The Sound of Actuators: Disturbance in Human -Robot Interactions?

Melanie Jouaiti, Patrick Henaff

▶ To cite this version:

Melanie Jouaiti, Patrick Henaff. The Sound of Actuators: Disturbance in Human -Robot Interactions?. Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on, Aug 2019, Oslo, Norway. hal-02144955

HAL Id: hal-02144955
https://hal.science/hal-02144955v1

Submitted on 11 Jul 2019

The Sound of Actuators: Disturbance in Human - Robot Interactions?

Melanie Jouaiti[1] and Patrick Henaff[2]

Abstract— **Human-Robot interactions promise to increase as robots become more pervasive. One important aspect is gestural communication which is quite popular in rehabilitation and therapeutic robotics. Indeed, synchrony is a key component of interpersonal interactions which affects the interaction on the behavioural level, as well as on the social level. When interacting physically with a robot, one perceives the robot movements but robot actuators also produce sound. In this work, we wonder whether the sound of actuators can hamper human coordination in human-robot rhythmic interactions.**

Indeed, the human brain processes the auditory input in priority compared to the visual input. This property can sometimes be so powerful so as to alter or even remove the visual perception. However, under given circumstances, the auditory signal and the visual perception can reinforce each other.

In this paper, we propose a study where participants were asked to perform a waving-like gesture back at a robot in three different conditions: with visual perception only, auditory perception only and both perceptions. We analyze coordination performance and focus of gaze in each condition. Results show that the combination of visual and auditory perceptions perturbs the rhythmic interaction.

I. INTRODUCTION

Involuntary movement coordination inevitably emerges from repetitive tasks when two humans interact with each other. They unconsciously influence each other's behaviour while also creating emotional links. Emergence of synchrony can be observed in rhythmic tasks but also for discrete movements such as "pick and place". Human-Robot coordination or imitation is a privileged research subject in socially assistive robotics and beyond proposing new frameworks for coordination [1], it has been widely used with a therapeutic goal, particularly for autistic children [2], [3], but also for motor rehabilitation [4] since rhythm therapy is one of the most effective therapies for Parkinson, cerebral palsy, autism... Most studies justify the use of robots with the enhanced engagement demonstrated by children with robots compared to the one with humans. There is, however, an important difference in movement generation between a human and a robot: a robot produces sound when it moves. Research in neuroscience shows that the human brain processes different sensory stimuli in different ways and that sensory dominance can appear. So, could the sound of actuators influence the interaction in any way? This is an important aspect of our research in assistive robotics for human-robot coordination as the therapeutic effect of the interaction might be diminished due to the confusion created by the sound of actuators. We are in no way saying that visual and auditory perceptions cannot be combined. Interpersonal interactions are inherently multimodal with speech, gestures... but in the case of the robot, the interaction partner receives two different pieces of information (visual and auditory) related to the same movement and this might be confusing if they are congruent with each other.

In this work, we study the influence of the sound of actuators in human-robot rhythmic interactions. We propose a study where participants were asked to perform a waving-like gesture at a robot in three conditions: with visual perception only, auditory perception only and both perceptions. We chose the waving-like motion because it is a rhythmic gesture reminiscent of a familiar social act. We analyze upper limb coordination performance and gaze in each condition. In the first section, we give an overview of knowledge on human sensory perception. In the second section, we present the evaluation method used and the experimental protocol. Then, in the third section, we present our results. Finally, the fourth section discusses and concludes this work.

II. HUMAN SENSORY PERCEPTION

Auditory and visual information are processed differently by the brain and depending on the context, the dominance can vary.

A. Vision may be dominant over audition in spatial localization

[5] argued that visual information does not have the strong alerting capacity of auditory signals and therefore people are predisposed to direct their attention towards visual stimulation, thus causing visual dominance. Moreover, a modality-appropriateness hypothesis suggests that "the human perceptual system is cognizant of the fact that vision is a more trustworthy modality for spatial localization than audition and proprioception and, for this reason, it is more closely attended" [6]. In summary, vision may dominate audition in spatial judgments.

B. Audition may be dominant for temporal judgments

[7] showed that the auditory perception can severely alter or even remove the visual perception when the stimuli correspond spatially and temporally. Indeed, white noise bursts presented through headphones degraded visual orientation discrimination performance. However, for that suppression effect to take place, the sound and visual target stimuli had to be presented in an ipsilateral, spatially congruent manner.

---

*This work was supported by CPER 2015-2020, plateform IT2MP-SCIARAT, region Nancy Grand-Est, France

[1] Melanie Jouaiti is with Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France `melanie.jouaiti@loria.fr`

[2] Patrick Hénaff is with Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France `patrick.henaff@loria.fr`

Besides, the auditory suppression effect mostly occurred when the sound and visual target stimuli were presented in a temporally congruent manner.

Besides, research has shown that short intervals (less than 2 s) are discriminated and reproduced with greater accuracy when the stimuli are auditory [8], [9], [10]. Likewise, discrimination and reproduction of rhythmic patterns are superior in the auditory modality [11]. [12] conducted an experiment on sensorimotor coordination. Participants were required to tap their finger in synchrony with an auditory or visual sequence. After each sequence, they had to report whether they had noticed a time-shifted event. In a second part, participants were subjected to both auditory and visual sequence and were instructed to ignore the auditory information. Results showed greater variability of movement, smaller involuntary phase correction response and poorer time-shifted event detection with a visual information than with an auditory information. In the second condition, variability was similar to the auditory sequences, and involuntary phase correction response depended more on auditory than on visual information, even though attention was always focused on the visual sequences. Moreover, people have greater difficulty synchronizing finger taps with visual than with auditory sequences [13], [14], [15]. This suggests that there are different timing mechanisms in the two modalities [16] and [15] has argued that the motor system is more responsive to auditory than to visual input. Moreover, simple reaction times are shorter to auditory than to visual stimuli, which suggests faster neural processing of auditory stimuli [17], [18].

C. *Sensory dominance is fragile*

However, sensory dominance is fragile and can switch with intensity changes [19]. More intense tones and lights are judged as longer than less intense signals [20], so the stimulus with the longer subjective duration might be dominant [21]. Auditory dominance has also been found in the perception of sequence rate at fast rates (greater than 4 Hz). A change in the rate of an auditory sequence causes the perceived rate of a constant visual sequence to change as well (auditory driving) [22], [23], [24], [6]. However, varying the rate of a visual sequence does not change the perceived rate of an auditory sequence.

On the other hand, in a Parkinson gait rehabilitation study, [25] showed that auditory and visual perception and both perceptions all improve different aspects of gait. So, in some cases, the auditory and visual perception can reinforce each other.

Sensory dominance, its fragility and the possible interference between modalities in human perception lead us to believe that the sound of actuators (See Figure 1) could interfere in human-robot coordination tasks.

III. MATERIAL AND METHODS

In this section, we present the experimental protocol and the error metric measurement employed to evaluate human-robot coordination.

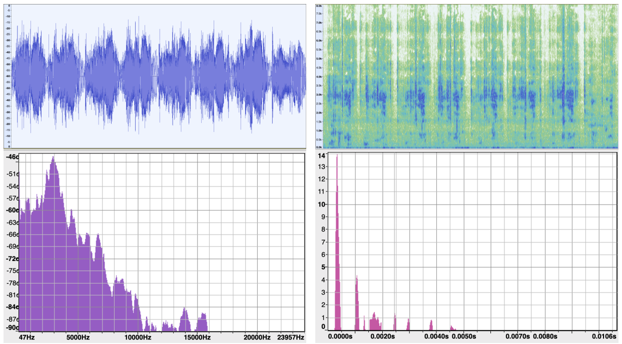

Fig. 1. Top Left: example of sound of the elbow actuator while performing a waving-like motion (in dB). Top Right: Spectrogram of the actuator sound calculated with Audacity. Bottom left: spectrum of the actuator sound calculated with Audacity. Bottom right: enhanced correlation of the actuator sound calculated with Audacity

A. *Experimental Protocol*

Seventeen volunteers (8 women, mean age: 32.71 ± 12.08) participated in the experiment. They were equipped with T-Sens [26] motion sensors on their right arm. The sensors allow us to record the movements of the participants, they are lightweight and should not hamper the participants' movements. They also wore a Tobii eye tracker at 100 Hz. Results from two participants were excluded because of technical issues and loss of data due to glasses. They were seated on a chair facing the Pepper robot (See Figure 3). Their right arm rested on a pillow so as to avoid muscle pain. They were instructed to perform a waving-like gesture at Pepper at the frequency which was comfortable for them. We were indeed interested in involuntary coordination and entrainment. We evaluated three conditions: eyes closed with only auditory perception, then with only visual perception using earplugs and finally with both auditory and visual perceptions (See Table I for an overview of the conditions). The experimenter could observe the subject thanks to a webcam to ensure that the movement was properly performed and that their eyes remained closed when necessary. For each condition, the participants performed the motion nine times. Each individual has its own movement natural frequency but observation of human waving showed us that most people tend to naturally wave at $1.0Hz$. The imposed frequency was randomly chosen between $0.9Hz$, $1.0Hz$ and $1.1Hz$ so as to avoid frequency acclimatization. Each waving motion lasted 10 seconds and was followed by 4 to 6 seconds of rest. Waving and resting periods were indicated by an auditory signal generated by the computer. See the associated video [1] for an example of experiment in the Visual-Auditory condition. Overall the experiment lasted roughly thirty minutes. Figure 2 represents the raw data from the motion sensor for one condition.

During the data analysis process, the videos were trimmed to consider only the interaction times and ignore the rest periods. Heat maps of fixation on the robot and fixation

---

[1] The associated video can be found at `https://members.loria.fr/mjouaiti/files/EPIROB19_2.mp4`.

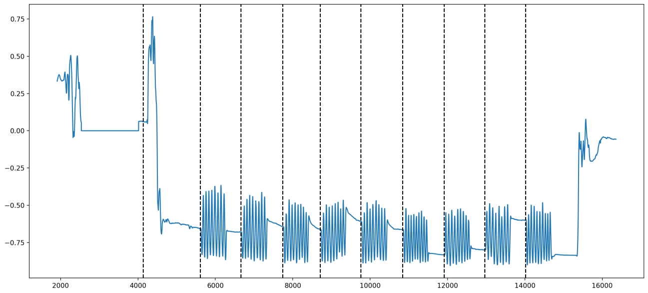

Fig. 2. Example of raw data obtained for quaternion z from the elbow motion sensor for the Auditory condition

Enabled Perception	Visual	Auditory
Condition VA	X	X
Condition A		X
Condition V	X	

TABLE I

SUMMARY OF THE THREE CONDITIONS SHOWING WHICH PERCEPTIONS ARE ENABLED

statistics were computed with the Tobii Pro Lab software.

B. *Phase Locking Value (PLV)*

The PLV is an error measurement metric for coordination between two signals [27]. It ranges from 0 (no coordination) to 1 (perfect coordination). The instantaneous PLV can be defined as:

$$PLV(t) = \frac{1}{N} \left| \sum_{i=0}^{N} e^{j(\phi_1(i) - \phi_2(i))} \right| \tag{1}$$

with N the sliding window size, $j = \sqrt{-1}$, ϕ_k the instantaneous phase of signal k computed with the Hilbert transform.

To evaluate motor coordination, we compute the average PLV between the human and robot elbow joints. The motion sensors give us a quaternion for each joint. The TEA Captiv-Software is able to reconstruct the joint motion from the quaternions. However, the PLV is amplitude-invariant so we only require a sinusoid for our analysis. For most people, the movement can be seen better on quaternion w which represents the abduction of the arm. For others who did not reproduce exactly the robot's movement, we will use quaternion z for the elbow flexion. The robot elbow joint values are also recorded for each interaction since it does not always perfectly obey commands.

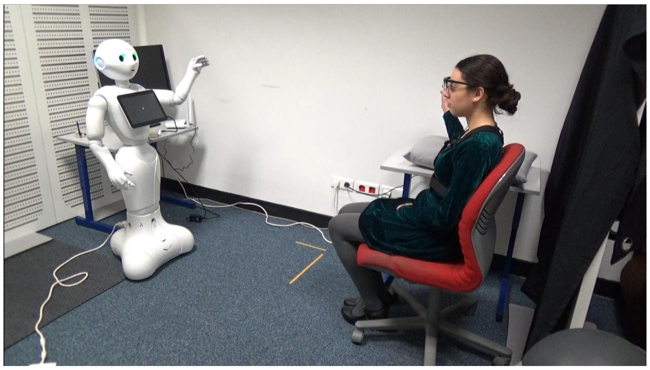

Fig. 3. Experimental setup. The robot is placed in front of the human and waves. The human waves back with its arm supported by a pillow.

IV. RESULTS

In this section, we present the results. First, we analyze the coordination evaluation results for the three conditions. Then the focus of gaze results for the visual-auditory and visual conditions are observed. Finally, we highlight the correlation between them.

A. *Coordination Evaluation*

It can be observed from the results that coordination performance is better in the visual condition (mean: 0.70 ± 0.19) than in the auditory condition (mean: 0.68 ± 0.22) or the Visual-Auditory condition (mean: 0.61 ± 0.22) (See Table II). A 3x3 repeated measure ANOVA was conducted to examine the effect of condition and frequency on coordination performance (PLV). There was a statistically significant interaction between the effects of condition and frequency on coordination performance, p = 0.017.

Running a simple effect test on the data (See Table III) yields that for $0.9Hz$, there is a significant difference between the Auditory condition and the other two conditions ($p = 0.001$ and $p < 0.013$). For $1.0Hz$, there is a significant difference between the Visual-Auditory condition and the other two conditions ($p < 0.0001$ and $p = 0.002$). There is no significant difference between the conditions at $1.1Hz$.

While the auditory and visual conditions yield similar results, the association of both signals seems to be perturbing for subjects at $0.9Hz$ and $1.0Hz$.

condition	frequency	mean	std error
VA	0.9 Hz	0.708	0.041
	1.0 Hz	0.597	0.045
	1.1 Hz	0.517	0.046
V	0.9 Hz	0.742	0.041
	1.0 Hz	0.733	0.038
	1.1 Hz	0.565	0.046
A	0.9 Hz	0.811	0.034
	1.0 Hz	0.733	0.038
	1.1 Hz	0.554	0.043

TABLE II

AVERAGE PLV AND STANDARD DEVIATION FOR THE VISUAL-AUDITORY (VA), VISUAL (V) AND AUDITORY (A) CONDITIONS ACCORDING TO FREQUENCY

Interestingly, there are also significant differences between the three imposed movement frequencies for the coordination performance. Participants performed better at $0.9Hz$ (mean: 0.75 ± 0.27) and $1.0Hz$ (mean: 0.69 ± 0.29) than at $1.1Hz$ (mean: 0.55 ± 0.31).

Comparing the frequencies for each condition (See Table IV) yields that in the Visual-Auditory condition, there is a significant difference between 0.9 Hz and 1.0 Hz and between 0.9 Hz ($p \leq 0.05$) and 1.1 Hz ($p \leq 0.001$). For the

frequency	condition	VA	V	A
0.9	VA		ns	***
	V	ns		*
	A	***	*	
1.0	VA		****	**
	V	****		ns
	A	**	ns	
1.1	VA		ns	ns
	V	ns		ns
	A	ns	ns	

TABLE III

SIGNIFICANT DIFFERENCE BETWEEN CONDITIONS FOR EACH FREQUENCY COMPUTED USING SIMPLE EFFECT TEST

condition	frequency	0.9	1.0	1.1
VA	0.9		*	***
	1.0	*		ns
	1.1	***	ns	
V	0.9		ns	**
	1.0	ns		****
	1.1	**	****	
A	0.9		**	****
	1.0	**		****
	1.1	****	****	

TABLE IV

SIGNIFICANT DIFFERENCE BETWEEN FREQUENCIES FOR EACH CONDITION COMPUTED USING SIMPLE EFFECT TEST

Visual condition, there is a significant difference between 0.9 Hz and 1.1 Hz ($p \leq 0.01$) and between 1.0 Hz and 1.1 Hz ($p \leq 0.0001$). In the Auditory condition, all the frequencies are significantly different ($p \leq 0.01$ between 0.9 Hz and 1.0 Hz, $p \leq 0.0001$ between 0.9 Hz and 1.1 Hz and $p \leq 0.0001$ between 1.0 Hz and 1.1 Hz).

Finally, five participants performed better in the auditory condition, seven performed better in the visual condition and only three performed slightly better with both the auditory and visual perceptions.

B. Eye Gaze Analysis

Engagement of the participants, i.e. their cognitive involvement in the interaction, can be evaluated thanks to a number of different metrics [28]: number of fixations, number of fixations on each area of interest, total number of fixations, fixation duration, total fixation duration, time to first fixation, fixation density... [29] showed that long fixation duration is correlated with high cognitive workload and higher cognitive effort.

To evaluate how much the participants really looked at the robot, we defined regions of interests including the robot's head and arm in the Tobii Pro Lab software and computed the statistics. Table V gives an overview of the average robot fixation duration and fixation count on the robot for each participant in the Visual-Auditory and Visual conditions. A repeated measures ANOVA reveals no significant difference for robot fixation duration or count between both conditions. This suggests that subjects were equally looking at the robot in the Visual and Auditory-Visual conditions.

Subject	Average Fixation Time VA (s)	Average Fixation Time V (s)	Fixation Count VA	Fixation Count V
1	**0.36**	0.30	109	**141**
2	**0.28**	0.26	165	**248**
3	0.22	**0.25**	196	**220**
4	**0.31**	0.28	221	**241**
5	0.24	**0.35**	**215**	122
6	**0.26**	0.25	223	**226**
7	0.16	**0.18**	**302**	295
8	0.18	**0.19**	198	**206**
9	0.10	**0.16**	7	**145**
10	0.16	**0.17**	167	**247**
11	0.22	0.22	**249**	207
12	0.28	**0.45**	**196**	163
13	0.32	**0.48**	**183**	143
14	**0.22**	0.16	99	**189**
15	0.48	**0.52**	**113**	86

TABLE V

AVERAGE FIXATION DURATION AND FIXATION COUNT ON THE ROBOT FOR THE VISUAL-AUDITORY (VA) AND VISUAL (V) CONDITIONS

Subject	Robot Fixation VA (%)	Robot Fixation V (%)	PLV VA	PLV V
1	42	**53**	0.51	**0.87**
2	**65**	59	0.95	**0.96**
3	49	**66**	0.7	**0.79**
4	71	**72**	0.61	**0.64**
5	**75**	41	0.54	**0.65**
6	**65**	59	**0.29**	0.18
7	49	**66**	**0.95**	0.92
8	**45**	43	0.88	**0.92**
9	10	**28**	**0.72**	0.66
10	35	**56**	0.43	**0.54**
11	**59**	50	0.3	**0.58**
12	61	**74**	0.65	**0.76**
13	78	**87**	0.23	**0.29**
14	**29**	28	0.57	**0.82**
15	72	**78**	0.62	**0.64**

TABLE VI

AVERAGE ROBOT FIXATION PERCENTAGE (ROBOT_FIXATION_TIME / INTERACTION_TIME) AND PLV FOR EACH PARTICIPANT IN THE VISUAL-AUDITORY (VA) AND VISUAL (V) CONDITIONS. THE MAXIMUM FIXATION AND PLV VALUES FOR EACH PARTICIPANT ARE IN BOLD

C. Visual Attention and Coordination Performance

Looking at all the results together, there is no statistical significance in the correlation between the percentage of fixation and the coordination performance (see Table VI for an overview).

However, since we find no significant difference in the visual attention of the subjects between both conditions, it is reasonable to assume that the coordination performance decreases in the Visual-Auditory condition compared to the Visual Condition because the additional signal confuses the participants.

V. DISCUSSION AND CONCLUSION

In this paper, we studied the influence of the sound of actuators in human-robot rhythmic interactions. We proposed a study where participants were asked to perform a waving-like gesture back at a robot in three conditions: with visual perception only, auditory perception only and both perceptions. We analyzed coordination performance and focus of gaze in each condition. Results showed that the combination of visual and auditory perceptions perturbs the rhythmic interaction. This was to be expected since both visual and auditory signals emanating from the robot are congruent both temporally and spatially, so they interfere with each other.

This study confirms that the auditory condition yields the best results for rhythmic interactions but the visual condition appears adequate too. However, the combination of both visual and auditory perceptions of a movement can be deemed perturbing while interacting with the robot. Most subjects noted that they tended to synchronize with the auditory signal and not to concentrate much on the robot in the Visual-Auditory condition. However, results show no significant difference in visual attention between the conditions. Only one subject fared worse in the auditory condition than the other condition and confessed being confused by this condition as he didn't know what to synchronize with, although he had no auditory deficit.

The data from the eye tracker was meant to show us where the participants were really looking, especially in the visual-auditory condition. Indeed, in a preliminary experiment without the eye tracker, most subjects reported that they were bored during the task and started thinking about something else and lost focus of attention on the robot and then remembered that they were supposed to look at the robot. They also said that this happened less in the visual condition since they had to focus more on the robot. In this preliminary experiment, we had no significant difference for any of the conditions. Consequently, our working hypothesis was that participants looked less at the robot in the visual-auditory condition and thus relied on their auditory perception. This would have explained the lack of difference. However, reproducing the experiment with the eye tracker data, we were unable to find such a difference in gaze and participants really looked as much at the robot in both conditions. And with the eye tracker, the difference is significant between the conditions.

It can be assumed that the eye tracker made participants self-conscious about where they were looking and only one subject made the observation that he lost focus of attention on the robot. This suspected "eye-tracker awareness" is confirmed by [30] who showed that participants feel the eye-tracker as a social presence and modulate their looking behaviour accordingly.

This paper is, in a way, also meant to be a cautionary tale: the sound of actuators can be confusing in human-robot interactions. This could be an important point to factor in, especially in therapeutic robotics. Even when the robot-therapy is effective, it might have been even better with the correct environmental settings. We unfortunately do not have a satisfactory solution to remove robot noise yet. One can easily see the technical limitations this could incur while attempting to maintain a natural interaction. Ideally, robots should be build with silent motors.

In future work, it could be interesting to observe if voluntary motor coordination is also perturbed by the sound of actuators.

ACKNOWLEDGMENT

We thank all the participants who eagerly volunteered to take part in this study.

REFERENCES

[1] E. Ansermin, G. Mostafaoui, N. Beaussé, and P. Gaussier, "Learning to synchronously imitate gestures using entrainment effect," in *International conference on simulation of adaptive behavior*, pp. 219–231, Springer, 2016.

[2] I. Fujimoto, T. Matsumoto, P. R. S. De Silva, M. Kobayashi, and M. Higashi, "Mimicking and evaluating human motion to improve the imitation skill of children with autism through a robot," *International Journal of Social Robotics*, vol. 3, no. 4, pp. 349–357, 2011.

[3] S. M. Srinivasan, K. A. Lynch, D. J. Bubela, T. D. Gifford, and A. N. Bhat, "Effect of interactions between a child and a robot on the imitation and praxis performance of typically devloping children and a child with autism: A preliminary study," *Perceptual and Motor Skills*, vol. 116, no. 3, pp. 885–904, 2013.

[4] A. Guneysu, R. D. Siyli, and A. A. Salah, "Auto-evaluation of motion imitation in a child-robot imitation game for upper arm rehabilitation," in *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pp. 199–204, IEEE, 2014.

[5] M. I. Posner, M. J. Nissen, and R. M. Klein, "Visual dominance: an information-processing account of its origins and significance.," *Psychological review*, vol. 83, no. 2, p. 157, 1976.

[6] R. B. Welch, "Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions," in *Advances in psychology*, vol. 129, pp. 371–387, Elsevier, 1999.

[7] S. Hidaka and M. Ide, "Sound can suppress visual perception," tech. rep., Sci. Rep. 5, 10483, 2015.

[8] S. Goldstone and W. T. Lhamon, "Studies of auditory-visual differences in human time judgment: 1. sounds are judged longer than lights," *Perceptual and motor skills*, vol. 39, no. 1, pp. 63–82, 1974.

[9] S. Grondin, G. Meilleur-Wells, C. Ouellette, and F. Macar, "Sensory effects on judgments of short time-intervals," *Psychological Research*, vol. 61, no. 4, pp. 261–268, 1998.

[10] W. T. Lhamon and S. Goldstone, "Studies of auditory-visual differences in human time judgment: 2. more transmitted information with sounds than lights," *Perceptual and Motor skills*, vol. 39, no. 1, pp. 295–307, 1974.

[11] R. H. Gault and L. D. Goodfellow, "An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them," *The Journal of General Psychology*, vol. 18, no. 1, pp. 41–47, 1938.

[12] B. H. Repp and A. Penel, "Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 28, no. 5, p. 1085, 2002.

[13] N. R. Bartlett and S. C. Bartlett, "Synchronization of a motor response with an anticipated sensory event.," *Psychological review*, vol. 66, no. 4, p. 203, 1959.

[14] K. Dunlap, "Reaction to rhythmic stimuli with attempt to synchronize.," *Psychological Review*, vol. 17, no. 6, p. 399, 1910.

[15] P. Fraisse, "Auditory rhythms and visual rhythms," *ANNEE PSYCHOLOGIQUE*, vol. 49, pp. 21–42, 1948.

[16] P. A. Kolers and J. M. Brewster, "Rhythms and responses.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 11, no. 2, p. 150, 1985.

[17] P. Jaśkowski, F. Jaroszyk, and D. Hojan-Jezierska, "Temporal-order judgments and reaction time for stimuli of different modalities," *Psychological Research*, vol. 52, no. 1, pp. 35–38, 1990.

[18] J. Vroomen and B. d. Gelder, "Sound enhances visual perception: cross-modal effects of auditory organization on vision.," *Journal of experimental psychology: Human perception and performance*, vol. 26, no. 5, p. 1583, 2000.

[19] S. Goldstone, W. K. Boardman, and W. T. Lhamon, "Intersensory comparisons of temporal judgments.," *Journal of Experimental Psychology*, vol. 57, no. 4, p. 243, 1959.

[20] S. Goldstone, W. T. Lhamon, and J. Sechzer, "Light intensity and judged duration," *Bulletin of the Psychonomic Society*, vol. 12, no. 1, pp. 83–84, 1978.

[21] J. T. Walker and K. J. Scott, "Auditory–visual conflicts in the perceived duration of lights, tones, and gaps.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 7, no. 6, p. 1327, 1981.

[22] J. Gebhard and G. Mowbray, "On discriminating the rate of visual flicker and auditory flutter," *The American journal of psychology*, vol. 72, no. 4, pp. 521–529, 1959.

[23] A. K. Myers, B. Cotton, and H. A. Hilp, "Matching the rate of concurrent tone bursts and light flashes as a function of flash surround luminance," *Perception & Psychophysics*, vol. 30, no. 1, pp. 33–38, 1981.

[24] T. Shipley, "Auditory flutter-driving of visual flicker," *Science*, vol. 145, no. 3638, pp. 1328–1330, 1964.

[25] M. Suteerawattananon, G. Morris, B. Etnyre, J. Jankovic, and E. Protas, "Effects of visual and auditory cues on gait in individuals with parkinson's disease," *Journal of the neurological sciences*, vol. 219, no. 1-2, pp. 63–69, 2004.

[26] "Captiv tea, france. http://teaergo.com/wp/,"

[27] J.-P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela, *et al.*, "Measuring phase synchrony in brain signals," *Human brain mapping*, vol. 8, no. 4, pp. 194–208, 1999.

[28] Z. Bylinskii, M. A. Borkin, N. W. Kim, H. Pfister, and A. Oliva, "Eye fixation metrics for large scale evaluation and comparison of information visualizations," in *Workshop on Eye Tracking and Visualization*, pp. 235–255, Springer, 2015.

[29] M. Shojaeizadeh, S. Djamasbi, and A. C. Trapp, "Density of gaze points within a fixation and information processing behavior," in *International Conference on Universal Access in Human-Computer Interaction*, pp. 465–471, Springer, 2016.

[30] E. F. Risko and A. Kingstone, "Eyes wide shut: implied social presence, eye tracking and attention," *Attention, Perception, & Psychophysics*, vol. 73, no. 2, pp. 291–296, 2011.