RESEARCH ARTICLE

Gender bending: auditory cues affect visual judgements of gender in biological motion displays

R. van der Zwan · C. MacHatch · D. Kozlowski ·
N. F. Troje · O. Blanke · Anna Brooks

Abstract The movement of an organism typically provides an observer with information in more than one sensory modality. The integration of information modalities reduces the likelihood that the observer will be confronted with a scene that is perceptually ambiguous. With that in mind, observers were presented with a series of point-light walkers each of which varied in the strength of the gender information they carried. Presenting those stimuli with auditory walking sequences containing ambiguous gender information had no effect on observers' ratings of visually perceived gender. When the visual stimuli were paired with auditory cues that were unambiguously female, observers' judgments of walker gender shifted such that ambiguous walkers were judged to look more female. To show that this is a perceptual rather than a cognitive effect, we induced visual gender after-effects with and without accompanying female auditory cues. The pairing of gender-neutral visual stimuli with unambiguous female auditory cues during adaptation elicited male after-effects. These data suggest that biological

R. van der Zwan · C. MacHatch · D. Kozlowski · A. Brooks (✉)
Laboratory of Cognitive Neuroscience and Behaviour,
Department of Psychology, Southern Cross University,
Coffs Harbour, NSW 2450, Australia
e-mail: anna.brooks@scu.edu.au

N. F. Troje
Department of Psychology and School of Computing,
Queen's University, Kingston, ON K7M 3N6, Canada

O. Blanke
Laboratory of Cognitive Neuroscience, Brain-Mind Institute,
Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne,
Switzerland

motion processing mechanisms can integrate auditory and visual cues to facilitate the extraction of higher-order features like gender. Possible neural substrates are discussed.

Introduction

An invaluable social skill is the ability quickly and accurately to identify the genders of those around us, particularly those with whom we are about to interact. Researchers are slowly exploring the perceptual processes and their neural substrates, mediating that ability in humans. For example, it has been demonstrated that it is possible to identify gender based solely on biological motion cues: using point-light walkers (PLWs; Johansson 1973) a number of groups have shown that the gender of a moving human readily can be identified by human observers even when the observers have no access to form-based features such as facial configuration, hairstyle, or clothing (Kozlowski and Cutting 1977; Mather and Murdock 1994; Runeson and Frykholm 1983; Troje 2002). That remarkable *visual* ability has parallels within the auditory modality. Using only recordings of the foot-strikes of a range of female and male actors as stimuli, Li et al. (1991) demonstrated that gender was correctly identified on 72% of trials. In the case of both visual and auditory tasks, sensitivity to gender cues seems to be independent of *observer* gender (Brooks et al. 2008; Li et al. 1991).

Sensitivity to biological motion cues in both the visual and auditory modalities is in keeping with the natural conditions under which such motion usually is observed; environmental cues about others are very often multisensory and the sight of someone walking frequently is paired with

auditory information about their motion. With that in mind, it is perhaps surprising that only recently the multisensory interactions have become a focus within the biological motion literature. For example, we recently reported that auditory information predictably influences vision-based biological motion sensitivity (Brooks et al. 2007). More specifically, auditory motion cues seem to affect observers' abilities to detect PLWs in noisy visual environments. Similar findings from other groups are now emerging (Burr et al. 2007; Meyer and Wuerger 2007) to support the contention that like perceptions of non-biological motion (Kawachi and Gyoba 2006; Sekuler et al. 1997; Watanabe and Shimojo 2001; Zhou et al. 2007) biological motion processes can also be mediated by audio-visual interactions.

Here we report an investigation into the effects of multisensory interactions, and in particular audio-visual interactions, on gender perceptions arising from biological motion processing. The aim of these experiments was to determine whether gender information in one domain can affect perceived gender in another. More specifically, our experiments were designed to investigate the potential effects of suprathreshold, gender-specific auditory cues (auditory walking, or foot-strike sequences) on visual biological motion perceptions. Our data suggest that auditory cues can affect visually perceived gender: auditory walking cues judged to be female make gender-ambiguous visual PLWs look more female. Further, visually adapting to such "perceptually" female walkers gives rise to gender aftereffects such that visually neutral PLWs look more male. We interpret that pattern of results as evidence for the existence of multisensory gender-tuned neurons located within multisensory cortex.

Experiment 1

Given that our perceptions of others are usually mediated simultaneously through more than one sense, it seems reasonable to speculate that multisensory cues to gender will interact to affect perceived gender. Before work can be done to determine where such interactions might occur in the cortex, it is necessary first to demonstrate that they do occur and do affect gender perceptions. Thus, the aims of "Experiment 1" were to identify an auditory walking sequence that was unambiguous in terms of its gender information, and then to measure, behaviourally, the effect on visual gender perceptions of pairing that auditory cue with visual PLWs.

Materials and methods

Participants

In this experiment, six observers completed the preliminary study of the auditory stimulus (three females and three males). Ten observers completed "Experiment 1" (six females and four males). Two observers were authors and those two authors completed all the three experiments. The remainder of the participants were naïve to the aims of this experiment, but all were experienced psychophysical observers. All had normal hearing and normal or corrected-to-normal vision.

Materials and design

Experiments were conducted in a light and sound attenuated cubicle. Visual stimuli were presented using a Pentium 4 processor paired with a Dell Triniton Flat Screen monitor (display resolution of 1,024 × 768 pixels) calibrated for luminance and set at a refresh rate of 100 Hz with 32 bit colour resolution. Auditory stimuli were presented using a Hitachi stereo amplifier HA330 attached to Athena loudspeakers (AS-B1-2B-1). The speakers were positioned adjacent to the visual display. Participants indicated their responses via keypresses on a Microsoft Wireless Multimedia Keyboard 1.0A. The monitor and the keyboard were situated so that the viewing distance from observer to screen was 57 cm.

Visual stimuli consisted of PLWs generated using custom-designed software (PointLightLab, v4.0.13) according to parameters established previously by Troje (2002, 2008). PLWs were shown in frontal orientation and were comprised of 15 individual point-lights positioned as if attached to the major joints of a human figure perambulating on a treadmill. Individual point-lights subtended a visual angle of 0.3°, with whole figures subtending on average 20.5° visual angle in height and 6.5° in width. Point-lights were darker than the background with a luminance of 0.4 cd/m^2 presented on a grey background of 12.0 cd/m^2. The Michelson contrast $[(L_{max} - L_{min})/(L_{max} + L_{min})]$ of the point-lights to the background was 0.94.

Using established techniques, the relative spatial position and trajectory of point-lights was manipulated such that individual PLWs occupied equidistant positions along a 'gender continuum'(see Troje 2002, 2008 and Fig. 1 below for a graphical representation). As described by Troje (2002, 2008) stimuli thus defined reflect a continuum of perceptual experience in terms of a linear classifier trained on a set of 50 male and 50 female individual walkers. Increasingly negative scores correspond to increasingly female walkers and increasingly positive scores to increasingly male walkers. More specifically, and as shown in Fig. 1, position 0 on the axis corresponds with the average or gender-neutral PLW. Position −6 corresponds to the PLW 6 standard deviations from the mean in the female range (i.e. perceived as being an exaggerated female walker), and position +6 with the PLW 6 standard deviations from the mean in the male range (i.e. perceived as being an exaggerated male walker). A total of 13 PLWs thus were generated.

Fig. 1 Pictorial representation of the 'gender' axis first described by Troje (2002). Stimuli were selected in ±1 standard deviation increments from an objectively average walker (stimulus 0). Details of the metric and associated assumptions are described elsewhere (Troje 2002, 2008). Also represented are static versions of 3 of the PLW groups: exaggerated female (−6), gender neutral or ambiguous (0) and exaggerated male (+6). Examples of some of the stimuli are provided in the supplementary material for this paper
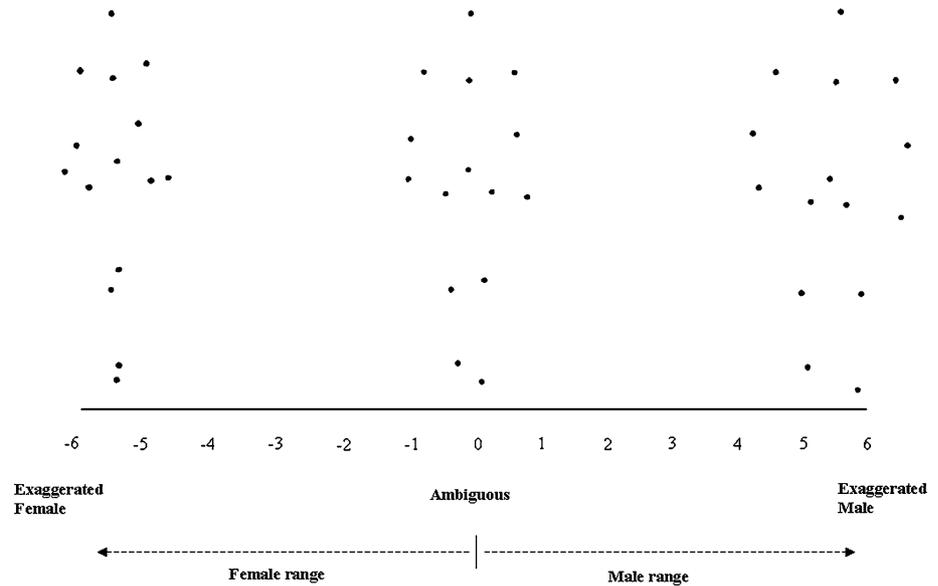
The temporal properties of the walking sequence of each PLW presentation were held constant across all conditions. Barclay et al. (1978) demonstrated that it takes an excess of 1,600 ms to extract the gender of a PLW if the walker completes two step cycles (four steps). With that in mind, the PLWs used in this investigation completed four steps of a walking cycle in 4,100 ms.

Auditory stimuli consisted of foot-strike sequences created from a stock library of high-resolution recordings of actual footsteps (supplied by the Laboratory for the Recognition and Organization of Speech and Audio, LabROSA, Columbia University). Each recording from the library depicted a single foot-strike (left or right) from a single individual (female or male), wearing a specified type of footwear, and on a specified surface. To standardise, as far as was possible, the quality of the auditory cues to be used in the multisensory experiments, audio sequences were constructed by selecting left/right foot-strike combinations only from concrete or wooden surfaces and only from walkers wearing hard-soled footwear. Matching left and right foot-strike pairs for individual walkers were spliced together to create walking sequences. Splicing and other manipulations of the recordings were done using Audacity 1.2.6 software to create four-step sequences as represented in Fig. 2. To maximise perceptual binding on *multisensory* presentations, the timing of the foot-strikes was matched to the footfall of the visual cues by temporally aligning the beginning of each foot-strike with the beginning of each footfall, the beginning of the low point of the path of each ankle dot. Such matching meant that inter-foot-strike interval was regular and held constant across conditions. Maximum loudness of the foot-strikes was matched across sequences (55 dB (A) fast scale) resulting in sequences that differed only in terms of their frequency composition and internal loudness profile. Aspects of both those dimensions likely carry gender information,

although further decomposition of the stimuli is required to elucidate which features are necessary and sufficient. Five such four-step sequences were created, three females and two males, and those sequences were labelled arbitrarily from 1 to 5 on the basis of a pilot study that surveyed perceived femaleness or maleness of each recording.

Procedure

In the preliminary auditory study, the five foot-strike sequences were presented to observers. Observers were required to report via a key press whether they perceived each of the walkers to be male (press "M") or not (press "Z"); the inverse task, female (press "M"), or not (press "Z"), yielded equivalent results as those reported here. Sequences were presented in random order in blocks of 50 presentations, which incorporated 10 repeats for each stimulus. Observers completed each block five times.

In the multisensory experiments, stimuli were presented to observers to measure the effects of the addition of auditory cues on visual perceptual experiences arising from the 13 point-light stimuli. Based on data from the preliminary study, only auditory foot-strike Sequence 1 was selected for use in "Experiment 1". Sequence 1 was most reliably identified as being female and was paired, in multisensory presentations, with each of the 13 PLWs. The task for observers in "Experiment 1" was to indicate the gender of the PLW. Observers performed five repeats of each audio-visual pairing in blocks consisting of a total of 130 trials.

Data analysis

In each experiment observers made a binary judgement, on each trial, about the gender of the point-light walker they
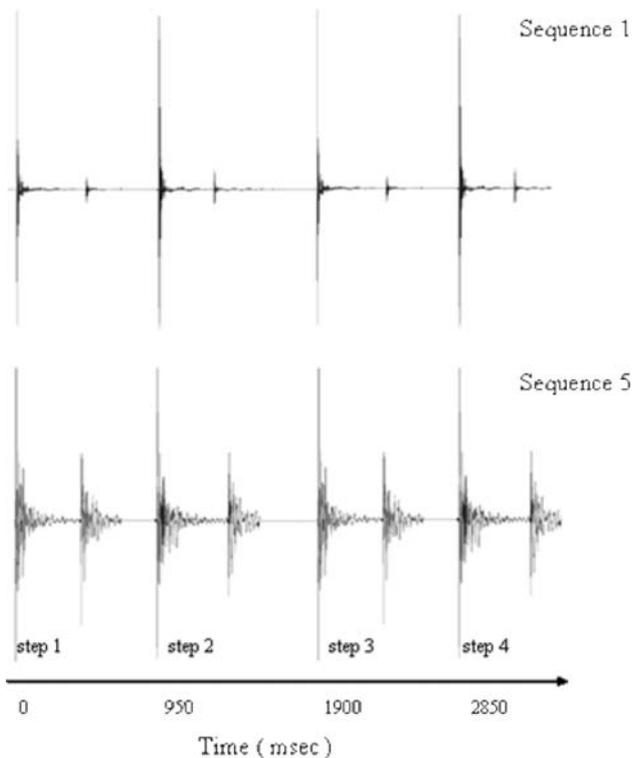
Fig. 2 Spectrograms of female and male foot-strike Sequences 1 and 5, respectively. Alternating left and right foot-strikes were spliced from individual recordings but always, within a sequence, from the same individual on the same surface in the same shoes. Each sequence was composed of four complete steps, each beginning with an initial impact sound that was always the loudest part of the foot-strike loudness profile. Maximum loudness equated for all auditory sequences. Inter foot-strike intervals were consistent (950 ms) and initial impact sounds were synchronised with the visual PLWs so that they always occurred at the point in the walking pattern that corresponded with the walkers' feet hitting the ground

could see. Judgements were then recorded, for each observer on each stimulus, as a proportion such that the number of times a given stimulus appeared to be male divided by the total number of presentations of that stimulus. Means across subjects were then calculated and those descriptive statistics are presented pictorially below (and as noted above, the same calculations were completed for judgements of femaleness). Comparisons of means were done using SPSS ver. 14.0 or using custom-written data processing functions in Microsoft Excel (2003: All functions were programmed according to the procedures described in Winer 1962). When deciding on the type(s) of analyses to conduct, an a priori decision was made to test statistically only performances at the statistical and the subjective mid-points. All previous work using the stimuli continuum used here suggested that the mean subjective mid-point for our participants would lie somewhere between those two values. Thus we tested both, expecting to see effects at the points of maximum ambiguity in each study. For this and the remaining experiments all reported analyses were repeated having discarded data from the authors with

no change in the pattern of results. Thus all analyses are reported for all observers.

Results

Preliminary data

The aim of the first study in this project was to obtain an auditory stimulus that was unambiguous in terms of its gender. Two criteria were set to establish such a stimulus: the auditory cue had to be generated by a walker that was the same gender as the perceived gender (a female walker judged by observers to sound female or a male walker judged by observers to sound male) and the stimulus had to be reliable, eliciting the same gender perception in all (or most) observers. Thus, this first study determined psychophysically which of five auditory walking sequences most reliably was judged to be female and which most reliably was judged to be male. As noted above, we categorised the five auditory sequences from most female-sounding (Sequence 1) to most male-sounding (Sequence 5) and report here the proportion of times each stimulus was judged by six observers to be male. Walking Sequence 1 perceptually was the most female, judged by all observers to be a female on all trials (mean proportion of "male" responses = 0.0 ± 0.0). Sequences 2–4 were judged incrementally to sound more and more male (Sequence 2: 0.06 ± 0.4; Sequence 3: 0.51 ± 0.15; Sequence 4: 0.81 ± 0.12). Finally, Sequence 5 was most often judged to be male but not always so (proportion of "male" judgements = 0.95 ± 0.01 SE). With that in mind, Sequence 1 was selected for use as the auditory female cue in multisensory presentations used in "Experiment 1".

Having established which auditory walking sequence was to be paired with the visually presented PLWs observers were presented, on each trial, with a stimulus selected randomly from one of two conditions: The first condition was unisensory and required observers to discriminate the gender of the visual PLWs presented without any corresponding auditory cue. The other condition was multisensory with the visual stimuli presented simultaneously with the female auditory walking sequence. Mean proportions of "male" judgements were calculated for each of the 13 visual stimuli for each condition and those data are shown in Fig. 3a. The unisensory visual stimuli reliably were judged to be either female or male at each end of the gender continuum (Fig. 3a, filled circles) and changed from one to the other, according to a linear integration function, as the visual stimuli incremented along the gender continuum. As reported elsewhere for these particular inducing stimuli, subjective gender ratings of the objective gender continuum show a slight "male" bias (Troje et al. 2006). That is, the statistically gender neutral walker (walker 0) was judged by observers significantly more often to be male (comparing the observed proportion of male responses = 0.732 to the point of subjective ambiguity 0.5; $t_9 = 2.20$, $P < 0.05$). Further, these
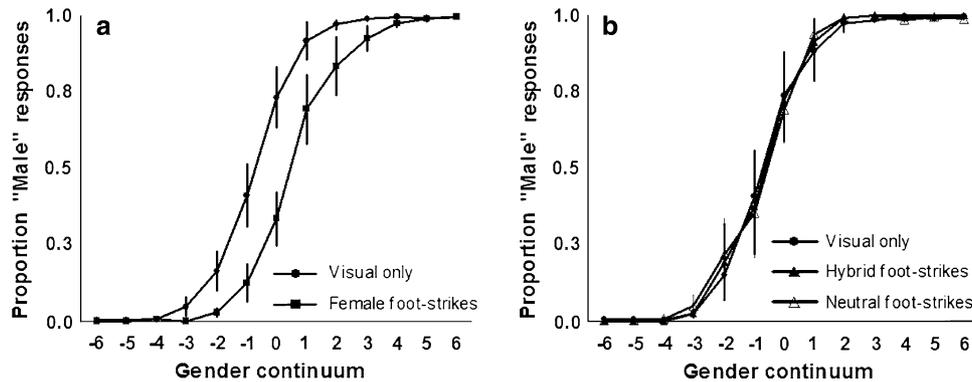
Fig. 3 **a** A comparison between visually perceived genders arising from uni- and multisensory presentations. Judgements are expressed as proportions of "male" responses to each PLW under each condition. Presentations of the visual PLWs alone gave rise to a systematic change in the perceived gender of the walkers such that the extreme female walkers never were judged to be male while the extreme male walkers were always judged to be male (*solid circles*). Between those extremes more ambiguous PLWs elicited a systematic change in the visually perceived gender as the walkers moved from the female to the male end of the continuum. The pairing of female auditory walking sequences with visual PLWs altered visual perceptions of the gender-

ambiguous walkers, but not of the extreme walkers (*solid squares*). The more gender-ambiguous the PLWs the more often they were perceived as female when paired with the auditory female walker. **b** Multisensory presentations in which visual PLWs are paired with hybrid male/female or female/male foot-strike sequences had no effect on the perceived gender of the walkers (*solid triangles*) when those judgements were compared to the unisensory judgements (*solid circles*). Similarly, the gender-ambiguous auditory walking sequence, 3, elicited no changes in visually perceived gender (*open triangles*). *Bars* show ±1 SE of the mean

data suggest that the walker one standard deviation below the objective mean of the gender continuum (walker −1) was subjectively the most gender-ambiguous walker, giving rise to equal numbers of female and male judgements (comparing the observed proportion of male responses = 0.412 to the point of subjective ambiguity 0.5; $t_9 = -0.83$, $P > 0.05$). In those respects these data replicate findings reported elsewhere (Brooks et al. 2008; Troje et al. 2006).

Pairing visual stimuli with the auditory female walker (Fig. 3a, filled squares) did affect observers' judgements of gender. The change from perceptually female to perceptually male matched the unisensory function, but this time the function was shifted such that the more perceptually ambiguous walkers (those in the middle of the continuum) were more often judged to be female. Tests for differences between the means for walker 0 show that the number of "male" judgements was significantly greater when walker 0 was shown without being paired with the auditory stimulus than when it was paired with the female auditory walking sequence ($t_9 = 4.18$, $P < 0.01$). Similarly for the subjectively ambiguous walker −1 the number of "male" judgements decreased when that walker was paired with the female auditory walking sequence ($t_9 = 2.91$, $P < 0.01$). That is, these data provide evidence that each of these most visually ambiguous PLWs was visually judged more often to be female, when paired with an unambiguously female auditory walking stimulus.

Discussion

In this experiment, we examined the effect of pairing an unambiguous auditory cue to gender with visual PLWs that

varied in the strength of their gender information. The reported gender of the most ambiguous visual walkers was affected by a simultaneously presented unambiguously female auditory walking sequence such that the visually ambiguous walkers more often were judged to be female. One interpretation of these data is that when higher-order features of a visual biological motion stimulus are ambiguous, relevant auditory information sometimes can be used to help resolve the ambiguity.

There are, however, a number of other explanations for these data that must be dismissed before that conclusion can be verified. For example, the pattern of results reported here might have arisen as a consequence of pairing an auditory walking cue with a visual stimulus depicting a human walker. That is, these results might be mediated by the integration of visual and auditory information alone and not arise at all from the processing of gender. Alternatively, these effects might reflect some type of response bias or priming effect arising from the presentation of the auditory cue. That is, these effects might be the result of integration of the auditory and visual information at the output level rather than from truly integrative sensory processes mediating perceptions. The next experiments tested those alternative accounts.

Experiment 2

In "Experiment 1" the auditory female walking stimulus seemed to have affected perceptions of gender-ambiguous PLW stimuli. That result may have arisen from a number of

possible interactions other than truly audiovisual interactions at a perceptual level. For example, the effects observed might be attributable to some properties of the multisensory presentation of auditory and visual stimuli (probability summation) and not to the integration of perceived gender in the auditory domain with that of the perceived gender of the visual stimulus.

With that in mind, observers in "Experiment 2" were presented with multisensory control stimuli. The first control stimulus was a so-called "hybrid" sequence that contained structural information from both female and male auditory cues: auditory walking sequences were composed of alternating perceptually female and perceptually male (or male and female) foot-strikes taken from auditory walking Sequences 1 and 5. We predicted that if it was some type of information in the auditory walking sequence other than gender information per se that gave rise to effects observed in "Experiment 1", visual perceptions of gender again would be affected by multisensory presentations with such auditory cues. Conversely, if the effects observed in "Experiment 1" were attributable to coherent gender information, the hybrid sequence would have no effect on visual perceptions. "Experiment 2" also examined the effect of a second gender-ambiguous auditory stimulus; auditory walking Sequence 3 from the preliminary auditory study. The preliminary study showed that, on average, observers were unable to decide whether Sequence 3 depicted a female or male walker (no significant difference from chance performance; $P > 0.05$) and inspection of individual data showed that all observers tested vacillated in their responses to this stimulus. We predicted that if the effects observed in "Experiment 1" arose from the combination of auditory walking sequences with visual PLWs via probability summation or other similar process, rather than from integration of specific gender information, those effects again should be observed here.

Materials and methods

Participants

Six observers completed "Experiment 2" (three females and three males). Two observers were authors. The remainder of the participants were naïve to the aims of the experiments but experienced psychophysical observers. All had participated in "Experiment 1", but had received no feedback as to their performances, and all had normal hearing and normal or corrected-to-normal vision.

Apparatus and stimuli

So-called hybrid auditory sequences were constructed for "Experiment 2". The hybrid sequences were composed of four alternating female and male foot-strikes spliced as if coming from a single walker. As the aim of this experiment was to determine whether any auditory cue was sufficient to generate the effects observed in "Experiment 1", rather than to test for gender effects, left and right foot-falls were taken from Sequence 1 and from the previously discarded Sequence 5 and arranged into recordings such that gender (female/male, male/female) and foot order (left/right, right/left) completely were counterbalanced. The gender-neutral walking Sequence 3 was the other auditory cue used in this experiment. All auditory cues were paired with the 13 visual PLWs. All other parameters were as for "Experiment 1".

Task and procedures

These were as described for "Experiment 1".

Results and discussion

Six observers were presented, in two different sessions, with the hybrid then gender-neutral auditory stimuli (or with the gender-neutral and then the hybrid stimuli) paired with the visual PLWs. Again observers were asked to indicate the gender of the PLW on each presentation, and Fig. 3b shows there was no effect on the visual judgements of the gender of the walkers induced by either the hybrid auditory sequences (solid line, filled triangles) or the gender-neutral auditory sequence (solid line, open triangles). These two sets of data provide evidence that the effects observed in "Experiment 1" do not arise simply as a product of the general integration of auditory and visual information and point to the importance of the gender loading of the auditory cue in the multisensory integration of gender perception in "Experiment 1".

Experiment 3

The results from "Experiment 2" suggest the effects observed in "Experiment 1" do not arise just from pairing visual PLWs with auditory walking sequences. Those data do not, however, preclude the possibility that the effects observed arose from processes mediating observers' responses rather than by processes mediating their perceptions. That is, the effects reported in "Experiment 1" might be some type of response bias at output rather than a truly perceptual integration of the visual and auditory cues.

To test for that possibility in "Experiment 3" we induced gender aftereffects like those described by Jordan et al. (2006) and by Troje et al. (2006). Both those groups demonstrated that prolonged viewing of a PLW of one gender can give rise to an aftereffect such that subsequently presented PLWs look more like the opposite gender. Those

results suggest gender perceptions are mediated by gender-tuned neurons or networks (Jordan et al. 2006). The location of such neurons or networks is unclear but both Jordan et al. (2006) and Troje et al. (2006) suggest extra-striate cortex as a possible location. Troje, however, does not dismiss the possibility that the adaptation effects might reflect adaptation of low-level motion processes, and the protocols used by that group do not preclude that explanation. That account seems unlikely, however, as Jordan did use so-called "dephased" stimuli to control for local motion effects and found no evidence to support the mediation of gender aftereffects at lower levels of processing (Jordan et al. 2006). Following that rationale, we reasoned that if the effects observed in "Experiment 1" are the result of perceptual (rather than cognitive) processing the point of multisensory integration would be, as argued above, in extrastriate cortex. If so, then adapting to pairings of auditory and visual stimuli that make an ambiguous PLW *look* more female should give rise to gender aftereffects such that test stimuli (ambiguous PLWs) should look more male. The aim of "Experiment 3" was to test that hypothesis.

Materials and methods

Participants

Ten observers completed "Experiment 3" (seven females and three males). Two were authors. The remaining participants, who were naïve to the aims of this experiment, were experienced psychophysical observers, and none had completed either of the other two experiments. All had normal hearing and normal or corrected-to-normal vision.

Apparatus and stimuli

The apparatus for this experiment was the same as that used in the two previous experiments. Visual adapting stimuli also were taken from PLW gender continuum used in the previous experiments. Three adapting stimuli were chosen from the continuum, the extreme female (−6), the extreme male (+6), and the gender neutral walker (−1). The auditory walking sequence was the same as that used in "Experiment 1".

Tasks and aprocedures

"Experiment 3" consisted of a total of 50 blocks of trials with each block comprised of five trials. Using timing and protocols described by Troje et al. (2006) each trial consisted of an adaptation phase during which the adapting stimulus was presented initially for 14 s. Then one of five test stimuli was presented and observers were required to signal the gender of the test walker. Following presentation

of each test stimulus, the adapting stimulus was again presented; this time for a 7-s top-up (van der Zwan and Wenderoth 1994) before a second test was presented and that protocol followed until the adapting stimulus had been used with each of the five test stimuli (see below). Presentations of adapting and test stimuli were separated by interstimulus intervals of 200 ms. Test stimuli were presented for 350 ms. Observers were required to fixate the centre of the hips of the adaptation stimulus and then to signal the gender of the test stimulus via a key press as in the previous experiments. The visual-only adaptation phase was comprised of 30 blocks, 10 blocks each for the three visual adaptation stimuli. Test stimuli were five walkers from the ambiguous range of the gender continuum (walkers −2, −1, 0, 1, 2). The audiovisual adapting stimulus was created by pairing the subjectively gender-neutral walker (−1) with the auditory cue used in "Experiment 1". On ten trials the adapting stimulus was only the gender neutral walker, and on ten others the adapting stimulus was the paired audiovisual stimulus.

Results

Figure 4 shows mean gender aftereffects for visual only and multisensory adapting stimuli. Using visual-only cues during adaptation, we were able to generate gender aftereffects like those described by Jordan et al. (2006) and Troje et al. (2006). As shown in Fig. 4a, adapting to an extreme female walker made the gender-ambiguous walkers appear more male (filled squares), when compared to adapting to a gender-neutral (PLW 0) walker. Similarly, adapting to an extreme male walker made the gender-ambiguous PLWs more female (filled triangles) compared to adapting to the gender-neutral walker. Inspection of the mean aftereffects as measured by the subjectively neutral test walker at −1 illustrates that point: following adaptation to the extreme female walker a strong male aftereffect was observed (mean proportion of "male" judgements = 0.85 ± 0.065), and that aftereffect was significantly "more male" than the aftereffect elicited by adapting to the objectively neutral (but perceptually slightly male) walker (PLW 0: mean proportion of "male" responses = 0.24 ± 0.050; $t_9 = 16.11$, $P < 0.01$). Similarly, adapting to the extreme male walker elicited a strong female aftereffect (mean proportion of "male" judgements = 0.03 ± 0.015), and that aftereffect was significantly "more female" than the aftereffect elicited by adapting to the objectively neutral walker (PLW 0; $t_9 = 3.71$, $P < 0.01$). This demonstrates that gender aftereffects can be elicited using the adaptation/test paradigm employed here.

Next, using the same adaptation/test paradigm, observers adapted to the gender neutral PLW paired with either no auditory cue or with the auditory cue used in
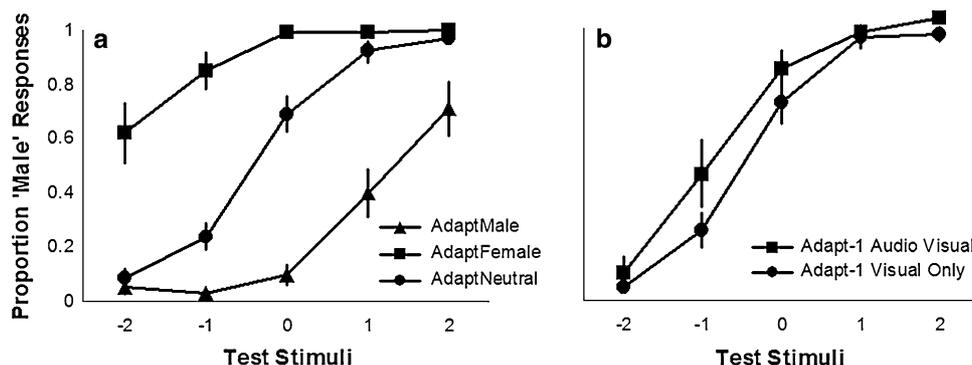
Fig. 4 Gender aftereffects induced with **a** visual-only adapting stimuli and **b** with audiovisual adapting stimuli. Both *graphs* show adapting to a gender-neutral walker gives rise to aftereffects such that "female" test PLWs mostly are judged to be "female" and "male" test PLWs are judged mostly to be "male" (*solid circles*). Figure 4a shows that adapting to an exaggerated female walker causes almost all the test PLWs to look more male (*solid squares*). Adapting to an exaggerated male walker causes almost all the test PLWs to look more female (*solid triangles*). Figure 4b shows that adapting to a gender-neutral PLW paired with an unambiguously female auditory walking sequence gives rise to aftereffects such that test PLWs are more often judged to be male (*solid squares*). Those aftereffects suggest that the unambiguous gender information in the auditory cue does affect the *perceived* gender of the PLW. *Bars* show ±1 SE of the mean

"Experiment 1" (Fig. 4b). We reasoned that if the change in reported gender really was perceptual then the pairing of the visual and auditory cues during adaptation should be equivalent to viewing a more female PLW, and so we should subsequently observe a male gender aftereffect. As shown in Fig. 4b, that was the case. Observers judged all five test stimuli as appearing more male following adaptation to the multisensory stimuli (the subjectively neutral PLW −1 paired with female footstrikes) than following adaptation to the visual stimulus alone. Again we did a mean comparison of the aftereffects elicited by the subjectively neutral PLW (−1); this time comparing aftereffects elicited by that visual stimulus alone and when paired with the female auditory cue. The aftereffect elicited by adapting to the neutral PLW was comparable in magnitude to that observed above (mean proportion of "male" responses = 0.25 ± 0.060). The aftereffect elicited by adapting to the same visual stimulus; this time paired with the unambiguous female auditory walking sequence was much more "male" (mean proportion of "male" responses = 0.45 ± 0.114) and significantly so ($t_9 = 2.30$, $P < 0.05$) than when adaptation was to the neutral PLW without the auditory cue.

Certainly the aftereffects induced by the multisensory stimuli are not as large as those induced by the extreme walkers presented in the visual-only domain. However, in these experiments we made no attempt to perceptually equate the strength of the uni- and multi-sensory adaptation stimuli. Nonetheless, these data do suggest that the audiovisual interactions between the unambiguous female walking sequence and the gender-ambiguous PLWs observed in "Experiment 1" are perceptual. Thus, while previously it has been shown that visually perceived gender can affect heard sounds (Johnson et al. 1999), we believe this to be

the first demonstration that gender information in the auditory domain can affect perceived gender in the visual domain.

General discussion

The primary aim of this work was to extend knowledge about the effects of auditory information on perceptions of visually defined biological motion stimuli. More specifically, we aimed to investigate the effects of unambiguous, gender-specific auditory information (perceptually *female* walking sequences) on visual judgements of PLW gender. In "Experiment 1", visual gender perceptions were compared across unisensory (vision-only) and multisensory (visual + auditory) conditions. The data reported for that experiment suggest visually perceived gender was significantly influenced by the auditory cues when visual gender was ambiguous. More specifically, when paired with auditory female walking sequences, gender-ambiguous PLWs more often were judged as representing female walkers than when those same PLWs were presented with no auditory stimulus. "Experiment 2" paired the PLW stimuli with two different types of gender-neutral walking stimuli. The results of the experiment showed that gender-neutral auditory cues did not significantly influence perceived PLW gender, suggesting that differences observed in "Experiment 1" are not simply the product of pairing auditory information per se with visually ambiguous walkers. Finally, "Experiment 3" suggests that the effects observed in "Experiment 1" are not some type of response bias or cognitive artefact. Gender aftereffects were induced first using extreme visual walkers. Then, using the same paradigm but with a multisensory adapting stimulus, a perceptually

female walker was used to induce a male-walker aftereffect. In summary, the results of these experiments provide support for the hypothesis that auditory cues carrying gender information can, under some conditions, influence visually perceived gender in biological motion displays, and particularly when there is gender-ambiguity in the visual stimulus.

Previous investigations have demonstrated audio-visual facilitation in lower-order tasks; detection of biological motion (Burr et al. 2007), and also in discriminating the direction in which a point-light figure moves (Brooks et al. 2007; Meyer and Wuerger 2007). The findings we report here are some evidence that the same types of interactions now seem to affect also representational qualities of figures in motion. While Troje et al. (2006) suggest that gender aftereffects might be generated by adaptation of low-level motion processes, Jordan et al. (2006) suggests an alternative interpretation. Jordan argues that gender aftereffects are evidence of extrastriate gender selective neurons or neural arrays, and the demonstration here of audiovisual interactions affecting perceived gender in such a way as to change aftereffects is support for their position. Certainly that interpretation is consistent with some models of the neural mechanisms mediating biological motion perceptions. For example, Giese and Poggio (2003) have proposed that the perception of biological motion arises via sequences of processes acting first to extract biological motion cues from the visual array, then discriminations of coherent directions of motion are made, and then finally more complicated analyses are completed. Those more complicated processes give rise to perceptions of representational properties of an actor, like gender.

In locating their proposed model onto a possible neural array, Giese and Poggio (2003) suggest that biological motion perception is associated with processes in specific cortical locations including the superior temporal sulcus (STS) and parts of the frontal lobe. STS and its posterior aspect in particular has been identified as being selectively involved in auditory motion processing (Howard et al. 1996), multisensory integration (Blake and Shiffrar 2007; Puce and Perrett 2003) and in visual and auditory biological motion processing (Bidet-Caulet et al. 2005). Similarly, Schubotz et al. (2003) reported that premotor cortex generally and lateral premotor cortex in particular are tuned for some aspects of auditory processing. More recently, Saygin (2007) demonstrated that premotor cortex plays a role too in biological motion perception. Thus, while the suggestion is purely speculative in the context of the data reported here, it does seem plausible that either or both STS and premotor cortex may contribute to these effects. Of course, confirmation of that hypothesis awaits further investigation, especially with neuro-imaging techniques.

In addition to neuro-imaging work additional psychophysical studies will also help to elucidate more about the types of neural processes mediating the effects reported here. We did not, for example, equate the gender-salience of the auditory walking sequences with the gender-salience of our visual stimuli. It is almost certainly the case that the sequence judged here to be most female lies some distance from a notional gender-neutral point in auditory space, and it would be informative to explore auditory gender space in the same way that Troje has explored visual-motion gender space (Troje 2002). Currently we are creating an audition-based gender continuum similar to the visual continuum used here. With those stimuli it will be possible to equate for salience not only our female and male auditory cues but also the salience of our intersensory (visual and auditory) cues. Then, issues such as whether the effects observed here (auditory gender information affecting visual gender perceptions) are reciprocal in nature (visual gender perceptions influencing auditory gender experience) can be addressed. Those issues are important because, while gender aftereffects have been reported by others, there are data to suggest that the effects are category specific (e.g. Kovacs et al. 2006) and do not generalise across modalities (Schweinberger et al. 2008). It is tempting to interpret those results as suggesting that gender is encoded only within modalities (visual, auditory, and so on) and even within categories within a modality. If that is so, then the findings reported here become difficult to explain. Such conclusions may be premature. The observed strength of any aftereffect is proportional to a number of factors including the salience of the cue being adapted, the duration of adaptation, and the proportion of all neurons tested that were also adapted. With that in mind, there are important differences between the methods used here and those used by others. For example, the effects reported here were elicited by adaptation periods almost three times longer than those used by others (Kovacs et al. 2006; Schweinberger et al. 2008). Similarly, the stimuli used here were abstract representations of gender categories (female walkers, male walkers, and so on), rather than specific examples from within categories. That difference may be a clue to understanding the nature of gender encoding in the cortex. Further exploration of those issues will help inform discussion as to what specific processes in extrastriate cortex are mediating the inter-sensory interactions that give rise to gender perceptions.

References

Barclay CD, Cutting JE, Kozlowski LT (1978) Temporal and spatial factors in gait perception that influence gender recognition. Percept Psychophys 23(2):145–152

Bidet-Caulet A, Voison J, Betrand O, Fonlupt P (2005) Listening to a walking human activates the temporal biological motion area. Neuroimage 28:132–159

Blake R, Shiffrar M (2007) Perception of human motion. Annu Rev Psychol 58:12.1–12.27

Brooks A, van der Zwan R, Billard A, Petreska B, Clarke S, Blanke O (2007) Auditory motion affects visual biological motion processing. Neuropsychologia 45:523–530

Brooks A, Schouten B, Troje NF, Verfaillie K, Blanke O, van der Zwan R (2008) Correlated changes in perceptions of the gender and the orientation of ambiguous biological motion figures. Curr Biol 18:R728–R729

Burr D, Arrihi A, Marini D (2007) Audio-visual integration in the perception of tap-dance. In: The international intersensory research symposium: perception and action, 2007

Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4:179–192

Howard RR, Brammer M, Wright I, Woodruff PW, Bullmore ET, Zeki S (1996) A direct demonstration of functional specialization within motion related visual and auditory cortex of the human brain. Curr Biol 6:1015–1019

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211

Johnson K, Strand EA, D'Imperio M (1999) Auditory-visual integration of talker gender in vowel perception. J Phon 27:359–384

Jordan H, Fallah M, Stoner GR (2006) Adaptation of gender derived from biological motion. Nat Neurosci 9:738–739

Kawachi Y, Gyoba J (2006) Presentation of a visual nearby moving object alters stream/bounce event perception. Perception 35:1289–1294

Kovacs G, Zimmer M, Banko E, Harza I, Antal A, Vidnyanszky Z (2006) Electrophysiological correlates of visual adaptation to faces and body parts in humans. Cereb Cortex 16:742–753

Kozlowski LT, Cutting JE (1977) Recognising the sex of walkers from a dynamic point light display. Percept Psychophys 21:575–580

Li X, Logan RJ, Pastore RE (1991) Perception of acoustic source characteristics: walking sounds. J Acoust Soc Am 90:3036–3049

Mather G, Murdock ML (1994) Gender discrimination in biological motion displays based on dynamic cues. Proc R Soc Lond B Biol Sci 258:273–279

Meyer G, Wuerger SM (2007) Auditory—visual integration for biological motion. Perception 36; ECVP Abstract Supplement

Puce A, Perrett D (2003) Electrophysiology and brain imaging of biological motion. Philos Trans R Soc Lond B Biol Sci 358:435–445

Runeson S, Frykholm G (1983) Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. J Exp Psychol Gen 112(4):585–615

Saygin AP (2007) Superior temporal and premotor brain areas necessary for biological motion perception. Brain 130:2452–2461

Schubotz RI, von Cramon DY, Lohmann G (2003) Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. NeuroImage 20:173–185

Schweinberger SR, Casper C, Hauthal N, Kaufmann JM, Kawahara H, Kloth N et al (2008) Auditory adaptation in voice perception. Curr Biol 18:684–688

Sekuler R, Sekuler AB, Lau R (1997) Sound alters visual motion perception. Nature 385(23):308

Troje NF (2002) Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. J Vis 2(5):371–387

Troje NF (2008) Retrieving information from human movement patterns. In: Shipley TF, Zacks JM (eds) Understanding events: how humans see, represent, and act on events. Oxford University Press, Oxford, pp 308–334

Troje NF, Sadr J, Henning G, Nakayama N (2006) Adaptation aftereffects in the perception of gender from biological motion. J Vis 6:850–857

van der Zwan R, Wenderoth P (1994) Psychophysical evidence for area V2 involvement in the reduction of subjective contour tilt aftereffects by binocular rivalry. Vis Neurosci 11:823–830

Watanabe K, Shimojo S (2001) When sound affects vision: effects of auditory grouping on visual motion perception. Psychol Sci 12:109–116

Winer BJ (1962) Statistical principles in experimental design. McGraw Hill, New York

Zhou F, Wong V, Sekuler R (2007) Multi-sensory integration of spatio-temporal segmentation cues: one plus one does not always equal two. Exp Brain Res 180:641–654