nyas13693 W3G-nyas.cls March 27, 2018 14:48

NYAS	nyas13693-2000079	Dispatch: March 27, 2018	CE:
Journal	MSP No.	No. of pages: 18	PE: Lia Zarganas

Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Special Issue: *Annals Reports*
ORIGINAL ARTICLE

Acceleration and deceleration at constant speed: systematic modulation of motion perception by kinematic sonification

Q Alfred O. Effenberg and Gerd Schmitz

Institute of Sports Science, Leibniz Universität Hannover, Hannover, Germany

Address for correspondence: Alfred O. Effenberg, Institute of Sports Science, Leibniz Universität Hannover, Hannover,
Q Germany. effenberg@sportwiss.uni-hannover.de

Many domains of human behavior are based on multisensory representations. Knowledge about the principles of multisensory integration is useful to configure real-time movement information for the online support of perceptuo-motor processes (motor perception, control, and learning). A powerful method for generating real-time information is movement sonification. Remarkable evidence exists on movement-acoustic real-time information being effective in behavioral domains (music training, handwriting acquisition, sports). Here, we investigate whether and how biological motion perception can be enhanced, substituted, or modulated by kinematic sonification, with a focus on pitch coding. We work with gross motor cyclic movements and investigate the effectiveness of pitch scaling and consistent transposition on audio–visual motor perception accuracy (Experiment A). Beyond that, a new kind of audiovisual stimulus with inconsistent pitch transposition is used to produce a directed modulation of the integrated audiovisual percept (Experiment B). Results from Experiment A indicate pitch being powerful for mediating kinematic information to enhance motor perception and substituting information between perceptual modalities, even exceeding visual performance. Beyond these findings, results from Experiment B indicate that visual estimations of movement velocity can be enhanced or reduced auditorily. Movement sonification used for reshaping intermodal adjustments should be a powerful new tool for subconsciously shaping human movement patterns in the future.

Keywords: biological motion perception; intermodal adjustment; motor rehabilitation; movement sonification; multisensory integration; multisensory representation

Introduction

Motor learning is based on motor perception and the emergence of adequate internal representations, the sensory-motor internal models.[1] Internal representations originate when appropriate movements are observed by others in mental simulations, via observational learning, and when new actions are executed more or less successfully by oneself.[2] A specific case of motor learning is given in musical training, where scholars benefit from the pure listening to a certain melody for motor performance, as soon as a functional linkage between actions and sounds had been acquired[3] or music-specific sensorimotor associations had been established.[4] Learning to play a musical instrument requires the fast integration of information from different perceptual modalities (kinesthetic, tactile, auditory, visual), as stated by Zimmerman and Lahav.[5] Even if the theory of internal models does not focus comprehensively on modality-specific questions, internal models rely—at least partially—on multimodal sensory streams and multisensory representations.[6–8] Extensive neurophysiological evidence on the integration of multisensory information down to the level of single neurons indicates a seamless integration of the senses, as well as a direct involvement of multisensory areas of the central nervous system (CNS) into motor regulation.[9–12] Even single multisensory convergence neurons in the deep layers of the superior colliculus integrate (afferent) visual, auditory, and proprioceptive input and affect orientation

doi: 10.1111/nyas.13693

and attention behavior via (efferent) motor output, as described by Stein and Meredith[10] for cats.

Behavioral research indicates a broad spectrum of effects based on multisensory integration or intersensory phenomena, such as the McGurk-effect[13] or, with regard to neurophysiological findings, the ventriloquism effect.[10] Also, common spatial references had been considered as a general principle for multisensory perception.[14] Besides fundamental audiovisual effects, more abstract audiovisual stimulus arrays had been used, as realized by Giard and Peronnet[15] with an object recognition task: Participants acted more accurately and rapidly when identifying audiovisual objects compared to a purely auditory or visual condition. Besides object recognition, multisensory learning can be more effective, as shown with an artificial direction detection task:[16] The audiovisual stimulus was beneficial, indicating the superiority of multisensory learning over unimodal settings.

Here, we focus on multisensory research using additional acoustic real-time information in the fields of music, sports, and rehabilitation to describe the coding and the emergent kind of information that is effective on human behavior. On that basis, we investigate the effectiveness of kinematic–acoustic information on movement velocity perception of observed gross motor cyclic—that is, breast-stroking—movements. To attain high external validity, a real-world-like setting was created. The precision of the velocity estimation was measured regarding relative movement velocity[a]—a perceptual reference that can be realized within a broad range of human behavior, for instance, observing others while walking, playing music, swimming, boxing, or playing badminton or volleyball.

The core idea is an auditory coding of movement kinematics, which has already been introduced and investigated by our workgroup.[17] Research on the inherent information of natural movement-attendant sounds indicates a rich spectrum of different kinds of information, such as for agent identification and discrimination with complex natural movement sounds[18] or even related to temporal deviations in tap dance sequences.[19] The used intermodal mapping and coding strategy was built on the basic natural relation between kinetic and acoustic event categories as described in the ecological approach to acoustic perception by Carello *et al.*[20] and as already adapted to movement sonification.[21,22] A well-known example of a supramodal fundamental feature category is energy, which is defined within the auditory domain by the amplitude of a sound and within the kinetic domain by the kinetic energy and the potential energy. Even though movement kinematics are usually perceived visually (also designated as "biological motion perception," see Troje[23]), selected kinematic parameters were transformed here into the auditory domain. This is realized to give more weight to these parameters and to enhance the subtlety and precision of (audiovisual) biological motion perception and emerging multisensory representations. If successful, an intermodal support of kinematic movement perception could be used in future to increase the efficiency of training methods in sports and motor rehabilitation by perceptual enhancement and substitution.

Research on multisensory integration

There exists a broad spectrum of research about multisensory integration related to a wide scope of different aspects of human behavior. Frassinetti *et al.*[24] adapted the paradigm of Stein and Meredith[10] on apes and cats to human behavioral research. The authors demonstrated that spatially–temporally coincident low-intensity sound enhances the visual detection rate of static low-intensity visual stimuli by an enhanced perceptual sensitivity in humans. In the study of Seitz *et al.*[16] a spatially moving sound (noise) significantly supported the learning of a visual direction detection task (moving dot-pattern) based on the auditory indication of the movement direction of the visual pattern. Bringing both studies together, multisensory integration is not only effective for the detection of static stimuli but also when learning a moving direction detection task. Further basic studies about multisensory integration deal with fundamental effects of multisensory perception,

[a]For the stimuli used, described in detail within the section "Stimulus material," the term "relative movement velocity" can be understood also as movement frequency. Though the center of the pelvis is used as the origin of the coordinate system, only relative movements can be observed in order to estimate the velocity of the movement, absolute (i.e., translational) movement of the swimmer does not take place.

such as on auditory effects of perceived acoustic event numbers, the "sound-induced flash illusion,"[25] or about an auditory enhancement of the temporal order judgment of time-dense sequential visual events as described by Hairston *et al.*[26] Such basic research on intersensory processing is important to understand the mechanisms of multisensory integration. As reported by Stein and Meredith,[10] certain basal temporal and spatial criteria have to be fulfilled to provoke a supra-additive activation enhancement of multisensory neurons. Visual and auditory stimuli have to emerge from nearly the same direction and within a temporal proximity window of about 100–150 ms to provoke clear behavioral effects. Besides this neurophysiologically oriented research on primates and basic behaviorally oriented research on humans, more recent studies dedicated to biological motion perception and motor control/motor learning should be taken into account.

A broad range of intermodal audiovisual effects have been reviewed by Shams and Kim,[27] indicating that visual perception can be significantly altered by synchronous perceptions of stimuli of other modalities (sound, touch). In addition, they discuss empirical evidence about crossmodal interactions that affect visual learning and adaptation in a statistically optimal manner, referencing the findings of Ernst and Banks.[28] Shams and Kim conclude: "Indeed, visual processing, while an important component of human perception, functions as part of a larger network that takes sensory measurements from a variety of sources and modalities, and tries to come up with an interpretation of the sensory signals that as a whole leads to least amount of error on average."[27]

Recently, a growing number of studies have referred to multisensory integration of audiovisual motion perception. Some of them offer direct support for the development of new efficient methods for sports and rehabilitation. Mendonca *et al.*[29] investigated the impact of the temporal order of visual and auditory gait stimuli in a velocity discrimination task. Based on the findings of Bidet-Caulet *et al.*[30] and Barraclough *et al.*[31] on the multimodal character of the posterior superior temporal sulcus (STSp) as being involved in human motion recognition, Mendonca *et al.*[29] confirmed the benefits of congruent audio–visual stimuli in terms of a reduced variability on audiovisual velocity discriminations. In this study, ecological gait sounds were combined with a visual biological motion pattern. Furthermore, the authors were able to show that information is integrated most efficiently within a temporal window of about 76 ms (with an asymmetric shape of –13 to +63 ms delay of the acoustic stimulus), resulting in the lowest variability of velocity discriminations.

The work of Young *et al.*[32] demonstrates that the kinetic and kinematic characteristics of walking sequences can be perceived and imitated in terms of stride lengths and cadences from walking sound sequences. The authors asked participants to listen to natural recordings of footsteps on a gravel path taken from different stride lengths and cadences and to discriminate differences in perceived stride lengths. Afterwards, participants were asked to adapt their own stride length (1) and cadence (2) according to the presented sound sequences. The participants were successful in both tasks (1 and 2); however, they were also successful when the natural footstep sounds were changed into synthesized sounds. These synthesized sounds were based on kinetic data (ground reaction force vectors) from the foot-ground contacts. Such findings are further supported by a considerable amount of research indicating the beneficial effects of rhythmic auditory stimulation on the cyclic movement of walking, with a particular relevance to rehabilitation, as recently shown in complementary studies by Murgia *et al.*[33] and Ghai *et al.*[34]

Obviously, not only the temporal but also the spatial attributes of action sounds can be discriminated and re-enacted during the perception of an auditory model—even when only basic kinetic features of the action are coded acoustically.

Growing evidence underlines the efficiency of audiovisual information for the perception and execution of complex movements. The use of sonification has been effective in different domains, such as music training,[35] the acquisition of handwriting,[36,37] motor learning in sports,[38] and even in motor rehabilitation.[39] Our own research was directed to noncyclic, not explicitly rhythmical or musical movements, such as acyclic everyday or sports movements.[21,39] Modes of efficient motor-acoustic mappings for sonification have just been preliminarily investigated for overt gestures by Kuessner *et al.*,[40] for the discrimination of similar everyday actions,[17] and for the motor learning of indoor rowing.[41] More recent studies report inconclusive results. Although Dyer *et al.*[42] found transient effects of concurrent rhythmic sonification on a bimanual 4:3 shape-tracing task resolving in a 24-h retention measure, Effenberg *et al.*[22] reported persistent effects of dynamic and kinematic real-time sonification on motor learning of indoor rowing of novices—even beyond effects of rhythmic adjustments.

Research question

Taken all together, the reported findings are valuable for developing more effective methods in sports and motor rehabilitation. It has become clear that biological motion perception is not confined to visual perception. Natural movement sounds are processed in STSp as well as in audio–visual mirror neurons in premotor areas of monkey brains,[35] indicating clearly auditory properties of the mirror neuron system. In addition, it has become evident that cyclic, as well as acyclic, movement patterns can be supported by additional acoustic movement information, and that the multimodal character of biological motion perception is a potential explanation for the observed effects. All the referenced studies deal with human motor behavior. At about 70–80 ms, Mendonca *et al.*[29] draw a closer temporal window for efficient audio–visual integration related to behavioral features compared to Stein and Meredith[10] related to single-neuron neurophysiology of primates. Young *et al.*[32] also demonstrated that synthesized footstep sounds are perceptually processed like ecological footstep sounds on stride length, cadence estimations, and adaptation, supporting the concept of real-time kinematic movement sonification used by Effenberg[21] and Effenberg *et al.*[22] Most of the referenced studies mapped the additional acoustics to distal segments or parts of the acting person (hand or hands,[3,39,40,42] feet,[29] pen-tip,[36,37] hands, and feet).[22] In addition, Vinken *et al.*[17] drafted a mapping-concept explicitly referencing the "effectors' endpoint trajectory" (p. 537) and stated: "Movement sonification was used to transform kinematic data of the distal end effector into the acoustic domain" (p. 539).

The present study investigates the quality of motor perception related to visual and auditory movement information. We attempt to prove if additional auditory information about the arm and leg movements of a swim avatar—animated using the kinematic data of a breast-stroking human—enhances the observers' estimation of velocity differences between two swimmers. Furthermore, we aim to investigate if this kinematic auditory movement information can substitute for visual information in the same task if designed properly. Accordingly, movement sonification might be usable to compensate via intermodal phenomena a partial loss of visual information, as described by Ladavas (p. 108)[43] with reference to multisensory integration: "(. . .) multisensory integration might improve the sensitivity of a unisensory modality in situations of deficit, and, again, favor a possible functional role for multisensory integration in ameliorating the performance deficits of perceptual systems."

Even though our own study seems to be completely in line with the referenced studies and especially with the first experiment (on the discrimination of perceived stride length) by Young *et al.*,[32] it is nevertheless quite different. Breast-stroking is a gross motor cyclic sports movement like walking; however, it does not generate analogously clearly structured natural acoustics. It is executed within the water while the surrounding water produces more complex forms of water sounds blurring the information about the movement. Water splashes cause sounds but water sounds are dependent on many factors, like the shape of the water surface, air bubbles in the water, the posture of the hand when dipping into the water, etc. This enhances the variability of the emerging sounds considerably and thereby reduces the amount of direct information on the movement pattern. We decided to work with breast-stroking because the real-time acoustic movement information (movement sonification) used here is based on selected kinematic parameters chosen by their biomechanically justified importance for the propulsion of the swimmer. In contrast to Young *et al.*,[32] we are not interested in generating a movement sound similar to natural water noises but a movement sound representing selected features of the kinematics continuously. This was realized in order to achieve a high degree of structural equivalence to correlated visual kinematic features. The idea behind this is to configure additional real-time information that is well suited to be integrated with visual biological motion information within multimodal brain areas (e.g., STSp). Although it is not possible to transfer the whole kinematics or body segments and joints, respectively, into the acoustic domain due to uncontrollable acoustic/auditory masking effects, the biomechanically most important references were selected for the sonification as described in the subsequent paragraph.

The kinematic data of a breaststroke movement executed by an expert have been used to animate a human swim avatar in front of a monochrome black background. Thereby, all information except the relative kinematics of the swim avatar were eliminated (see section "Stimulus material"). In our study, we explore the amount of information mediated by the auditory kinematics (movement sonification based on the mapping of the relative distance of the center of both metacarpi and both ankles to the center of the pelvis on sound frequency) compared to visual and audiovisual kinematic information (see H1a below). The amount of information was determined by the estimation of velocity differences between two consecutive breaststroke sequences, whereby the estimates between four different treatments (visual, auditory, audiovisual congruent, and audiovisual divergent) are compared. Furthermore, we changed the scaling of the velocity-dependent global pitch transposition systematically to explore the effect of different scales on estimation accuracy (see H1b below). Besides these two scientific issues (H1a and H1b), we are interested in exploring a potential substitution of visual–kinematic information by auditory–kinematic information (see H1c below). For that, we used a visual treatment and compared the performance under all conditions, including a divergent audiovisual control condition. Finally, with Experiment B, we aim to investigate the effect of an inconsistent pitch transposition in terms of systematic under-/overtranspositions of the sonification on the audiovisually based velocity estimations. If the kinematic sonification is integrated with visual information into a multimodal representation, a systematic change of the estimates in the direction of the under-/overtransposition should be expected. This interrelation is evaluated with Experiment B and operationalized with H2 drafted below.

Four hypotheses are tested with Experiment A and Experiment B:

H1a: Pitch-coded kinematic movement sonification of cyclic gross motor patterns can enhance motor perception/motor estimation (Experiment A).

H1b: Different mapping scales of pitch coding change the effect of the kinematic movement sonification of cyclic gross motor perception/motor estimation (Experiment A).

H1c: Kinematic movement sonification of cyclic gross motor patterns can partially substitute for visual kinematic information (Experiment A).

H2: Global under-/overtransposition of kinematic movement sonifications of cyclic gross motor patterns result in analogously directed changes in motor perception of motor estimation (Experiment B).

Materials and methods (Experiment A)

Participants

A total of 36 female and 36 male students (24.8 ± 3.8 years) participated in Experiment A. They all had normal vision (except for corrective lenses) and hearing abilities as confirmed by a standardized vision (Oculus) and hearing test (HTTS Audiometry). None of them exhibited overt sensory or motor deficits. All participants were able to breaststroke at a nonprofessional level.

This study was carried out in accordance with the recommendations of the Central Ethics Committee of the Leibniz Universität Hannover with written informed consent of all participants and the Declaration of Helsinki 2008.

Stimulus material

Unimodal (visual or auditory) stimuli, as well as bimodal (audiovisual) stimuli, were used. A visual stimulus (component) consisted of two subsequent animation sequences of a breast-stroking avatar based on kinematic data of a former world champion, who was recorded with a three-dimensional video-capture system (PEAK Performance Motion Analysis System, 50 Hz, resolution 768 × 576) in a counterflow system. Video data of 19 optical markers attached to the head, shoulders, elbows, wrists, metacarpi, pelvis, hip joints, knees, ankle joints, and toes were digitized, yielding two-dimensional Cartesian coordinates for each marker. These coordinates were normalized to the coordinates of the pelvis: Thus, all bodily movements were presented as relative-motion to the pelvis, which represented the basis of a Cartesian coordinate system, resulting in fixation of the swim avatar at the middle of the video frame.

nyas13693 W3G-nyas.cls March 27, 2018 14:48

Figure 1. **Visual breaststroke avatar performing human motion. Movements of the avatar were driven by kinematic data of a former breaststroke world champion. One trial (here No. 1) of a visual stimulus—breaststroke sequence 1—represents the velocity reference and should be set as 100%, and the velocity of sequence 2 should be estimated against sequence 1 when the green screen occurs.**

A visual swim avatar (see Fig. 1 below) was created with Simba Software[44] (version 2.0). With the software, the movement data of the swim expert were transformed into a visual volume model. In addition, a stepwise elevation of the frame rate was realized with the built-in interpolation algorithms, enabling a stepwise reduction of the swim frequency in 2% steps when playing back the video sequences to the participants. The use of a human avatar in front of a monochrome black background allows to restrict visual perception to the relative kinematics of the motion, to biological motion perception, respectively: No additional information like, for instance, the dynamics of the surrounding water or the use of pool tiles as a background scale, was given. The elimination of such additional, swimming-specific perceptual references enables a broader transferability of the results to other fields of sports and motor rehabilitation because biological motion information is available in most kinds of sport and rehabilitation settings. Body position at the beginning of a stimulus was varied in order to avoid the recognition of a certain stimulus. At the original velocity, one swim cycle took 1120 ms. To get breaststroke sequences of different velocities, the original 1120 ms sequence was systematically stretched with the factors of 2%, 6%, 8%, 10%, and 12%, resulting in durations of 1142, 1187, 1210, 1232, and 1254 ms, which are 98%, 94%, 92%, 90%, and 88% of the original velocity.

One trial consisted of two consecutive breaststroke sequences. Between both breaststroke sequences, the relative swimming velocities were varied pseudorandomly within a range from 0 ms (both with same velocity) up to a maximum of ± 134 ms per single breaststroke cycle (both sequences with maximum difference). Each breaststroke sequence had a length of 6000 ms. Thus, the stimulus with the highest velocity contained 5.36 breaststroke cycles, and the stimulus with the lowest velocity had 4.78 breaststroke cycles.

In order to configure the auditory stimulus (component) for all congruent audiovisual stimuli, kinematic data were mapped onto sound with the software Sonifikation-Tool (Version 1.0)[b]. A congruent auditory stimulus (*AV_con*) was based on the sonification of two movement parameters. One parameter was the relative distance of the metacarpi to the pelvis, a second parameter the relative distance between ankles and pelvis. These two submovements—the arm stroke and the leg strike—were chosen because these are key elements for generating a high propulsion. The metacarpi distance was mapped onto the amplitude and frequency of the electronic sound "Fairlight Aahs," within a pleasant range of amplitude of 40–74 dB and a pitch range between fis' and e'' (Helmholtz pitch notation). The ankle distance was mapped onto the sound "Pop Oohs" with a pitch range from contra B′ to D′. Figure 2 illustrates the mapping of the two kinematic parameters to both sounds. Both sounds are part of the sound library of the synthesizer E-MU E4K (E-MU Systems, Inc., Scotts Valley, CA). This mapping resulted in a rising sound

[b]Becker, A. 1999. Echtzeitverarbeitung dynamischer Bewegungsdaten mit Anwendungen in der Sonification. Unpublished thesis, Rheinische Friedrich-Wilhelms-Universität Bonn. This reference for the Sonifikation-Tool software is an unpublished thesis. The software is not publicly available, but it can be requested from the author.

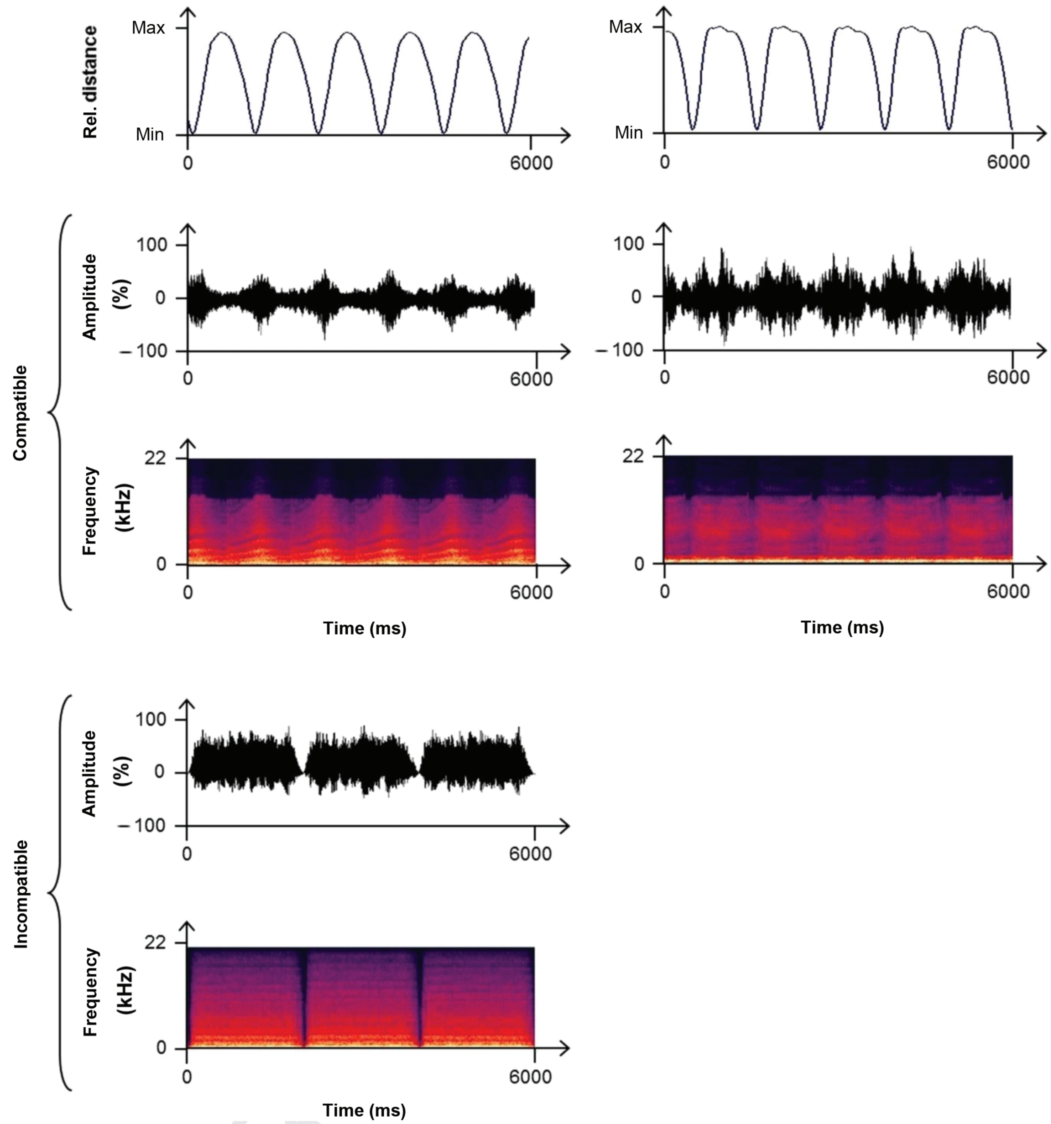

Figure 2. Kinematic–acoustic mapping of the breaststroke sonification. The relative distances of the metacarpi (left) and ankles (right) were mapped onto sound amplitude and frequency in the compatible condition. In the incompatible condition, neither frequency nor amplitude was related to kinematic parameters.

with increasing pitch and volume for the arm stroke. The more energetic the arm stroke got, the louder and more vigorous the arm sound became. It also resulted in a lower sound with decreasing pitch for the leg strike—the more energetic the leg strike got, the louder and more vigorous the leg sound became. Example files are provided as Files S1 and S2 (online only).

The auditory component of a divergent auditory stimulus (*AV_div*) was a combination of two chords of the same timbre and frequencies as *AV_con*. Chords of each stimulus changed twice (A–B–A) after 2000 and 4000 ms. Chord changes were not related to kinematic parameters and the divergent stimulus was not providing any information about a certain kinematic movement feature. It was created as an auditory control stimulus (Fig. 2).

For all six different velocities of the visual breaststroke sequences, three different kinematic–acoustic mappings were realized to test hypotheses

Table 1. Visual and auditory stimuli of six different velocities were created: although the duration was kept constant, the pitch mapping was changed from 0% to 1% and 2% related to a 2% velocity difference of the visual stimulus

Visual stimuli			Auditory stimuli					
			(1) Constant_pitch		(2) Half_transposition (1%)		(3) Full_transposition (2%)	
Duration (ms)	Stretch factor (%)	Velocity (%)	Duration (ms)// Pitch (%)		Duration (ms)// Pitch (%)		Duration (ms)// Pitch (%)	
1120	0	100	1120	100	1120	100	1120	100
1142	2	98	1142	100	1142	99	1142	98
1187	6	94	1187	100	1187	97	1187	94
1210	8	92	1210	100	1210	96	1210	92
1232	10	90	1232	100	1232	95	1232	90
1254	12	88	1254	100	1254	94	1254	88

H1a and H1b. Although the temporal durations of both the visual and auditory stimuli had been the same (with 1120, 1142, 1187, 1210, 1232, and 1254 ms as described above), the pitch mapping was varied threefold: (1) the pitch was kept constant (*constant_pitch*); (2) transposed to 99%, 97%, 96%, 95%, and 94% (*half_transposition*); or (3) transposed to 98%, 94%, 92%, 90%, and 88% (*full_transposition*) of the original sound pitch (100%). Therefore, in the condition *full_transposition*, the alteration of the auditory stimulus was congruent to the alteration of the visual stimulus. An overview of the three different auditory stimuli is given in Table 1.

Stretching an audiovisual swim cycle by 2% corresponded to a lowering of pitch frequency by 0% (*constant_pitch*), 1% (*half_transposition*), or 2% (*full_transposition*). Modifications of the visual and the auditory stimuli were performed with Version 2.0 of the Simba Software and Version 2.0 of Cool Edit Pro 2.0.

Procedure

Participants sat 4.0 m in front of a screen (2.30 m × 1.70 m), wore headphones (beyerdynamic DT 100), and had an unrestricted view during all treatments. They were instructed to estimate the velocity differences of a swim avatar presented within one trial of two consecutive stimuli. The stimulus was presented as a video clip of 18.5 s length. The clip illustrated first a trial number for 1 s and then two consecutive stimuli (each 6 s) interleaved by a gray screen for 0.5 second. The trial ended with a green screen of 5 s length for the participants to state their estimate.

Experiment A contained four different treatments: purely visual (*V*), purely auditory (*A*), audiovisual congruent (*AV_con*), and audiovisual divergent (*AV_div*) (as the control condition). To evaluate if pitch transposition (i.e., pitch scaling) between two consecutive stimuli with different swimming velocities enhances the subjects' perceptual accuracy, 24 subjects heard auditory stimuli without pitch transpositions (*constant_pitch*), 24 with half (*half_transposition*) and 24 with full pitch transpositions (*full_transposition*). Each treatment consisted of 26 trials. Velocity differences were balanced across treatments. The order of treatments was balanced in a Latin square design.[45,46] To familiarize subjects with the auditory and/or visual stimuli, feedback about perceptual accuracy was provided in four practice trials prior to each treatment.

Data analysis

Each judgment (limited by instruction to ± 14%) was converted into an error (ms) between judged and given velocity difference with respect to the length of one breaststroke cycle. To measure perceptual performance, two error terms were calculated as:

$$\mathrm{AE} = \frac{\sum |j_t - \Delta v|}{n}, \tag{1}$$

$$\mathrm{CE} = \frac{\sum (j_t - \Delta v)\{\Delta v\,|\Delta v \geq 0\} + \sum (\Delta v - j_t)\{\Delta v\,|\Delta v < 0\}}{n}. \tag{2}$$

Note that Δv is the difference within a pair of two breaststroke sequences (one trial), j_t is the subject's individual estimate of this difference in a given trial, and n is the number of trials. AE represents an absolute error and CE a constant error. Note that according to this definition, the constant error

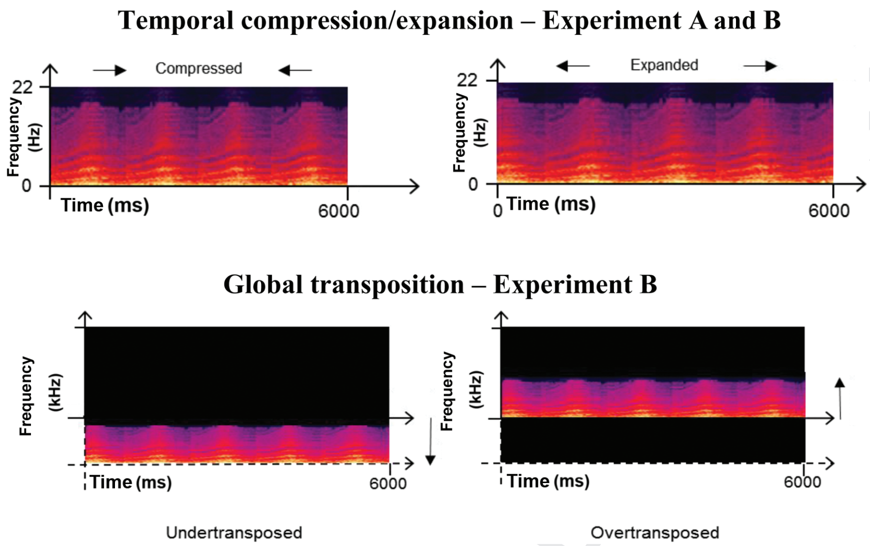

Figure 3. Schematic illustration of the over- and undertransposition of the kinematic–acoustic mapping of the breaststroke sonification: Consistent transpositions as used in Experiment A against inconsistent transpositions used in Experiment B.

provides information about biased estimations that are constantly lower (negative) or larger (positive) than the given differences.

Dependent variables were submitted to repeated measures ANOVAs with the between-factor group (*constant_pitch*, *half_transposition*, *full_transposition*) and the within-factor treatment (*V*, *A*, *AV_div*, *AV_con*). Significant effects were decomposed with Newman–Keuls *post hoc* tests. Sphericity was analyzed with Mauchley's test, homogeneity of variances with Levene's test. Only significant results of sphericity or heterogeneity are reported.

Materials and methods (Experiment B)

Participants

Twelve female and 12 male students (24.8 ± 3.4 years) participated in Experiment B. They all had normal vision (except for corrective lenses) and hearing abilities as confirmed by a standardized vision ("Oculus") and hearing test ("HTTS Audiometry"). None of them exhibited overt sensory or motor deficits. All participants were able to breaststroke at a nonprofessional level.

This study was carried out in accordance with the recommendations of the Central Ethics Committee of the Leibniz Universität Hannover with written informed consent of all participants and the Declaration of Helsinki 2008.

Stimulus material

For Experiment B, the stimulus material was the same as used in Experiment A but only the audio-visual stimuli were used. Experiment B contained two treatments in a first step: Subjects heard audio–visual congruent stimuli (AV_con) with half (half_transposition) and full pitch transpositions (full_transposition). A third treatment was based on the same stimuli as the full_transposition treatment, but with a significant modification: In addition to full pitch transpositions, varying inconsistent global over-/undertranspositions of pitch characterized the treatment varying_transposition. The pitch of one stimulus was enhanced by 2% or 4%, whereas the pitch of the other stimulus was reduced by 2% or 4%, resulting in a reduction or in an enlargement of the auditory interval of a stimulus pair of ±4% or ± 8% compared to full_transposition treatment. Figure 3 illustrates the temporal compression and expansion of the auditory stimulus as used in Experiment A (upper section) and the global transposition used in Experiment B (lower section).

The durations of the over-/undertransposed stimuli were not affected by the transposition, resulting in a congruent temporal relation of acoustic and optical stimulus components. The half and full_transposition treatments had 24 trials each. The varying_transposition treatment consisted of 48 trials (12 trials for each of the four transpositions) and was therefore presented in two blocks of 24 trials each. Each block contained the same number of over-/undertranspositions in randomized order. The order of treatments was balanced in a Latin square design.[46]

Procedure

The procedure was the same as in Experiment A. Only the decision time of a single trial was reduced from 5 to 4 s as a consequence of the participants' performance in Experiment A. To familiarize subjects with auditory and/or visual stimuli, feedback about perceptual accuracy was provided in four practice trials prior to each treatment.

Data analysis

Once again the absolute error (AE) and the constant error (CE) were calculated. In Experiment B, dependent variables were submitted to repeated measures ANOVAs with the within-factor Treatment (half_transposition, full_transposition, varying_transposition) or to a repeated measures ANOVA with the within-factor Interval Size (–8%, –4%, +4%, +8%). Significant effects were decomposed with Newman–Keuls *post hoc* tests. Sphericity was analyzed with Mauchley's test, homogeneity of variances with Levene's test. Only significant results of sphericity or heterogeneity are reported.

Results (Experiment A)

Absolute (AE) and constant errors (CE) are illustrated in Figure 4. Figure 4 illustrates that perceptual performance differed between treatments and these differences were significant for both dependent variables (AE: $F_{(3,207)} = 21.17$, $P < 0.001$, $\eta^2 = 0.23$; CE: $F_{(3,207)} = 29.32$, $P < 0.001$; $\eta^2 = 0.30$). Errors were significantly lower in *A* and *AV_con* than in *V* and *AV_div* ($P < 0.001$ at both dependent variables). They did not differ between *V* and *AV_div* (all *P's* > 0.05). For the dependent variable CE, audiovisual congruent stimuli enhanced the performance compared to a purely auditory stimulus ($P < 0.001$), which was not the case for variable AE ($P > 0.05$).

Figure 5 illustrates that frequency distributions of CE in *V*, *A*, and *AV_div* are shifted toward negative values, but they are not narrower than in *AV_con*. Thus, CE reflects a misalignment of velocity estimates in *V*, *A*, and *AV_div* in terms of an underestimation. The frequency distribution of *AV_con* is not misaligned anymore and nearly symmetrically distributed around zero. Accordingly, CE in *AV_con* did not differ significantly from zero ($t(71) = 0.22$, $P > 0.05$), whereas all other values did significantly differ (lowest $t(71) = -3.44$, $P < 0.001$).

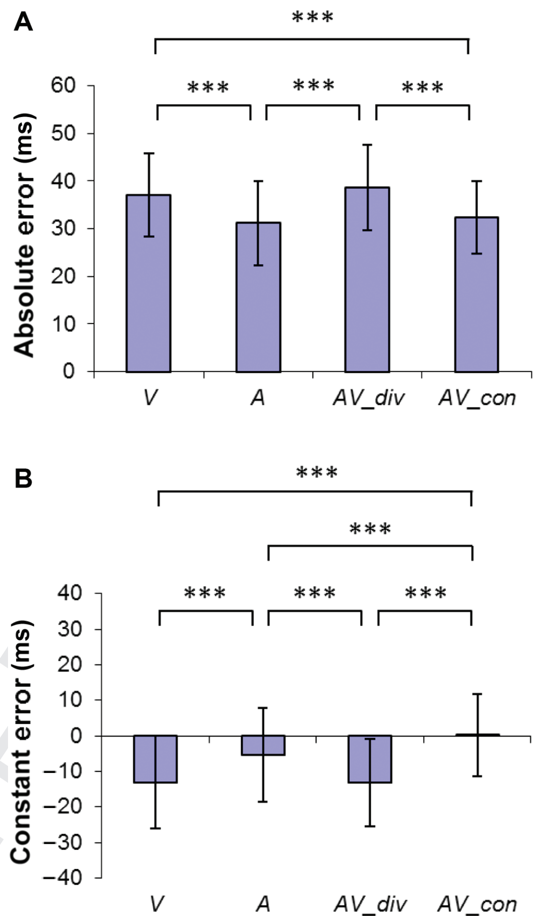

Figure 4. Absolute and constant errors of Experiment A. Means and standard deviations of participants observing visual (*V*), auditory (*A*), audiovisual divergent (*AV_div*), or audio–visual congruent (*AV_con*) stimuli. Significant differences are indicated by: $^*P < 0.05$, $^{**}P < 0.01$, or $^{***}P < 0.001$.

Treatment effects differed between groups (Treatment × Group: AE: $F_{(6,207)} = 8.49$, $P < 0.001$, $\eta^2 = 0.20$; CE: $F_{(6,207)} = 4.34$, $P < 0.001$, $\eta^2 = 0.11$). *Post hoc* analyses confirmed significantly greater CE in treatment *A* than *AV_con* for the group *constant_pitch* ($P < 0.05$). In both groups, *half_* and *full_transposition*, CE and AE were greater in treatments *V* and *AV_div* compared to *A* and *AV_con* (at least $P < 0.01$), with one exemption: the AE of group *half_transposition* did not differ significantly between the treatments *V* and *AV_con* ($P > 0.05$). Thus, when pitches were kept constant,

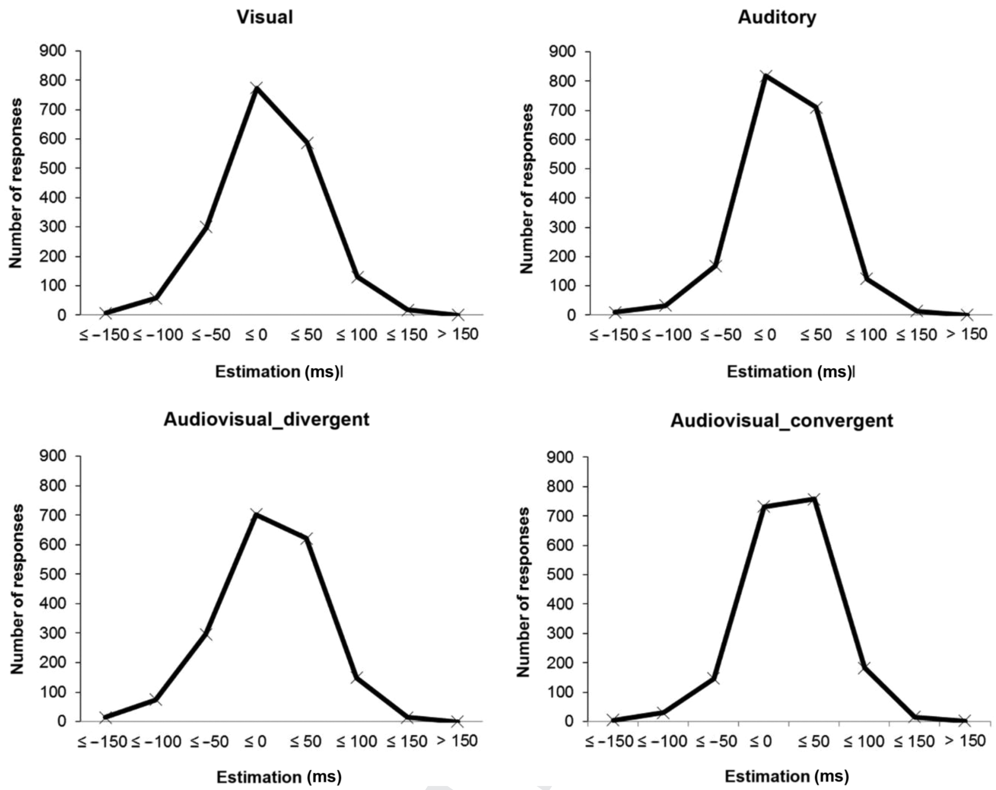

Figure 5. Frequency distributions of the constant error. The numbers of responses in the four treatments are illustrated. Means and standards deviations are: visual –12 ± 12 ms, auditory –5 ± 13 ms, audiovisual divergent –12 ± 13 ms, and audio–visual convergent 1 ± 13 ms. The abscissa shows the upper boundary of 50-ms intervals.

performance during the perception of auditory stimuli was not significantly different from performance using visual stimuli. However, when the pitch was modified, the perception became more accurate with the auditory stimuli (Fig. 6).

The overall performance differed between both groups (AE: $F(2,69) = 3.61$, $P < 0.05$, $\eta^2 = 0.09$; CE: $F(2,69) = 6.42$, $P < 0.01$, $\eta^2 = 0.16$). AE and CE were significantly larger in group *full_transposition* than in group *half_transposition* (AE: $P < 0.05$, CE: $P < 0.01$), and CE was significantly larger in group *full_transposition* than in group *constant_pitch* ($P < 0.01$). Thus, the effects of pitch transposition were not directly compared between groups in Experiment A.

Results (Experiment B)

The effects of pitch transpositions were investigated in more detail in Experiment B. The results are illustrated in Figure 7(A). Perceptual performance depended on the size of pitch transformation and constancy of pitch transposition, as confirmed by one-way ANOVAs (AE: $F_{(2,46)} = 21.93$, $P < 0.001$, $\eta^2 = 0.49$; CE: $F_{(2,46)} = 3.81$, $P < 0.05$, $\eta^2 = 0.14$). AE in treatment *full_transposition* was lower than in treatment *half_transposition* ($P < 0.01$), and better in treatments *full_* and *half_transposition* than in treatment *varying_transposition* ($P < 0.001$).

CE is illustrated in Figure 7(B). It was close to zero in all treatments. It was significantly lower in treatment *full_transposition* than treatment *varying_transposition* ($P < 0.05$), but it did not differ between other treatments (all P's > 0.05). This result is quite surprising because the mean transposition in *varying_transposition* was similar to these in *full_transposition.*

To scrutinize how varying pitch transpositions calibrate perception of relative movement velocities, all the inconsistent trials of the treatment *varying_transposition* were classed into –8%, –4%,

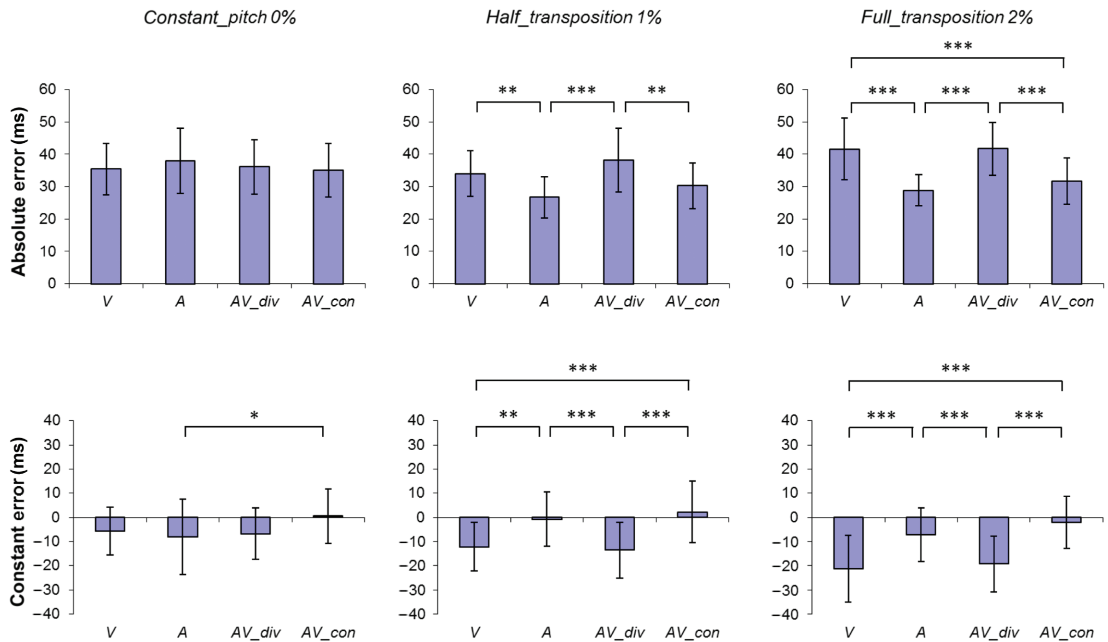

Figure 6. **Absolute and constant errors of each group from Experiment A. Means and standard deviations of participants observing visual (*V*), auditory (*A*), audiovisual divergent (*AV_div*), or audio–visual congruent (*AV_con*) stimuli. Significant differences are indicated by: *$P < 0.05$, **$P < 0.01$, or ***$P < 0.001$.**

+4%, and +8% clusters. On the one hand, as illustrated in Figure 8A, CE was systematically lowered when pitch transpositions indicated lower velocity differences (–8% and –4%) than the visual volume model and the consistently auditory sonification. On the other hand, CE was systematically enhanced when pitch transpositions indicated larger velocity differences (+4% and +8%) than the other stimulus components. The magnitude of the deviation from the reference treatment *full_transposition* was proportional to the inconsistent pitch transposition. Most interestingly, CE was nearly linearly scaled by the magnitude and direction of the pitch transposition. These observations were statistically significant. A repeated measures ANOVA with the within-factor Interval Size yielded a significant effect ($F_{(3,69)} = 74.31$, $P < 0.001$, $\eta^2 = 0.76$). Newmann–Keul's *post hoc* test confirmed significant differences between all four variable transpositions ($P < 0.001$).

In order to exclude the fact that the participants had exclusively based their estimations merely on the acoustic stimulus component (the pitch differences between two consecutive swimmers of a trial) and disregarded the other stimulus components, we compared each of the constant errors illustrated in Figure 8A with the reference data from Experiment A (Fig. 8B) in a control analysis. We averaged constant errors from *treatment A full_transposition* (purely auditory trials) across trials with the same pitch differences as in each of the four conditions from Experiment B and compared them statistically across groups.

Each of the conditions from Experiment B differed significantly from the corresponding reference data of Experiment A (interval –8%: $F(1,46) = 59.93$, $P < 0.001$, $\eta^2 = 0.55$; interval –4%: $F(1,46) = 19.56$, $P < 0.001$, $\eta^2 = 0.30$; interval +4%: $F(1,46) = 19.67$, $P < 0.001$, $\eta^2 = 0.30$; interval +8%: $F(1,46) = 86.74$, $P < 0.001$, $\eta^2 = 0.65$). Thus, the participants from Experiment B estimated velocity differences between swimmers not only based on the auditory information.

Discussion

Up to now, little is known about the specific effectiveness of different mapping designs of auditory stimulus features. The aim of the present study was to investigate whether and how motion perception can be enhanced, substituted, or modulated by

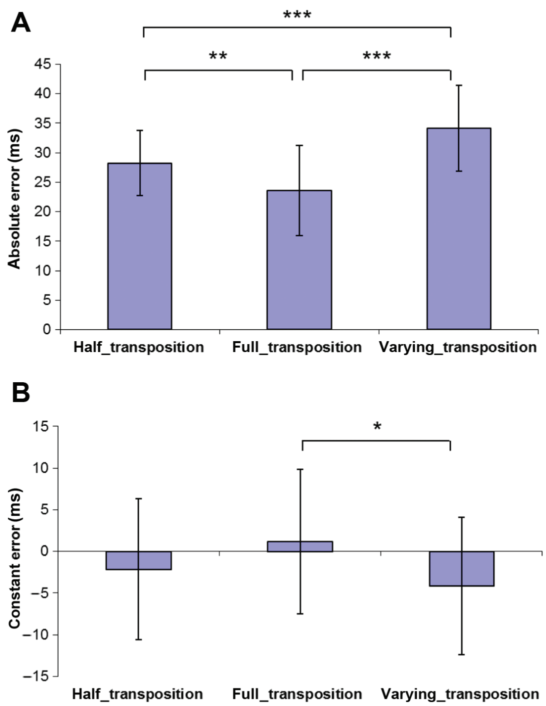

Figure 7. Absolute (A) and constant error (B) of Experiment B. Between-subject means and standard deviations are illustrated. Significant differences are indicated by: $^{*}P < 0.05$, $^{**}P < 0.01$, or $^{***}P < 0.001$.

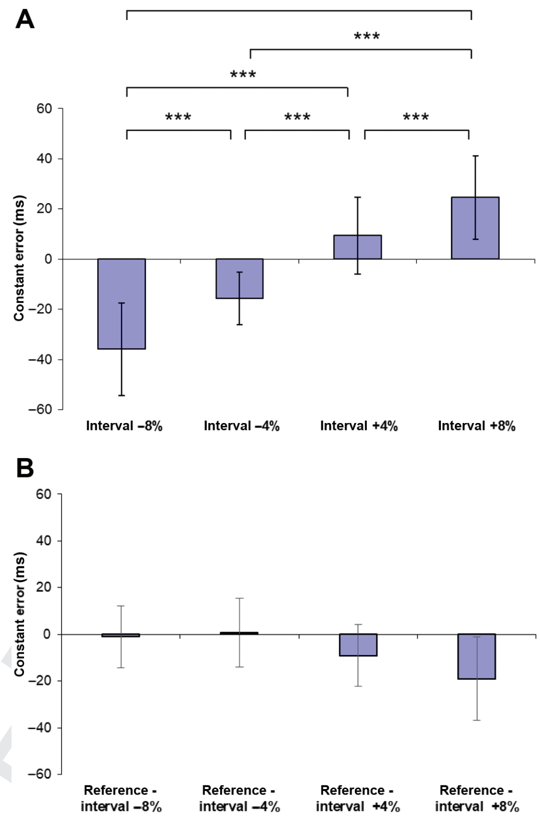

Figure 8. (A) Constant error of the treatment "varying transposition," showing between-subject means and standard deviations. Interval size results from the over-/undertransposition of a stimulus pair, consisting of two successive breaststroke sequences, are shown. Significant differences are indicated by: $^{*}P < 0.05$, $^{**}P < 0.01$, or $^{***}P < 0.001$. (B) Data from Experiment A representing reference values that are achieved if participants just listen to the auditory stimulus and neglect the visual stimulus components in Experiment B. Between-subject means and standard deviations of the constant errors from treatment A (auditory) are illustrated.

kinematic sonification with a focus on pitch coding (scaling and consistent transposition in Experiment A; inconsistent transposition in Experiment B). The sonification model was based simply on two kinematic parameters of a swim avatar with the kinematics of a human breaststroke world champion. This means a huge reduction of information for the auditory treatment compared to the visual treatment, consisting of biological motion scenarios. We would like to emphasize that we aligned the range of transposition between 0% and 2% for a 2% velocity interval of two consecutive stimuli and, thereby, within a linear range of intermodal relation. We deliberately avoided creating an artificially enhanced acoustic indicator for small differences of the selected kinematic parameters and maintained the consistency of the basic kinematic–acoustic framework, the mapping of the relative distance to the frequency of the sound (Fig. 1). Explicitly, we created kinematic acoustics of selected movement parameters with a maximum of structural equivalence regarding the related features of the visual swim avatar to provoke an integration of auditory and visual perceptual streams in areas of multisensory integration in the CNS.

Experiment A

The perceptual auditory accuracy in the auditory task was as good as, or even better, as in the purely visual treatment, depending on the scaling factor of the global transposition of the sonification. Obviously, the chosen parameters contained enough information to solve the task. The absolute error was reduced under auditory, as well as audio–visual, congruent treatments compared to visual, as well as audio–visual, divergent treatments, indicating that the sonified movement sound was an efficient information carrier. Therefore, H1a is confirmed by these findings. While the sonification proved to transmit much information, the additional effect of the audiovisual treatment was restricted to a reduction of the constant error merely, as discussed below.

Our results suggest that the frequency of the sound, perceived as pitch, can be an effective carrier of distance- or velocity-based information. Although intended *a priori*, a between-group comparison had been avoided in Experiment A due to different overall performances across groups. The within-group comparisons indicate that visual–auditory congruent movement acoustics with inherent pitch transpositions of 1% or 2% result in better performance compared to visual or visual–auditory divergent treatments, whereas this effect cannot be shown for visual–auditory congruent movement acoustics without pitch transposition. This might indicate the validity of H1b, but our results from Experiment A cannot sufficiently address this hypothesis.

When focusing on the 0% transposition condition (Fig. 4), where no differences between treatments became evident, it is interesting to note that the performance under the auditory condition was as good as it was under the visual condition. That is remarkable since movement sonification was completely new for the participants. Nevertheless, the kinematic movement acoustics alone can obviously provoke even more precise judgments than the related visual kinematic. This is a highly established source of information for motor perception. These results confirmed H1c and also support the idea that kinematic sonification may be suitable to substitute for another perceptual modality with limitations or even that is missing, as for blind people or in case of the loss of proprioception after stroke. These findings are in line with currently published results from Danna and Velay[47] indicating that real-time sonification supports handwriting character acquisition of proprioceptively deafferented subjects.

The results of Experiment A provide only restricted evidence for multisensory enhancement. With respect to the absolute error performance, the results in the audiovisual trials were not better than in the purely auditory trials. On the one hand, this is surprising and in contrast to other studies, which found enhanced perceptual performance when visual and auditory stimulus components were spatially and/or temporarily congruent.[16,48] On the other hand, there exists further research with differing findings as presented by Sors *et al.*:[49] Although early auditory information was supportive for the prediction of visual ball motion of volleyball smashes, for the prediction of the visual ball motion of soccer penalties, additional early auditory information was not more effective. Furthermore, Allerdissen *et al.*[50] did not report any effects of additional auditory information on the prediction of attack movements in fencing.

This finding suggests that the impact of movement acoustics as well as of movement sonification might change with the particular experimental demands. Possible explanations for this discrepancy relate to the movement information itself, the method of providing this information to the subject, and particularly the paradigm or the kind of the task. Because the mapping of parameters to pitch and loudness has been proved to be effective in the present study and in former studies, the nature of the sonification technique might not be the reason.[51] Thus, the key might be the movement information itself. Visual and auditory stimulus components were based on kinematic movement parameters and provided information about positions and positional changes of body parts. Other studies provided information about dynamic parameters of complex human movements[21] and it might be possible that sonification of kinematic and dynamic parameters result in different perceptual effects.

Accordingly, several studies suggest that neuronal activation differs with respect to the type of sonified movement parameters. Scheef *et al.*[48] investigated the neuronal responses during observation of audiovisual countermovement jumps with sonified ground-reaction force. This force is the counterpart to the vertical components of forces produced by the moving subject and reflects an integrating dynamic movement parameter. The authors reported activation of a widespread network including the superior temporal sulcus, the cerebellum and inferior parietal cortex. However, most importantly, they reported a supra-additive activation of area V5/MT in response to audiovisual compared to the summed activation of purely visual and purely auditory stimuli. This kind of supra-additivity was interpreted as reliable evidence of multimodal integration. In contrast, Schmitz *et al.*[51] investigated central activations during observation of identical breaststroke stimuli as in our study. They found enhanced activations of the medial and superior temporal sulcus, inferior parietal cortex, premotor regions, and subcortical structures, representing the mirror neuron system and key players of the striato–thalamo–frontal motor loop.

Thus, both sonifications seemed to address several areas of the brain that were identical, but others that were clearly different. However, it should be considered that the behavioral tasks and type of fMRI analysis differed between both studies, which might also explain partial differences. Up to now, there is only functional evidence for multisensory enhancement (integration) on the perception of dynamic movement sonification[48] but not for kinematic movement sonification. Therefore, it is necessary to investigate this issue in a future study.

In the present study, another parameter indicates an intermodal fusion effect: The constant error was significantly lower in audiovisual congruent trials than in audiovisual divergent or unimodal trials. The constant error informs about constant over- or underestimations of velocity differences. The results show that subjects tended to underestimate velocity differences. A comparison of frequency distributions (Fig. 5) suggests that estimations are biased based on visual movement information and less biased by purely auditory information. Congruent audiovisual movement information removed the bias completely but the underlying mechanisms of this effect are not clear. In the case of multisensory integration, an effect on the absolute error would also have been expected.

Experiment B

The hypothesis of an intermodal calibration effect was investigated in Experiment B. It was supposed that when velocity differences between two breaststroke samples were auditorily coded as large, the visually perceived velocity difference was enhanced, and when it was auditorily coded as low, the visual perceived difference was reduced. With the control analysis, it could be verified that the reported distributions had been based indeed on bimodal processing of audiovisual information. Therefore, when the auditory component of an audiovisual stimulus was systematically manipulated, subjects' estimations mirrored these intermodal manipulations nearly perfectly. Perceptually based judgments were systematically increased or inverted depending on the mapping rule between the kinematic parameters and the pitch frequencies. With these findings, H2 is confirmed. In addition, the participants were not even aware of the impact of the auditory stimulus component. All of them reported having based their judgments essentially on the visual stimulus and most of them described the movement sound as negligible or even as distracting.

These results suggest that pitch changes shape velocity estimations. Furthermore, pitch changes seem to enhance velocity estimations. As Experiment B was designed as a within-group comparison, it was possible to show that 2% transpositions of the global pitch result in significantly lower errors than 1% transpositions, which supports H1b. The close relationship between the perception of movement velocities and pitch might relate to ecological perception: An increasing velocity is usually associated with an increasing frequency, for instance, while sawing wood, rasping metal, or enhancing the revolutions per minute of a motor. These aspects indicate a possible limitation of our sonification strategy because within a single stimulus changes of the pitch indicated relative velocities and not velocity changes.

Conclusion

Additional movement sonification can be generated to address mechanisms of multisensory integration target specifically. However, it is still widely unknown how to map movement parameters to sound to provoke an optimal effect despite the lack of an adequate theoretical background. The present work confirms a significant impact of kinematic auditory movement information on motor perception and estimation but it is not possible to allocate all reported effects clearly to certain neurological mechanisms in terms of multisensory integration, intermodal calibration, or others.

Neurophysiological research confirms that congruent perceptual streams of different sensory modalities are integrated and, beyond that, directly influence perceptual and motor processes. Based on current neurophysiological findings, with some originating from our own workgroup,[48,51] we have drafted plausible explanations. With further research in the behavioral as well as in the neurophysiological domain, it should be possible to enhance the effectiveness of adequately configured auditory movement information stepwise.

The present study scrutinizes an efficient mapping of movement kinematics on sound features of sonification with a special focus on the role of pitch frequencies. We found that consistent transpositions can enhance perceptual accuracy and velocity estimation (Experiment A) and that inconsistent transpositions modify the integrated audiovisual percept systematically (Experiment B). Therefore, the factor pitch needs to be controlled carefully when movements are mapped onto sound. When estimating velocity differences of a swim avatar, audiovisual stimulus congruency had a significant influence on perceptual biasing or intermodal adjustments in terms of a reduced constant error, but not on the absolute error. Different mechanisms for audiovisual integration in the human brain might be an explanation here, being effective in generating behaviorally relevant information without the need for conscious information processing that has been already observed for rhythmical information by Tecchio *et al.*[52] and Thaut *et al.*[53]

The stationary fixation of the visual volume model was realized to restrict visual perception to the relative kinematics of the motion and to prevent the use of additional information. Avoiding such additional, swimming-specific perceptual references enables a broader transferability of the results to other fields of sport and motor rehabilitation. Movement sonification, if coded in a suitable ecologically oriented fashion, can be used like highly established visual information—initially and without the need of prior learning. Even if the extent of such modifications might be clearly limited, the findings on the systematic modifiability of the integrated audiovisual percept (Experiment B) is interesting with respect to a bundle of further implementations in sports, motor rehabilitation, and therapy. The calibration of the body scheme could be modified or disturbances of motor patterns like gait asymmetries could be made perceivable more clearly and, thereby, reducible or even resolvable.

Establishing an additional real-time auditory kinesthesia in a first step and recalibrating it with the objective to shape a special behavior in a second step might be a new and powerful approach in the field of motor learning and relearning—at least if intermodal adjustment will be effective for motor control and motor learning in a comparable extend. The referred findings on the enhanced activation of some key players of the striato–thalamo–frontal motor loop provoked by the movement sonification give support to such ideas,[51] as well as empirical evidence on motor relearning in rehabilitation and motor learning in sports.

We feel confident living in a visual world but it has been shown here that visual breast-stroking with a fixed constant velocity is perceived as faster when combined with an accelerated kinematic sonification with higher global pitch—and vice versa it is perceived as lower when combined with a decelerated kinematic sonification with lower global pitch. Even though the duration of two visual and two auditory consecutive breaststroke sequences had been the same, participants' estimations have been distorted by inconsistent transformation of global pitch.

Acknowledgment

Parts of this work were supported by EC H2020-FETPROACT-2014 No. 641321.

Supporting Information

Additional supporting information may be found in the online version of this article.

File S1: Varying_global_transposition_&_full_trans position_2%_.mp4.

File S2: Comment_to_Video_examples.docx.

Author contributions

A.O. Effenberg developed the method of real-time movement sonification and the experimental paradigm. G. Schmitz created the stimulus material and conducted the experiment. The statistical analyses as well as the writing of the paper have been performed by both authors together.

Competing interests

The authors declare no competing interests.

References

1. Wolpert, D.M., Z. Ghahramani & J.R. Flanagan. 2001. Perspectives and problems in motor learning. *Trends Cogn. Sci.* **5:** 487–494.
2. Wolpert, D.M., J. Diedrichsen & J.R. Flanagan. 2011. Principles of sensorimotor learning. *Nat. Rev. Neurosci.* **12:** 739–751.
3. Lahav, A., A. Boulanger, G. Schlaug, *et al.* 2005. The power of listening: auditory–motor interactions in musical training. *Ann. N.Y. Acad. Sci.* **1060:** 189–194.

4. Pfordresher, P.Q. 2012. Musical training and the role of auditory feedback during performance. *Ann. N.Y. Acad. Sci.* **1252:** 171–178.
5. Zimmerman, E. & A. Lahav. 2012. The multisensory brain and its ability to learn music. *Ann. N.Y. Acad. Sci.* **1252:** 179–184.
6. Zmigrod, S., M. Spape & B. Hommel. 2009. Intermodal event files: integrating features across vision, audition, taction, and action. *Psychol. Res.* **73:** 674–684.
7. Imamizu, H. & M. Kawato. 2012. Cerebellar internal models: implications for the dexterous use of tools. *Cerebellum* **11:** 325–335.
8. Lacquaniti, F., G. Bosco, S. Gravano, *et al.* 2014. Multisensory integration and internal models for sensing gravity effects in primates. *BioMed Res. Int.* **2014**. https://doi.org/10.1155/2014/615854.
9. Calvert, G.A. 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* **11:** 1110–1123.
10. Stein, B.E. & M.A. Meredith. 1993. *The Merging of the Senses.* The MIT Press.
11. Stein, B.E. & T.R. Stanford. 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9:** 255–266.
12. Chabrol, F.P., A. Arenz, M.T. Wiechert, *et al.* 2015. Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nat. Neurosci.* **18:** 718–727.
13. McGurk, H. & J. MacDonald. 1976. Hearing lips and seeing voices. *Nature* **264:** 746–748.
14. Jamal, Y., S. Lacey, L. Nygaard, *et al.* 2017. Interactions between auditory elevation, auditory pitch and visual elevation during multisensory perception. *Multisens. Res.* **30:** 287–306.
15. Giard, M.H. & F. Peronnet. 1999. Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* **11:** 473–490.
16. Seitz, A.R., R. Kim & L. Shams. 2006. Sound facilitates visual learning. *Curr. Biol.* **16:** 1422–1427.
17. Vinken, P.M., D. Kröger, U. Fehse, *et al.* 2013. Auditory coding of human movement kinematics. *Multisens. Res.* **26:** 533–552.
18. Kennel, C., T. Hohmann & M. Raab. 2014. Action perception via auditory information: agent identification and discrimination with complex movement sounds. *J. Cogn. Psychol.* **26:** 157–165.
19. Murgia, M., V. Prpic, P. McCullagh, *et al.* 2017. Modality and perceptual–motor experience influence the detection of temporal deviations in tap dance sequences. *Front. Psychol.* **8:** 1340.
20. Carello, C., J.B. Wagman & M.T. Turvey. 2005. Acoustic specification of object properties. In *Moving Image Theory: Ecological Considerations.* J. Anderson & B. Anderson, Eds.: 79–104. Southern Illinois University Press.
21. Effenberg, A.O. 2005. Movement sonification: effects on perception and action. *IEEE Multimedia* **12:** 53–59.
22. Effenberg, A.O., U. Fehse, G. Schmitz, *et al.* 2016. Movement sonification: effects on motor learning beyond rhythmic adjustments. *Front. Neurosci.* **10:** 219.
23. Troje, N.F. 2013. What is biological motion? Definition, stimuli and paradigms. In *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention.* M.D. Rutherford & V.A. Kuhlmeier, Eds.: 13–36. MIT Press.
24. Frassinetti, F., N. Bolognini & E. Làdavas. 2002. Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* **147:** 332–343.
25. Shams, L., Y. Kamitani & S. Shimojo. 2000. Illusions: what you see is what you hear. *Nature* **408:** 788.
26. Hairston, W.D., D.A. Hodges, J.H. Burdette, *et al.* 2006. Auditory enhancement of visual temporal order judgment. *NeuroReport* **17:** 791–795.
27. Shams, L. & R. Kim. 2010. Crossmodal influences on visual perception. *Phys. Life Rev.* **7:** 269–284.
28. Ernst, M.O. & M.S. Banks. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415:** 429–433.
29. Mendonça, C., J.A. Santos & J. López-Moliner. 2011. The benefit of multisensory integration with biological motion signals. *Exp. Brain Res.* **213:** 185–192.
30. Bidet-Caulet, A., J. Voisin, O. Bertrand, *et al.* 2005. Listening to a walking human activates the temporal biological motion area. *Neuroimage* **28:** 132–139.
31. Barraclough, N.E., D. Xiao, C.I. Baker, *et al.* 2005. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.* **17:** 377–391.
32. Young, W., M. Rodger & C.M. Craig. 2013. Perceiving and reenacting spatiotemporal characteristics of walking sounds. *J. Exp. Psychol. Hum. Percept. Perform.* **39:** 464–476.
33. Murgia, M., F. Corona, R. Pili, *et al.* 2015. Rhythmic auditory stimulation (RAS) and motor rehabilitation in Parkinson's disease: new frontiers in assessment and intervention protocols. *Open Psychol. J.* **8:** 220–229.
34. Ghai, S., I. Ghai, G. Schmitz, *et al.* 2018. Effect of rhythmic auditory cueing on Parkinsonian gait: a systematic review and meta-analysis. *Sci. Rep.* **8:** 506.
35. Lahav, A., E. Saltzman & G. Schlaug. 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* **27:** 308–314.
36. Danna, J. & J.-L. Velay. 2015. Basic and supplementary sensory feedback in handwriting. *Front. Psychol.* **6:** 169.
37. Effenberg, A.O., G. Schmitz, F. Baumann, *et al.* 2015. SoundScript—supporting the acquisition of character writing by multisensory integration. *Open Psychol. J.* **8:** 230–237.
38. Sigrist, R., G. Rauter, R. Riener, *et al.* 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychon. Bull. Rev.* **20:** 21–53.
39. Schmitz, G., D. Kroeger & A.O. Effenberg. 2014. A mobile sonification system for stroke rehabilitation. In *Proceedings of the 20th International Conference on Auditory Display (ICAD2014)* Georgia Tech SMARTech repository, New York, NY. http://hdl.handle.net/1853/52045.
40. Küssner, M.B., D. Tidhar, H.M. Prior, *et al.* 2014. Musicians are more consistent: gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Front. Psychol.* **5:** 789.

41. Sigrist, R., G. Rauter, L. Marchal-Crespo, *et al.* 2015. Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Exp. Brain Res.* **233:** 909–925.
42. Dyer, J., P. Stapleton & M. Rodger. 2017. Transposing musical skill: sonification of movement as concurrent augmented feedback enhances learning in a bimanual task. *Psychol. Res.* **81:** 850–862.
43. Làdavas, E. 2008. Multisensory-based approach to the recovery of unisensory deficit. *Ann. N.Y. Acad. Sci.* **1124:** 98–110.
44. Spahr, M. 1999. Biomechanics: the use of a three-dimensional real-time movement analysis system in sport. *New Stud. Athlet.* **14:** 43–56.
45. Langdridge, D. & G. Hagger-Johnson. 2009. *Introduction to Research Methods and Data Analysis in Psychology*. Pearson Education.
46. Bortz, J. 2006. *Statistik: Für Human-und Sozialwissenschaftler*. Springer-Verlag.
47. Danna, J. & J.-L. Velay. 2017. On the auditory–proprioception substitution hypothesis: movement sonification in two deafferented subjects learning to write new characters. *Front. Neurosci.* **11:** 137.
48. Scheef, L., H. Boecker, M. Daamen, *et al.* 2009. Multimodal motion processing in area V5/MT: evidence from an artificial class of audio–visual events. *Brain Res.* **1252:** 94–104.
49. Sors, F., M. Murgia, I. Santoro, *et al.* 2017. The contribution of early auditory and visual information to the discrimination of shot power in ball sports. *Psychol. Sport Exerc.* **31:** 44–51.
50. Allerdissen, M., I. Güldenpenning, T. Schack, *et al.* 2017. Recognizing fencing attacks from auditory and visual information: a comparison between expert fencers and novices. *Psychol. Sport . Exerc.* **31:** 123–130.
51. Schmitz, G., B. Mohammadi, A. Hammer, *et al.* 2013. Observation of sonified movements engages a basal ganglia frontocortical network. *BMC Neurosci.* **14:** 32.
52. Tecchio, F., C. Salustri, M.H. Thaut, *et al.* 2000. Conscious and preconscious adaptation to rhythmic auditory stimuli: a magnetoencephalographic study of human brain responses. *Exp. Brain Res.* **135:** 222–230.
53. Thaut, M., B. Tian & M. Azimi-Sadjadi. 1998. Rhythmic finger tapping to cosine-wave modulated metronome sequences: evidence of subliminal entrainment. *Hum. Mov. Sci.* **17:** 839–863.

nyas13693 W3G-nyas.cls March 27, 2018 14:48

Graphical Abstract & Image

Trial No. (1 s) | Sequence 1 (6 s) set as 100% | Gray screen (0.5 s) indicating seq. 2 | Sequence 2 (6 s) | Green screen (5 s Exp. A/4 s Exp. B) estimating velocity difference

Many domains of human behavior are based on multisensory representations. Knowledge about the principles of multisensory integration is useful to configure real-time movement information for the online support of perceptuomotor processes. Here, we investigate whether and how biological motion perception can be enhanced, substituted, or modulated by kinematic sonification, with a focus on pitch coding.