Audio Engineering Society

Convention Paper 9703

Presented at the 142nd Convention
2017 May 20–23 Berlin, Germany

Evaluation of Training to Improve Auditory Memory Capabilities on a Mobile Device Based on a Serious Game Application

Hunor Nagy[1], and György Wersényi[1]

[1]*Széchenyi István University*

Correspondence should be addressed to Hunor Nagy (hunor.nagy@sze.hu)

ABSTRACT

Capabilities of the auditory memory system were tested in a serious game application developed for the Android mobile platform. Participants played the well-known game of finding pairs by flipping and remembering objects on cards arranged in a matrix structure. Visual objects were replaced by iconic auditory events (auditory icons, earcons). Total time and different error rates were recorded and the effect of training was also evaluated. Results indicate that training contributes to a better performance and human voice samples are the easiest to remember.

1 Introduction

Serious gaming offers the possibility of (perceptual) training, testing and evaluation of scientifically relevant data in an entertaining and motivating user friendly environment. Competitiveness and entertainment help users staying motivated during data collection if carefully designed testing scenarios and procedures are covered in games and subjects enjoy the interaction with the device [1-3]. Especially, younger generations prefer the use of handheld mobile devices for their activities, thus, serious game applications developed for tablets and smartphones can target and reach a large group of possible subjects.

On the field of psychoacoustics, assistive technologies, response times and the capabilities of the auditory memory are key issues and have been investigated for a long time. Speaking of assistive technologies, one major field is the development of devices assisting navigation and safety of the visually impaired. A large number of software applications and the access to relatively small, lightweight mobile devices introduced new challenges in designing auditory interfaces, in audio modelling, in HCI technologies etc.

The capability of our auditory memory determines how we can remember auditory events. This of course depends on many parameters: type of the sounds, length, loudness, number of sources, even directional information. In case of auditory displays, where sounds are typically used for feedback to the user, e.g. for helping navigation, giving information or alerts, remembering the meaning (and/or the direction) of the sounds can be critical. As long our capacity is limited, it is an important question to find out, how many sound sources can we remember, what type of sounds are better suited for use etc. Another key point is training: how far can this process be improved. In a side project of an ongoing Horizon2020 research called „Sound of Vision", auditory memory measurements were conducted [4, 5]. This project is aimed at developing a navigational device for the visually impaired, incorporating tactile (haptic) and auditory feedback.

The latter is based on classical iconic sound events, such as auditory icons, earcons, speech samples. Sounds should provide information of the environment for the blind, including warnings, extended information or direction [6-8]. Sonification is applied to provide deeper information of movement or other changes of the visual objects identified. The number of sound events and sonification methods are a central problem of the user interface and the audio modelling as well [9-11].

Connected to the project, a serious game based application was developed for the Android platform. This game is the auditory version of the well-known board memory-game, where the player has to find pairs of cards in a matrix-arranged board (e.g. Concentration, Matching Pairs, Find the Pair etc.) [12-15]. Cards are facing down and the player flips two of them until he or she finds a pair. A single player can play against the clock or against the computer. In two-player-mode, players compete against each other. This game tests and trains the visual memory, and of course, the winner sometimes benefits from lucky constellations.

The auditory memory is the ability to remember words and sounds and to recall information which was received verbally [16-18]. Considering the hypothesis that, granted the centrality of memory to the understanding of language, memory has partially different functions in reading and in listening. Attention to memory describes the process of attending to memory traces when the object is no longer present [19]. It has been studied primarily for representations of visual stimuli with only few studies examining attention to sound object representations in short-term memory.

In [19] the interplay of attention and auditory memory is reviewed, including short-term and long-term memory guided modulation of attention based on characteristic features, location, and/or semantic properties of auditory objects, and propose that auditory attention to memory pathways emerge after sensory memory.

Studies show that recognition memory for sounds is inferior to memory for visuals [20-25]. Four experiments were conducted to examine the nature of auditory and visual memory, including evaluation the role of experience in auditory and visual memory [20]. Similarly to our experiments, here participants received a study phase, followed by a recognition memory test and auditory training, and finally a second memory test. Despite auditory training the visual memory was superior to auditory memory. It was found that it is possible to improve auditory memory, but only after 3 days of specific auditory training. Visual memory appeared to have a larger capacity, while auditory memory was more enduring. These results also indicate that visual and auditory memory are different memory systems.

2 Measurement Setup

The auditory version of the game uses the same scenario of cards, but instead of visual patterns or pictures, short auditory events will be played back. Players have to remember the location of the sounds and find the pairs. So far, the single player option is available in nine different levels, where 2x5 is the easiest and 6x7 is the most difficult. The game has to be played over a headphone and stereo panning is applied to simulate left-right displacements according to the cards' locations on the screen. The following parameters are recorded in a database for each completed game: user ID, gender, age, date and time, level, total time, number of flips for cards/pairs, error rates for each sound.

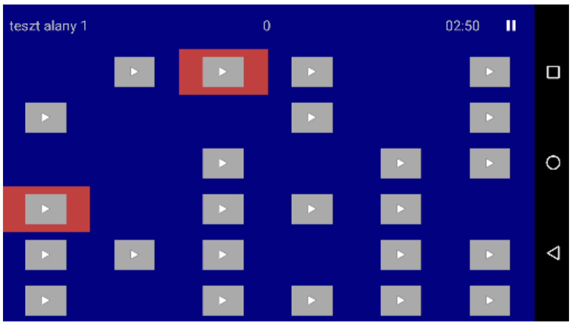

Figure 1. Screenshot of the application (6x7).

Furthermore, a special algorithm is responsible for reducing the factor of luck. Time and date are important for repeated games and testing the effect of training. Sounds were selected to represent a wide variety of iconic sounds, such as impulse, white noise, pink noise, click-train, sinusoid signals, different environmental sounds. The sound database is built up in a hierarchic way, so the sounds of the easiest level (5 sounds on the 2x5 grid) is part of the next level and so on.

In the measurement setup, recruited untrained participants played several games under supervision. That included three levels (2x5, 4x5, 6x6). A clear description of the task was given, the application was installed on their smartphones and data was collected from their devices afterwards. Following this, a training plan was provided for home use: a minimal number of games distributed over weeks. In the second setup, they repeated the same task, so the effect of training could be evaluated. The tests consists of the following steps:

- At the beginning of the game the player has to listen to all the sounds to check the locations of each sound.
- In one single test each level has to be repeated several times: five times for 2x5 and 4x5, and two times for 6x6.

3 Results and Discussion

Best performance was measured on almost every level for male voice and female voice. On the other hand, the worst results were recorded for white noise and pink noise, based on the error rates. At level 2x5 pink noise and female voice were not included to the sound set.

Table 1 shows a ranking of sounds based on the cumulative number of errors. Higher number indicates more flips, thus, more errors.

Noise samples are hard to remember, especially if more of them are present in the same setup. Although white and pink noise do sound different, users usually remember "something noisy" and do not recognize the difference and/or are not able to recall the location on the screen. The database of results does not contain information about white vs. pink noise, but informal conversations revealed that users often click white instead of pink or the other way around.

The same is true for periodic signals, such as sine or square signals. Although they may have different frequency or base frequency, if more of them are included to the same setup, they will be confused.

This phenomenon is supported also by the numbers gathered with earcons and auditory icons. Iconic sounds representing known events from our everyday life such as instruments, phone ringing etc. are easier to remember. This is also due to the variety: there were no similar phone rings or similar guitar tunes to make subjects confused.

As expected, human voice samples are good signal types and could be used more frequently on auditory displays as well. This is also true for meaningless speech samples or very short commands ("left", "click" etc.) that avoid problems with languages.

No.	Type of the	Number of flips
1.	white noise	1113
2.	pink noise	1074
3.	sine 1 kHz	1052
4.	square 1 kHz	1039
5.	impulse	1015
6.	sine 5 kHz	984
7.	click-train	935
8.	sweep linear	924
9.	phone ringing	873
10.	flute	820
11.	guitar	808
12.	toy train	792
13.	whistle long	737
14.	drums	724
15.	violin	710
16.	bells	691
17.	male voice	671
18.	female voice	663

Table 1. Ranking of sounds based on the cumulative number of flips needed for finding the pair.

Although many test results were collected, some of the subjects did not take the test seriously. Some of them missed the second test run or just have not practiced enough. The effect of training could not be tested scientifically deep, but all results obtained with subjects participating in all tests as planned, showed significant improvement. Subjects also reported the test to become more easier after practicing and playing over again. They would also prefer a desktop version or an internet-based version. Based on our experience with crowdsourcing methods to obtain scientific data, such versions could contribute to collect more data.

4 Conclusions

This paper presents first results of a serious game for testing auditory memory capabilities on mobile devices. The main goal is to study different types of sounds and to examine the effect of training. Results showed that from of all the auditory events considered, human voice samples are the best, followed by different auditory icons. It is suggested to use sounds that are different in frequency, timing etc., to avoid any similarities. Sinusoid signals or filtered noise bursts are more easily to be confused if they are present in the same scenario.

Training sessions did contribute to better results, but for a deep statistical analysis a more detailed measurement setup and evaluation are needed, and they are put for future work. The game will be available later in the Google Playstore.

5 Acknowledgement

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 643636 “Sound of Vision”.

This research was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 „National Excellence Program – Elaborating and operating an inland student and researcher personal support system” The project was subsidized by the European Union and co-financed by the European Social Fund.

References

[1] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, and R. Berta, “Assessment in and of Serious Games: An Overview,” *Advances in Human-Computer Interaction*, vol. 2013, article ID 136864, 11 pages, 2013

[2] J. B. Hauge, I. O. Stanescu, S. Antoniu, M. B. Carvalho, T. Lim, S. Louchart, S. Arnab, “Serious Game Mechanics and opportunities for Reuse,” in *Proc. of The International Scientific Conference eLearning and Software for Education*, Bucharest, pp. 19-27, 2015

[3] S. Deterding, S. L. Björk, L. E. Nacke, D. Dixon, and E. Lawley, “Designing gamification: creating gameful and playful experiences,” *in Proceeding of CHI '13 Extended Abstracts on Human Factors in Computing Systems*, Paris, pp. 3263-3266, 2013

[4] O. Balan, A. Moldoveanu, and F. Moldoveanu, “Navigational audio games: an effective approach toward improving spatial contextual learning for blind people,” *International Journal on Disability and Human Development,* vol. 14, no.2, pp. 109-118, 2015

[5] O. Balan, A. Moldoveanu, and F. Moldoveanu, “Multimodal Perceptual Training for Improving Spatial Auditory Performance in Blind and Sighted Listeners,” *Archives of Acoustics*, vol. 40, no.4., pp. 491-502, 2015

[6] W. W. Gaver, “Auditory icons: Using sound in computer interfaces,” *Human-Computer Interaction*, vol. 2, pp. 167–177, 1986

[7] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, “Earcons and icons: Their structure and common design principles,” *Human-Computer Interaction*, vol. 4, pp. 11–44, 1989

[8] J. Sterkenburg, M. Jeon, C. Plummer, ”Auditory Emoticons: Iterative Design and Acoustic Characteristics of Emotional Auditory Icons and Earcons,” *Human-Computer Interaction. Advanced Interaction Modalities and Techniques,* vol. 8511 of the series Lecture Notes in Computer Science, pp. 633-640, 2014

[9] Á. Csapó, Gy. Wersényi, and M. Jeon, “A Survey on Hardware and Software Solutions for Multimodal Wearable Assistive Devices Targeting the Visually Impaired,” *Acta Polythenica Hungarica*, vol. 13, no.5, pp. 39-63, 2016

[10] H. Nagy, Á. Csapó, and Gy. Wersényi, "Contrasting Results and Effectiveness of Controlled Experiments with Crowdsourced Data in the Evaluation of Auditory Reaction Times," *in Proc. 7th IEEE Conference on Cognitive Infocommunications*: CogInfo-Com, Wrocław, pp. 421-426, 2016

[11] Á. Csapó, Gy. Wersényi, H. Nagy, and T. Stockman, "A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research," *J. on Multimodal User Interfaces*, vol. 9, no.4., pp. 275-286, 2015

[12] https://en.wikipedia.org/wiki/Concentration_(game)

[13] http://mypuzzle.org/find-the-pair

[14] https://play.google.com/store/apps/details?id=com.escogitare.memory

[15] https://itunes.apple.com/us/app/matching-pairs-card-match/id1020612987?mt=8

[16] R. G. Crowder, "Thinking in sound: The cognitive psychology of human audition," in Auditory memory. McAdams, Stephen (Ed); Bigand, Emmanuel (Ed)., New York, Clarendon Press/Oxford University Press, pp. 113-145, 1993

[17] W. Ritter, D. Deacon, H. Gomes, D. C. Javitt, and H. G. Vaughan, "The Mismatch Negativity of Event-Related Potentials as a Probe of Transient Auditory Memory: A Review," *Ear and Hearing*, vol. 16., no. 1., pp.52-67, 1995

[18] R. G. Crowder, "The relationship between speech and reading" in Visual and auditory memory, in Kavanagh, J. F. (Ed); Mattingly, I. G. (Ed)., Oxford, England: Massachusetts Inst. of Technology, 1972.

[19] J. F. Zimmermann, M. Moscovitch, and C. Alain, "Attending to auditory memory," *Brain Research*, vol. 1640, part B1, pp. 208–221, 2016

[20] M. E. Gloede, E. E. Paulauskas, and M. K. Gregg, "Experience and information loss in auditory and visual memory," *The Quarterly Journal of Experimental Psychology* 18 May 2016, pp. 1344-1352, 2016

[21] K. C. Backer, and C. Alain, "Attention to memory: orienting attention to sound object representations," *Psychol. Res. – Psychol. Forsch.*, vol. 78, no. 3, pp. 439-452, 2014

[22] J. L. Burt, D. S. Bartolome, D. W. Burdette, and J. R. Comstock, "A psychophysiological evaluation of the perceived urgency of auditory warning signals," *Ergonomics*, vol. 38, no. 11, pp. 2327-2340, 1995

[23] M. A. Cohen, T. S. Horowitz, and J. M. Wolfe, "Auditory recognition memory is inferior to visual memory," *Proc. Natl. Acad. Sci. USA*, vol. 106, no.14, pp. 6008-6010, 2009

[24] N. Cowan, "Visual and auditory memory capacity," *Trends Cogn. Sci*, vol. 2, no.3, pp. 77-78, 1998

[25] B. G. Shinn-Cunningham, "Object-based auditory and visual attention," *Trends Cogn. Sci*, vol. 12, no.5, pp. 182-186, 2008