Tenth International Congress
on Sound and Vibration
7-10 July 2003 • Stockholm, Sweden

PERCEPTUAL ASPECTS OF AUDITORY-TACTILE ASYNCHRONY

M. Ercan Altinsoy

Institut für Kommunikationsakustik,
Ruhr University Bochum,
D-44780 Bochum, Germany
altinsoy@ika.ruhr-uni-bochum.de

Abstract

This paper addresses the temporal factors involved in the integration of auditory and tactile information. Perceptual threshold values for auditory-tactile asynchrony were measured using a broadband noise for the auditory stimulus (presented to the subjects via headphones) and a sinusoidal wave for the tactile stimulus (presented at the tip of the index finger via a shaker). In the first experiment, subjects were asked to make a three-alternative forced-choice judgment whether the audio stimulus and the tactile stimulus were synchronous, the audio stimulus preceded the tactile stimulus, or the tactile stimulus preceded the audio stimulus. Stimuli with audio delays in the range of –26 to 51 ms were judged synchronous. In the second experiment, the judgement was whether the auditory stimulus and the tactile stimulus were synchronous or asynchronous. The results showed that stimuli with audio delays in the range of –24 to 50 ms were judged synchronous. In the third experiment the subjects had to judge whether the audio stimulus preceded the tactile stimulus, or the tactile stimulus preceded the audio stimulus. Stimuli with audio delays in the range of –13 to 28 ms were judged synchronous.

INTRODUCTION

We obtain information from different sensory modalities by interacting with our natural world. The integration of the information from the main sensory modalities, namely, auditory, visual and tactile is one of the most important functions of our brain. Temporal factors, particularly synchronicity between different modalities is a important cue for the brain to integrate multi-modal information.

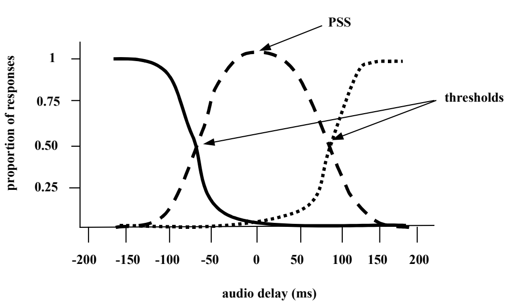
Multi-modal systems have been developed to take advantage of the multi-sensory nature of humans (Dix, Finlay, Abowd & Beale, 1993). They communicate with users by giving and also getting information from different sensory channels. Integration of multi-modal information is a important task for multi-modal system designers to produce compelling multi-modal display. Synchronization of different modalities in multimedia applications is a big problem. The processing time for different modalities is quite high with todays technology and causes high latencies and therefore delayed feedback reproduction. As the asynchrony between different modalities increases, the sense of presence and realism for the multi-media applications will decrease. Therefore an understanding of the perceptual aspects of multi-modal asynchrony is necessary prerequisite for multi-media designers.

Several studies have discussed the perceived simultaneity of the multi-modal stimuli. A multi-modal synchronisation threshold has been defined by Altinsoy et. al.(2000) as the maximum tolerable temporal separation of the onset of two stimuli, one of which is presented to one sense and the other to another sense, such that the accompanying sensory objects are perceived as being synchronous. In order to measure this threshold different psychophysical measurement methods have been applied. The schematic response patterns of different methods are shown in Fig.1. One response method asks the subject to make a three-alternative forced-choice judgment as to whether the stimuli are synchronous or which one was presented first (for auditory-visual (AV) asynchrony, Van de Par and Kohlrausch, 2000). The response pattern of the three-alternative forced-choice judgment method is shown in Fig.1a. The intersection between the curves of the "audio stimulus preceded the tactile stimulus" and "they were synchronous" defines the threshold for detecting asynchrony in the direction of delayed audio. The intersection between the curves of the "tactile stimulus preceded the audio stimulus" and "they were synchronous" defines the threshold for detecting asynchrony in the direction of delayed tactile stimulus. The maximum point on the synchronous curve indicates the point of subjective simultaneity (PSS).

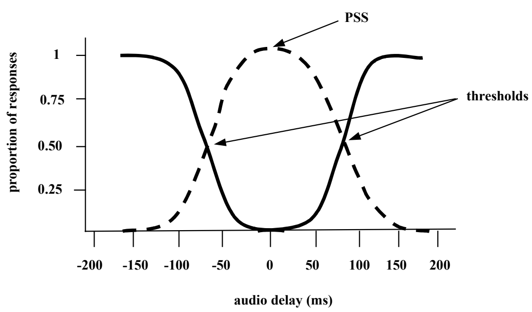
Another measurement method, is to ask the subject to judge whether the audio and the tactile are synchronous or asynchronous. Fig.1b shows the response pattern of this measurement method. In this method, two intersections between the curves of the "synchronous" and "asynchronous" indicate the thresholds for detecting asynchrony in the direction of delayed audio and delayed tactile stimuli. Again, the maximum point of synchronous curve indicates the point of subjective simultaneity (AV, Dixon and Spitz, 1980, Miner and Caudel, 1998, Van de Par, Kohlrausch and Juola, 2002).

The temporal order judgments (TOJ) is one whereby the subject has to judge the temporal order of an auditory and tactile stimuli, which results in the minimal multi-modal delay, for which subjects are able to indicate in which temporal order the two different sense are being stimulated (for auditory-tactile (AT) asynchrony; Hirsh and Sherrick, 1961, for AV asynchrony; Jaskowski, Jaroszyk, and Hojan-Jezierska, 1990). The response pattern of the TOJ method is shown in Fig.1c. The intersection between "tactile first" and "audio first" curves gives us the point of subjective simultaneity. The proportion of responses, being 25 % and 75 %, indicates the thresholds for detecting asynchrony in the direction of delayed audio and delayed tactile stimulus.

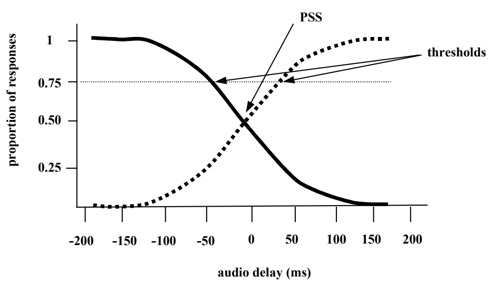
Methodological aspects for measuring asynchrony detection in audio-visual stimuli have been reported by Van de Par, Kohlrausch and Juola (2002). They found that the point of subjective equivalence in auditory-visual synchrony is shifted towards audio delays by about 35 ms compared to the point of objective equivalence. TOJ method allows for different decision strategies (for determining whether audio or video was leading even if the stimulus perceived as synchronous) and therefore results of the TOJ method depend on which strategy the subjects chooses. The other methods are rather robust and in agreement with each other.
In the last 10 years, haptic researchers have made significant progress in developing haptic feedback devices for multimedia applications, particularly virtual reality applications. Haptic feedback brings the sense of touch (tactile sense) to multi-media applications besides the mostly utilised modalities e.g. auditory and visual. Due to increasing usage of the tactile modality in multi-modal systems, the perceptual aspects of the interaction between tactile modality and other modalities is becoming more important, and consequently so is the temporal sensitivities between tactile and other modalities. This paper addresses the topic of temporal perception of auditory-tactile asynchrony.



a)



b)



c)

Fig1. The response patterns
a) 3 categories: tactile first, synchronous, audio first
b) 2 categories: asynchronous,synchronous
c) 2 categories: tactile first, audio first (*the response patterns a and c are adapted from Kohlrausch, 2000)*

If we observe our daily environment, we notice the strong physical relationship between sound and vibration. Sound is usually produced by vibrations of the objects and mostly the result of the our tactile interaction with objects in our environment, e.g. knocking a door or playing a piano etc. By knocking a door, we perceive tactile information from our hand and hear a knocking sound related to our action. After our interaction with the door, and until the tactile information arrives the brain, it takes a transduction time along the somatosensory pathway. Until the auditory information arrives the brain, it takes some time to generate sound, some time to reach the ears that is related to sound velocity and a transduction time along the auditory pathway. That means this strong relationship between sound and vibration and physiological realities play important role in the perception of simultaneous events.

In this paper data is presented on the perceptual threshold values for auditory-tactile asynchrony as determined by three different measurement methods. Furthermore, reaction time data for auditory and tactile stimulation is also presented. The results of the different measurement methods will be compared and discussed to establish the point of subjective simultaneity for auditory-tactile asynchrony.

EXPERIMENTS

Subjects

The same six subjects participated in the experiments. They were four right-handed men and two right-handed women with self reported normal-hearing ability. Their ages ranged between 22 and 32 years.

Stimuli and Set-up

The tactile stimulus was a sine wave and presented at the tip of the index finger of the participant via a B&K Type 4810 mini-shaker which is of the electrodynamic type with a permanent field magnet, with a maximum stroke 6 mm and force rating 10 Newton sine peak in the vertical direction. The shaker delivered the stimuli to the skin via a vibrating probe. The probe was 4 mm in diameter. The shaker was located inside a wooden box, which contained a circular hole on which the participant placed their index finger. A further necessity of the box was to mask the visual information, which occurs from the shaker. To minimize the structural vibrations which were generated by shaker, the floor was isolated from the shaker by using some vibration damping materials.

The auditory stimulus was a burst of white noise presented from a PC. The noise was amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones which has a very high sound isolation level and therefore masked the background noise of shaker when it generated the signal. The experiments were conducted in sound-attenuated room.

The durations of the stimuli were 25 ms. It is possible that the intensities of auditory & tactile stimuli has a important influence on perceptual asynchrony. Therefore a cross-modal intensity matching experiment (Stevens 1975) was conducted to determine a suitable sound pressure level and vibration intensity level. In this level-matching experiment the participants task was to match the apparent loudness of the burst of white noise to the apparent strength of the vibration on their finger. The tactile stimulus were presented randomly at six different levels, 35 dB – 65 dB re 1 micron (0.07, 0.18, 0.6, 0.75, 1.3, 1.6mm) (each stimulus was presented 20 times) and subjects adjusted the level of the sound by using an amplifier until its apparent loudness seemed as great as the strength of the vibration on their finger. In Fig.2, the medians of the sound pressure level are plotted against the vibration amplitude (RMS).

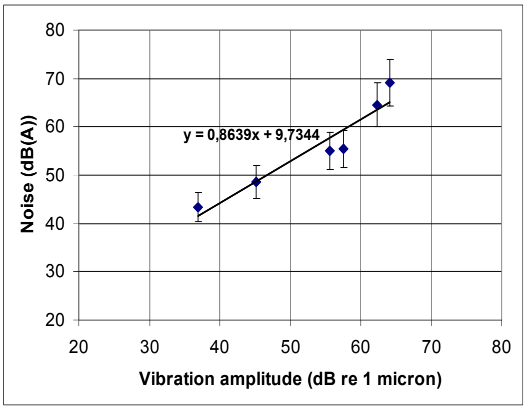


Figure 2. Equal sensation functions relating sine vibration on the finger tip to the intensity of a burst of the white noise.

The power equations according to Stevens (1975) for the two modalities;

$$\psi_s = \phi_s{}^m \quad (1)$$

$$\psi_v = \phi_v{}^n \quad (2)$$

where ψ is subjective magnitude, Φ is stimulus magnitude, m is characteristic exponent for noise, n is characteristic exponent for vibration, s indicates auditory modality and v indicates vibration modality. If the participant equates subjective magnitudes by cross-modal matching experiment

$$\psi_s = \psi_v \quad (3)$$

$$\phi_s = \phi_v{}^{n/m} \quad (4)$$

The obtained exponent from the equal sensation function which was determined by the results of the cross-modal matching experiment is n/m = 0.86.
The sound pressure level was set to 56 dB which was shown to match a vibration amplitude of 58 dB, which is a displacement of 0.6 mm.

Methods

Altogether four different experiments were conducted. In the first three experiments synchronisation thresholds and point of subjective simultaneity of auditory-tactile presentations were measured. Stimuli were presented in a random order with an audio delay ranging from -200 to 200ms with varying step sizes (-200 to150 ms, 50 ms steps; -120 to –60 ms, 20 ms steps; -60 to 60 ms, 10 ms steps; 60 to 120 ms, 20 ms steps;150 to 200 ms, 50 ms steps). Each condition was presented twelve times. Negative delay values indicate that the auditory stimulus was presented first, and positive delay values indicate that the tactile stimulus was presented first.
In the first experiment response categories were “tactile first”, “synchronous” or “audio first”. In the second experiment response categories were “synchronous” or “asynchronous. In the third experiment response categories were “tactile first” or “audio first”. Condition-order of the three experiments was counter balanced across the subjects according to a Latin square.
The fourth experiment was carried out to measure auditory and tactile reaction times. Subjects were asked to respond as quickly as possible to the stimulus by pressing a button. A warning signal was presented to the subject before each trial. As in the study of Jaskowski, Jaroszyk and Hojan-Jezierska (1990), the stimulus followed the warning signal after a random fore-period. The fore-period was a sum of a fixed interval of 1s and an interval sampled from an exponential distribution with mean equal to 1s. Each modality was stimulated alone. One session consisted of 100 trials, and each subject joined four sessions.

RESULTS

The point of subjective simultaneity and synchronisation threshold values are shown according to measurement method in Table 1. The proportions of the responses of all three experiments for each response alternatives are shown in Figure 4. The mean reaction times to auditory stimulus and tactile stimulus are presented in Figure 3.
The results of the first experiment are depicted by thick lines and the filled black symbols. The black triangles indicate “tactile first” responses, the black diamonds indicate “synchronous” and the black squares indicate “audio first”. The synchronous curves seem to peak for slightly positive audio delays. This shift can be seen especially clearly in the results of the Subjects S1, S2 and S6. In the first experiment stimuli with audio delays in the range of –26 to 51 ms were judged synchronous.

	PSS’s			Synchronisation Thresholds					
	Exp.1	Exp.2	Exp.3	Exp.1		Exp.2		Exp.3	
S1	10	10	2	-10	36	-10	30	-15	10
S2	8	8	-5	-15	40	-15	35	-14	12
S3	10	10	-2	-60	60	-45	75	-8	10
S4	0	10	12	-35	70	-20	50	0	18
S5	5	-5	-2	-20	53	-28	54	-10	30
S6	2	12	0	-20	52	-23	57	-28	90

Table 1: PSS’s and synchronisation thresholds in milliseconds for six subjects.

The results of the second experiment are depicted by the thin lines and the white symbols. White diamonds indicate the “synchronous” responses, and the white circles indicate “asynchronous” responses. As similar as first experiment, the PSS is shifted toward positive audio delays. The synchronisation thresholds which are found in the second experiment are also very similar to the synchronisation thresholds which are found in the experiment. Only subject S4 has lower threshold values in the second experiment than in the first experiment. In both experiment (first and second) the transition between “audio first” and “synchronous” responses is sharper than between “tactile first” and “synchronous”. The results of the second experiment showed that stimuli with audio delays in the range of –24 to 50 ms were judged synchronous.

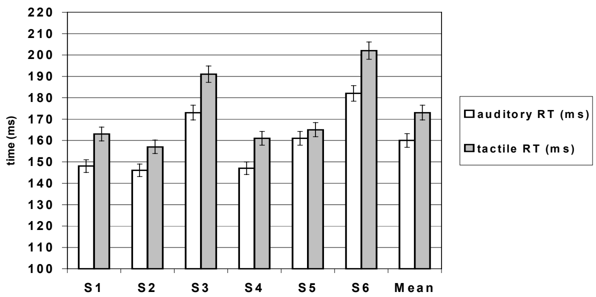


Figure 3: Auditory and tactile reaction times (ms)

The results of the third experiment are depicted by the dotted lines. The multiplication signs indicate “tactile first” responses and the plus signs indicate “audio first” responses. Three subjects S2, S3 and S5 show negative PSS, two subjects S1 and S4 show positive PSS, and one subject S6 has a PSS that is zero. For the subjects S3, S4, S5 and S6, intersection of the curves for “tactile first” and “audio first” is in the area of where subjects responded with “synchronous” in the first two experiments. For subject S1 and S2, the transition coincides with the intersection of “audio first” curve with the “synchronous” curve of the first experiment. The results of the reaction time (RT) experiment show that the participants react 13 ms (SEM 2.17ms) quicker with an auditory stimuli than a tactile stimuli (Fig.3). A paired sample t-test shows that RT's are significantly improved with a noise burst compared to tactile stimulation, $t(5) = 5.878$, $p<0.01$.

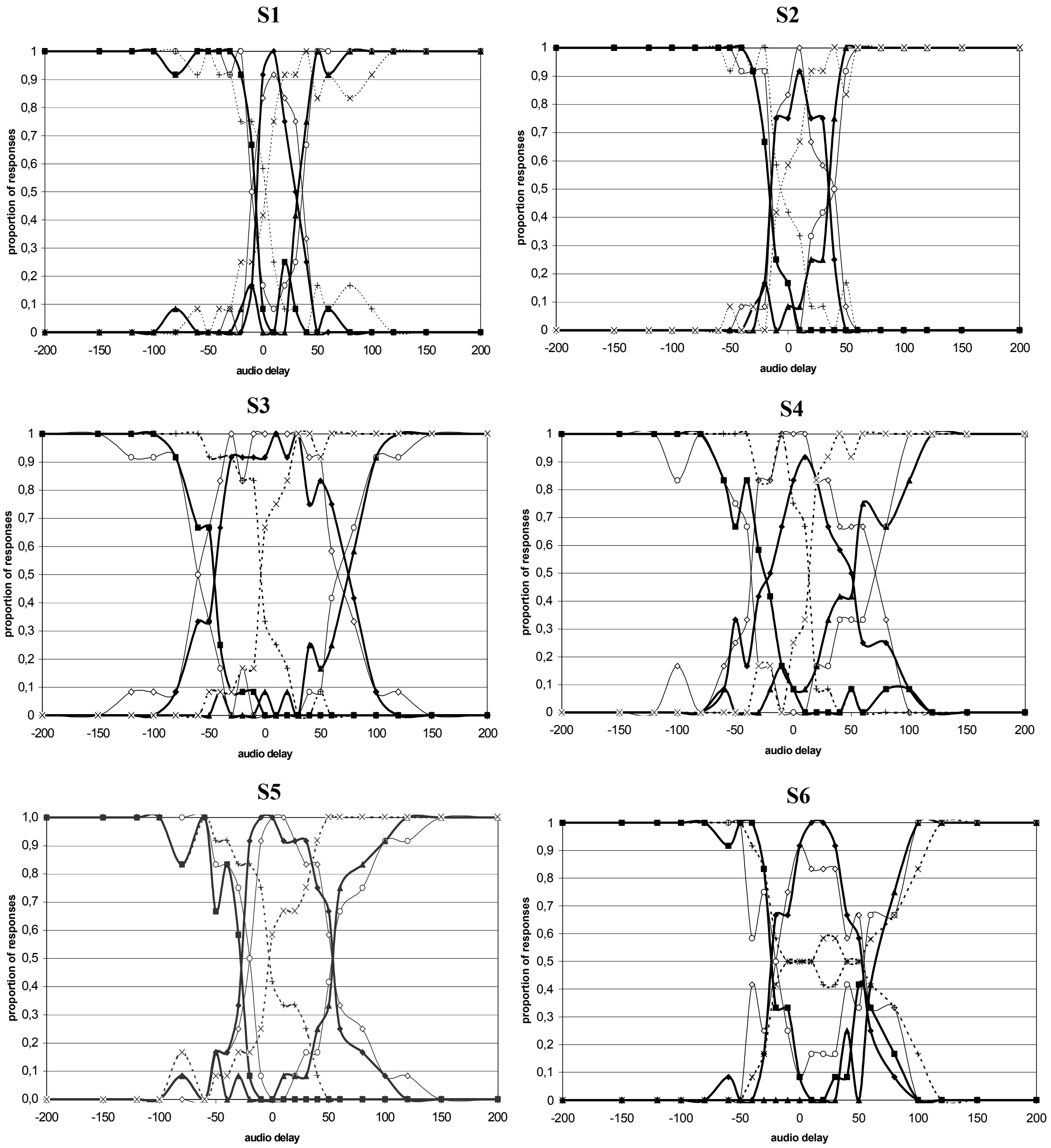


Figure 4. The results which are a proportion of responses as a function of audio delay, of the three experiments are shown for each subject in a separate panel.

DISCUSSION

The results of the all three experiments show that point of subjective simultaneity does not coincide with the point of objective simultaneity (0 ms). The PSS is found at an audio delay of about 7 ms. The most interesting finding is that audio advances are detected better than audio delays. This facts may be linked to the physical rules, e.g. speed of sound. The distance between our hands and ears is about 1 m, therefore sound would take about 3 ms longer to reach us than tactile stimulus. Also physiological realities, the transduction time along the auditory neural pathway and somatosensory neural pathway is different. The reaction time experiments show this difference. The reaction times are 13 ms shorter to auditory stimulus than tactile stimulus. Therefore it is possible that human perceptual system is adapted to tolerate larger audio delays than tactile delays, as suggested for audio-visual asynchrony by van de Par, Kohlrausch, 2000 and Dixon, Spitz, 1980.

The results of the first and second experiments are in agreement each other. However in the third experiment (TOJ) there is a inconsistency between subjects. The possible reason can be that subjects adopted different decision criteria for determining whether audio or tactile leading, as suggested for audio-visual asynchrony by van de Par, Kohlrausch and Juola, 2000.

CONCLUSIONS

This study investigated perceptual threshold values for auditory-tactile asynchrony. It was found that an asynchrony is easily detected when an auditory stimulus precedes a tactile stimulus. The point of subjective simultaneity is shifted towards audio delays by about 7 ms compared to the point of objective simultaneity.

ACKNOWLEDGEMENTS

This work was supported by the International Graduate School for Neuroscience (**IGSN**) at the Ruhr University Bochum.

REFERENCES

Altinsoy, M.E., Blauert, J. & Treier, C., "Inter-Modal Effects of Non-Simultaneous Stimulus Presentation", In A. Alippi (Ed.), Proceedings of the 17 th International Congress on Acoustics. Rome, Italy, (2001)

Dix, A., Finlay, J., Abowd, G. & Beale, R., *Human-Computer Interaction* (New York. Prentice Hall., 1993)

Dixon, N.F., and Spitz, L. "The detection of auditory visual desynchrony", Perception, Vol.9., pp.719-721

Hirsh I.J. & Sherrrick C.E. "Perceived order in different sense modalities", J. Exp. Psychol., 62, 423-432 (1961)

Jaskowski, P., Jaroszyk, F., and Hojan-Jezierska, D., "Temporal-order judgment and reaction time for stimuli of different modalities", Psychol. Res., Vol.52, pp.35-38, (1990)

Kohlrausch A., „Perceptual consequences of asnchrony in audio-visual stimuli", IPO Ann. Progr. Rep.35, 140-149, (2000)

Miner N., Caudell T., "Computational requirements and synchronization issues for virtual acoustic displays", Presence, 7, 396-409, (1998)

Stevens, S.S., "*Psychophysics: Introduction to its perceptual, neural and social prospects",* (New York: Wiley, 1975)

Van de Par, S., and Kohlrausch, A., "Sensitivity to Auditory-Visual Asynchrony and to Jitter in Auditory-Visual Timing", Human Vision and Electronic Imaging V, Bernice E. Rogowitz, Thrasyvoulos N.Pappas, Editors, Proceedings of SPIE, Vol. 3959, pp. 234-242, (2000)

Van de Par, S., Kohlrausch, A., and Juola, J.F., „Some methodological aspects for measuring asynchrony detection in audio-visual stimuli", European Acoustics Association (Hg). in: Proceedings of Forum Acusticum Sevilla 2002. 2002 Jg. Nr. Vol. XXXIII. Madrid: Sociedad Espanola de Acustica / Hirzel, (2002).