Using Sonification to Enhance Perception and Reproduction Accuracy of Human Movement Patterns

ALFRED O. EFFENBERG

Abstract– **In the field of motor control and motor learning sonification could be used to enhance the performance of the perceptual system. To transform data into sound there is an almost endless amount of options technically available. In kinesiology an efficient way to create movement sonification is to focus the ecological approach to perception: The essential relation between kinetic events and sound events becomes evident as does the specification of sound parameters by parameters of the kinetic event. The sound of a bouncing table tennis ball is specified by the height of fall and the consistency of the surface. Extending those natural kinetic-acoustic relations to mute phases of human movement should lead to additional information about movement patterns, which should be usable to support processes of motor control and learning. Further features of the perceptual system are discussed such as multisensory integration and conscious and unconscious information processing with regard to effective interactive sonification. Own research is presented indicating that convergent audiovisual information is actually enhancing perception and reproduction accuracy of sport movements.**

Index Terms– **Motor Control, Motor Learning, Multisensory Integration, Sonification, Perceptual Dissociation**

I. Introduction

In the field of motor control and motor learning sonification is used to enhance the performance of the perceptual system [1], [2]. Watching and reproducing human movement patterns are key functions on motor control and motor learning, in sports as well as in motor rehabilitation. When using sonification for generating additional information supporting the functionality of the perceptual system as well as the origin of sound is essential. Three adjacent areas of research are touched: The ecological approach to acoustic perception (II.), multisensory integration research (III.) and perceptual dissociation (IV.). Own research is presented subsequently (V.) showing that perception accuracy of sport movements can be enhanced by movement-sonification as well as reproduction-accuracy.

II. Intermodal Data Mapping: Everything is Possible - What is Efficient?

To transform data into sound there is an almost endless amount of options technically available. Different hierarchical classifications of audifications and sonifications are discussed [3], [4]. In the field of motor control and motor learning an efficient way to create movement sonification is to focus the ecological approach to perception [5], [6]. Structural equivalences of music and movement are apparently in dancing, figure skating or rhythmic sport gymnastics. Where natural movement noises and sounds emerge, they are used by motor control as shown on tennis [7]. How does sound emerge, and how is information encoded? First of all: A kinetic event is absolutely essential to generate a sound event. If there is no movement at all, there is nothing to hear. Hitting the tennis ball with the racket or a bouncing ball on a wooden surface will initiate audible sound waves. The impact of the moving objects on the racket string or upon the wooden surface initiate vibration of the materials, generating sound waves. On human movement only the contact of the body with it`s surrounding or to another object creates hearable sound, since movement frequencies are beyond the human hearing range: Hearing ranges from about 20 Hz to 20 000 Hz, fastest movement frequencies are about 15 Hz among pianists.

Once a kinetic event generates a sound within the human hearing-range, not only the existence of the kinetic event is perceivable via sound. On a wooden surface the sound of a bouncing table tennis ball differs from that of a bouncing tennis ball: Features of the kinetic event as well as features of involved materials (racket, tennis ball or ball and wooden surface) are specified by different parameters of the sound. Sound parameters are generally specified into two distinct categories [8]:

(A) *Material category*: Sound event is specified via physical parameters of related materials (tennis racket, string, tennis ball: specific density of the material, tension of the racket string, material and pressure of the tennis ball etc.) and related media (air).

(B) *Kinetic category*: Sound event is specified via kinematic and dynamic parameters (energy resp. velocity, direction and spin of approaching tennis ball, point of contact; size and density of the wooden surface, size, pressure and kinetic energy of the ball etc.).

The more energy the approaching balls have received, the louder, the longer the sound will be. Also spectral features like timbre or the hardness of the sound will vary as functions of kinetic parameters. Figure 1 shows a first-order sonification used in the studies described below: Horizontal component of ground reaction force was mapped to amplitude and frequency of the sound, an electronically sampled 'vocal a'.

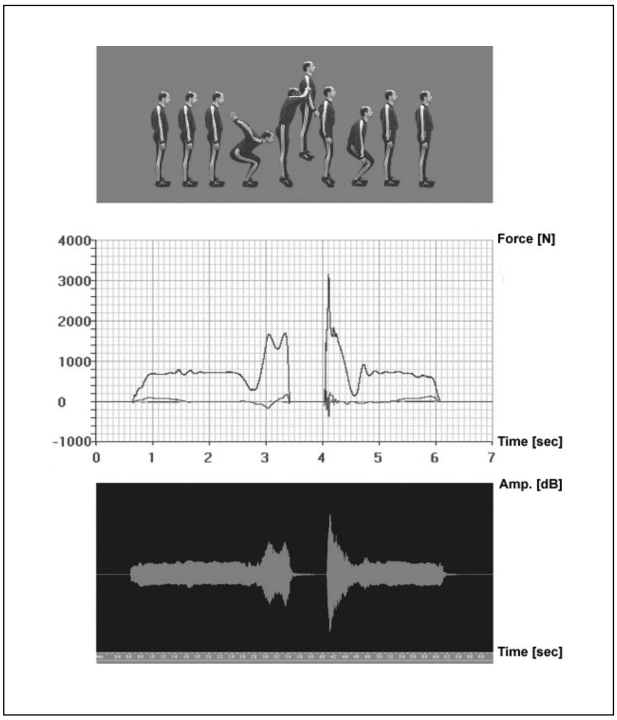


Fig. 1: A counter-movement jump of 48 cm, executed on a KISTLER force plate. Vertical component of ground reaction force reaches about 3200 N. The parameter triggers and modulates amplitude and frequency of an electronic sound in real-time (delay ~ 15 msec.). Sound pressure level is shown below with peak values at 60 dB and within a frequency range between 200 Hz – 2000 Hz.

III. Interactive Sonification and Multisensory Integration – Temporal and Spacial Constraints

On the one hand sonification is used as a tool to mediate information via the auditory system. Specific abilities of the auditory system can be used to perceive time-critical structures and to modify motor behavior unconsciously and even more precisely than would be possible when consciously intended ('subliminal entrainment' on finger-tapping, [9]). Additionally, the ear integrates multiple rows of sounds as in hearing polyphon music, two kinds of integration are discriminable [10]: 'Grouping', meaning vertical resp. spectral integration in auditory perception when hearing an integrated orchestral sound. And 'streaming', meaning horizontal resp. temporal integration, when hearing a rhythm, a phrase, a melody.

But task-specific activation of auditory integration mechanisms by creating appropriate sound sequences is not the only way sonification is effective to mediate information: The ear is only a part of the perceptual system; the different senses work together and are integrated within the perceptual system [11]. There is a lot of research on behavioral effects of multisensory integration [12]-[14]; and there is also supporting neurological evidence on the existence of subcortical/cortical areas integrating sensory input from different modalities [15], [16]. Speech perception is probably the area of research in which audio-visual integrations are investigated most comprehensively.

When perceiving talking faces, speech comprehension is more precise compared to hearing alone [17]. Equally, principles of multisensory integration as well as activation of appropriate subcortical/cortical areas had been examined by neurological studies: Temporal and spatial constraints of evoking additional audiovisual activation on stimulus detection had been revealed (light/sound events, detected by cats and monkeys, [16], [18]) as well as on stimulus identification (speech comprehension by humans, [17], [19]). Research on multisensory integration is essential in order to determine constraints for real-time sonification when used together with visual input: Additional activation of multisenory integrative areas – as well as correlating behavioral benefits – are bound to criteria of temporal coincidence and spatial contiguity: Spatial contiguous stimuli cause explicit enhancing bimodal effects on stimulus detection up to a temporal divergence of about 250 msec. [16]. Concerning humans the limits of consciously perceived simultaneity had been determined to about 100 msec. maximum and spatial divergence of about 3° [20]. On speech perception (stimulus identification) the potential influence of the information content as well as the temporal microstructure of bimodal stimuli on multisensory integration had been additionally discussed [21][1]:

> "Another factor that could determine integration is the information content of the different sensory inputs. .... For information-rich stimuli, especially those with complex temporal microstructure, simultaneous onset or spatial contiguity might be less critical for integration of the inputs to occur. Instead, time-varying similarities in the patterning of information might prove a more salient feature for binding." (see [21], p. 248).

Summarising the work on multisensory integration in the field of speech perception resp. comprehension [17], [19], [21], [22] some constraints for integration multisensory cues can be extracted:

A. Event/object *detection* is bound to:
- temporal stimulus coincidence
- spatial stimulus contiguity

B. Event/object *identification* is bound to additional constraints:
- similar stimulus duration
- similar stimulus intensity
- equivalent temporal (micro-)structure of bimodal stimuli

IV. Sonification: Conscious and Unconscious Information Processing

Beside the constraints of distal stimulus convergence and the internal principles of multisensory integration there is another feature of the perceptual system which has to be taken into account when sonifying kinetic data: The dissociation of perceptual processes resp. the existence of two perceptual streams. The dorsal stream,

[1] In the article the authors also discuss the idea of some observed discrepancies in spatial and temporal constraints affecting multisensory integration possibly being attributed to perceptual dissociation (see [21], p. 248).

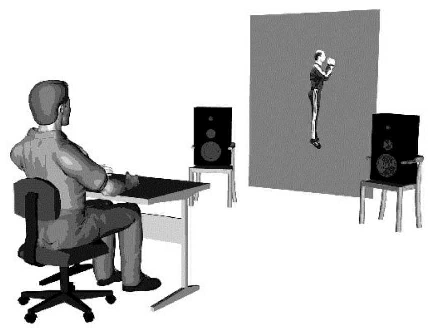
responsible for object location ('where') and the ventral stream, responsible for object identification ('what') [23], [24]. Even discussion on behavioral consequences is still continuing (pro: [25]; contra: [26], [27]) there is valuable evidence for behavioral relevance of unconscious perceptual entrainment: Several studies on visual perception [28], [29] led to the theory of 'Direct Parameter Specification': Even consciously non visible cues (prime-stimulus masked by target-stimulus) reduced time of decision-making on choice-reaction tasks.

And perceptual dissociation is obviously not restricted to vision: Several references indicating that there are dual streams ('what'- and 'where') also separable within the auditory system [30]-[32]. Even all details of the functional range of each stream are not yet completely known, for purposes of sonification it is important to realize that there is conscious and non-conscious information processing within the perceptual system. There is a lot of evidence that ecological motion perception is also based on unconscious perceptual functions [33]-[35] as well as perception of biological motion [36]-[38]. Currently, many functions of motor control and motor learning are attributed to unconscious perception resp. information-processing (review in [39]). If movement-sonification should also serve unconscious perception, intermodal mapping should be realized according to natural essential kinetic-acoustic relations to enclose functions of multisensory integration.

V. Sonification in the Field of Motor Control and Motor Learning

Previous neurophysiological and behavioral research has revealed different areas of multisensory integration in the CNS. Behavioral consequences of multimodal integration are not limited to speech perception; they are also essential for cortical representation of space and control of human movement [40], [12]. We are investigating whether perception and reproduction of complex grossmotor movement patterns (sport movements) under real-world like conditions could be enhanced by bimodal convergent, audiovisual information in comparison to unimodal - visual or auditory - information. Figure 2 shows investigations (A) and (B) in the motor lab.

(A) Percpetion accuracy and jump height assess-ment of different counter-movement jumps



(B) Reproduction resp. imitation accuracy of counter-movement jumps of different heights

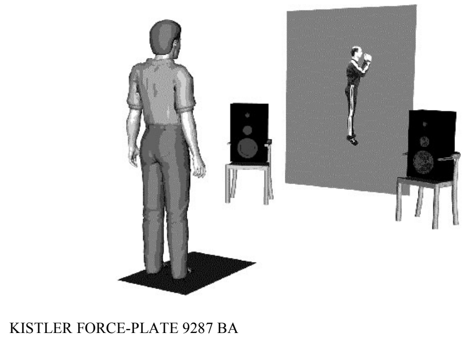


Fig. 2: Subject sitting in front of the 1.90 m x 1.42 m visual, auditory resp. audiovisual projection (A); on reproduction of counter-movement jumps (B) treatments were reduced to visual and audiovisual condition: Preliminary investigation had revealed effects of exhaustion running all three conditions.

A. Perception Accuracy of Sport Movements

Method: Subjects (sport students, n = 40, 22 male, 18 female, 20-34 years old, average: 23.8) were asked to judge the differences of two consecutive counter-movement jumps (CMJ) heights under three conditions (intra-subjective design): Visual treatment (VT), a video of a sportsman performing jumps between 20 cm and 45 cm, graded in 5cm steps. Audio-visual treatment (AVT) consists of VT with movement-sonification added and auditory treatment (AT) is the sonification solely. All subjects heard movement sonification for the first time. There was no feedback given during the investigation. Order of treatments was randomised, each treatment included 24 trials.

Results: Absolute error (AE) of VT and AT reveals no significant difference, though accuracy of perception was less in auditory treatment: AE_{VT} was 4.41 cm and AE_{AT} was 4.70 cm, t-test fails level of significance with $p = .146$. Perception of audiovisual convergent stimuli enhanced precision of height-assessment: AE_{AVT} decreases to 3.48 cm, differences to both single modal conditions reaches level of statistic significance with $p = .000$ both.

B. Reproduction Accuracy of Sport Movements

Method: In a second investigation subjects (male students, n = 40, 20-33 years old, average: 25.0) were asked to watch CMJs of different heights on video performed by a sportsman. CMJs of 60 % - 90 % of the individual maximum level of each subject were presented. Individual maximum level was tested beforehand for each single subject. Two different conditions had been run (intra-subjective design): Visual treatment (VT, video, no sound) and audio-visual treatment (AVT, video and sonification). Subjects heard movement-sonification for the first time, they got no feedback during the investigation. Order of treatment blocks was randomised. Subjects were required to reproduce the height of each single CMJ seen just before. They performed CMJ on a Kistler force plate. After each trial the superelevation of the centre-of-mass (COM) was computed.

Results: Differences of demonstrated and reproduced COM-superelevations were computed for each of the 24 trials for both treatments. AE shows a clear difference for both conditions: AE_{VT} was 3.17 cm and AE_{AVT} was 2.54 cm; t-test reaches level of significance with p = .000. The accuracy of reproduction was higher under convergent audiovisual condition and variance decreases from $s^2_v = 0.896$ to $s^2_{av} = 0.342$.

VI. Summarising Discussion

Experiment (A) reveals that additional convergent auditory information enhances perception accuracy of sport movements. Even perception accuracy was high under both unimodal conditions with an AE < 5 cm, AE decreases significantly if additional convergent auditory information (movement-sonification) was available to the perceptual system. But efficiency of the movement-sonification had not been realized by most of the subjects: Only 15 out of 40 judged audiovisual treatment subjectively as best supporting, but 30 out of 40 had been most precise under audiovisual condition. Experiment (B) demonstrates that additional convergent auditory information also enhances reproduction accuracy of sport movements in comparison to the visual condition. None of the subjects had even heard a movement-sonification before, precision of reproducing CMJ of different unknown heights was enhanced significantly: AE decreases from 3.17 cm to 2.54 cm, which means a reduction of nearly 20 %.

Efficiency of movement-sonification was surprising since some subjects had mentioned during the investigation they did not like the kind of sound at all. Efficiency of movement-sonification had not been realized by many subjects again: 20 out of 40 voted for the audiovisual treatment as the best, but 32 had in fact been most precise audiovisually. Furthermore, another detail became obvious supporting the hypothesis of using unconscious perceptual functions as well as unconscious control functions: There was only little time for judging the difference of to consecutive CMJ of about 2 sec., and for reproducing a single CMJ of 500 msec. - 1000 msec. (maximal). In other words: less time for additional conscious cognition.

Even it was not possible to measure subcortical/cortical activation measured behavioral data are corresponding to results of reported effects of audiovisual integrations in the field of speech perception. Further research is needed to demonstrate contrary effects of divergent audiovisual input as well as neurological studies for detecting sites of multisensory integration in the CNS while perceiving audiovisually presented sport movements. In a further step, research on motor learning should be initiated, because an enhanced informational base should also support processes of motor learning.

Acknowledgement

We thank the Deutsche Forschungs-gemeinschaft (DFG), (German Research-Association) for financially supporting research on 'Acoustical Movement-Transformation', Project Nr.: ME 1526 1-1/1-5.

References

[1] Effenberg, A. O. (2001). *Multimodal convergent information enhances perception accuracy of human movement patterns.* Proceedings of the 6th Annual Congress of the European College of Sport Science (ECSS), Cologne, 122.

[2] Effenberg, A. O., & Mechling, H. (2003). Multimodal convergent information enhances reproduction accuracy of sport movements. Proceedings of the 8th Annual Congress of the European College of Sport Science (ECSS), Salzburg, 196-197.

[3] Kramer, G. (1994). An Introduction to Auditory Display. In G. Kramer (Ed.), *Auditory Display*. New York, 1-77.

[4] Scaletti, C. (1993). Sound synthesis algorithms for auditory data representations. SIGGRAPH '93. An Introduction to Data Sonification. Course Notes 81, Anaheim, 2.2-2.25.

[5] Carello, C., Wagman, J. B., & Turvey, M. T. (in press). Acoustic Specification of Object Properties. In J. Anderson & B. Anderson (Eds.). *Moving Image Theory: Ecological Considerations*. Illinois.

[6] Gibson , J. J. (1966). *The senses considered as perceptual systems*. Boston.

[7] Takeuchi, T. (1993). Auditory Information In Playing Tennis. *Perceptual and Motor Skills,* (76), 1324-1328.

[8] Effenberg, A. O., & Mechling, H. (1999). Zur Funktion audiomotorischer Verhaltenskomponenten. *Sportwissen-schaft*(2), 200-215.

[9] Thaut, M. H., Tian, B., & Azimi-Sadjadi, M. (1998). Rhythmic finger-tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science*(17), 839-863.

[10] Williams, S. M. (1993). *Perceptual Principles in Sound Grouping*. SIGGRAPH '93. An Introduction to Data Sonification. Course Notes 81, Anaheim, 4.66-4.91.

[11] Stoffregen, T. A., & Bardy, B. G. (2001). On specification and the senses. *Behavioral Brain Sciences, 24*(2), 195-213; discussion 213-261.

[12] Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature, 391*(6669), 756.

[13] McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*(264), 746-748.

[14] Vroomen, J., & de Gelder, B. (2000). Sound Enhances Visual Perception: Cross-Modal Effects of Auditory Organization on Vision. *Journal of Experimental Psychology: Human Perception and Performance*(26), 1583-1590.

[15] Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*(5), 473-490.

[16] Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge.

[17] Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*(10), 649-657.

[18] Stein, B.E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research,123*(1-2), 124-135.

[19] Calvert, G. A. (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*(11), 1110-1123.

[20] Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport, 12*(1), 7-10.

[21] Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal Identification. *Trends in Cognitive Sciences, 2*(7), 247-253.

[22] Massaro, D. W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. Cambridge, Massachusetts.

[23] Goodale, M. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 20-25.

[24] Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle & M. A. Goodale & R. J. W. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA, 549-586.

[25] Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5*, 679-685.

[26] Donkelaar, P. v. (1999). Pointing movements are affected by size-contrast illusions. *Experimental Brain Research, 125*, 517-520.

[27] Smeets, J. B. J., & Brenner, E. (2001). Perception and Action Are Inseparable. *Ecological Psychology, 13*(2), 163-166.

[28] Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research*(52), 207-215.

[29] Neumann, O., & Klotz, W. (1994). Motor responses to nonreportable, masked stimuli: Where is the limit of direct parameter specification? In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing*. Cambridge, Massachusetts, 123-150.

[30] Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceeding of the National Academy of Science U S A, 98*(21), 12301-12306.

[31] Kaas, J. H., & Hackett, T. A. (1999). "What" and "where" processing in auditory cortex. *Nature Neuroscience, 2*(12), 1045-1046.

[32] Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is 'where' in the human auditory cortex? *Nature Neuroscience, 5*(9), 905-909.

[33] Duffy, C. J., & Wurtz, R. H. (1997). Multiple temporal components of optic flow responses in MST neurons. *Experimental Brain Research, 114,* 472- 482.

[34] Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience, 3*(12), 1322-1328.

[35] Warren, W. H., Jr., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience, 4*(2), 213-216.

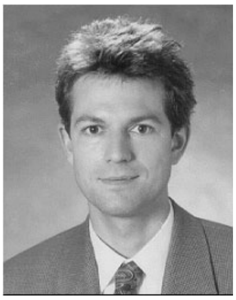
[36] Berenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*(5), 221-225.

[37] Grossman, E. D., & Blake, R. (2002). Brain Areas Active during Visual Perception of Biological Motion. *Neuron, 35*(6), 1167-1175.

[38] Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception, 31*(7), 837-853.

[39] Effenberg, A. O. (2003). Unbewusste Wahrnehmungsfunktionen bei der Bewegungsregulation. In H. Mechling & J. Munzert (Eds.), *Handbuch Bewegungslehre*. Schorndorf, 197-217.

[40] Andersen, R. A. (1997). Multimodal integration for the representation of space in the posterior parietal cortex. *Philosophical Transactions of the Royal Society of London: Biological Sciences*(352), 1421-1428.



Alfred O. Effenberg is working as a Research Associate at the Institute of Sport Science and Sport at the German Sport University Cologne and the University of Bonn. He received a Ph.D. with the thesis 'Sonification - ein akustisches Informationskonzept zur menschlichen Bewegung' in 1996 at the University of Hamburg and his Postdoctoral Lecture Qualification with the thesis 'Synergien der Sinne' in 2002 at the University of Bonn. He teaches in Kinesiology, Training Science and Perception Psychology at the German Sport University Cologne / University of Bonn. He has earned a few awards on Sonification- and Multisensory integration-Research. He is author of a number of national and international publications. His current research interests are on Motor Control/Motor Learning, Movement Sonification, Multisensory Integration and Perception-Action Coupling.