



Indiscriminable sounds determine the direction of visual motion

Maori Kobayashi¹, Wataru Teramoto^{1,2,*}, Souta Hidaka³ & Yoichi Sugita^{4#}

SUBJECT AREAS:

BEHAVIOUR

SENSORY SYSTEMS

DEVELOPMENT

PLASTICITY

¹Research Institute of Electrical Communication, Tohoku University, 2-1-1, Karahira, Aoba-ku, Sendai, Miyagi, 980-8577, Japan, ²Department of Computer Science and Systems Engineering, Muroran Institute of Technology, 27-1, Mizumoto-cho, Muroran, Hokkaido, 050-8585, Japan, ³Department of Psychology, Rikkyo University, 1-2-26, Kitano, Niiza, Saitama, 352-8558, Japan, ⁴Department of Psychology, Waseda University, 1-24-1, Toyama, Shinjuku-ku, 162-8644, Japan.

Received
26 October 2011

Accepted
19 March 2012

Published
17 April 2012

Correspondence and requests for materials should be addressed to M.K. (te11001@meiji.ac.jp) or Y.S. (y.sugita@waseda.jp)

* Current address:
Department of
Computer Science and
Systems Engineering,
Muroran Institute of
Technology, 27-1
Mizumoto-cho,
Muroran, Hokkaido,
050-8585 Japan.

#Current address:
Department of
Psychology, Waseda
University, 1-24-1
Toyama, Shinjuku-ku,
162-8644 Japan.

On cross-modal interactions, top-down controls such as attention and explicit identification of cross-modal inputs were assumed to play crucial roles for the optimization. Here we show the establishment of cross-modal associations without such top-down controls. The onsets of two circles producing apparent motion perception were accompanied by indiscriminable sounds consisting of six identical and one unique sound frequencies. After adaptation to the visual apparent motion with the sounds, the sounds acquired a driving effect for illusory visual apparent motion perception. Moreover, the pure tones with each unique frequency of the sounds acquired the same effect after the adaptation, indicating that the difference in the indiscriminable sounds was implicitly coded. We further confirmed that the aftereffect did not transfer between eyes. These results suggest that the brain establishes new neural representations between sound frequency and visual motion without clear identification of the specific relationship between cross-modal stimuli in early perceptual processing stages.

The brain constantly receives signals from multiple sensory modalities, and has to evaluate whether or not these signals come from a common external event or object. One possible strategy to accomplish this task efficiently and effectively is to form associations between signals from different senses on the basis of past experiences, such as the frequent occurrence of these signals. Indeed, there have been several examples of rapidly induced changes in multimodal perception. We recently reported that a new association forms rapidly between a sound sequence containing no spatial or motion information and visual motion^{1,2}. In this sound-contingent motion aftereffect, two white circles placed side by side were presented in alternation. The onsets of the two circles were synchronized to a tone burst of high- and low frequency, respectively (Fig. 1A). After exposure to the visual apparent motion with tone bursts for a few minutes, a circle blinking at a fixed location was perceived as lateral motion in the same direction as the previously exposed apparent motion, when the flash onset was synchronized to the tones (Fig. 1B). These studies suggest that the brain can easily establish a strong association between a sound sequence and visual motion within a short period and that, after forming the association, sounds are able to trigger visual motion perception for a static visual stimulus.

However, it is not clear whether the aftereffect is caused by direct audio-visual interaction with perceptual learning mechanisms or by top-down controls, such as explicit recognition for a specific relationship of cross-modal stimuli. In Teramoto et al.¹, two easily discriminable tones were presented in conjunction with visual left-right apparent motion so that the participants could explicitly relate each tone to each location of the visual stimuli. This manipulation allowed participants to control their intention and/or attention to the special pairs of the audio-visual stimuli. It has been reported that multimodal interactions are strongly influenced by such top-down control^{3,4}. In contrast, several previous studies have also shown that motion-contingent aftereffects in visual domain⁵ and the multisensory integration of visual and auditory motion information⁶ can occur automatically, indicating the involvement of implicit perceptual processes.

The aim of the present study was to elucidate whether this sound-contingent visual motion aftereffect is caused by direct audio-visual interactions in perceptual level or mediated by top-down controls based on explicit recognition over audio-visual relationships. For this purpose, we presented a pair of sounds whose pitches were hard to discriminate, in conjunction with a left-right alternating visual stimulus. Two types of sounds were tested: complex tones (Fig. 1C) and band-pass noise bursts. The results showed that even indiscriminable sounds could acquire the driving effect for illusory visual motion and determine the direction of the motion, after prolonged exposure to visual apparent motion with these sounds. This finding suggests that new neural representations between sound and visual motion can be established through direct, bottom-up, audio-visual interactions.

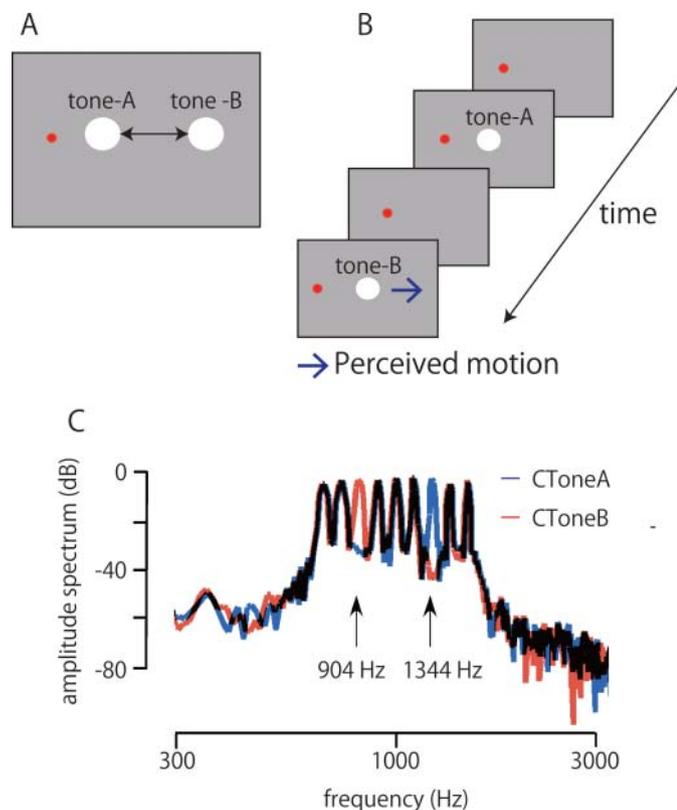


Figure 1 | Experimental design. (A) In an adaptation phase, participants were exposed to visual apparent motion in which two white circles placed side by side were presented in alternation. The onset of the leftward circle was accompanied by a tone A and the rightward circle a tone B. (B) In a test phase, a white circle was presented twice. The circle was perceived as lateral motion in the same direction as the previously exposed apparent motion, when the onset of the circle was synchronized to tones of alternating frequencies. When the onset of the first circle was synchronized to the tone A and the second circle the tone B, the circle appeared to move from left to right. (C) The actual amplitude spectrums of the two stimulus tones recorded digitally at the surface of the headphones in Experiment 1.

Results

Experiment 1: Visual motion aftereffect contingent on a complex tone. In order to measure the magnitude of the aftereffect, we compared a point of subjective stationarity (PSS) before and after the adaptation phase. In the adaptation phase, participants were repeatedly presented with visual apparent motion produced by a pair of white circles placed side by side. Each of the two circles was synchronously accompanied by a unique complex tone. The tones were created by removing a different sound frequency of an original complex tone consisting of 9 frequency components (Fig. 1C). A component at 904 Hz was removed for one complex tone and 1344 Hz for the other. The strength of the illusory visual motion was quantified by a motion-nulling procedure before and after a 15-minute exposure to the apparent motion with the tone bursts (i.e., pre- and post-test sessions, respectively). The task of the participants was to determine the perceived direction of motion of the visual stimulus, which shifted in horizontal direction at various distances (0.12° , 0.24° , 0.48° , or 0.96°) from left to right or vice versa. Based on the psychometric functions obtained, we determined the amount of visual displacement that corresponded to the PSSs; the 50% point (the point of subjective equality) was estimated by fitting a cumulative normal-distribution function to each individual's data using a maximum likelihood curve fitting technique.

The results revealed that the tones acquired driving effects for visual motion after the 15-minute exposure (Fig. 2A, B). The PSS

shifted in the direction of the leftward visual motion when, during the exposure, the first visual stimulus was synchronized with the tone accompanied by the leftward stimulus, and the second stimulus with the tone accompanied by the rightward stimulus (rightward sound condition). However, the PSS shifted in the direction of the rightward visual motion when the sound sequence was reversed (leftward sound condition). A two-way analysis of variance (ANOVA) showed that the interaction between the exposure and the sound condition was significant ($F_{2,18} = 4.12$, $P = 0.032$). Post hoc tests (Tukey's HSD) revealed the significant differences in PSS between the right-

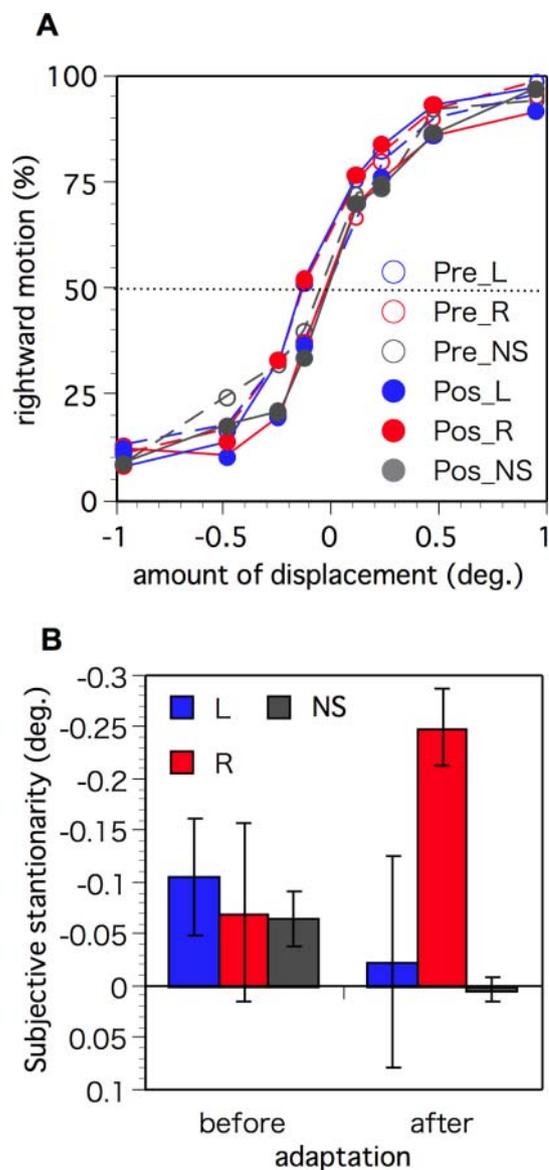


Figure 2 | Sound-contingent visual motion aftereffects in Experiment 1. (A) The proportion of rightward motion perception of visual stimuli as a function of the amount of physical displacements of visual stimuli. Positive values indicate rightward displacement in the horizontal axis, and negative values indicate leftward displacement. Blue lines represent the results for the leftward sound condition, and red lines for the rightward sound condition. Open circles represent the results obtained before adaptation. Filled circles represent the results obtained immediately after adaptation. (B) The points of subjective stationarity (PSS). Positive values represent the shift of PSS in the direction of the leftward visual motion. Blue bars represent the results for the leftward sound condition, red bars for the rightward sound condition, and gray bars for no sound condition. The error bar denotes the standard error of the mean.



ward and leftward sound conditions ($P = 0.002$). These results indicated that, after prolonged exposure to visual apparent motion with the complex tones, the tones became drivers for illusory motion perception.

A follow-up investigation was conducted to test whether the effect of adaptation occurred particularly for adapted sound frequency or not in the sound-contingent visual motion aftereffect. While two tones used in Experiment 1 were presented for the adaptation, new two tones were presented for the pre- and post-tests. The new tones were generated by removing a component at 819 Hz for one tone (CtoneA') and 1217 Hz for the other tone (CtoneB') from the originally generated complex tone. We confirmed that the aftereffect could not transfer between these tone pairs (Supplementary Figure S1). This result suggests that the the sound-contingent visual motion aftereffect could occur particularly for the adapted sound frequency.

Pre- and post-pitch discrimination between two complex tones.

Discriminability in pitch for complex tones was examined before and after the main experiment (Fig. 3). We used a two-alternative forced-choice procedure to measure the correct percentage of pitch discrimination for two complex tones used in the main experiment. The two complex tones were sequentially presented, and participants determined the interval in which the higher complex tone containing the 1344-Hz component (higher in pitch) was presented. We set nine conditions for the amplitudes of the components on both sides of the removed component by reducing in steps of 1 dB from 0 to 8 dB. This was because the attenuation of the target frequencies was maximum in amplitude for the undiscriminable sounds used in the adaptation and test phases. As can be seen in Fig. 3, the performance for the tones with 0 dB reduction, which was consistent with those used in the main experiment, did not improve at all after the exposure ($t(9) = 0.014$, $P = 0.91$). The correct answer rate for these tones did not exceed chance level before and after the exposure (two-tailed binomial test, before: $P = 0.27$, after: $P = 0.23$). This result clearly indicated that, even though the two tones were mutually indistinguishable in pitch, these tones were able to determine the direction of illusory visual motion after the adaptation.

Cross-adaptation. Although the two complex tones were indistinguishable in pitch, these tones did affect visual motion perception. One potential underlying mechanism could be that the presence or absence of the removed components between the tones

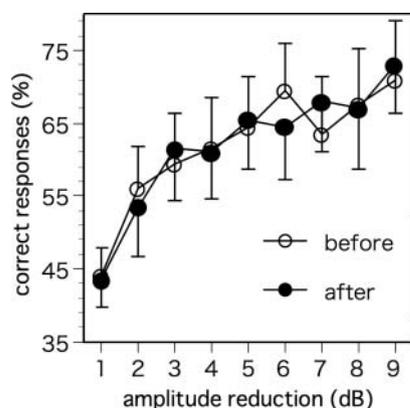


Figure 3 | Discriminability of complex tones. Percentages of correct responses were plotted against the magnitude of amplitude reduction for two neighboring pure tones of the deleted component. Open circles indicate results obtained before exposure while closed circles after exposure. The performance is almost identical before and after the exposure. The error bar denotes the standard error of the mean.

was implicitly coded in the brain so that the tones exerted influence on the visual motion perception. To examine this possibility, we presented two pure tones, each of which were uniquely contained in the adapted complex tones, before and after the exposure to visual apparent motion with the complex tones. Before the exposure, the pure tones did not affect visual motion perception. However, the pure tones acquired driving effects for visual motion after the exposure (Fig. 4). It should be noted that, although the presence or absence of the specific components in the complex tones were not perceived explicitly during the exposure, they could determine the direction of the illusory visual motion. The PSS shifted in the direction of the leftward visual motion, when the first visual stimulus was synchronized with the pure tone accompanied by the leftward stimulus during the exposure, and the second stimulus with the tone accompanied by the rightward stimulus (rightward sound condition). However, the PSS shifted in the direction of the rightward visual motion, when the sound sequence was reversed (leftward sound condition). A two-way ANOVA showed the PSSs were significantly different for sound conditions ($F_{2,10} = 4.156$, $P < 0.05$), and that the interaction between the exposure and the sound condition was significant ($F_{2,10} = 12.08$, $P < 0.005$). Post hoc tests (Tukey's HSD) revealed the significant differences in PSS between the rightward and leftward sound conditions ($P < 0.05$). These results showed that, after prolonged exposure to visual apparent motion with complex tones, the pure tones also became drivers for illusory motion perception.

Experiment 2: Indiscriminable noise-contingent visual motion aftereffect.

Although participants could not discriminate the tones containing the specific frequency component in the pitch discrimination test, it could be that they used clues to differentiate between these two tones. We further investigated whether the driving effect could be replicated by using highly complex, indiscriminable sounds. The new sounds were created by applying peak and notch filters centered at 500 Hz and 2000 Hz (or vice versa), respectively, to white noise.

The results revealed that the noise acquired driving effects for visual motion after the exposure (Fig. 5A, B). The PSS shifted in the direction of the leftward visual motion, when the first visual stimulus was synchronized with the noise accompanied by the leftward stimulus during the exposure, and the second stimulus with the tone accompanied by the rightward stimulus (rightward sound condition). However, the PSS shifted in the direction of the rightward visual motion, when the sound sequence was reversed (leftward sound condition). A two-way ANOVA showed that the interaction

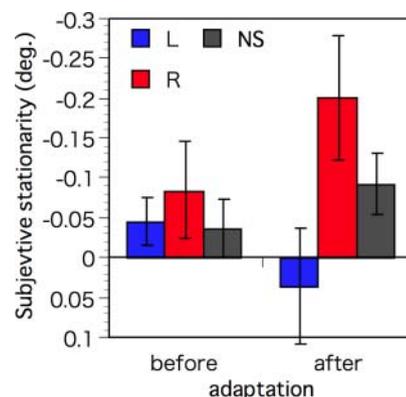


Figure 4 | Cross adaptation. Positive values represent the shift of PSS in the direction of the leftward visual motion. Blue bars represent the results for the leftward sound condition, red bars for the rightward sound condition, and gray bars for no sound condition. The error bar denotes the standard error of the mean.

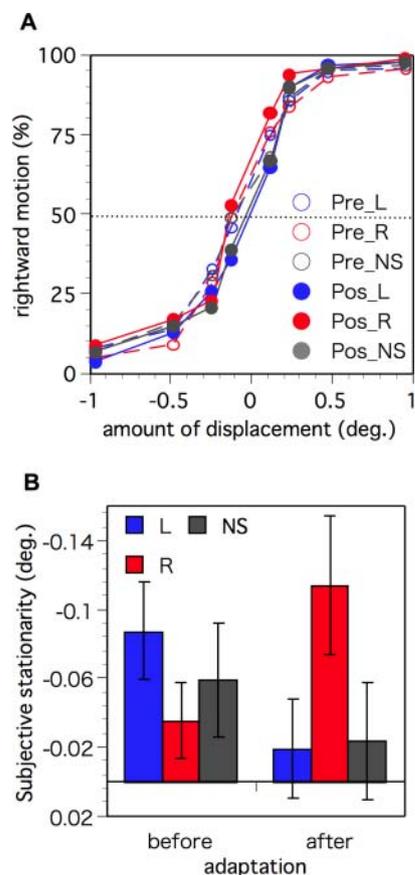


Figure 5 | Sound-contingent visual motion aftereffects in Experiment 2. (A) The proportion of rightward motion perception of visual stimuli as a function of the amount of physical displacements of visual stimuli. Positive values indicate rightward displacement in the horizontal axis, and negative values leftward displacement. Blue lines represent the results for the leftward sound condition, and red lines for the rightward sound condition. Open circles represent the results obtained before adaptation. Filled circles represent the results obtained immediately after adaptation. (B) The points of subjective stationarity (PSS). Positive values represent the shift of PSS in the direction of the leftward visual motion. Blue bars represent the results for the leftward sound condition, red bars for the rightward sound condition, and gray bars for no sound condition. The error bar denotes the standard error of the mean.

between the exposure and the sound condition was significant ($F_{2,18} = 4.86, P = 0.021$). Post hoc tests (Tukey's HSD) revealed significant differences in PSS between the rightward and leftward sound conditions ($P = 0.01$), and PSS between the leftward and no-sound conditions after the adaptation phase ($P = 0.014$). These results indicated that sounds can be implicitly contingent on visual motion perception even when these sounds were highly complex and indistinguishable.

Discriminability for the noise was examined before and after the main experiment. The method of constant stimuli was used with a two-interval forced choice procedure. Two noises with different peak/notch filters were presented successively in either the first or the second interval. In the second interval, two noises with the same peak/notch filter were presented successively. The task of the participants was to determine which interval contained noise with the same filter. The performance did not improve after the exposure ($t_{10} = 0.58, P = 0.46$). The correct answer rate did not exceed chance level (corrects answer, before: 51.1 %, after: 55.5 %; binomial test, before: $P = 0.92$, after: $P = 0.32$). These results confirmed that the noises were indistinguishable even after the exposure.

Experiment 3: Eye selectivity in the sound-contingent visual motion aftereffect. The sound-contingent visual motion aftereffect could be established implicitly, indicating that relatively lower perceptual processing would be involved. Here, we examined whether the aftereffect can transfer across eyes. The participants' eye was covered and the eye exposed to the stimuli was different or not between the adaptation and test phases. The auditory and visual stimuli were the same as the one in Experiment 1.

The results revealed that there was no interocular transfer of the aftereffect (Fig. 6A, B). Whereas the shift in PSS was observed when the eye exposed to the stimuli was consistent between the adaptation and test phases (same-eye condition), PSS did not change when the eye exposed to the stimuli was inconsistent between the phases (different-eye condition). A three-way ANOVA showed that the interaction among adaptation (before or after), sound and eye (same or different) condition was significant ($F_{2,10} = 4.64, P = 0.038$). The post hoc test showed the significant differences in PSS between the rightward and leftward sound conditions ($P = 0.013$), and PSS between the leftward and no-sound conditions in the same-eye condition after the adaptation phase ($P = 0.021$).

The existence of eye selectivity in the sound-contingent visual motion aftereffect would suggest that the implicit association between audio and visual signal could be established at very early perceptual stage in visual processing.

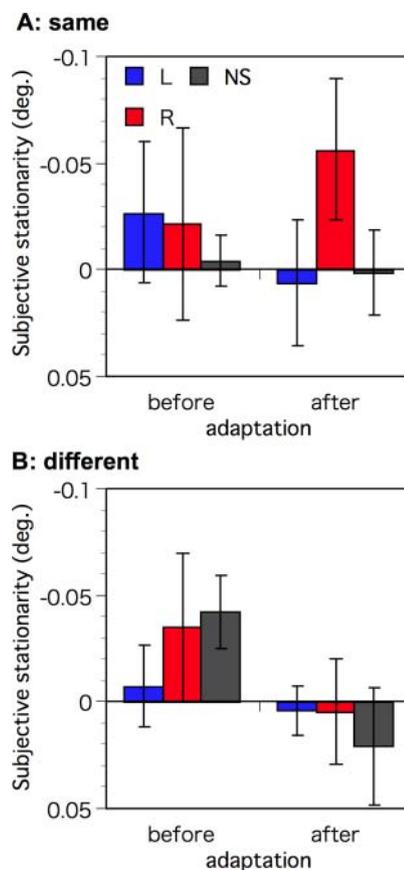


Figure 6 | Sound-contingent visual motion aftereffects in Experiment 3. Positive values represent the shift of PSS in the direction of the leftward visual motion. Blue bars represent the results for the leftward sound condition, red bars for the rightward sound condition, and gray bars for no sound condition. The error bar denotes the standard error of the mean. (A) The PSS in the same-eye condition. (B) The PSS in the different-eye condition.



Discussion

The present study demonstrates that prior adaptation to visual apparent motion paired with two alternating and indiscriminable complex tones results in illusory apparent motion of a static visual stimulus, where the perceived direction depends on the order in which the sounds are replayed. We also ensured that the uniquely contained frequency-component in each complex tone was coded and associated with visual motion perception. One might assume that the difference between the two sounds was learned during the exposure. For example, the discrimination between complex tones becomes better remembered by perceptual learning^{7,8}. It was also shown that perceptual learning can occur as a result of the exposure to a subliminal stimulus in the visual domain^{9–11}. However, discriminability for the sounds was not improved even after the exposure. These results indicate that activation of the human auditory system without reaching consciousness can not only drive illusory visual motion but also determine the direction of motion.

In our previous study, two easily discriminable tones were presented in conjunction with visual left-right apparent motion so that the participants could explicitly relate each tone to each location of the visual stimulus¹. This manipulation allowed participants to control their intention and/or attention. The present results indicate that such top-down controls are not necessary to observe the sound-contingent motion aftereffect. Recent studies have reported that perceptual learning can occur in situations that lack attention, awareness, and reinforcement e.g.^{9,10,12,13}, as well as under explicit training conditions. These findings suggest that a key to the establishment of new representations is sensory stimulation that can sufficiently drive the neural system past the point of a learning threshold¹⁴, but not top-down controls. In addition to the undiscriminability of our auditory stimuli, the aftereffect had a selectivity such that the effect of the adaptation did not transfer across the eyes. This implies the involvement of the perceptual processing before the integration of interocular information in the sound-contingent visual motion aftereffect. These findings could also exclude the possible engagement of top-down processes including response bias. Consistent with these previous studies, the current findings indicate that new neural representations between sound and visual motion perception can be established by unconscious perceptual learning.

Sensory processing stages responsible for cross-modal integration are a matter of debate across psychophysical studies. Some claim that multisensory interactions can be explained by decisional processes that occur after extensive unisensory processing^{15–17}, whereas others claim that interactions take place during early sensory processing^{6,18,19}. It is not possible at this point to specify the exact processing level in each sensory system at which this association takes place. However, the present findings and earlier studies using adaptation provide some clues concerning the processing stage by which audio-visual integration is created. For instance, our previous study reported that the sound-contingent visual motion aftereffect was well observed at the retinal position that was previously exposed to apparent motion with tone bursts¹. This finding indicates that the association between audio and visual modalities involves retinal position selective process in the visual system. In addition, the present findings also demonstrated that the aftereffect selectively occurred only for the adapted eye, suggesting that the association of audio-visual inputs was created at early perceptual processing stages. In line with a magnetoencephalographic study showing that indiscriminable complex tones elicit different brain activities²⁰, the present results reveal that audiovisual integration occurs at processing stages without a conscious perception of the auditory stimulus. Taken together, these findings suggest that audio-visual interactions and the establishment of new associations between auditory and visual information can occur at very early processing stages in both the visual and auditory sensory systems. Further research on the present effect may be promising.

Methods

All participants, including the study authors, reported normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained after the nature and possible consequences of the studies were explained. All procedures were approved by the local ethics committee of Tohoku University.

In a dark and quiet room, the participants wore headphones and were seated at a distance of 1 m from a 24 inch CRT display (refresh rate: 60 Hz).

In all the experiments, white circles (5.12 cd/m², 1.0° in diameter) were presented as visual stimuli on a black background. A red circle (0.4° in diameter; 17.47 cd/m²) was also presented for fixation. The auditory stimuli were sound bursts (sampling frequency 44.1 kHz, 85 dB SPL, 50 ms in duration with 5 ms rise and fall time) delivered to both ears through the headphones. The type of auditory stimuli were different for each experiment (see below). We confirmed that the onset of the visual and the auditory stimuli were synchronized using a digital oscilloscope.

To determine the amount of visual displacement that corresponded to a point of subjective stationarity (PSS), we estimated the 50% point (the point of subjective equality) by fitting a cumulative normal-distribution function to each individual's data using a maximum likelihood curve fitting technique. These PSSs were measured before and after participants were exposed to the visual apparent motion with tone bursts.

Procedure in experiment 1 (indiscriminable pitch sounds). 10 participants took part in experiment 1. For the auditory stimuli, two complex tones with eight components were presented. Two stimulus tones were made by removing one of the nine components, spaced at 1/7 octaves apart from 672 to 1638 Hz on the logarithmic scale. A component at 904 Hz was removed for one complex tone (CToneA) and 1344 Hz for the other (CToneB).

(i) *Adaptation and test.* In order to measure the magnitude of aftereffect, we compared a PSS before and after the adaptation phase. During adaptation, two white circles were placed side by side and presented in alternation at 7.5° and 12.5° to the right of the red fixation. The duration of each circle was 400 ms and stimulus onset asynchrony was 500 ms. For half of the participants, the onset of the leftward circle was synchronized to a tone burst of CToneA and the rightward circle to CToneB. For the remaining half, the onset relationship was reversed. Participants were asked to keep looking at the fixation and were exposed to the visual apparent motion with tone bursts for 15 minutes.

Each test block was preceded by a 6-minute top-up adaptation in order to maintain aftereffect through all the test trials. During the test phase, two circles (with 400 ms duration in each) were presented with 500 ms of a Stimulus Onset Asynchrony (SOA), synchronized with two tone bursts. In a rightward sound condition, the first visual stimulus was synchronized with a tone that was accompanied with the leftward stimulus during the exposure to apparent motion and the second stimulus with a tone that was accompanied with the rightward stimulus. In a leftward sound condition, the order was reversed. The no-sound condition was also included. The visual stimulus was displaced 0.12°, 0.24°, 0.48°, or 0.96° from left to right or vice versa. The amount of displacement and the sound condition were randomized from trial to trial. The observers were asked to judge whether the visual stimulus moved leftward or rightward. Twenty responses were obtained for each condition.

(ii) *Pitch discrimination.* To test the discriminability for complex tones, an additional 8 pairs of complex tones were generated. The amplitudes of the components on both sides of the removed component were reduced in steps of 1 dB from 0 to 8 dB. The reduction rate was decided by measuring sound pressure levels for a single pure tone. The complex tones of the selected pair were presented in random order for 50 ms with the SOA of 500 ms. The subjects were asked to judge which tone was higher in pitch. Twenty responses were obtained for each pair.

(iii) *Cross adaptation.* In another test session, two pure tones that were deleted from the complex tones were presented instead of the complex tones.

Procedure in experiment 2 (indiscriminable noise). 10 participants took part in experiment 2. The auditory stimuli were noises with peak and notch filters. The noises peaked at 500 Hz and notched at 2000 Hz, or vice versa. The peak or notch was about 1/8 octave wide and more than 8 dB above or below the base line, which was kept constant during stimulus presentation.

(i) *Pre-test, adaptation, and test.* We used the same procedure in experiment 1 to measure amplitude of the aftereffects and PSS shift, but we used the above noise as auditory stimuli.

(ii) *Same-different sound discrimination test.* To test the discriminability for these noises, we measured the percentage for the same-different discrimination task by using the constant method with a two-interval forced choice. Pairs of notch and peak noise were presented successively in either the first or the second interval in random order for 50 ms with the SOA of 500 ms. In the other interval, two similar noises (either the peak or the notch noise) were presented successively. Participants had to indicate the interval containing both peak and notch noises.

Procedure in experiment 3 (Eye selectivity). 6 participants took part in experiment 3. The visual stimuli were presented monocularly; whereas the stimuli were presented



to the participants' right eye in the adaptation session, they were presented to the right (same-eye condition) or left (different-eye condition) in the test phases. Except for these, the stimuli, apparatus, and procedures were the same with Experiment 1.

1. Teramoto, W., Hidaka, S. & Sugita, Y. Sounds move a static visual object. *PLoS ONE* **5**, e12255 (2010)
2. Hidaka, S., Teramoto, W. & Sugita, S. Alternation of sound location induces visual motion perception of a static object. *PLoS ONE* **4**, e8188 (2009)
3. Vroomen, J., Bertelson, P. & de Gelder, B. in *Out of Mind: Varieties of Unconscious*, eds de Geber BLMF, de Gelder, de Haan E, Heywood C (Oxford University Press), pp 140–150 (2001)
4. Jack, C. E. & Thurlow, W. R. Effects of degree of visual association and angle displacement on the “ventriloquism” effect. *Percept Mot Skills* **37**, 967–979 (1973)
5. Blaser, E., Papathomas, T. & Vidnyánszky, Z. Binding of motion and colour is early and automatic. *Eur J Neurosci* **21**, 2040–2044 (2005)
6. Soto-Faraco S., Spence C. & Kingstone A. Assessing automaticity in the audiovisual integration of motion. *Acta Psychol* **118**, 71–92 (2005).
7. Irvine, D. R., Martin, R. L., Klimkeit, E. & Smith, R. Specificity of perceptual learning in a frequency discrimination task. *J Acoust Soc Am* **108**, 2964–2968 (1992).
8. Demany, L. & Semal C. Learning to perceive pitch differences. *J Acoust Soc Am* **111**, 1377–1388 (2002).
9. Watanabe, T., Náñez, J. E. & Sasaki, Y. Perceptual learning without perception. *Nature* **413**, 844–848 (2001).
10. Seitz, A. R. & Watanabe, T. Is subliminal learning really passive? *Nature* **422**, 36 (2003)
11. Tsushima, Y., Sasaki, Y. & Watanabe, T. Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science* **314**, 1786–178 (2006).
12. Godde, B., Stauffenberg, B., Spengler, F. & Dinse, H. R. Tactile coactivation-induced changes in spatial discrimination performance. *J Neurosci* **15**, 1597–1604 (2000).
13. Pleger, B. *et al.* Shifts in cortical representations predict humans discrimination improvement. *Proc Natl Acad Sci USA* **98**, 12255–12260 (2001).
14. Seitz, A. R. & Dinse, H. R. A common framework for perceptual learning. *Curr Opin Neurobiol* **12**, 148–153 (2007).
15. Meyer, G. F. & Wuerger, S. M. Cross-modal integration of auditory and visual motion signals. *Neuroreport* **12**, 2557–2560 (2001).
16. Wuerger, S. M., Hofbauer, M. & Meyer, G. F. The integration of auditory and visual motion signals at threshold. *Percept Psychophys* **65**, 1188–1196 (2003).
17. Alais, D. & Burr, D. No direction-specific bimodal facilitation for audiovisual motion detection. *Brain Res Cogn Brain Res* **19**, 185–194 (2004).
18. Kitagawa, N. & Ichihara, S. Hearing visual motion in depth. *Nature* **416**, 172–174 (2002).
19. Vroomen, J. & de Gelder, B. Visual motion influences the contingent auditory motion aftereffect. *Psychol Sci* **14**, 357–361 (2003).
20. Hoshiyama, M., Okamoto, H. & Kakigi, R. Priority of repetitive adaptation to mismatch response following undiscriminable auditory stimulation: a magnetoencephalographic study. *Eur J Neurosci* **25**, 854–862 (2007).

Acknowledgements

This research was supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Specially Promoted Research (No. 19001004). We also gratefully appreciate Stephen Machnik, and anonymous two reviewers' constructive comments on earlier versions of the manuscript.

Author contribution

Conceived and designed the experiments: MK WT SH YS. Performed the experiments: MK WT SH YS. Analyzed the data: MK YS. Contributed reagents/materials/analysis tools: MK WT YS. Wrote the paper: MK WT SH YS.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

How to cite this article: Kobayashi, M., Teramoto, W., Hidaka, S. & Sugita, Y. Indiscriminable sounds determine the direction of visual motion. *Sci. Rep.* **2**, 365; DOI:10.1038/srep00365 (2012).