SCIENTIFIC REPORTS

natureresearch

OPEN

Cross-modal size-contrast illusion: Acoustic increases in intensity and bandwidth modulate haptic representation of object size

Maiko Uesaki[1,2,3], Hiroshi Ashida[4], Akiyoshi Kitaoka[2,5] & Achille Pasqualotto[1,6,7]

Changes in the retinal size of stationary objects provide a cue to the observer's motion in the environment: Increases indicate the observer's forward motion, and decreases backward motion. In this study, a series of images each comprising a pair of pine-tree figures were translated into auditory modality using sensory substitution software. Resulting auditory stimuli were presented in an ascending sequence (i.e. increasing in intensity and bandwidth compatible with forward motion), descending sequence (i.e. decreasing in intensity and bandwidth compatible with backward motion), or in a scrambled order. During the presentation of stimuli, blindfolded participants estimated the lengths of wooden sticks by haptics. Results showed that those exposed to the stimuli compatible with forward motion underestimated the lengths of the sticks. This consistent underestimation may share some aspects with visual size-contrast effects such as the Ebbinghaus illusion. In contrast, participants in the other two conditions did not show such magnitude of error in size estimation; which is consistent with the "adaptive perceptual bias" towards acoustic increases in intensity and bandwidth. In sum, we report a novel cross-modal size-contrast illusion, which reveals that auditory motion cues compatible with listeners' forward motion modulate haptic representations of object size.

Being able to accurately estimate object size is essential in guiding our interaction with the environment. However, the perceptual representations of object size are not inherently determined by one source of information; rather, they are constructed by integrating multiple sources of information.

For example, within vision, the size of an object is computed based not only on its projected retinal image size, but also on other sources of information provided by the environment. Much of the evidence for the roles different sources of information play in the estimation of object size comes from sensory illusions whereby the size of an object is misperceived. One of such illusions is the Ebbinghaus illusion[1,2]. In the Ebbinghaus illusion, the central target object appears smaller when surrounded by an array of similar objects that are larger, as compared to when surrounded by an array of the similar objects of the identical size. The Ebbinghaus illusion demonstrates how the surrounding context can influence the visually perceived size of the target object.

The information potentially relevant to the determination of object size can also be generated by multiple modalities. Evidence is accumulating of the interaction of multiple sensory modalities, in construction of the perceptual representations of object size. The effect of looming sounds on visual perception of object size, for example, illustrates how the context provided by auditory information influences the visually perceived size of objects. It has been shown that being exposed to auditory information indicative of approaching sound sources (i.e. looming sounds) leads to visually perceived increases in the size of the target object[3–6]. Not only does the effect of looming sounds on visual perception of object size indicate that inputs from multiple modalities can be integrated in order for perceptual representations of object size to be constructed, but it also demonstrates the prominence of sensory motion cues.

[1]School of Social Sciences, Nanyang Technological University, Singapore, Singapore. [2]Open Innovation & Collaboration Research Organization, Ritsumeikan University, Osaka, Japan. [3]Japan Society for the Promotion of Science, Tokyo, Japan. [4]Graduate School of Letters, Kyoto University, Kyoto, Japan. [5]College of Letters, Ritsumeikan University, Kyoto, Japan. [6]Faculty of Arts and Social Sciences, Sabanci University, Istanbul, Turkey. [7]School of Psychology, Faculty of Science and Engineering, University of Nottingham Malaysia, Selangor Darul Ehsan, Malaysia. Correspondence and requests for materials should be addressed to M.U. (email: maiko.uesaki@ntu.edu.sg)

Sensory information indicative of approaching objects, either due to the observer moving towards the objects or to the objects moving towards the observer, provide critical information for our navigation through the environment. One important sensory cue to the change in distance between the observer and the objects in the environment derives from the observer's self-motion, and is primarily represented in the retinal size of the objects. For example, an increase in the retinal size of trees on the sides of a road would indicate the observer's forward motion along the road. Although the dominant cue to self-motion is the pattern of retinal motion resulting from the observer's own motion (i.e. optic flow[7,8]) under typical circumstances, this moving pattern is often confounded by eye movements[9–12] and motion of other objects in the visual field[8,13,14], which render sensory information from other modalities such as audition highly relevant. Auditory cues to self-motion tend to be less compelling compared to those of the visual modality, however, characteristics of auditory modality such as high temporal sensitivity and acuity are instrumental in the evaluation of self-motion, and the resulting change in distance between the moving individual and the stationary objects in the environment[15].

Auditory cues to self-motion are even more critical in visually impaired individuals deprived of visual cues to self-motion. In fact, sensory substitution tools have been developed to convert images into forms that the visually impaired can process with sensory modalities other than vision, in order to assist the visually impaired in using visual information as supplement to other sensory cues in guiding their navigation. One of the sensory substitution devices that transforms visual information to auditory information is software called "the vOICe." The vOICe scans visual images from left to right, converts them into greyscale images, and subdivides them into pixels; each of which is then converted to a sound based on its position within the scanned image, and its luminance. The horizontal positions of the pixels are expressed in terms of order (i.e. sounds representing pixels on the left are played before those representing pixels to their right), the vertical positions in terms of pitch (i.e. sounds representing pixels towards the top of the image have a higher pitch than those towards the bottom), and the luminance in terms of loudness (i.e. sounds representing pixels of higher luminance are louder than those of lower luminance). Although originally developed to assist the visually impaired, sensory substitution tools have since been used in research conducted with both visually impaired as well as sighted individuals[16–18]. In this study, the snapshots of visual input during linear motion along a tree-lined road were converted into their auditory equivalents, using the vOICe, in order to investigate what impact auditory cues to self-motion may have on perception of object size.

Evidence from the effect of looming sounds on the perceived size of objects suggests that auditory cues increasing in intensity and bandwidth, compatible with forward self-motion, would lead to overestimation of the size of an object. Sutherland and colleagues[19] explain that when sounds increasing in intensity are associated with static visual objects they bias the visual perception of the size of those objects, as in-depth auditory motion cues are re-mapped onto visual features. Furthermore, it has also been shown that visual perception of size can be modulated by auditory cues to distance[20]. It is interesting to note, however, that Serino et al.[21] reported differential responses to stimuli that are further away and to those that are close enough to interact with the body. This raises the question: Do auditory motion cues modulate the construction of perceptual representations of object size, only when the auditory cues are re-mapped onto the *target* object (i.e. the object, of which size is to be estimated) as to the looming sound effect, or also when the sounds are re-mapped onto the *environment* (i.e. the context) in which the listener and the target object are placed within the interactable distance? And if so, how?

In this article, we report on a novel cross-modal size-contrast illusion that demonstrates the contribution of auditory motion cues in the modulation of haptic representation of object size.

Results

The aim of this experiment was to determine how haptic size perception is affected by the exposure to auditory information compatible with forward motion, backward motion, and incompatible with any linear motion. Results reported below are based on the performance of all 66 participants. Blindfolded participants were asked to estimate the lengths of three wooden sticks through haptic exploration while exposed to the auditory stimulus compatible with either their forward motion (Forward Condition) or backward motion (Backward Condition) along a tree-lined path, or the auditory stimulus without any motion components (Scrambled Condition).

Before analysing the difference between the three conditions, we determined whether there was any effect of the order in which the sticks of three different lengths were presented. A three-by-three factorial ANOVA with Condition (Forward vs. Backward vs. Scrambled) and Order (first vs. second vs. third) as factors showed no significant main effect of Order ($F_{(2, 126)} = 0.27$, $p = 0.76$), and no significant interaction between Condition and Order ($F_{(4, 126)} = 0.41$, $p = 0.80$).

We also confirmed that there were no significant differences between male and female participants in any of the three conditions with a three-by-two factorial ANOVA with Condition and Sex (male vs. female) as factors ($F_{(1, 60)} = 0.16$, $p = 0.69$), and that there was no interaction between Condition and Sex ($F_{(2, 126)} = 0.40$, $p = 0.68$). Although sex differences have been reported in the literature of looming sounds and their influence on perception[22,23], since we found no differences between sexes, this variable will not be discussed.

Table 1 describes the average length estimates for the three wooden sticks and standard errors (SEs) of the means for each condition. For Forward Condition, the lengths of all three objects were consistently underestimated. In contrast, for Backward Condition and Scrambled Condition, the direction of error (i.e. over- or under-estimation) was not consistent across the three lengths.

As Table 1 shows, there was a consistent margin of error in the length estimations regardless of the conditions; however, the errors were notably larger for all lengths for Forward Condition than for Backward Condition and Scrambled Condition. The percentage of error for the 15-, 25-, 40-cm lengths was 22%, 15% and 17%, respectively, for Forward Condition, as compared with 3%, 2% and 3% for Backward Condition and 9%, 0.3% and 2% for Scrambled Condition.

	15 cm	25 cm	40 cm
ForwardCondition	11.73 cm (0.86)	21.14 cm (1.01)	33.14 cm (1.35)
BackwardCondition	14.56 cm (0.85)	25.47 cm (1.10)	41.23 cm (1.76)
ScrambledCondition	13.68 cm (0.98)	25.1 cm (1.71)	39.1 cm (2.17)

Table 1. Average length estimates for the 15-, 25-, 40-cm long sticks and standard errors (SEs) of the means reported in brackets (n = 22 per condition).
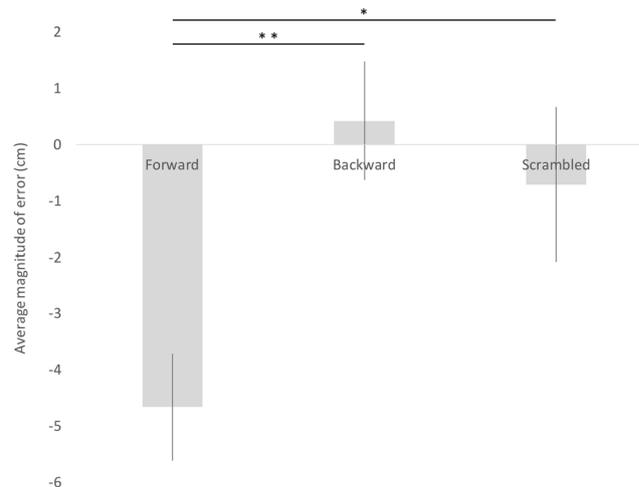
Figure 1. Magnitude of size-estimation error in centimetres for Forward Condition, Backward Condition and for Scrambled Condition (* significant at p < 0.05; ** significant at p < 0.01). Error bars indicate +/−1 SE.

In Fig. 1, the magnitude of error in size estimation averaged across the three objects is plotted for each auditory condition. The average magnitude of error was 4.67 cm (SE = 0.95) for Forward Condition; 0.42 cm (SE = 1.05) for Backward Condition; and 0.71 cm (SE = 1.38) for Scrambled Condition. There was a statistically significant difference between conditions as determined by one-way ANOVA (with Condition as a factor; $F_{(2, 63)} = 5.47$, p < 0.01). Subsequent Tukey's honestly significant difference (HSD) post hoc test showed that Forward Condition differed significantly from Backward Condition at p < 0.01 and Scrambled Condition at p < 0.05. There was no significant difference between Backward Condition and Scrambled Condition (p = 0.74).

Length estimates in Forward Condition and Backward Condition were also analysed with regards to Scrambled Condition (as opposed to the actual lengths of the objects) as a baseline. Two-tailed one-sample t-test revealed that the difference between Forward Condition and Scrambled Condition was significantly larger than 0 (t (65) = −3.37, p < 0.01, two-tailed), despite the small error observed in Scrambled Condition in the same direction as in Forward Condition (Fig. 1). This provides further evidence for the extent of underestimation in Forward Condition. Contrarily, the magnitude of error in Backward Condition compared with that in Scrambled Condition was not significantly different from 0 (t (65) = 0.86, p = 0.40, two-tailed). In line with the results of analyses whereby the actual lengths of the objects were considered a baseline, an independent-samples t-test yielded that the magnitude of error in size estimation differed significantly between Forward Condition and Backward Condition (t (130) = −2.88, p < 0.01, two-tailed) when the estimates in Scrambled Condition were used as a baseline.

Discussion

This study investigated the effect of auditory cues to self-motion on the construction of haptic representations of object size, and we discovered a novel cross-modal size-contrast illusion that demonstrates the contribution of auditory motion cues in the modulation of haptic representations of object size.

Results showed that the lengths of the objects were consistently underestimated during the exposure to auditory motion cues increasing in intensity and bandwidth, and therefore compatible with forward self-motion. This phenomenon resembles the dynamic Ebbinghaus illusion, which was recently reported by Mruczek and colleagues[24,25]. They modified the classic Ebbinghaus illusion so that the target object was surrounded by an array of similar objects that expanded and contracted in size and bandwidth. In the dynamic Ebbinghaus illusion, the concept of motion is introduced as another factor that can influence the construction of the perceived size of the target object. Mruczek and colleagues[24,25] showed that the target object was perceived to be smaller as the surrounding objects increased in size, much like what was observed in our experiments. The dynamic Ebbinghaus illusion[24,25] reveals the prominent role of motion information in determining visual representations of object size. Our results suggest that the mechanism underlying the dynamic Ebbinghaus illusion may be present across modalities. By contrast, our results may seem contradictory to the evidence from the looming sound effect, which is defined by the perceived increase in the size of the target object due to the acoustic increase in intensity[3–6]. However, the different effects of auditory motion cues on the perceived size of the target object observed in the

case of the looming sound effect and in our experiment can be accounted for by the objects to which the auditory cues are assigned. In the case of the looming sound effect, the auditory cues are re-mapped onto the *target* object itself, whereas in our study, the association was established between the sound and the object that constituted the *environment* (i.e. the context) within which the listener and target object were placed, prior to the experiment.

The findings indicate that the phenomenon observed here is a novel cross-modal size-contrast illusion, that reveals that auditory motion cues compatible with forward motion contribute to modulation of the perceived size of objects. As mentioned above, this illusion may share some aspects with the dynamic Ebbinghaus illusion[25], whilst it is unique in other aspects.

One of the most intriguing aspects shared between the cross-modal size-contrast illusion observed here and the dynamic Ebbinghaus illusion, is the motion component in the context in which the target object is placed. The auditory stimuli used in this experiment were generated by converting visual images simulating snapshots during linear motion along a tree-lined road into their auditory counterparts. Therefore, acoustic increases and decreases in intensity and bandwidth were compatible with the listener's forward and backward motion, respectively. Such auditory motion cues to self-motion through the environment over time, are indicative of the listener's shift in both time and space. As explained by a theory of magnitude (ATOM[26]), connections between conceptual elements such as time and quantity are a well-established phenomenon[27], and are also echoed in other parallels such as space in quantity[28]. There is also neurological evidence linking time, space, and size perception[29], derived from observations of those with parietal cortex lesion. Based on previous findings[30–34], one of the functions of the parietal cortex is to encode information regarding magnitudes in the environment (e.g. how far, how fast, how long), with respect to action (i.e. motor planning). ATOM may, at least in part, explain the mechanism underlying the size-contrast illusion described here. As the auditory stimuli used in this experiment were compatible with the listener's linear motion along a path, they were indicative of the speed of motion, the distance travelled, and how close the stationary objects in the environment were to the listener; and those auditory cues were processed with respect to the task that involved proprioceptive perception and judgement (i.e. estimation of object size). The illusion highlights how important sensory cues to self-motion are, as they enable us to estimate the direction and velocity of our movement through the environment, to recognise our position relative to the objects in the environment, to calculate the time-to-contact with those objects, and to conduct motor planning[35,36].

On the other hand, certain aspects of the illusion observed here differ from the Ebbinghaus illusion. The most fundamental of those is the fact that one is cross-modal whilst the other is uni-modal. Most of size-contrast illusions can be described as sources of information from one modality (e.g. vision in the case of the Ebbinghaus illusion) contributing to the modulation of perceptual representations of object size within the same modality. Contrarily, the illusion observed here is the result of the effect of acoustic changes in intensity and bandwidth on haptic representations of object size. However, it should be noted that, in this case, it may not be as straightforward as the auditory input modulating haptic perception. This is due to the application of sensory substitution in the rendering of auditory stimuli in our study, and the pre-experimental training during which "calibration" of the auditory stimuli and their visual origins took part. As demonstrated in Pasqualotto and Esenkaya[37], visual stimuli can be re-mapped onto auditory dimension through training (see also[17,38]). Although participants were blindfolded during the experiment, they had learnt beforehand the association between the visual and auditory representations of the object that constituted the surrounding in which the target object was placed (i.e. the pine tree). According to Nanay[39] (see also[40]), sensory substitution assisted perception (e.g. vision) may not be the same as vision as typically defined, or perception through the substituting sensory modality (e.g. audition); but instead, it may be multimodal mental imagery[41,42]. This is corroborated by a body of research that has shown that activity in the visual cortex can be elicited in the blind or blindfolded, by corresponding sensory stimulation in other sensory modalities (for examples, see[43–45]). The notion that the cross-modal size-contrast illusion observed here was induced by visual mental imagery triggered by auditory sensory stimulation, cannot be rejected until further investigated.

Results also showed that the perceived lengths of the same objects were little influenced by the auditory motion cues decreasing in intensity and bandwidth, compatible with backward self-motion. In fact, estimation of lengths was as accurate during the exposure to sounds compatible with backward motion as during the exposure to the same set of sounds presented in a scrambled order (i.e. auditory stimuli with no motion component). This may seem counter-intuitive as it might be expected that acoustic decreases in intensity and bandwidth would have the opposite effect to acoustic increases in intensity and bandwidth, and thus lead to overestimation of object size. Although the phenomenon observed is unidirectional; i.e. the effect was only present when subjects experienced the auditory stimulation compatible with forward motion, but not when they experienced auditory stimulation compatible with backward motion, it would be premature to discard the effect completely considering how consistent and robust the effect was during auditory stimulation with increasing intensity and bandwidth. It is possible that the illusion only occurs during exposure to auditory stimulation compatible with forward motion, because acoustic increases in intensity and bandwidth tend to be more relevant and salient compared to decreases in intensity and bandwidth, because they are often indicative of objects moving further from the listener[6,46–48]. For example, an approaching car is potentially more threatening than a car that is moving away. Likewise, stationary objects towards which the observer is moving should be more relevant to the observer's safety and navigation than objects from which they are moving away, since the former would require the observer to take appropriate measures in order to avoid colliding with them. This is described in the "adaptive perceptual bias" reported by Neuhoff[6], which suggests that increases of acoustic intensity are perceptually salient as they can indicate an approaching sound source in the environment, which can potentially be threatening. This bias towards approaching sounds provides a temporal margin of safety[48], by allowing for the time necessary for defensive motor action (e.g. avoidance). The different effects auditory motion cues of increasing and decreasing intensity and bandwidth had on the construction of haptic representations of object size are also in line with previous findings that sounds of increasing intensity and bandwidth are processed differently from sounds of decreasing intensity and
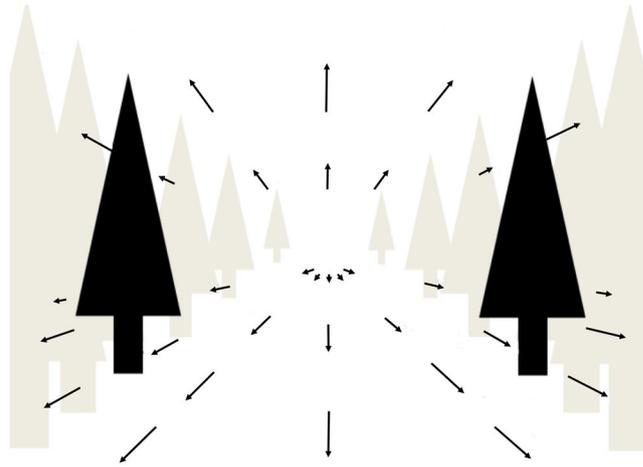
Figure 2. Visual representation of the path with pine trees. Auditory stimuli were either compatible with forward or backward motion along this path, or not compatible with any linear motion.

bandwidth in the auditory domain[4,5,49,50], and that sounds of increasing intensity and bandwidth elicit greater physiological responses (i.e. skin conductance and phasic alertness) compared to sounds of decreasing intensity and bandwidth[3].

Our findings introduce a novel cross-modal illusion of relative size perception. It illustrates that exposure to acoustic increases in intensity and bandwidth compatible with forward motion lead to underestimation of the size of objects by touch. The fact that error of such magnitude was not induced by acoustic decreases in intensity and bandwidth compatible with backward motion suggests that "adaptive perceptual bias"[6] may influence the illusory effect. In order to determine whether it is indeed the potential threat posed by the sources of acoustic increases in intensity and bandwidth that lead to the underestimation of object size, future study should systematically manipulate the level of threat posed by the sound sources and assess whether the level of threat is reflected in the magnitude of error in haptic size estimation (i.e. magnitude of illusory effect). This may be achieved by manipulating components of the motion cues, as to indicate varying distances to, or varying speeds at which the observer is approaching the objects in the environment (e.g. smaller volume indicating greater distance to the objects, ergo less immediate threat; more rapid increases indicating less time-to-contact, ergo more immediate threat). Another aspect of the illusion that this study was unable to clarify due to the confounding nature of the stimuli used, is the role mental imagery plays in the illusion. Future study should address this by comparing those who receive a training to establish the association between the auditory stimuli and the substituted visual stimuli with those who do not receive such a training and therefore are unaware of the original visual images used to render the auditory stimuli. Furthermore, in order to have a more comprehensive understanding of the nature of this new illusion, it is also important to investigate other factors on which the magnitude of this effect depend, such as the number of inducers (i.e. sound sources), and similarities between the target of which size is to be estimated and inducers.

On a more global scale, it would be interesting to see if this illusion generalises to other modalities. For example, whether optic-flow stimulation compatible with forward motion similarly modulates the haptically perceived size of objects.

This illusion opens new avenues to studying cross-modal influences of sensory cues compatible with self-motion, and multisensory interaction, but also has implications in its application for multisensory virtual reality, as well as safety systems that are dependent on sensory perception such as vibrotactile warning signals[51].

Materials and Methods

Participants.   Sixty-six healthy volunteers (33 males and 33 females; of the ages between 19 and 23 years) recruited amongst students at Sabanci University participated in the study. Each participant was randomly assigned to one of the three conditions: Forward Condition, Backward Condition, and Scrambled Condition (see Procedure). Mean age (standard deviation; SD) of participants was 21.32 (SD = 1.25), 21.23 (SD = 1.11), 21.23 (SD = 1.51) years for Forward Condition, Backward Condition, and Scrambled Condition, respectively. All had normal or corrected-to-normal vision, and normal hearing. All participants gave written informed consent to take part in this study, which was conducted in accordance with the ethical standards stated in the Declaration of Helsinki and approved by Sabanci University Research Ethics Committee. Participants were compensated for their time with course credits or meal vouchers.

Auditory stimuli.   A series of six images were converted into an auditory format with sensory substitution software called the vOICe (freely available at: www.seeingwithsound.com). Each image represented a visual snapshot during forward motion along a path with a pine tree on each side (Fig. 2).

In the first image of the sequence (the leftmost image in Fig. 3a), the pine trees appear small and close to the centre of the screen; in the following images, the pine trees appear bigger and further from the centre of the screen as compared with the preceding image, simulating the observer's forward motion along the path (Fig. 3a).
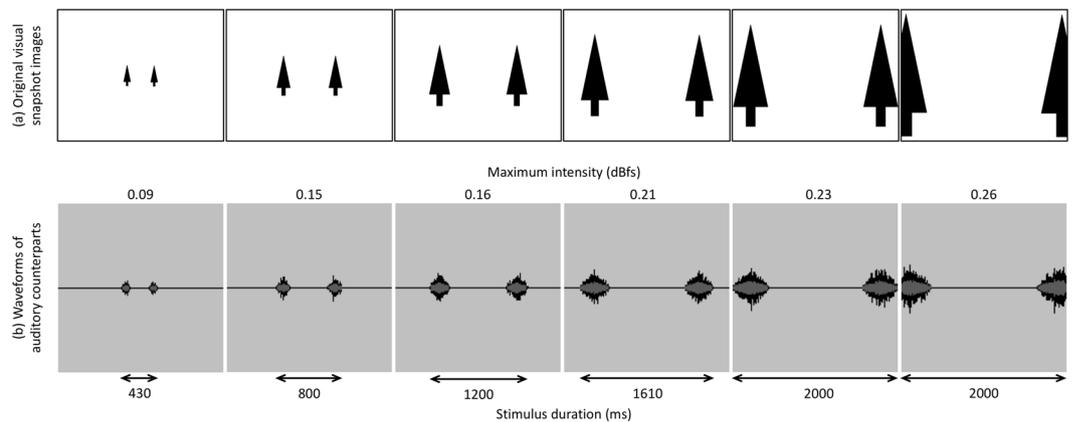
Figure 3. Coupling of (**a**) the original visual images of pine trees and (**b**) waveforms of their auditory counterparts rendered using the vOICe, in the sequence compatible with the observer's forward motion. The maximum intensity (dBfs) of each auditory stimulus reflected the height of the pine trees as they appear in the corresponding original visual image, and the stimulus duration (ms) the size of the pine trees and the distance between them.

Figure 3 demonstrates the coupling of the original visual images and the waveforms of their auditory counterparts generated with the vOICe.

Each of the six resulting auditory stimuli represented a pair of pine trees in the auditory modality (Fig. 3b). The duration of each stimulus reflected the size of the pine trees and the distance between them, as they appear in the corresponding original visual image. Stimulus durations were; 430 ms, 800 ms, 1200 ms, 1610 ms, 2000 ms, and 2000 ms. In the last image of the sequence (the rightmost image in Fig. 3a), the pine trees are so far into the peripheral of the "visual field" that a portion of each of them is not "seen". This is also reflected in its auditory counterpart (the rightmost image in Fig. 3b). The peak loudness intensity of each stimulus represented the height of the pine trees as they appear in the corresponding original visual image. Maximum intensities as expressed in dB full scale (dBfs; ranging from −1 to 1) were: 0.09 dBfs, 0.15 dBfs, 0.16 dBfs, 0.21 dBfs, 0.23 dBfs, and 0.26 dBfs.

Those auditory stimuli were presented using Audacity (www.audacityteam.org) in three different orders, depending on the condition to which the participant was assigned. In one condition, the stimuli were presented in an ascending sequence (i.e. increasing in intensity and bandwidth conveying the pair of pine trees becoming larger and shifting towards the peripheral) compatible with snapshots of the scene during forward motion (Forward Condition); in another, in a descending sequence (i.e. decrease in intensity and bandwidth conveying the trees becoming smaller and shifting towards the centre) compatible with snapshots of the scene during backward motion (Backward Condition). In the control condition, the same set of auditory stimuli was presented in a scrambled order (Scrambled Condition). The total duration of stimulus presentation was 16655 ms in all conditions.

The duration between any two auditory stimuli (i.e. inter-stimulus interval: ISI) was 1400 ms.

Haptic test material. Three smooth, round-surface wooden sticks (1 cm diameter) of varying lengths (15 cm, 25 cm, and 40 cm) were used as the test objects (i.e. target objects), the lengths of which were estimated with haptics by the participants.

Procedure. Before the experimental session, participants put on headphones and saw a short PowerPoint presentation on a computer. The presentation briefly introduced the concept of sensory substitution, that sensory information that is normally conveyed through one modality (i.e. vision) is translated into the information compatible with another sensory modality (i.e. audition). It then presented an image of one sample pine tree along with its auditory counterpart that was generated using the vOICe, in order for the participants to learn the coupling between the visual and auditory representations of the pine tree (i.e. calibration). Note that the sample pine tree was identical in its shape as any of the pine trees in the original images used to generate the auditory stimuli for the experiment. The sound of the pine tree was repeated until participants moved onto the next slide of the presentation. Participants were encouraged to look at the image and to listen to the sound for about one minute.

Subsequently, participants were blindfolded (Mindfold™, Colorado, USA) and asked to listen to the auditory stimuli: Those in Forward Condition heard the auditory stimuli presented in an ascending sequence, those in Backward Condition in a descending sequence, and the others in a scrambled order.

After one minute of attending to the auditory stimuli, whilst the sound remained playing, blindfolded participants were handed the three wooden sticks one by one for free haptic exploration using both hands. Each wooden stick was presented once. The presentation order of the three sticks was randomised amongst participants. A verbal estimate (in cm) of the perceived length was made once for each wooden stick by each participant. In total, each participant in each condition generated three estimates. The estimates were reported within a few seconds from the beginning of haptic exploration.

Data Availability

The dataset generated and analysed in this study is available as supplementary information.

References

1. Burton, G. The tenacity of historical misinformation: Titchener did not invent the Titchener illusion. *History of Psychology* **4**(3), 228 (2001).
2. Thiery, A. Aristote et la psychologie physiologique. *The New Scholasticism* (1896).
3. Bach, D. R., Neuhoff, J. G., Perrig, W. & Seifritz, E. Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology* **74**(1), 28–33 (2009).
4. Cappe, C., Thut, G., Romei, V. & Murray, M. M. Selective integration of auditory-visual looming cues by humans. *Neuropsychologia* **47**(4), 1045–1052 (2009).
5. Cappe, C., Thelen, A., Romei, V., Thut, G. & Murray, M. M. Looming signals reveal synergistic principles of multisensory integration. *Journal of Neuroscience* **32**(4), 1171–1182 (2012).
6. Neuhoff, J. G. Perceptual bias for rising tones. *Nature* **395**(6698), 123–124 (1998).
7. Gibson, J. J. *The Perception of the Visual World*. (Houghton Mifflin, UK, 1950).
8. Gibson, J. J. The visual perception of objective motion and subjective movement. *Psychological Review* **61**(5), 304 (1954).
9. Banks, M. S., Ehrlich, S. M., Backus, B. T. & Crowell, J. A. Estimating heading during real and simulated eye movements. *Vision Research* **36**(3), 431–443 (1996).
10. Crowell, J. A., Banks, M. S., Shenoy, K. V. & Andersen, R. A. Visual self-motion perception during head turns. *Nature Neuroscience* **1**(8), 732–737 (1998).
11. Royden, C. S., Banks, M. S. & Crowell, J. A. The perception of heading during eye movements. *Nature* **360**(6404), 583–585 (1992).
12. Royden, C. S., Crowell, J. A. & Banks, M. S. Estimating heading during eye movements. *Vision Research* **34**(23), 3197–3214 (1994).
13. Royden, C. S. & Hildreth, E. C. Human heading judgments in the presence of moving objects. *Attention, Perception, & Psychophysics* **58**(6), 836–856 (1996).
14. Warren, W. H. Jr. & Saunders, J. A. Perceiving heading in the presence of moving objects. *Perception* **24**(3), 315–331 (1995).
15. Väljamäe, A. Auditorily-induced illusory self-motion: A review. *Brain Research Reviews* **61**(2), 240–255 (2009).
16. Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B. & Scadden, L. Vision substitution by tactile image projection. *Nature* **221**(5184), 963 (1969).
17. Bach-y-Rita, P. (Ed.). *Brain Mechanisms in Sensory Substitution*. Academic Press. (1972).
18. Proulx, M. J., Brown, D. J., Pasqualotto, A. & Meijer, P. Multisensory perceptual learning and sensory substitution. *Neuroscience & Biobehavioral Reviews* **41**, 16–25 (2014).
19. Sutherland, C. A., Thut, G. & Romei, V. Hearing brighter: Changing in-depth visual perception through looming sounds. *Cognition* **132**(3), 312–323 (2014).
20. Jaekl, P., Soto-Faraco, S. & Harris, L. R. Perceived size change induced by audiovisual temporal delays. *Experimental Brain Research* **216**, 457–462 (2012).
21. Serino, A., Canzoneri, E. & Avenanti, A. Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: an rTMS study. *Journal of Cognitive Neuroscience* **23**(10), 2956–2967 (2011).
22. Neuhoff, J. G., Planisek, R. & Seifritz, E. Adaptive sex differences in auditory motion perception: looming sounds are special. *Journal of Experimental Psychology: Human Perception and Performance* **35**(1), 225 (2009).
23. Grassi, M. Sex difference in subjective duration of looming and receding sounds. *Perception* **39**(10), 1424–1426 (2010).
24. Mruczek, R. E., Blair, C. D. & Caplovitz, G. P. Dynamic illusory size contrast: a relative-size illusion modulated by stimulus motion and eye movements. *Journal of Vision* **14**(3), 2–2 (2014).
25. Mruczek, R. E., Blair, C. D., Strother, L., & Caplovitz, G. P. The Dynamic Ebbinghaus: motion dynamics greatly enhance the classic contextual size illusion. *Frontiers in Human Neuroscience*, 9 (2015).
26. Walsh, V. A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences* **7**(11), 483–488 (2003).
27. Walsh, V. & Pascual-Leone, A. *Transcranial Magnetic Stimulation: A Neurochronometrics of mind*. (MIT press, USA, 2003).
28. Seno, T., Taya, S., Ito, H. & Sunaga, S. The mental number line in depth revealed by vection. *Perception* **40**(10), 1241–1244 (2011).
29. Critchley, M. *The Parietal Lobes*. (Williams and Wilkins, UK, 1953).
30. Cohen, L. & Dehaene, S. Cerebral networks for number processing: Evidence from a case of posterior callosal lesion. *Neurocase* **2**(3), 155–174 (1996).
31. Bjoertomt, O., Cowey, A. & Walsh, V. Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain* **125**(9), 2012–2022 (2002).
32. Ashbridge, E., Walsh, V. & Cowey, A. Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia* **35**(8), 1121–1131 (1997).
33. Rushworth, M. F., Paus, T. & Sipila, P. K. Attention systems and the organization of the human parietal cortex. *Journal of Neuroscience* **21**(14), 5262–5271 (2001).
34. Göbel, R., Muckli, L., Zanella, F. E., Singer, W. & Stoerig, P. Sustained extrastriate cortical activation without visual awareness revealed by fMRI studies of hemianopic patients. *Vision Research* **41**(10), 1459–1474 (2001).
35. Harris, L. R., Sakurai, K. & Beaudot, W. H. A. Tactile flow overrides other cues to self motion. *Scientific Reports* **7**, 1059 (2017).
36. Huang, R.-S., Chen, C.-F. & Sereno, M. I. Spatiotemporal integration of looming visual and tactile stimuli near the face. *Human Brain Mapping* **00**, 1–21 (2018).
37. Pasqualotto, A., & Esenkaya, T. Sensory substitution: the spatial updating of auditory scenes "Mimics" the spatial updating of visual scenes. *Frontiers in Behavioral Neuroscience*, 10 (2016).
38. Meijer, P. B. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering* **39**(2), 112–121 (1992).
39. Nanay, B. Sensory Substitution and Multimodal Mental Imagery. *Perception*, 0301006617699225 (2017).
40. Ward, J. & Wright, T. Sensory substitution as an artificially acquired synaesthesia. *Neuroscience & Biobehavioral Reviews* **41**, 26–35 (2014).
41. Amedi, A., Malach, R. & Pascual-Leone, A. Negative BOLD differentiates visual imagery and perception. *Neuron* **48**(5), 859–872 (2005).
42. Ward, J. & Meijer, P. Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition* **19**(1), 492–500 (2010).
43. Murphy, M. C. *et al*. Top-down influence on the visual cortex of the blind during sensory substitution. *Neuroimage* **125**, 932–940 (2016).
44. Poirier, C. *et al*. Specific activation of the V5 brain area by auditory motion processing: an fMRI study. *Cognitive Brain Research* **25**(3), 650–658 (2005).
45. Renier, L. *et al*. Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage* **26**(2), 573–580 (2005).
46. Hall, D. A. & Moore, D. R. Auditory neuroscience: The salience of looming sounds. *Current Biology* **13**(3), R91–R93 (2003).
47. Maier, J. X., Neuhoff, J. G., Logothetis, N. K. & Ghazanfar, A. A. Multisensory integration of looming signals by rhesus monkeys. *Neuron* **43**(2), 177–181 (2004).
48. Neuhoff, J. G. Looming sounds are perceived as faster than receding sounds. *Cognitive Research: Principles and Implications* **1**(1), 15 (2016).

49. Maier, J. X. & Ghazanfar, A. A. Looming biases in monkey auditory cortex. *Journal of Neuroscience* **27**(15), 4093–4100 (2007).
50. Tyll, S. *et al.* Neural basis of multisensory looming signals. *Neuroimage* **65**, 13–22 (2013).
51. Ho, C., Spence, C., & Gray, R. Looming auditory and vibrotactile collision warnings for safe driving. In *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (No. 83, pp. 551-557) (2013, December).

Acknowledgements

Author Contributions

All authors (A.K., A.P., H.A., M.U.) reviewed and edited the manuscript. M.U. wrote the manuscript with feedback from the other authors. A.P. and M.U. analysed the data. A.P. performed the experiments. A.P., H.A. and M.U. contributed to the design of the study.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-50912-8.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.