the specificity of the effect to stimuli that produce the perceptual version of the illusion and suggest that the effect cannot merely be explained by some form of response bias.

The findings presented in these three experiments are especially striking given the inherent limitation of not being able to verify directly whether the participants were actively engaged in imagery, the vividness of their mental image, nor their ability to imagine the mental image at the correct time. Future research will be useful in determining the relative contribution of these limitations to the strength of imagery-percept multisensory illusions.
Supplemental Experimental Procedures

Participants

Ninety-six participants took part in experiment 1A (n = 22), experiment 1B (n = 12), experiment 2A (n = 21), experiment 2B (n = 18), and experiment 3 (n = 23). All participants were recruited from within the student population in the Stockholm area, were healthy, reported no history of psychiatric illness or neurologic disorder, and had no problems with hearing or vision (or had corrected to normal vision). All participants gave their written informed consent before the start of the experiment, which was approved by the Regional Ethical Review Board of Stockholm.

Experiment 1A Materials and Procedures: Imagery Cross-Bounce

First, the participants were given instructions regarding the forms of imagery they would be asked to perform during the experiment. For the imagined auditory stimulus, a 200 ms ‘clink’ sound was played for the participant several times. The participant was instructed to “try to imagine the sound as vividly as possible,” and the sound was replayed as needed during the instructions until the participant felt comfortable vividly replicating it in his or her mind. To make our investigation more complete, we also included a condition in which the participants imagined tapping their fingertip against the support surface at the moment of circle coincidence. The rationale for including this extra condition stems from an earlier observation that tactile stimulation near the moment of coincidence will lead to an increase in the perception that the two moving objects bounce, rather than cross [1, 2]. This shows that dynamic sensory events in non-auditory modalities can also influence the cross-bounce illusion, and we expected the same rule to hold true for imagined dynamic events in other modalities. For the finger-tap condition, the participants were instructed to rest their hand on the desk with their palm facing down. Then, the participants were asked to move their finger in a quick (∼500 ms) movement up and back to its original resting place on the table (i.e., a tapping movement). Once a participant had demonstrated a consistent movement, he or she was then asked to imagine executing the same finger tapping motion as vividly as possible. The participant could repeat the action as many times as necessary during the instructions to establish a vivid motor-somatic image of the tapping movement.

After receiving the instructions, the participants had one practice trial of each of the five conditions (i.e., view only, imagine sound before, imagine sound during, imagine sound after, and imagine finger tap). In the three ‘imagine sound’ conditions, a prompt appeared on the screen instructing the participant to “Imagine HEARING SOUND when the circles are HERE” presented either above the fixation cross (corresponding to ≈ 500 ms prior to coincidence), in the center of the screen (at the moment of coincidence), or below the fixation cross (corresponding to ≈ 500 ms after coincidence). In the passive-view condition, a “VIEW ONLY” prompt appeared in the center of the screen. As in Sekuler et al.’s original experiment [3], the view only condition served to assess participants’ baseline susceptibility of the illusory bouncing percept. In this way, the proportions of perceived bounce in all other conditions can be interpreted in terms of their relationship to participants’ perceptual baseline. In the imagine-finger-tap condition, the participant saw the instruction “Imagine MOVING YOUR FINGER when the circles are HERE” appear in the center of the screen. Following the condition prompt, the visual display during each trial was the same throughout the experiment. First, a white fixation cross appeared in the center of the screen with a gray background. Next, two blue circles (each with a 7.5 mm radius) appeared from the top right and left corners of the screen and moved (125 mm/second) at a 45° angle toward the opposite corner of the screen, crossing in the middle at the location of the crosshair (see Figure S1A for a schematic overview). The trial ended once the circles disappeared from view in their respective corners. After each trial, the participants were prompted by the computer to indicate whether they saw the circles ‘bounce off’ or ‘stream
through one another by pressing ‘b’ or ‘s’ on the keyboard, respectively. A total of 100 trials (20 per condition) were conducted. Each condition was presented randomly across all 100 trials. The stimuli were presented via PsychoPy software [4, 5] on a 30.48 cm MacBook computer. At the conclusion of the experiment, the participants filled out a funneled debriefing that probed their knowledge of the hypothesis of the experiment.

Experiment 1B Materials and Procedures: Imagery Cross-Bounce Control

The experiments took place in a soundproof testing room (40 decibel noise reduction). The experiment was divided into two separate blocks. In the first block, the participants were given instructions regarding the forms of imagery they would be asked to perform during that block. The participants were given the same instructions as in experiment 1A regarding how they should imagine the sound during the experiment. For the finger-lift imagery, the participants were instructed to rest their hand on the desk in a relaxed position with their palm facing up. Then, the participants were asked to move their finger in a quick (≈ 500 ms) movement, lifting their finger upward while keeping the back of their hand flat on the table, then allowing it to retract back to its original resting place. In contrast to the imagine-finger-tap condition in experiment 1A, this finger movement was specifically selected for both its similarity in movement and its lack of a punctate tactile component. Whereas in experiment 1A, in which the participants imagined a tapping movement that included the salient tactile component associated with touching their finger to the table, in this experiment, the participants' imagery was focused solely on the motor act of moving their finger. Once the participant had demonstrated a consistent movement, the participant was then asked to imagine the motion as vividly as possible. The participant could repeat the action as many times as necessary during the instructions to establish a vivid mental image of the action.

Following the instruction period, the participants had one practice trial of each of the three conditions (i.e., view only, imagine sound during, and imagine finger lift). In the imagine-sound-at condition, an “Imagine HEARING SOUND when the circles are HERE” prompt was presented at the moment of coincidence. In the view-only condition, a “VIEW ONLY” prompt appeared in the center of the screen, and in the imagine-finger-lift condition, the participant saw the instruction “Imagine MOVING YOUR FINGER when the circles are HERE” appear in the center of the screen.

Following the condition prompt, the visual display during each trial was the same throughout the experiment. First, a white fixation cross appeared in the center of the screen with a gray background. Next, two blue circles (each with a 7.5 mm radius) appeared from the top right and left corners of the screen and moved at a 45° angle toward the opposite corner of the screen, crossing in the middle at the location of the fixation cross. The trial ended once the circles disappeared from view in their respective corners.

The participants indicated whether they perceived the circles to stream through or bounce off one another as in experiment 1A. A total of 60 trials (20 per condition) were conducted. The conditions were presented randomly across all 60 trials.

In the second (perceptual) block, the participants were instructed that now, in trials for which they had previously imagined the stimuli, they should actually perform the finger lift or would actually hear a sound. For each trial in this block, the participant saw the following: a prompt from one of the three conditions was presented at the center of the screen in random order, instructing the participant to either “LIFT YOUR FINGER when the circles are HERE” for the finger-lift condition, “HEAR SOUND when the circles are HERE” for the sound-at condition, or “VIEW ONLY” in the view-only condition. Following the condition prompt, the participants watched the visual display (described above) while they lifted their finger, heard the ‘clink’ sound play, or simply viewed the stimuli.

The stimuli were presented via PsychoPy software [4, 5] on a 30.48 cm MacBook computer. At the conclusion of the experiment, the participants filled out a questionnaire...
Experiment 2A Materials and Procedures: Imagery Ventriloquism Experiment

The experiments took place in a soundproof testing room (40 decibel noise reduction). The stimuli in the experiment were presented via PsyScope software [6] using a 27-inch iMac computer. The computer screen was projected onto a wall from an Optima EP 1080 DLP overhead projector (3600 lm) 79 cm in front of the participant, resulting in a 211 x 105 cm projected image with a 1 pixel to 1.0 mm correspondence between the original and projected displays. In this way, the visual stimuli were presented on a black background (the wall in the darkened room). The auditory stimuli in the experiment consisted of 110 ms ‘beep-like’ sounds comprised of a 3000 Hz and 4000 Hz mixed tone (sine waveforms with amplitudes of 0.5 N/m²). The sound was generated using Perfect Tone software (Line of Sight Software, 2004). To ensure the presentation of ecologically valid stereo sound, we pre-recorded our ‘beep-like’ auditory stimulus from 16 locations along a horizontal plane (eight in each hemifield, starting at 22 cm from the center and one location every 11.4 cm to the end of the projected display) using studio-quality microphones inside the ear canals of a dummy head (KU 100 dummy head audio system; Neumann artificial head stereo microphone system). The sound played from each location was recorded relative to the dummy head, which was placed in the exact position of the participant’s head during the experiment (107 cm from the ground). To ensure consistent presentation of the stimuli, each participant was seated with his or her head in a chin rest 107 cm above the ground and 79 cm from the wall for the duration of the experiment. The pre-recorded auditory stimuli were presented via state-of-the-art high-fidelity Sennheiser HD 600 headphones at a comfortable volume. Thus, when the sounds were played through the headphones, the participants experienced them as if they were coming from the locations on the wall from which they were originally recorded.

The experiment was split into two parts: an imagery part, in which the participants imagined the visual stimuli, and a perceptual part, in which the participants actually saw the visual stimuli. Before each trial in the imagery version of the experiment, the participant saw a prompt that said, “get ready…” (3 s), allowing the participant to fixate on the fixation cross (0°) and prepare for the events of the trial. Next, a white circle (radius = 9 cm) appeared (1000 ms) at 33.4 cm or 79 cm from the center of the participant’s left or right hemifield. This position cued the location at which the participant should imagine the visual stimulus during that trial. Then, a countdown from 3 appeared on the screen (3 s) just above the fixation cross. Immediately following the countdown, the participants imagined the visual stimulus at the location of the cue as vividly as possible while maintaining fixation on the fixation cross. At the same time, an auditory stimulus was presented 15° away, 30° away, or at the same location as where the participant was imagining the visual stimulus. In another condition, rather than a circle, the participants saw a cue appear just above the fixation point that said “Imagine Nothing.” This cue instructed the participant not to imagine anything during that trial. In these trials, the auditory stimuli were presented in the same locations as the circles were imagined in the other conditions (33.4 cm or 79 cm from the center in the participant’s left or right hemifield). Thus, there were four different stimulus combinations: imagery-auditory same, imagery-auditory 15° disparity, imagery-auditory 30° disparity, and auditory only (see Figure S1B for schematic overview). Each stimulus combination was presented relative to each of the four imagined visual stimuli three times, resulting in a total of 48 trials presented in random order. Following every trial, a horizontal line appeared that extended the length of the visual display—vertically aligned with the location at which the sounds were pre-recorded and from which they were played. Using a mouse, the participants indicated where they perceived the sound to come from on that trial by navigating the cursor to the perceived location and clicking. Once the participant clicked where he or she perceived the sound to come from, the next trial began.
The perceptual version of the experiment came next. All of the procedures, stimuli, and repetitions were the same for the perceptual version of the experiment, except that rather than imagining the visual stimuli at the end of the countdown, a visual stimulus was actually presented. For consistency, there was also a cue before the countdown as in the imagery version of the experiment.

For both the imagery and perceptual versions of experiment 2A, signed auditory localization errors (see Figure S2 for mean absolute errors) were calculated for each trial and averaged by location and condition across hemifields, resulting in mean localization errors for the four conditions: spatially congruent audiovisual stimuli (AVC), spatially incongruent audiovisual stimuli (AVI) of 15° (AVI15) and 30° (AVI30), and auditory only (AOnly). Following previous studies on the ventriloquism effect [11, 12], the percent visual bias (%VB) \[100^* (AVI-AOnly)/\text{Actual AV Difference}\] was calculated for audiovisual conditions, in which positive percentages represent the extent to which the perceived location of the auditory stimulus was biased toward the visual stimulus [11, 12]. Similarly, a multisensory enhancement index (MEI) \[(AVC-AOnly)/AOnly\] was calculated for the audiovisual conditions to determine the extent to which auditory localization benefitted from congruent visual and auditory stimuli (negative values represent more accurate localization in the audiovisual conditions compared to the auditory only condition, whereas positive values represent the opposite) [11, 13].

**Experiment 2B Materials and Procedures:** *Psychophysical Staircase Experiment*

In this experiment, we made use of a psychophysical staircase procedure and closely followed the methodology of Bertelson and Aschersleben [7]. The benefit of such a paradigm is that the visual bias of auditory localization can be measured in a manner that removes the possibility that the participants will base their responses on post-perceptual decisions rather than their genuine perception [7, 8]. The auditory stimuli were presented from two randomly selected staircases that began at extreme left or right positions (48°) and gradually converged as the participant made dichotomic judgments of whether the sound came from the left or right of fixation (0°). The participants indicated the location of the sound by pressing the ‘left’ or ‘right’ key on the keyboard. For the left staircase, the stimuli moved one step to the right when the participant indicated that the sound came from the left and one step to the left when he or she indicated that the sound came from the right. That is, as the participants made ‘left’ judgments in the left staircase, the location of the next sound moved centrally in steps of 4° until a threshold of 21° was met; thereafter, the staircase proceeded at 2° increments. However, if the participant made a ‘right’ judgment in the left staircase then the location of the next sound moved one step outward. The reverse pattern was followed for the right staircase. This meant that while the participant was certain about location of the sounds, each staircase moved from its respective left/right starting points towards the center. Then, at some point in each staircase, the participant became uncertain about the location of the sound and began to make responses that were different from the previous response (i.e., response reversals). This procedure continued until the participant made eight response reversals in each staircase. All analyses were subsequently conducted on these eight response reversals. Before the presentation of each auditory stimulus, there was a countdown from 3 (3 s) with a fixation cross at 0°. The fixation cross disappeared at the end of the countdown, and an auditory stimulus was delivered at the next beat in the countdown.

The participants experienced the above stimulus presentation in two separate counterbalanced blocks. In one block (imagine-circle condition), the participants imagined seeing a white circle (radius = 3.5 cm) appear at fixation (0°) at the end of the countdown, and in the other block, the participants performed the localization task while maintaining fixation but without imagining a circle (no-circle condition). In the imagine-circle condition, the participants first viewed five example trials in which a circle appeared at the end of the countdown, demonstrating how they should imagine the appearance of the circle.
(temporally paired with a non-spatial auditory stimulus) during the experiment. All of the auditory stimuli (41 total; 20 per hemifield and one at 0°) were pre-recorded in the same manner as in experiment 2A. All of the auditory and visual stimuli were presented to the participants using the same setup as in experiment 2A, except that stimulus presentation was controlled using PsychoPy software [4, 5] to implement the aforementioned staircasing procedure.

**Experiment 3 Materials and Procedures: Imagery in Modified McGurk Experiment**

Similar to the cross-modal effects of real auditory stimuli on the perception of visual stimuli found in the ‘reverse McGurk effect’ [7, 8], we sought to examine whether an imagined auditory stimulus could interact with a visual speech stimulus to produce an illusory multisensory speech percept. The experiment took place in a soundproof testing room (40-decibel noise reduction) and was split into three separate blocks. Each participant experienced all three blocks, and the visual stimuli in all three blocks were identical. In all three blocks, the participants saw six pre-recorded videos (4 s each) of six different people saying “ga” three times at a rate of 40 bpm. All of the videos were silent. Following each video, the participants were prompted by the computer to indicate whether they perceived the person in the video to be saying “ga” or “da” by pressing the g or d key on the keyboard, respectively. All of the stimuli and questions were presented using PsyScope Software [6] on a 27-inch iMac computer.

During the first block (baseline condition), the participants were instructed to passively view the videos. During the second and third blocks, the participants were either instructed to “imagine hearing the sound ‘ba’” (experimental condition) or to “imagine hearing the sound ‘ka’” (control condition) at the same pace that the person in the video was moving his or her lips. Similar to experiment 1, the view only condition served to assess participants’ baseline susceptibility of the illusory percept. In this way, the proportions of reported ‘da’ in all the imagine ‘ba’ and imagine ‘ka’ conditions can be interpreted in terms of their relationship to participants’ perceptual baseline. In both imagery blocks, the participants were trained to imagine the sound (ba or ka) by first saying the sound aloud at the pace of 40 beats per minute, measured using a metronome. Next, the participants were instructed to imagine hearing the sound at that pace, again using a metronome. Finally, once the participants were comfortable with this task, they were asked to imagine the sound at the pace of 40 beats per minute without the use of the metronome. During the trials, the participants were instructed to begin imagining the sound at that rate the moment they saw the film appear on the screen. Blocks 2 and 3 were counterbalanced across subjects to eliminate the possibility of order effects.

After all of the trials were completed, the participants were tested for their susceptibility to the perceptual version of the McGurk illusion using a standard McGurk video, in which a person’s mouth is saying ‘ga’ while ‘ba’ is dubbed over his or her voice. The participants watched the video three times, and were asked to report freely what they perceived the person in the video to be saying after each presentation. If the participant reported the fused illusory percept (i.e., ‘da’) at least once, he or she was included in the main analysis as a perceiver. Our reported percentage of McGurk-perceivers is consistent with recent replications and investigations of the McGurk illusion [9, 10].
Supplemental References


