Université d'Ottawa · University of Ottawa

PERCEPTION OF SYNCHRONY BETWEEN AUDITORY AND VISUAL STIMULI

Catherine Lee

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Experimental Psychology

University of Ottawa

July 2002

0-612-76600-4

I dedicate, with my entire heart, this thesis to my beloved Trinity; God of mercy, Jesus of love, and Holy Spirit the invisible guide. The dedication also goes to the feet of my very beloved Mother Virgin Mary of tenderness.

I would also like to dedicate this thesis to my parents, Noel and Stephen, and my sisters-in-law, Christina and Najung, and my brothers, Andrew and Reno.

ACKNOWLEDGMENTS

I would like to thank my supervisor George Fouriezos for his advice and supervision in the development, execution, and writing of this thesis.

I would also like to thank professor Catherine Bielajew and professor Catherine Plowright who sat on my thesis committee for their advice and suggestions.

Thanks also to Michelle Wesley for assisting in making graphics of three Figures.

ABSTRACT

The literature has fairly consistently reported a difference in how well humans perceive synchrony depending on the order of auditory and visual stimuli. When the auditory stimulus occurs first and the visual stimulus follows, subjects are more sensitive and so perceive asynchrony with smaller time delay between the stimuli. On the other hand, when the auditory follows the visual stimulus, the subjects are more tolerant and perceive stimuli with larger time delays as synchronous. Thresholds of synchrony perception in these two conditions are thus asymmetrical.

The present study attempts to test the Lewkowicz Model, by which the asymmetrical thresholds are explained as a result of arrival-time differences between auditory and visual stimuli to the brain, such that visual stimulus takes longer in processing to be perceived versus auditory one. Reaction-times to these stimuli were measured to determine the arrival-time difference and plotted with synchrony perception. On the basis of Lewkowicz Model we predicted that reaction-time difference between the two stimuli correlate with subjective synchrony. The results did not support the Lewkowicz Model. The expected tendency of 30-40ms of subjective synchrony was not shown. The subjects took, in average, only 7.7ms to detect asynchrony when the auditory stimulus followed the visual stimulus. That the subjects did not tolerate greater temporal gap when the auditory followed versus when it preceded the visual stimulus was a very different result from majority of previous studies. Different factors in perceiving synchrony are discussed in this paper, as well as the application of the research in telecommunications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TALBES

INTRODUCTION

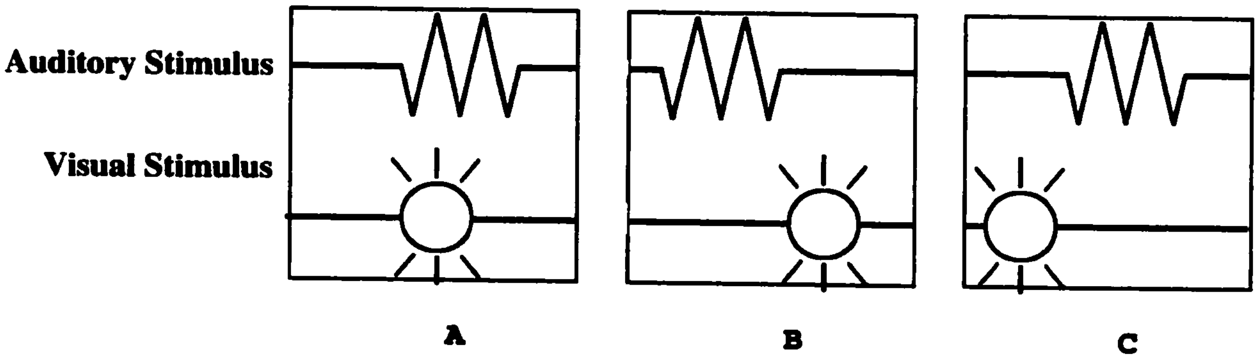
We constantly receive simultaneous or nearly simultaneous pieces of information from diverse sensory modalities and experience them together as a comprehensive event. We can integrate stimuli from separate modalities to perceive them as a single event. In singing, for instance, visual lip movements and auditory-utterance signals occur close to each other in time, and this allows the listener to unite these two stimuli as a single event. Conversely, the listener may detect separate stimuli if the singer's lip movements are not close enough in time to the song's lyrics. So, temporal synchrony of stimuli is essential in understanding and interacting with the world.

Today, telecommunication has become a part of everyday life, and its success highly depends on synchronized audiovisual experience. Using teleconferencing systems businesspeople attend the meetings in their own offices with their partners in other parts of the country or the world. High technology adventurers enjoy a totally different world in a simulated virtual environment (VE). These people all benefit from synchronization of multisensory stimuli. If the businesspeople saw their partners on the screen and heard them speak at different points in time, the meting would not be effective as they should be. Adventurers' enjoying a realistic and compelling virtual experience relies on their perceiving sensory stimuli in synchrony. As these examples illustrate, our synchronization threshold — or the amount of delay we can experience between the onset of a first-occurring stimulus and that of a later-occurring stimulus while still perceiving them as being in synchrony — plays a vital role in our experience of the world.

Research on the perception of the temporal synchrony of auditory and visual inputs has for a long time been conducted using psychophysical techniques. Psychophysicists have been studying the perception of the temporal synchrony and asynchrony by measuring a threshold. This threshold is the amount of time needed between auditory and visual stimuli for subjects to detect asynchrony. Temporal asynchrony, as in the lip-synch-singing example, is detected as soon as the time discrepancy of stimuli exceeds a certain threshold. It is of value to know how the threshold changes as some aspects of the stimuli change, such as their nature or intensity.

To study the detection of temporal asynchrony, experimenters present multimodal stimuli at diverse points in time, and the researchers in the literature discuss the amount of time separating multimodal stimuli in terms of delay between the two stimuli. Figure 1 illustrates the presentation of stimuli over time: Panel A shows auditory and visual stimuli occurring simultaneously; Panel B shows auditory preceding visual stimuli; and Panel C shows visual preceding auditory stimuli. When we discuss delay, or interval, then in referring to a small delay of 50ms, we mean that 50ms separates the onsets of the two stimuli. We designate the visual stimulus as the point of reference in relation to the auditory stimulus. For convenience, negative delays, such as – 100ms, –200ms, designate the amount of time by which the auditory precedes the visual stimulus (AV condition). One knows, for instance, that –200ms of delay means that an auditory stimulus occurred 200ms before a visual stimulus. On the other hand, if the auditory follows the visual stimulus (VA condition), one usually indicates with positive delays, such as

Figure 1. Diagrams of the presentation of auditory and visual stimuli; A: synchronous pair; B: auditory stimulus preceding (AV); C: auditory following (VA).



100ms, 200ms. For example, 200ms of delay means that an auditory stimulus is presented 200ms after a visual stimulus.

There are four basic methods used to evaluate auditory–visual synchrony. First, in Order method the auditory and visual stimuli are presented one after the other in random sequence, with varying delays. This method requires subjects to judge which one of two heteromodal stimuli occurs first. So, the task is to indicate which stimulus came first, the auditory or the visual. If the subject is not biased in favour of synchronous auditory and visual stimuli (zero delay), 50% of the judgments should be Auditory First, and the other 50% should be Visual First. The threshold for the AV condition is obtained at 75% of Auditory First responses. The threshold for the VA condition, on the other hand, is computed at 25% of judgments.

Second method, In-Synchrony or Out-of-Synchrony involves presenting synchronous and asynchronous auditory and visual stimuli at random. The subject's task is to judge whether the two stimuli occurred together or separately. When delays are large, such as –300ms or 300ms, one can readily detect the asynchrony of stimuli. As the delays become smaller, however, a subject's ability to detect asynchrony decreases, and eventually, at a certain delay point, the asynchrony is not accurately detectable at all. In theory, the graph should appear as a U-shape distribution. The thresholds between synchrony and asynchrony are measured where stimuli are judged as being In-Synchrony in half the trials (50%) and Out-of-Synchrony in the other half (50%).

A third method is that of Adjustment. In this method, the delay between stimuli is adjusted after each trial. Subjects adjust for themselves the magnitude of the delay between stimuli. After a pair with auditory preceding is presented first, subjects can press a button repeatedly to edge the stimuli of the next pair come closer together, with goal of attaining the temporal synchrony. When subjects judge that synchrony is attained for a particular pair of stimuli, they can press another button to continue to next.

The fourth method, Better Synchrony (Figure 2) is an elaboration of the In-Synchrony or Out-of-Synchrony method. In this method, subjects are presented with two pairs of auditory and visual stimuli and are to judge which of the two pairs come more closely than the other. One of the pairs is always in synchrony, while the other comes with varying delays. The subjects are asked to judge which pair is in better synchrony. The Better Synchrony method is used in the present study and further described in the Method section.

Using all the four methods, AV and VA thresholds can be estimated. The interval between the two thresholds refers to the delays of uncertainty in which subjects cannot differentiate synchrony against asynchrony. Subjective synchrony, the centre between the two thresholds indicates a subject's point of perceived synchrony between the two stimuli. It can be estimated in theVA condition, where in reality auditory follows visual stimulus and is referred to as positive deviation of subjective synchrony. Or, it can be estimated in the AV condition, where auditory precedes visual stimulus and referred to as negative deviation (Figure 3). In the next section, we give an overview of the literature on the perception of synchrony where researchers reported using one or more of the methods described above.

Figure 2. Method of Better Synchrony

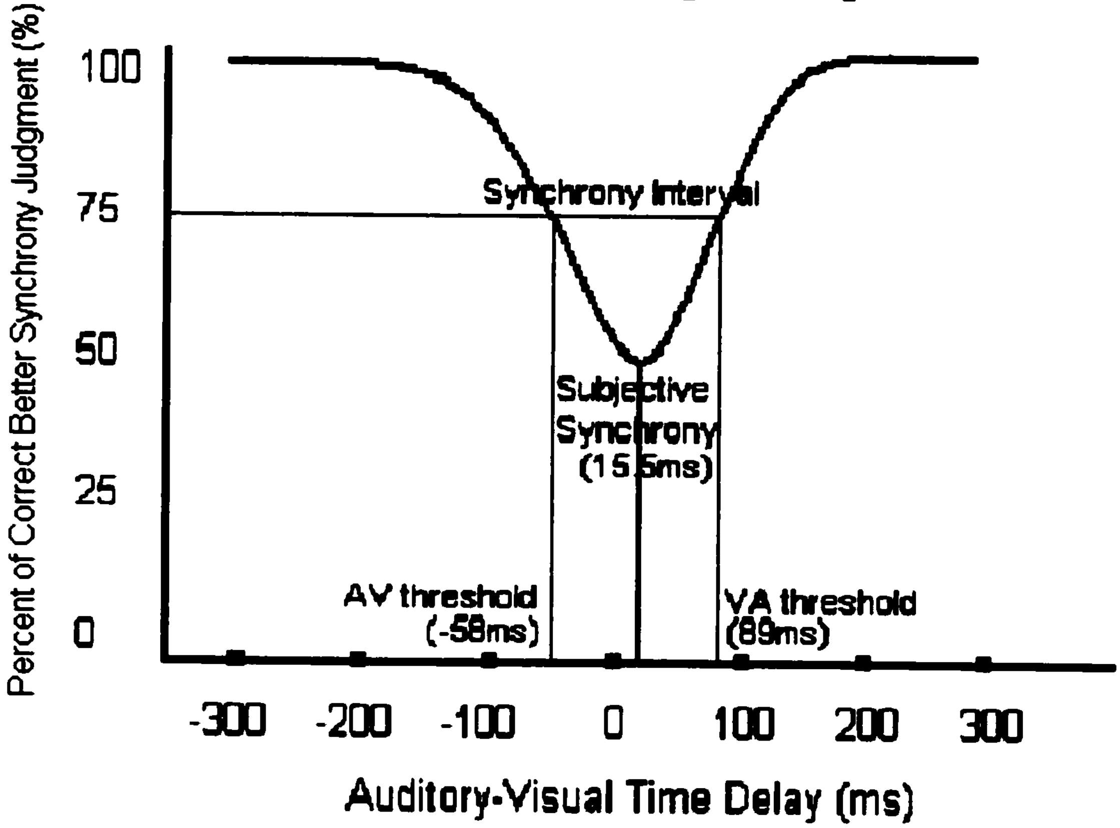
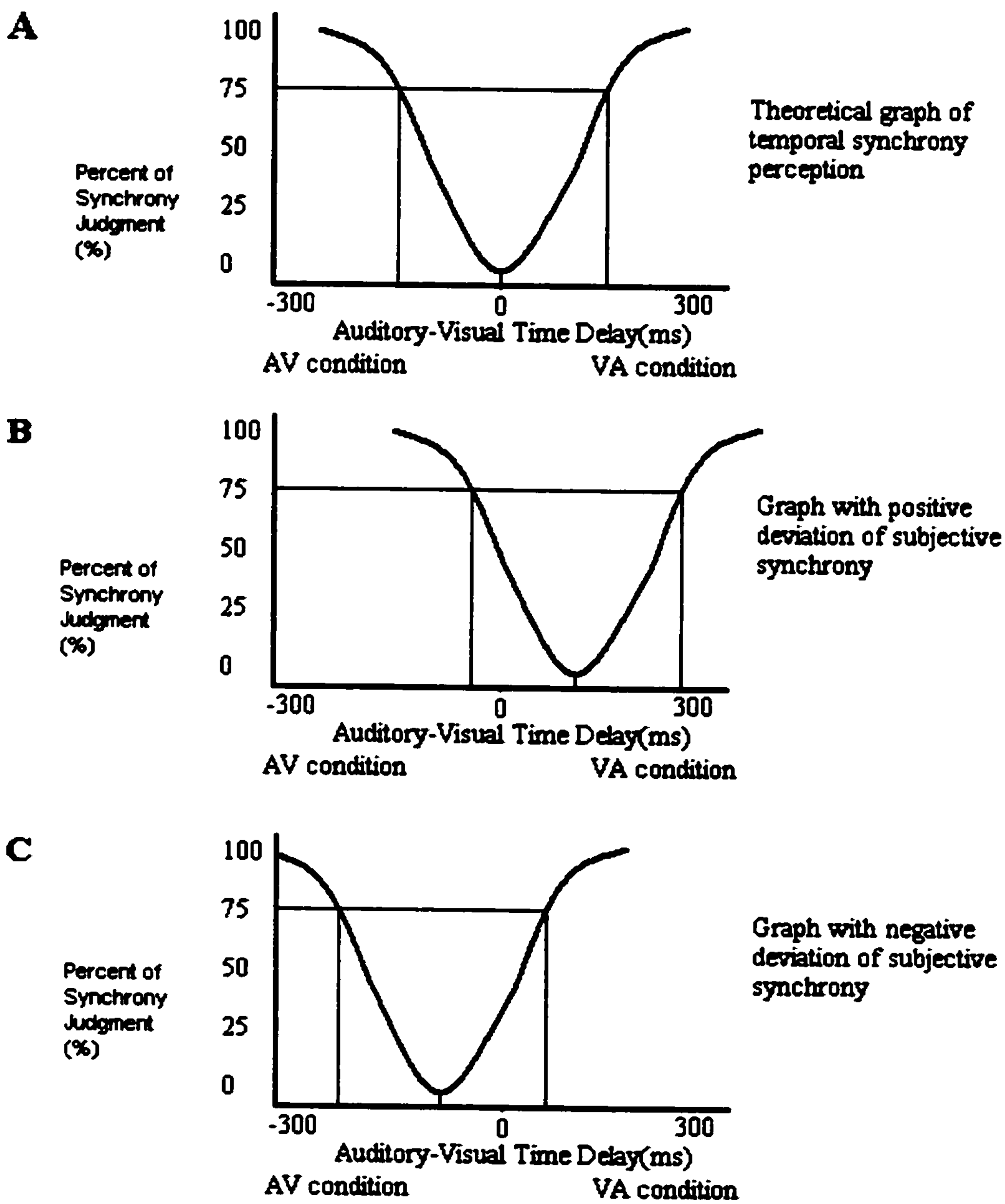


Figure 3. Positive and negative deviations of the subjective synchrony. A: Theoretical graph of temporal synchrony perception; B: Positive deviation indicates the shift of the subjective synchrony towards VA condition in relation to 0 ms; C: Negative deviation refers to that one's subjective synchrony is moved to the AV condition compared to the physical synchrony.



Literature Review

The studies included in the Literature Review date from as early as Hirsh and Sherrick (1961) to as recently as Bushara et al. (2001). Each study examined the perception of temporal synchrony between auditory and visual stimuli. Most found a temporal asymmetry of thresholds, an effect in which VA threshold is larger than AV threshold, and thus a positively deviated subjective synchrony. The findings from the literature reported here will be found to contrast with the present study's results. Some of the studies in the literature explained the cause of this positive deviation of subjective synchrony, which will be useful in understanding the difference in results between most of the literature and the present study.

The extensive investigations of temporal perception conducted by the pioneer psychophysicist, Exner (1875), were reviewed by Hirsh and Sherrick in 1961. Exner studied the thresholds for detection of separate stimuli, determining how close together in time two stimuli may occur and yet be perceived as separate. He explored temporal perception in both the same and various modalities. In an experiment using two brief flashes of light, he found the threshold for perception of the non-simultaneity of stimuli at 44ms. When subjects were presented with two separate flashes of light, one after the other within less than 44ms, they perceived the two stimuli as one long flash. Only when 44ms or more occurred between the two flashes of light did subjects start to detect the presence of two flashes.

In determining the auditory threshold, Exner (1875) used the successive clicks of a Savart wheel to provide stimuli. He found that to give a judgment of asynchrony subjects needed a minimum temporal gap between clicks of about 2ms. This indicates the superior acuity of the auditory system than visual system. Subjects required much less delay between auditory signals than between visual signals to perceive them as separate events.

Exner's experiment on the thresholds in diverse sense modalities was also reviewed in Hirsh and Sherrick (1961). According to Exner's findings, the thresholds were not symmetrical; that is, the value of temporal separation required for asynchrony perception depended on which of the two sense modalities carried the first stimulus. Also, subjects were less certain in their perception of the intervening time. Whether the auditory or visual stimulus came first affected the judgment of asynchrony, although the delay was in each case identical. Exner found the smallest perceptible time delay at –6ms when the auditory preceded the visual stimulus but found it at 16ms when the auditory followed the visual stimulus.

Hirsh and Sherrick (1959) themselves examined the perception of the temporal order of stimuli within the same modality, starting with auditory stimuli. They varied pairs of sounds in time of occurrence, frequency, spectrum, and duration, and they asked the subjects to judge which sound came first, such as the sound with higher or lower pitch or the longer or shorter sound. The threshold between the two sounds the subjects required to make correct judgments of asynchrony in 75% of the trials was 20ms, regardless of the kinds of sounds.

In Hirsh and Sherrick's (1961) experiment to examine visual temporal perception, they used two flashes of light. These were identical except in appearing in different places in relation to either a horizontal or a vertical line. To obtain the threshold between the two visual stimuli, Hirsh and Sherrick asked the subjects to report on whether the light located to the left of or to the right of or above or below the line came first. Their results were of interest because the delay required for subjects to judge which of two similar stimuli occurred first was very close to that found in the auditory experiment. In other words, for the perception of temporal order in the same modality, the delay between two auditory stimuli is similar to that between two visual stimuli. Hirsh and Sherrick (1961) then investigated the perceived order of stimuli in two sense modalities. They varied the amount of time by which a visual flash preceded or followed an auditory tap and asked the subjects to judge which of the two sensations came first. They found the same value for the threshold, 20ms, whether the stimuli were auditory and visual or auditory alone, and they concluded that for both intermodal or unimodal stimuli, the threshold needed for subjects to make correct judgments of asynchrony in 75% of trials is around 20ms. In either case, judging the order of stimuli requires subjects to organize two pieces of information in time (Hirsh & Sherrick, 1961). However, according to Hirsh and Sherrick's results, the threshold for discrimination is not affected by whether the two pieces of information come in the same or in different modalities.

In some of the more recent literature, Dixon and Spitz (1980) reported a positive deviation in subjective synchrony. Using a film and its soundtrack (given through a video recorder), these researchers manipulated the amount of physical asynchrony between picture and sound. Subjects were shown two film presentations; one was of a man reading prose and the other of a hammer hitting a peg. In half the trials for both film presentations, the sound of the voice or hammer was gradually advanced in relation to the picture, and in the other half the sound was gradually delayed. Using the method of Adjustment, they asked the subjects to respond by releasing a button, which they had been instructed to press as soon as they detected asynchrony between the sound and the picture.

In both film presentations, a greater sensitivity to "desynchrony" (Dixon and Spitz's term for asynchrony) was found when a sound preceded, rather than following, a picture. When the voice stimulus was used, the AV threshold was −131ms; the VA threshold was 258ms; and the subjective synchrony, the average of the two thresholds, was 63ms. When the hammer stimulus was used, the AV threshold was −75ms; the VA threshold was 188ms; and the subjective was synchrony 51ms. Subjects all had positive deviations, in which they were more tolerant to the sound following the visual stimuli. Each subject was quicker to detect asynchrony when the sound of the voice or the hammer came first, that is, when the picture of the man reading or the hammer hitting the peg came later.

Positive deviation and larger threshold in the VA condition were also reported in McGrath and Summerfield's (1985) study of auditory and visual speech perception. Subjects were presented with an lip-like Lissajou interference pattern, mimicking the opening and the closing of lips, as a visual stimulus, and rectangular pulses at various delays, either before or after the opening of the lips. The temporal asynchrony between the Lissajou figure and the tone was detected at −79ms when the tone preceded (AV condition), but it was detected at 138ms when the sound lagged (VA condition). The subjective synchrony was 30ms. When the tone actually came 30ms after the movement, McGrath and Summerfield's subjects still perceived the Lissajou motion and sound as belonging to one event.

Using asynchronous auditory and visual stimuli, Smeele (1994) examined the relation between bimodal speech perception and the ability to detect asynchrony between auditory and visual speech stimuli. Smeele (1992 as cited in 1994) earlier found that delays between stimuli (or "desynchronizations") that did not affect the ability of subjects to identify syllables ranged from −160 ms when the auditory stimulus came first to 280 ms when the visual stimulus came first. This finding was similar to those of other studies reviewed here.

Smeele then sought to know the origin of this asymmetry. In an earlier study, Smeele (1992 as cited in 1994) suggested that humans do not perceive small desynchronizations, and therefore they have no effect on speech perception. As the size of desynchronizations becomes larger, asynchronous speech and lip movement can be detected more easily, and the asymmetrical thresholds become more apparent. It is manifested in a positive deviation of subjective synchrony. Smeele speculated that the asynchronous thresholds could stem from differences between the processing of auditory and that of visual information, either in the early stages of perception or in short-term information storage. And so they hypothesized that the detection of asynchrony and the perception of asynchronous visual and auditory speech are reflections of the same underlying perceptual mechanism for auditory and visual stimuli. Accordingly, they carried out two experiments.

In the first experiment, the subjects judged the order of speech sources, that is, whether the visual signal preceded the auditory signal or vice versa. For the stimuli, they presented 10 CVC (Consonant-Vowel-Consonant) syllables, such as *bax* and *maf*, both synchronously and asynchronously, within a range of −550ms to 350ms. For each syllable, the subjects' subjective synchrony had a negative deviation. That is, the subjects' subjective judgments of synchrony were at points where in fact the auditory signal occurred before the visual signal. The average of the subjects' subjective synchrony was −105ms (with a range of −225ms to −14ms). These deviations to negative values were not in line with the positive deviations reported in other studies.

This kind of asymmetrical thresholds effect was demonstrated earlier in Hirsh and Sherrick's (1961) study, in which the time between successive stimuli required for the judgment of temporal order was much longer than that for a judgment of synchrony or asynchrony. To explain this difference in performance in temporal-order and asynchrony judgments, Sternberg and Knoll (1975) argued that the perception of order is based on information different from, and to some extent independent of, the information used in the perception of synchrony and asynchrony. As well, they noted that the perception of asynchrony does not always lead to a correct perception of order. Thus, the negative deviation found in the first experiment in Smeele's (1994) study may have been a result of subjects performing a temporal-order task. This led Smeele to undertake a second experiment to determine whether the subjects' judgments in a temporal-order task may differ from those of asynchrony detection.

In this second experiment, subjects judged whether visual and auditory stimuli were in synchrony or out of synchrony. The average thresholds, determined at 50% accuracy of judgment for asynchrony in the AV condition, were −207.8ms (with a range of −319ms to −216ms); and in the VA condition, 184.4ms (range of 57ms to 295ms). The average subjective synchrony for 50% accuracy of In-Synchrony judgment, for the 10 syllables, was −11.7ms (with a range of −131ms to 78ms). Five syllables had negative subjective synchrony, and the other five had positive subjective synchrony. As such, the subjective synchrony differed in the first and second experiments. This is in line with Hirsh and Sherrick's (1961) findings and supports the suggestion made by Sternberg and Knoll (1975) that one makes a decision using different information in the two tasks. Having compared the results of the two experiments, Smeele (1994) explained that there were asynchronies between audition and vision that could be detected but that did not affect speech perception. She concluded that judging the temporal order of auditory and visual stimuli and judging the synchrony and asynchrony of stimuli are two distinct tasks, relying on different perceptual mechanisms.

The difference in the two tasks was attributed to a difference in the strategies of the subjects (Smeele, 1994). Specifically, subjects pay attention to the difference in arrival times of the visual and auditory stimuli in the temporal-order task, whereas they attempt to integrate two speech stimuli in the asynchrony-detection task. Smeele inferred that detection of asynchrony is not directly associated with the perception of asynchronous auditory and visual speech.

In 1999, Kohlrausch and von de Par argued that in speech processing, the auditory signal provides enough information for a good understanding of speech. But, if auditory information is unclear, visual information from a speaker's lip movements can help in understanding speech. Having information from both sources of stimuli improves the perception of speech, and it is possible to integrate lip movement and speech of poor auditory quality as synchronous if they occur within the thresholds of –79ms when the auditory stimulus precedes and 138ms when it follows. However, the researchers found that their subjects lost the improvement in their understanding of speech when the auditory followed the visual stimulus by more than 160ms. When the auditory followed the visual stimulus by 160ms or more, the subjects no longer perceived the temporal synchrony of the auditory and visual stimuli but saw them as two events occurring at different times.

Later, in 1996, a work that forms the special focus of the present study appeared. Lewkowicz (1996) investigated perception of auditory and visual temporal synchrony not only in adults but also in infants. To determine whether changes in tolerance to delay when perceiving intersensory temporal synchrony occur during the first year of life, the researcher investigated developmental differences in the perception of intersensory temporal synchrony and asynchrony in infants versus adults, ranging in age from 2 to 8 months. The experimental method involved, first, habituating the infants to a combination of a small disk that bounced up and down on a video monitor and a short sound that occurred in synchrony. After the infant subjects met a predetermined habituation criterion, the researcher administered three asynchrony trials and one habituated-synchrony trial. During the asynchrony trials, the disk bounce and the sound occurred at different times. To determine whether the infants detected the asynchrony in any of these asynchrony trials, Lewkowicz compared the degree of response in a given asynchrony trial with that of the habituated-synchrony trial.

To establish whether the methods used in the experiments with infants would give rise to asynchrony thresholds in adults comparable to those reported in other studies, Lewkowicz carried out an experiment using the same stimulus materials and similar testing methods for adult subjects. Accordingly, adult subjects participated in seeing and hearing the same synchronous disk bounce and sound. Following the synchrony-habituation phase, the adult subjects were tested in AV and VA asynchrony trials. During the AV asynchrony trials, Lewkowicz presented the subjects with four asynchrony delays of −50, −80, −110, and −140ms. During the VA asynchrony trials, subjects were presented with the four asynchrony delays of 110, 140, 170, and 200ms. The average AV asynchrony threshold was −65ms, and the average VA asynchrony threshold was 112ms. Subjective synchrony was 23.5ms.

The experiment with infant subjects showed that the infants detected asynchronies at greater than −350 ms when auditory stimulus preceded and that they detected asynchronies at greater than 450ms when auditory stimulus followed. In the infants, the minimum delay needed to detect the VA asynchrony was longer than that needed to detect the VA asynchrony. This was consistent with the researcher's finding for adult subjects.

Lewkowicz then explained the difference in the thresholds between the AV and the VA conditions in terms of the difference in neural transduction times. Another 30–40ms is required for transduction of stimuli in the VA condition, and this leads to temporal asymmetry of thresholds. He used the term "psychological interval" for time passing between perceptual registration of the two stimuli, and his schematic representation of the psychological interval is shown in Figure 4. The upper panel illustrates the concept of a psychological interval in relation to the physical interval in the AV condition, and the lower panel displays the relation between psychological and physical intervals in the VA condition. The psychological interval for AV asynchrony comprises physical delay (−100ms) plus the 30–40ms needed for transduction of the visual stimulus. When the visual follows the auditory stimulus by 100ms, subjects perceive the stimuli as if the visual followed the auditory stimulus by 130–140ms, as a result of the extra neural transduction time for the visual stimulus.

Figure 4. Schematic representation of the role of the psychological interval and the physical interval (Lewkowicz, 1996).

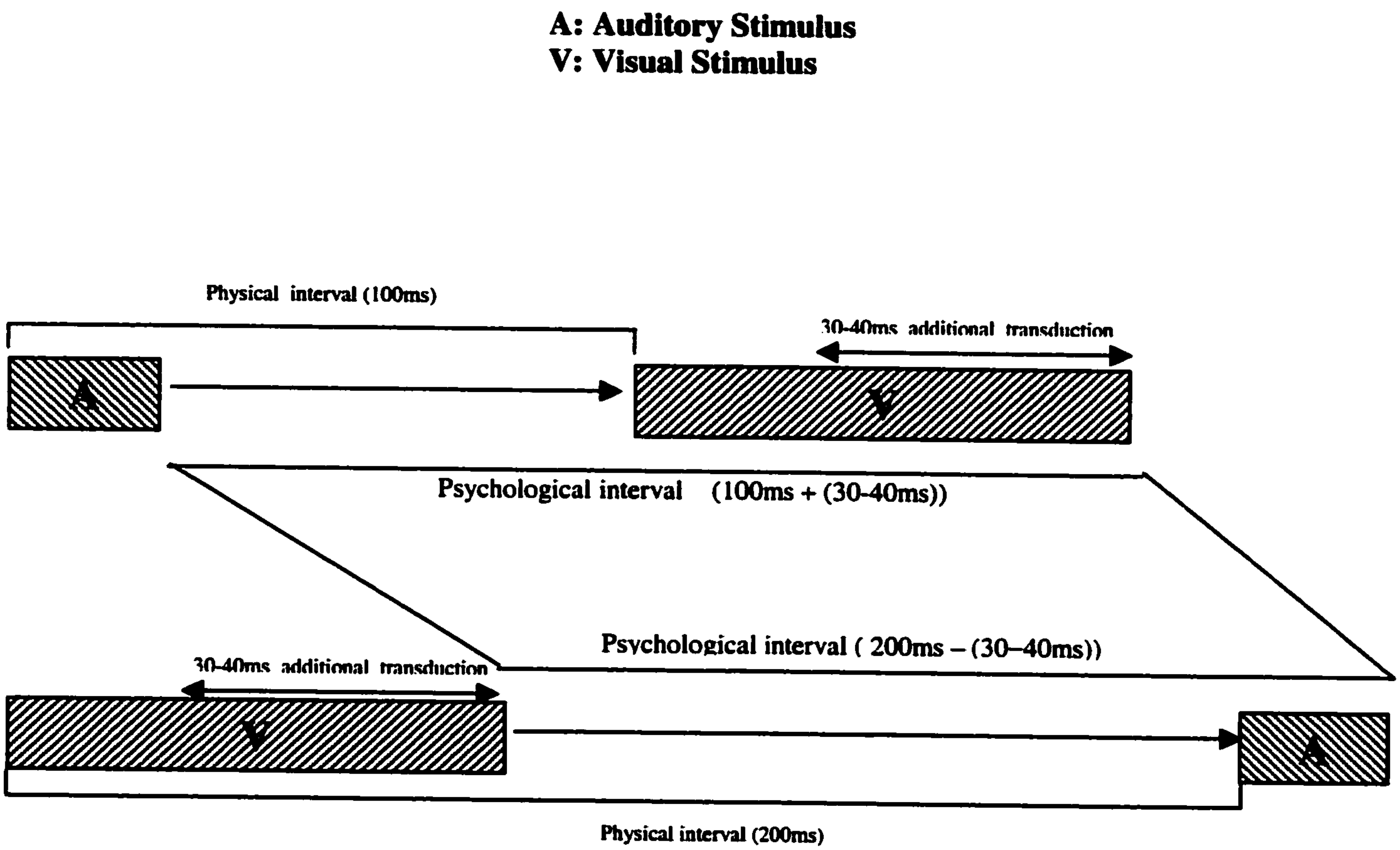
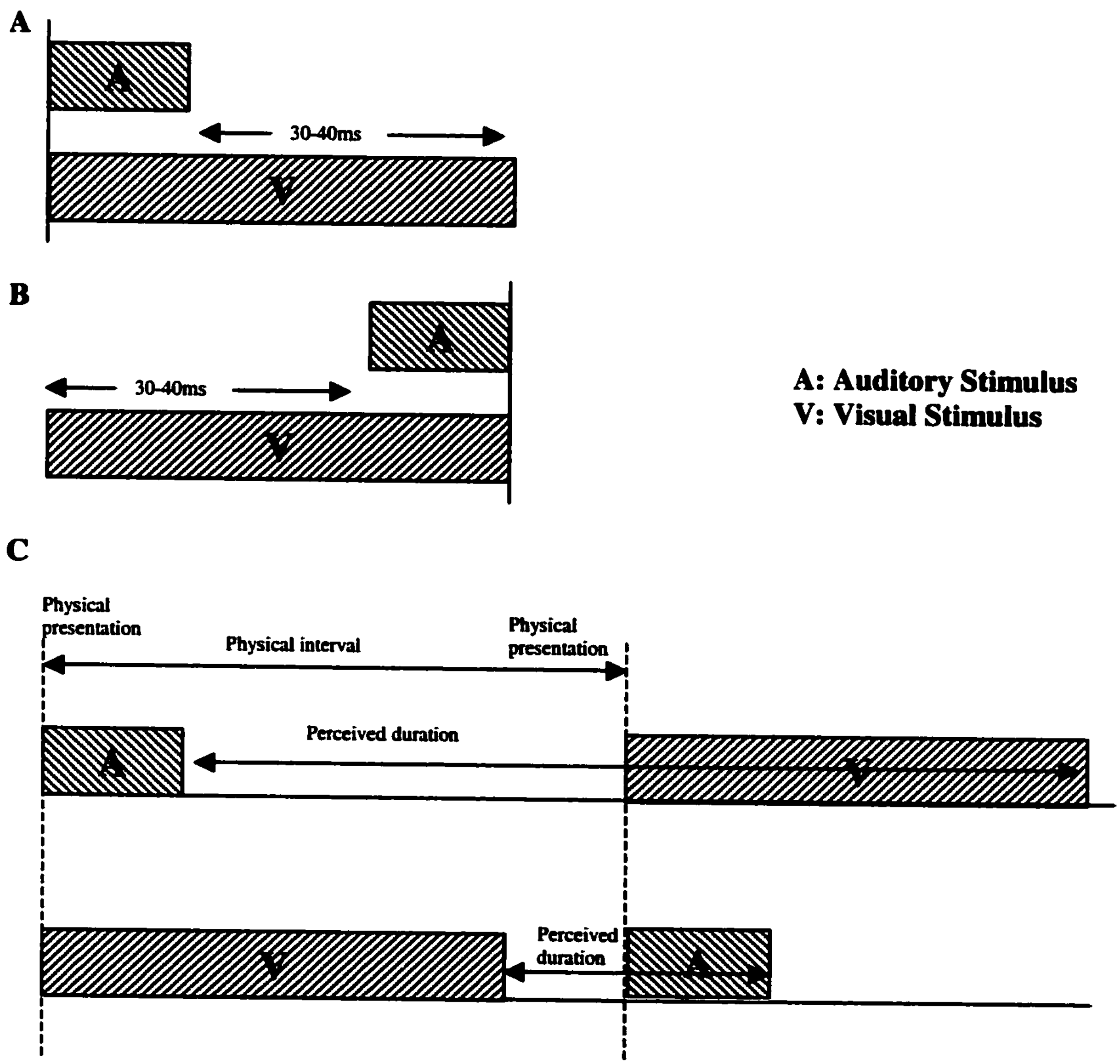


Figure 5. Diagrams of the auditory and visual stimuli pair based on Lewkowicz's postulate. The left end of box indicates the physical presentation of stimuli and the right end marks the mental representation. Length of the box represents time from physical presentation of stimulus to arrival at its corresponding central nervous system (CNS); A: Synchronous presentation of the stimuli. However, asynchrony is hypothesized to be perceived as auditory preceding (AV 30-40ms) due to longer transduction of visual signal; B: Asynchronous presentation of the stimuli designed to be mentally represented to be synchronous. Although auditory follows visual by 30-40ms, it reaches the CNS at the same time with visual stimulus; C: Two pairs of the stimuli with different order are presented with physically identical delay. However, due to the differential in transduction of the stimuli, the delay between stimuli in the upper pair (AV) is perceived longer than the that of the lower pair (VA).



In the lower panel, the psychological interval for VA asynchrony comprises the physical delay (200ms) minus 30–40ms visual tranduction time. Although the auditory followed the visual stimulus by 200ms, subjects falsely perceived the delay as being one of only 160–170ms. This reduced the psychological interval between the two stimuli and consequently made subjects more likely to perceive the two stimuli closer to each other.

The positive deviation of the subjective synchrony might result from the processes involved in the passage of auditory and visual stimuli from the external source into the brain. Auditory stimulus takes less time to arrive at the central nervous system (CNS) in the auditory cortex than the visual stimulus takes to arrive at the visual cortex. However, the two stimuli have to get to the same place to be compared for the judgment of synchrony. Then the difference in time that it takes for the two stimuli to arrive to some common point in the brain seems to be involved in the subjective synchrony. If the auditory stimulus is presented a little later in the external world but reaches the common processing destination in the brain at the same time as the visual stimulus, synchrony between the two stimuli is still perceived. Lewkowicz postulated that the visual stimulus needs extra neural transduction time and that the psychological interval explains why we perceive stimuli as synchronous when they are really asynchronous.

Synchrony effects affected by additional visual transduction are schematized in Figure 5. The left end of the box indicates physical presentation of the stimuli and the right end marks time that the stimuli are mentally represented. The length of the box represents time from physical presentation of stimulus to arrival at the CNS. In panel A, auditory and visual stimuli are presented simultaneously. Due to longer transduction of visual signal, asynchrony is hypothesized to be perceived as auditory preceding. In panel B, auditory stimulus occurs 30–40ms after visual stimulus, but the subject perceives them as synchronous. Panel C shows the results where two differently ordered pairs of stimuli were presented, each with a physically identical delay of 150ms. However, as a result of the differential in transduction of the stimuli, the subjects perceived the delay between stimuli in the AV condition (shown in the upper area of the panel) as being longer than that in the VA condition (shown in the lower area of the panel).

Hollier and Rimell (1998) undertook a study to measure sensitivity to synchrony. Subjects were asked whether they observed asynchrony (which the researchers referred to "synchronization error"). The researchers designated negative delay values for VA stimuli and positive values for AV stimuli, the opposite of the convention employed in other studies. For the systematic designative reference between stimulus conditions and value designations, the delays in this particular study were turned around to bring their sign values in line with those of other studies (and those of the present paper's, later on). The researchers used nine delays, ranging from −150ms to 300ms, and three types of test stimulus; a short nonspeech visual cue (a pen), a long nonspeech visual cue (an axe), and a speech stimulus (a talking head). They made the range of delays wider for the VA than for the AV condition. They gave each subject the whole range of stimuli in random sequences and asked them to respond with *yes* if they did perceive asynchrony and *no* if they did not perceive asynchrony.

The researchers' results supported the same kind of asymmetry of thresholds between the AV and VA conditions as that reported in most of the other studies. Although Hollier and Rimell neglected to average thresholds across all stimulus types, we can estimate these results from their graph. These values would be equal to the delay at which 50% of the subjects indicated an asynchrony. These estimated thresholds were −100ms for the AV condition and approximately 175ms for the VA condition. The subjective synchrony computed using these two thresholds is 37ms.

Bushara et al. (2001) also found a positive deviation of the subjective synchrony. Subjects were to detect intermodal temporal mismatch between simple stationary auditory and visual stimuli in the AV and VA conditions. The five asynchronous delays were −200, −150, −100, −75, and −50ms in the AV condition and 100, 150, 200, 250, and 300ms in the VA condition. Subjects were measured in their reaction time (RT) to the stimuli, which the researchers measured as the time from the onset of the second stimulus to the point of the subject's pressing a button. The researchers found thresholds similar to those obtained in other studies reviewed here. The estimate of the negative threshold in the AV condition was −120ms, and that of the positive threshold in the VA condition was 170ms. The subjective synchrony was 25ms, which constituted the positive deviation. As expected, the longer the delay of asynchrony, the more readily the subjects detected it. The percentage of correct responses had a positive correlation with increasing delays. As for RT, it decreased as a function of delay. Subjects took a shorter time to respond to longer asynchronous delays. Bushara et al. also reported that at the common delays of 100, 150 and 200ms, regardless of positive or negative signals in both conditions, the RT was quicker and the response was more accurate in the AV condition than in the VA condition. As well, Bushara et al. determined that the neural correlate specifically involved in the temporal synchrony- and asynchrony-detection processes was the right insula.

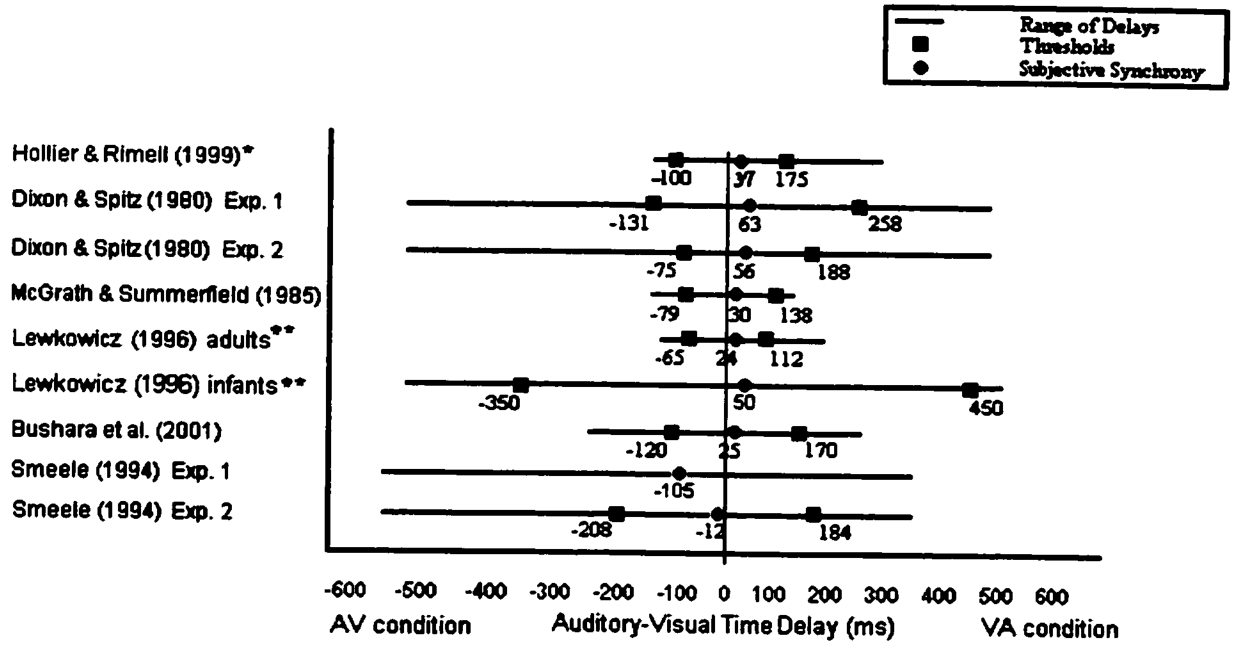
Table 1. Sample size, method, thresholds, subjective synchrony in studies (All values are in ms).

Study	Number of Subjects	Method	AV Threshold	VA Threshold	Subjective Synchrony
Hollier & Rimell (1998)	18	In-Synchrony or Out-of-Synchrony	-100*	175*	37*
Dixon & Spitz (1980) experiment 1	18	Method of Adjustment	-131	258	63
experiment 2	18	Method of Adjustment	-75	188	56
McGrath & Summerfield (1985)	16	Better Synchrony	-79	138	30
Smeele & Sittig (1994) experiment 1	6	Order	N/A	N/A	-105
experiment 2	6	In-Synchrony or Out-of-Synchrony	-108	184	-12
Bushara et al. (2001)	12	In-Synchrony or Out-of-Synchrony	-120	170	25
Lewkowicz (1996) Adults	10	In-Synchrony or Out-of-Synchrony	-65	112	24**
Lewkowicz (1996) Infants	96	In-Synchrony or Out-of-Synchrony	-350	450	50**

* Thresholds and subjective synchrony were averaged across all stimulus types and estimated according to the graph presented in the study.

** Subjective synchrony was calculated manually using the AV and VA thresholds.

Figure 6. Visual presentation of range of delay, AV threshold, VA threshold, and subjective synchrony in studies (Wesley, 2002 reprduced with permission).



* Thresholds and subjective synchrony were averaged across all stimulus types and estimated according to the graph presented in their study.

** Subjective Synchrony was calculated manually using the AV and VA thresholds.

Table 1 gives a summary of the AV and VA thresholds, the subjective synchronies, and synchronous intervals found in the studies so far described. Figure 6 provides a visual presentation of the values in Table 1. It is noteworthy that these studies gave similar subjective synchronies of asymmetrical thresholds. Most of the studies showed a positive deviation of 30–60ms from physical synchrony (0ms).

The research topic addressed in the present study concerns the consistent temporal asymmetry of the AV and VA thresholds, indicated by a positive deviation of subjective synchrony (30–40ms). To explain this effect, Lewkowicz (1996) considered that a visual stimulus requires an average of 30–40ms longer than an auditory one to reach the CNS.

As noted earlier, Lewkowicz proposed that the difference in neural transduction between the two stimuli is the reason for the positive deviation of subjective synchrony reported in much of the literature. When the difference is taken into account, the subjective synchrony should reflect the physical synchrony (Lewkowicz, 1996). The present study attempts to test the Lewkowicz Model by trying to determine whether the positive deviation in subjective synchrony takes place because auditory stimuli reach the comparator before visual stimuli do.

RT is used to estimate the difference in neural transduction speeds of auditory and visual stimuli. RT is the time it takes for stimuli to be sensed, brought into the nervous system, and processed and for motor commands to bring about muscle contractions. Therefore, RT of auditory and visual stimuli should reflect the neural transduction speeds required for auditory and visual stimuli in their respective central processing mechanisms. In other words, the RT difference between auditory and visual stimuli must reflect the difference in their transduction speeds. RT was thus used to measure transduction speeds for the Lewkowicz Model.

Following Lewkowicz Model, the present study hypothesized that the difference in RT for auditory and visual stimuli is positively correlated with subjective synchrony. Accordingly, two experiments were designed in which subjective synchrony was measured to observe whether the predicted 30–40ms of subjective synchrony would be replicated, and RT of subjects to auditory and visual stimuli were measured to examine the relationship with the subjective synchrony value obtained. If the Lewkowicz Model is correct, then the differences in RT to visual and auditory stimuli should match the subjective synchronies.

For the experiment in the present study, the method of Better Synchrony was used, which is regarded as a better psychophysical method than In-Synchrony or Out-of-Synchrony. In the method of In-Synchrony or Out-of-Synchrony, the psychophysical methodology is flawed because the subject is more likely to judge the stimuli as Out-of-Synchrony when heteromodal stimuli are presented synchronously or with zero delay. The subject tends to make an Out-of-Synchrony judgement even when unsure, which falsely leads to a correct response. That is, one is more open to bias in favour of the correct Out-of-Synchrony response. This tendency yields unreliable results. On the other hand, the method of Better Synchrony is not flawed in that respect. When the subject does not know which pair of the stimuli is in Better Synchrony, the response is likely to be the first pair presented for 50% of the trials, and the second pair for the other 50% of the trials. In other words, when the subject is unsure there is a 50% chance of either a correct and incorrect response.

METHODS

Subjects

Fifty-two volunteers (14 men, 38 women) recruited from University of Ottawa and the community participated as subjects. The median age of all subjects was 25 (range of 11–62). Most subjects had normal or corrected-to-normal vision and unimpaired hearing. Two subjects reported minor hearing problems, and one subject indicated a vision problem. None of these subjects' perceptual problems were considered great enough to object to their participation in the experiment. Three female and one male subject were on medication which interacted with the nervous system. As underlying effects of the drugs were unknown, the data of the four subjects were excluded. A total of 48 subjects were used for the data analysis. Of these, 11 subjects volunteered to be tested repeatedly for five sessions within 10 days. This secondary experiment was to examine whether there were any changes in the thresholds and the subjective synchrony over time.

Apparatus and Stimuli

The apparatus from which auditory and visual stimuli originated was a metal case that measured 10cm × 14.5cm × 8cm. The apparatus was situated at about 40cm in front of the subject and was connected to a computer. The auditory stimulus was a 500Hz square wave with intensity of 65dB and was delivered from Optimus NOA 408 headphones connected to the apparatus. Instead of being worn by the subjects, the headphones were placed on each side of the apparatus, about 40cm in front of the subject. The visual stimulus consisted of a 5mm circular red light-emitting diode (LED) and was fixed at the centre surface of the apparatus. The LED was activated with 17mA current. Both auditory and visual stimuli were 10ms in duration. Below the LED were three touch buttons for the responses.

Method for the Better Synchrony Experiment

The Better Synchrony method presents two pairs of auditory and visual stimuli. One of these is always in synchrony, while the other occurs with various delays between the light and the sound. The subjects' task is to judge which pair of the two pairs is more in synchrony. Since one pair occurs always in synchrony, the task is really to judge which of the two pairs of stimuli is in synchrony against the other pair out of synchrony. When the pair that is out of synchrony has a large delay, the judgment is easy, whereas a small delay may interfere with correct judgment. Obtaining thresholds involves examining the same U-shape distribution as in the In-Synchrony or Out-of-Synchrony method. In the Better Synchrony method, 75% of Better Synchrony judgments is used to locate the corresponding thresholds. When both pairs are in synchrony, the Better Synchrony judgment is likely to be correct 50% of the time. So, the lowest percentage of Better Synchrony judgments is 50% of trials on the ordinate, which should be used in relation to highest judgment to measure thresholds. Therefore, for the thresholds, the researcher selects the middle point between judgments in half of the trials (50%) and all of the trials (100%), and this is

75% of trials. Thresholds are obtained by finding the delay values that cross the graph at 75% of judgments, one for the AV condition and the other for the VA condition. In Figure 2, the AV threshold is −58ms, and the VA threshold is 89ms.

Procedure

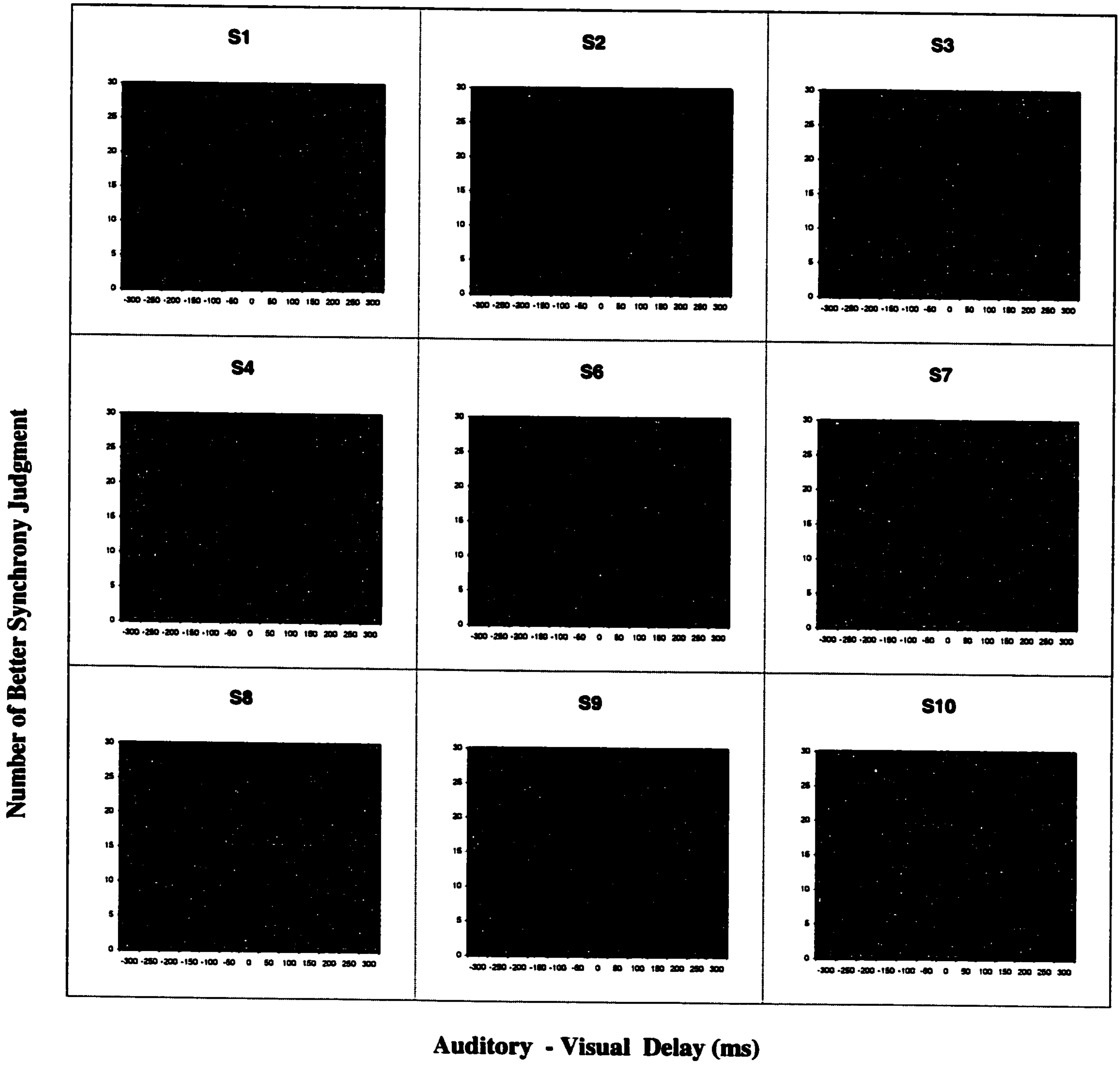
The subjects were seated in a dimly illuminated, quiet room during the testing session, which lasted approximately 50min. Each subject was tested individually, and the experimenter was in the next room. The subjects were informed that the session comprised two experiments: Better Synchrony and RT. The order of the two experiments was counterbalanced among subjects. For the Better Synchrony experiment, the subjects were told that two pairs of auditory and visual stimuli would be presented, and one pair was always in synchrony and the other was out of synchrony, with varying delays between the two stimuli. The subjects' task was to judge which of the two pairs of auditory and visual stimuli was better in temporal synchrony. The subjects indicated this by pressing the appropriate button on the test apparatus (left button for the first pair of stimuli, and right for the second). For the asynchronous pair there were 13 different delays, ranging from −300ms sound leading to 300ms sound following, at 50ms intervals. The subjects were presented with 25 repetitions of trials. From trial to trial, the 13 trials of the asynchronous pair were intermixed, so as not to enable anticipation of what delay of stimuli would occur next. As well, the two pairs, in-synchrony and out-of-synchrony occurred in random order in each trial. At zero delay, the pair of stimuli was in synchrony, just as the other pair that was always in synchrony. The first and the second pair were presented one after the other and separated by 1 to 1.5seconds set at random interval. The next trial started 1 to 2 seconds following the subject's response to the prior trial. A practice set of trials was administered before the testing session began. The subjects were permitted to repose, between trials, by withholding their press on the button as long as they wanted.

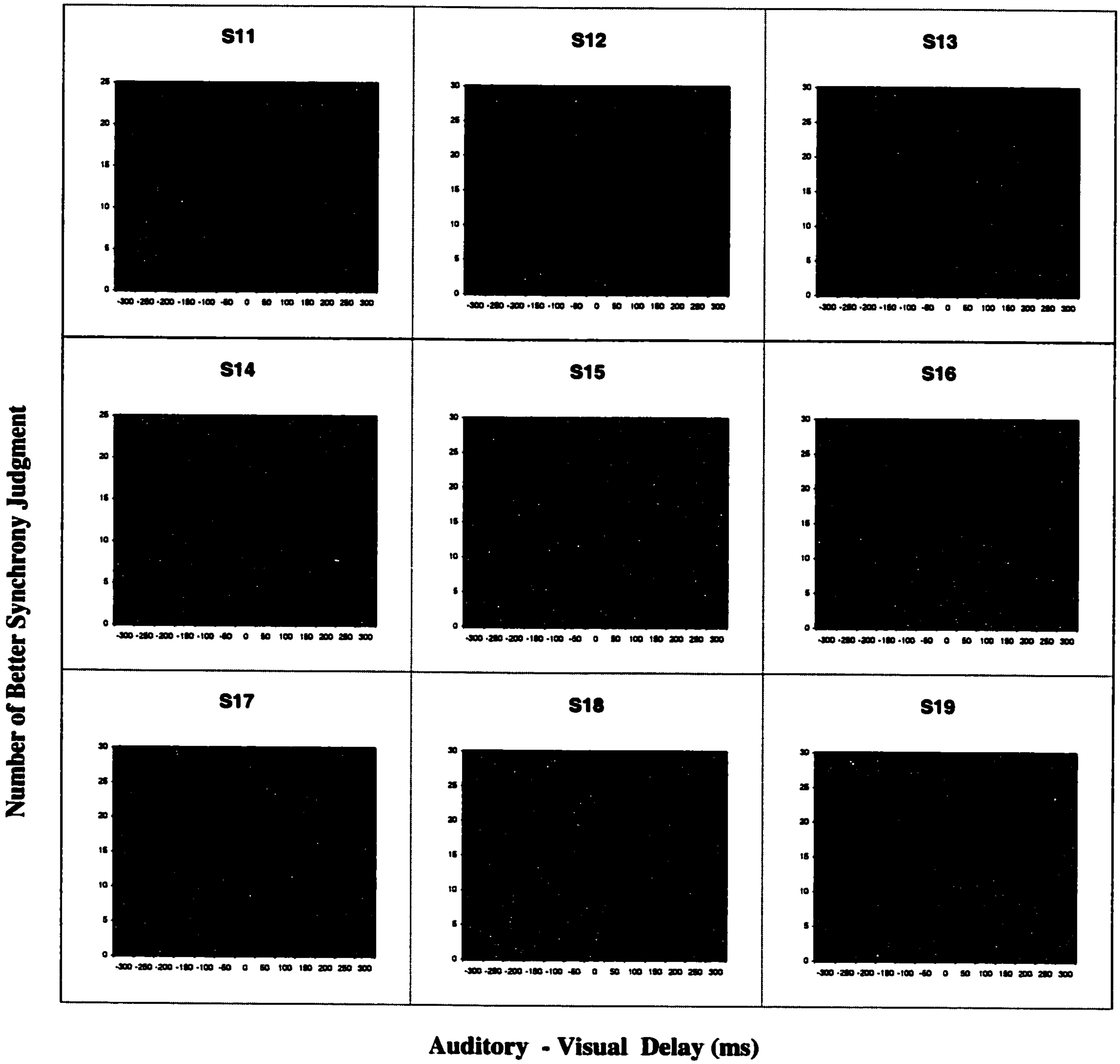
For the RT experiment, the subjects were told they would be measured on the time it took for them to perceive and respond to stimuli. The stimuli were either auditory alone, visual alone, or both auditory and visual stimuli. The same apparatus in Better Synchrony was used. The subjects were asked first to gently maintain contact with a button with the index or most comfortable finger and then to respond to auditory, or visual stimuli, or both, as soon as possible by immediately releasing the button. The order of the three possible stimuli was randomized. Each of the three stimuli was presented 25 times, and their order was randomized.

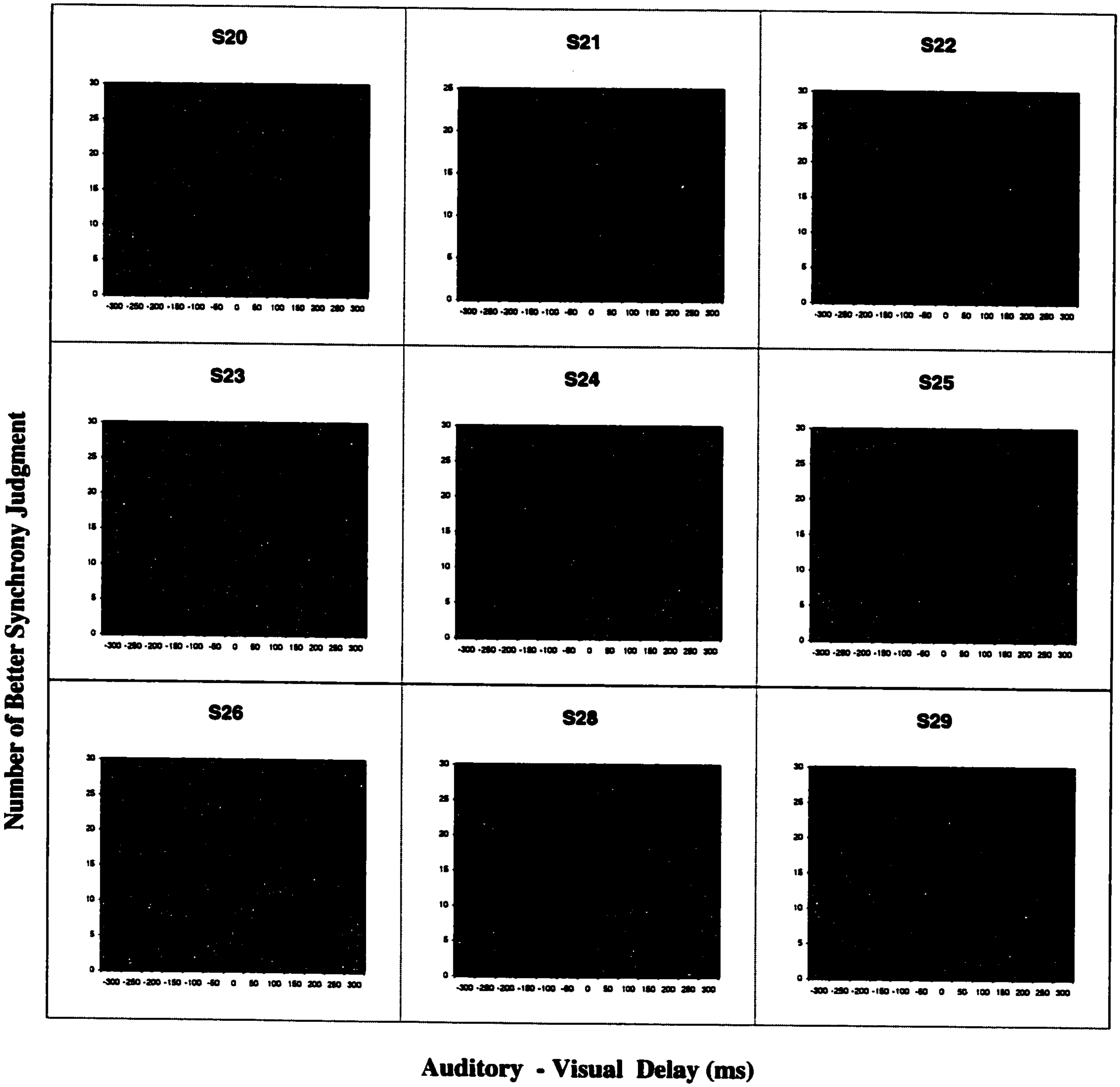
RESULTS

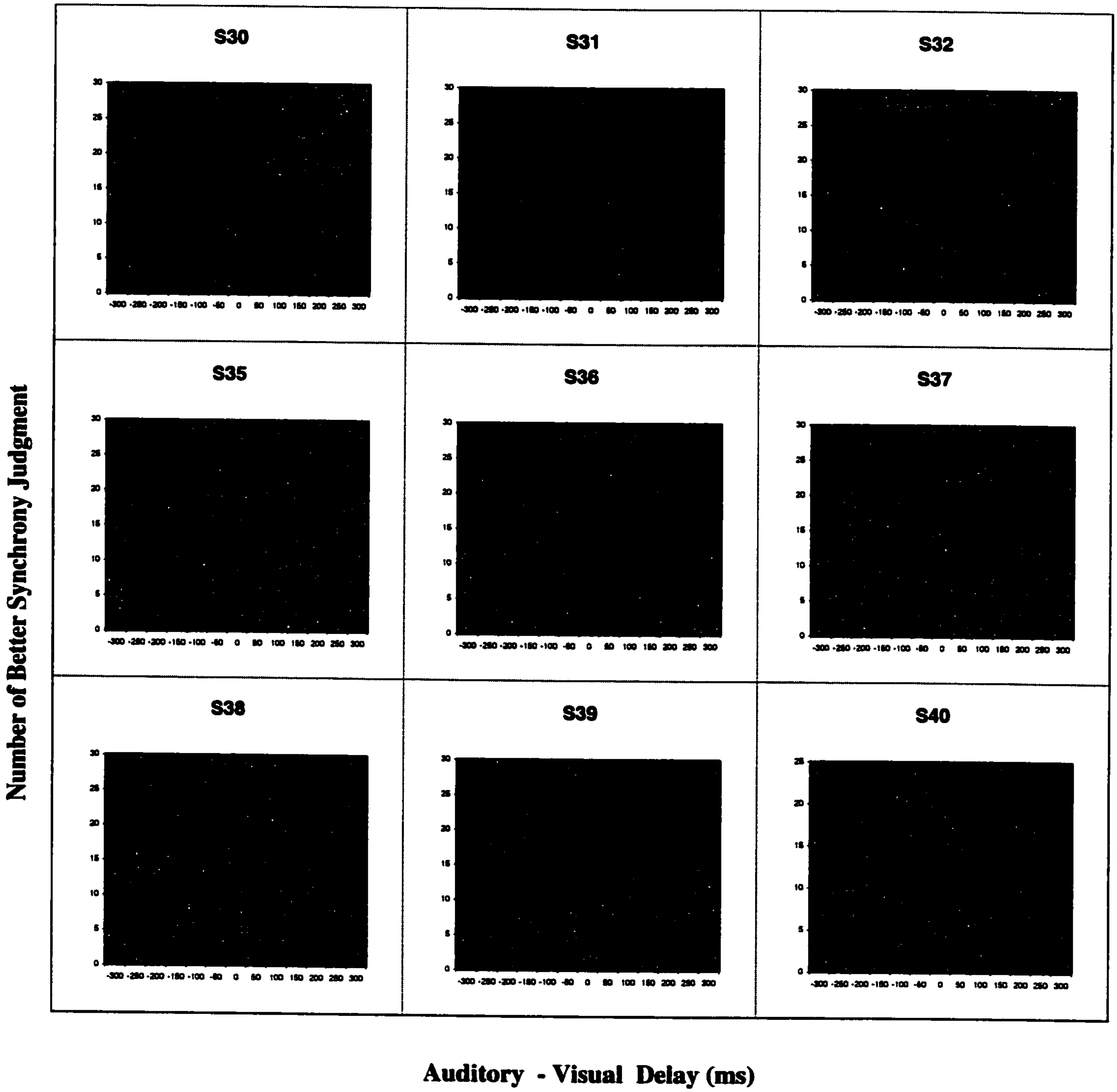
Figure 7 illustrates the results of the all subjects in graphs. In judging which stimulus pair was in Better Synchrony, most subjects' responses gave rise to a general U-shaped graph. For example, subject 2 made 25 correct judgments when delays were from –300 to –150ms, and the rate of correct judgments lowered toward 0ms. The subject made 12 correct Better Synchrony judgments when the two pairs of stimuli were both in synchrony. As the delay increased, again with auditory following, the rate of correct judgments increased as well and from 150ms on the subject made 25 correct judgments. The thresholds in the AV and VA conditions were obtained by finding the delays that crossed the graph at 75% of trials, that is, 18.75 out of 25 trials. Table 2 summarizes all subjects' thresholds, synchrony intervals, and subjective synchronies. For subject 2, the AV threshold was –70ms, and the VA threshold was 63ms. This indicates that at –70ms of delay, or when the auditory stimulus occurred 70ms before the visual stimulus, subject 2 judged that the pair of stimuli was as in synchrony. When the auditory stimulus was presented less than 70ms before the visual stimulus, the subject perceived the asynchronous pair as synchronous and so made an incorrect judgment. The subject needed at least 70ms of time gap between auditory leading and visual following to perceive the stimuli as separate. As well, when the auditory followed the visual stimulus by 63ms or more, the subject made a Better Synchrony judgment in 75% of trials and an incorrect judgment in 25% of trials. When the auditory stimulus followed less than 63ms after the visual stimulus, the asynchronous stimuli could not be

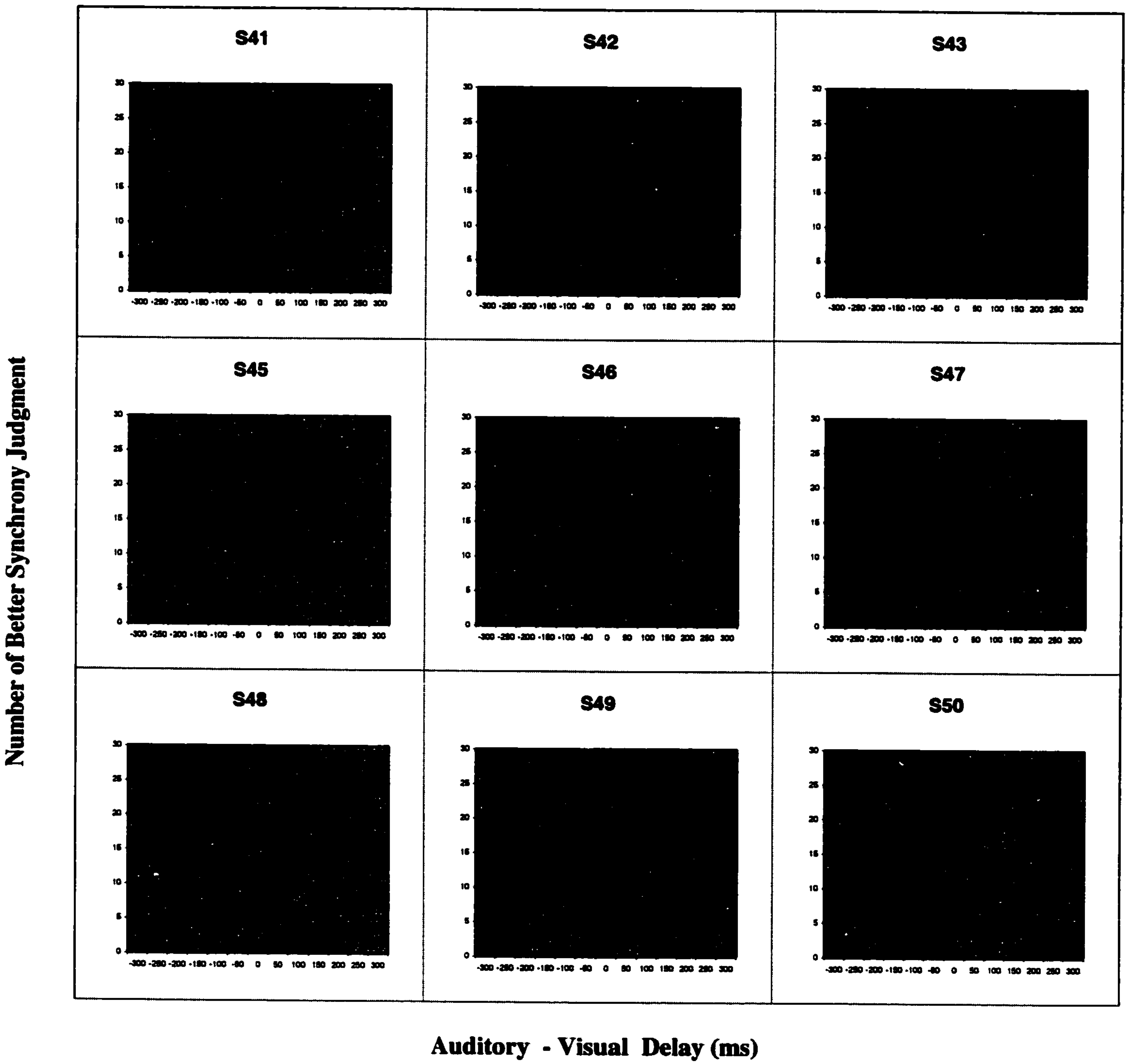
Figure 7. Graphs of results of Better Synchrony experiment

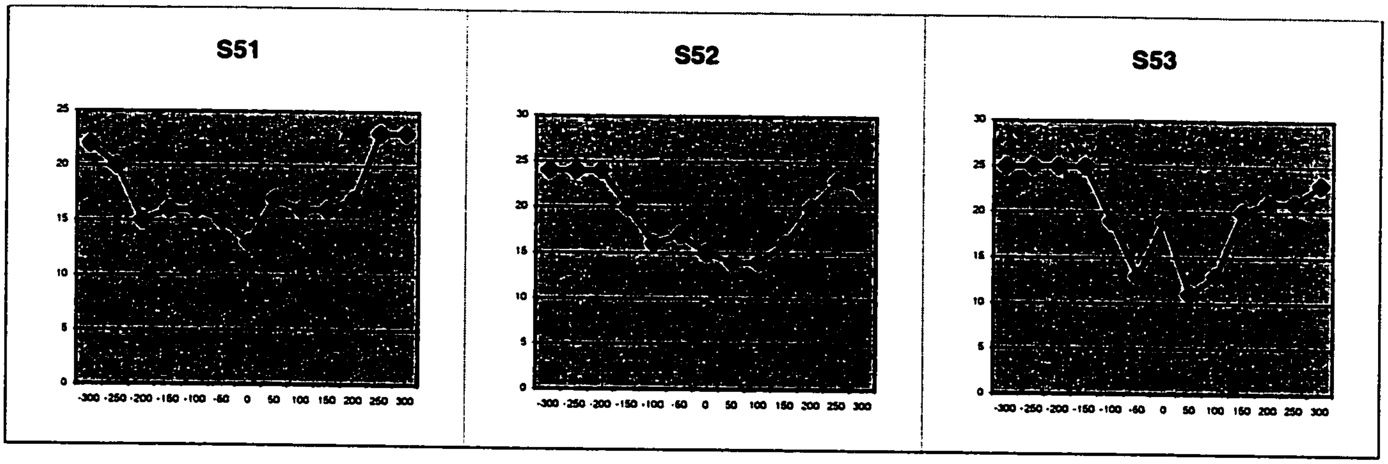



S11
S12
S13
S14
S15
S16
S17
S18
S19
Number of Better Synchrony Judgment
Auditory - Visual Delay (ms)


S20
S21
S22
S23
S24
S25
S26
S28
S29
Number of Better Synchrony Judgment
Auditory - Visual Delay (ms)


S30
S31
S32
S35
S36
S37
S38
S39
S40
-300 -250 -200 -150 -100 -50 0 50 100 150 200 250 300
Number of Better Synchrony Judgment
Auditory - Visual Delay (ms)


S41
S42
S43
S45
S46
S47
S48
S49
S50
0 5 10 15 20 25 30
-300 -250 -200 -150 -100 -50 0 50 100 150 200 250 300
Number of Better Synchrony Judgment
Auditory - Visual Delay (ms)



*N= 48 as subjects 5,27, 33, and 34 were excluded because they were on medication and subject number 44 was not assigned to anyone.

Auditory - Visual Delay (ms)

Table 2. Results of thresholds, subjective intervals, subjective synchrony, and RT.

	Thresholds				RT Medians			
Subject	AV Condition	VA Condition	Subjective interval	Subjective synchrony	Visual	Auditory	Both	Visual-Auditory
1	-48	48	96	0	221	185	173	36
2	-70	63	133	-4	176	181	145	-5
3	-94	73	167	-11	243	181	166	62
4	-127	148	275	11	264	194	176	70
6	-79	74	153	-3	188	177	147	11
7	-115	111	226	-2	160	116	122	44
8	-86	89	175	2	210	154	148	56
9	-79	34	113	-23	190	142	144	48
10	-84	95	179	6	198	161	137	37
11	-91	138	229	-24	265	179	171	86
12	-22	44	66	11	197	159	147	38
13	-128	223	351	48	233	242	192	-9
14	-198	246	444	24	265	179	171	86
15	-44	146	190	51	202	173	157	29
16	-98	77	175	-11	226	186	186	40
17	-88	55	143	-17	225	163	149	62
18	-156	190	346	17	214	200	176	14
19	-70	106	176	18	259	276	199	-17
20	-129	169	298	20	217	199	195	18
21	-108				203	167	153	36
22	-98	78	176	-10	207	179	166	28

	Thresholds				RT Medians			
Subject	AV Condition	VA Condition	Subjective interval	Subjective synchrony	Visual	Auditory	Both	Visual-Auditory
23	-108	80	188	-14	199	174	160	25
24	-88	79	167	-5	244	197	186	47
25	-73	77	150	2	229	214	209	15
26	-58	67	125	5	224	201	148	23
28	-56	48	104	-4	238	219	201	19
29	-81	115	196	17	219	212	157	7
30	-84	58	142	-13	247	198	179	49
31	-85	73	158	-6	268	228	205	40
32	-129				240	231	208	9
35	-119	168	287	25	217	169	148	48
36	-113	81	194	-16	222	173	158	49
37	-88	213	301	63	336	350	282	-14
38	-159	99	258	-30	228	188	192	40
39	-56	77	133	11	198	147	144	51
40	-214	284	498	35	275	247	207	28
41	-98	176	274	39	229	186	169	43
42	-126	208	334	41	244	246	194	-2
43	-170	137	307	-17	272	264	233	8
45	-91	48	139	-22	257	207	208	50
46	-79	78	157	-1	224	206	161	18
47	-117	86	203	-16	210	183	169	27
48	-86	109	195	12	171	140	138	31
49	-96	96	192	0	224	184	165	40

	Thresholds				RT Medians			
Subject	AV Condition	VA Condition	Subjective interval	Subjective synchrony	Visual	Auditory	Both	Visual-Auditory
50	-156	282	438	63	211	169	162	42
51	-238	215	453	-12	212	187	168	25
52	-134	184	318	25	253	186	189	67
53	-98	141	239	22	215	161	153	54

* N = 48 as subjects 5, 27, 33, and 34 were excluded because they were on medication and subject number 44 was not assigned to anyone.

distinguished from the synchronous. Accordingly, the synchrony interval, the length of time between the two thresholds, in which the subject cannot tell which pair is in better synchrony, was 133ms. The subject showed –4ms of subjective synchrony.

As shown in Table 3, the mean AV threshold is –104ms (SD = 42.11) with a range of –238 to –22ms. Subjects required, on average, a delay of –104ms or more to be rather certain of their judgements of asynchrony when the auditory stimulus occurred first. If, however, the auditory preceded the visual stimulus by less than 104ms, then the subjects became uncertain and their judgments became ambiguous. A histogram of the AV thresholds appears in Figure 8. It looks somewhat negatively skewed (sk = – .79).

Table 3 indicates that the mean of the VA thresholds is 119.3ms (SD = 64.6), with a range of 34 to 284ms. On average, when the visual preceded the auditory stimulus, the subjects needed at least 119.3ms of delay to perceive and judge asynchrony between the two stimuli. If the auditory followed the visual stimulus within 119.3ms, the subjects were more likely to judge the two stimuli as synchronous. Figure 9 shows a histogram of the VA thresholds. It is positively skewed (sk = .97). When the visual preceded the auditory stimulus, a few subjects needed delays of as much as 289ms.

The AV threshold and the VA threshold were compared in their absolute value. For this the AV threshold data were treated with positive sign as the VA threshold data. The paired t-test showed that VA thresholds are bigger than AV thresholds (t =2.324, df =45, p= .025). Because two subjects did not show VA thresholds, the degrees of freedom was based on sample size of forty six.

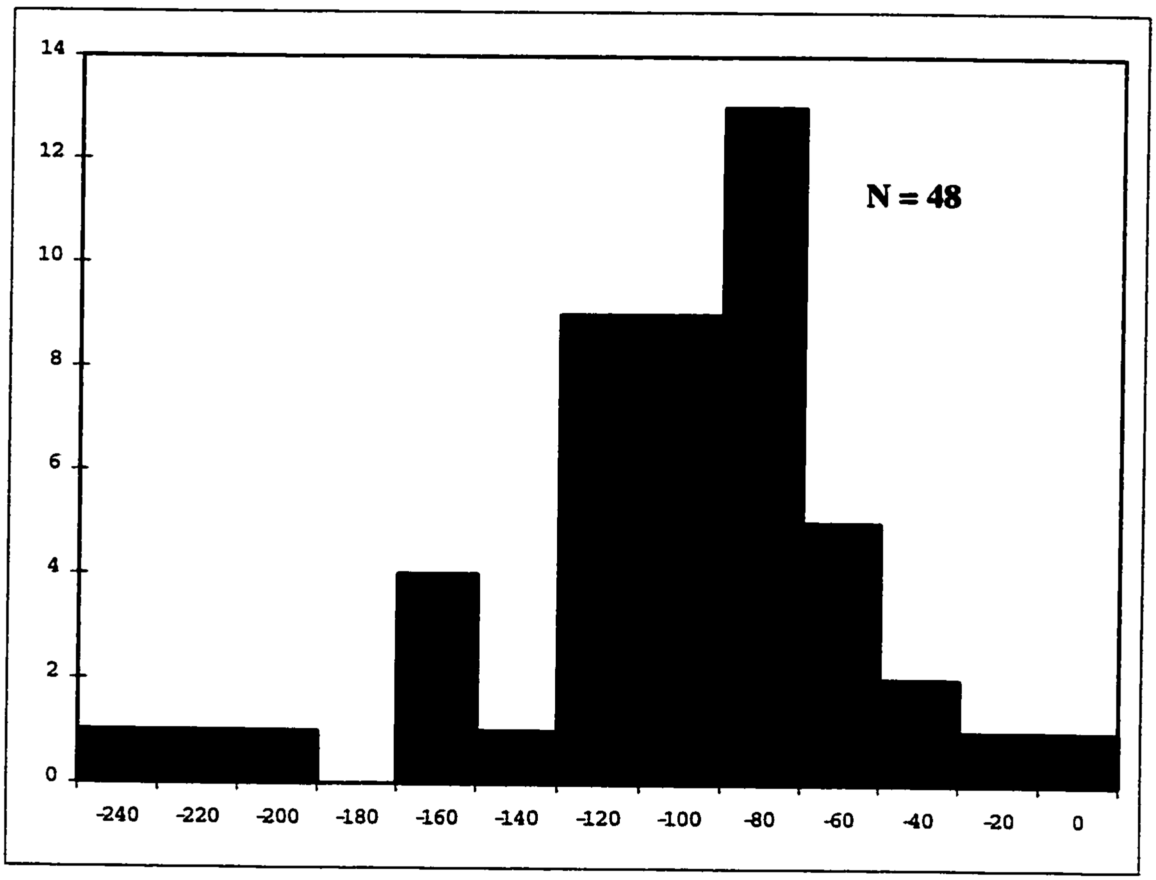
Table 3. Summary of the distribution of the variables in Better Synchrony and RT experiments. All values are in ms.

Variable	Mean (SD; Range)
AV threshold	-104 (42.11; -238 to -22)
VA threshold	119.26 (64.6; 34 to 284)
Synchrony interval	222.85 (100.1; 66 to 498)
Subjective synchrony	7.72 (22.3; -30 to 63)
RT to visual stimulus	233* (31.42; 160 to 336)
RT to auditory stimulus	202* (39.3; 116 to 350)
RT difference	31 (23.65; -17 to 86)

* the mean of the medians

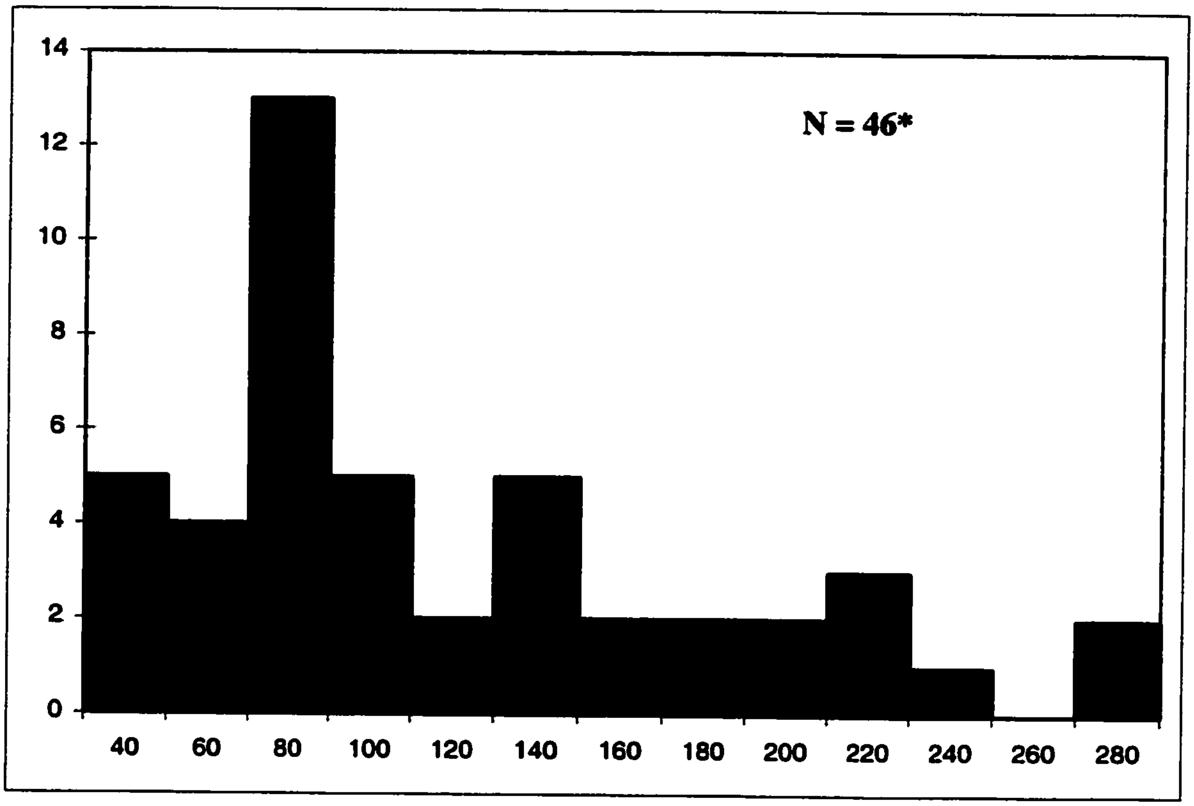
Figure 8. Histogram of AV thresholds

Frequency (number)



AV threshold(ms)

Figure 9. Histogram of VA thresholds

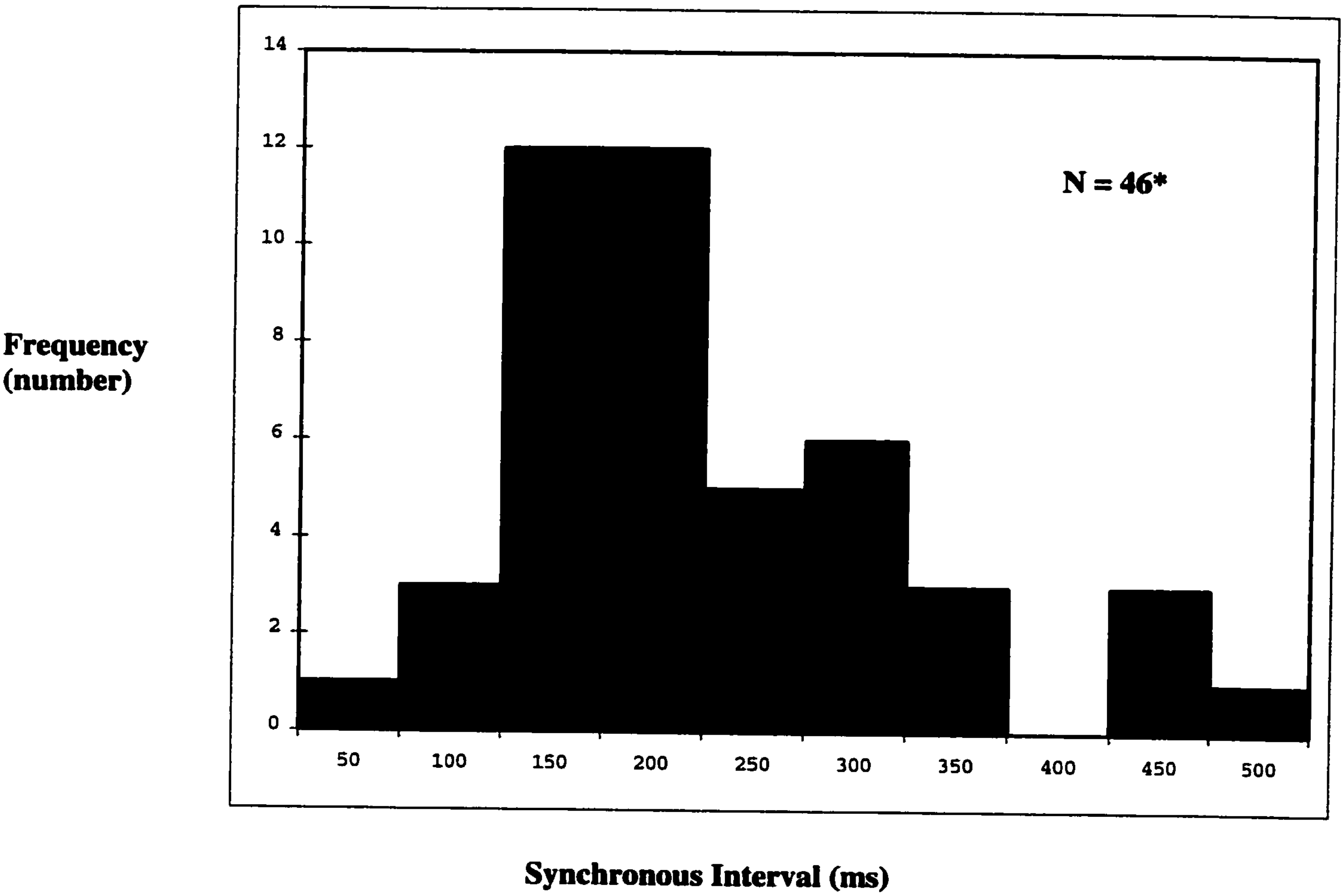


Frequency (number)

VA threshold(ms)

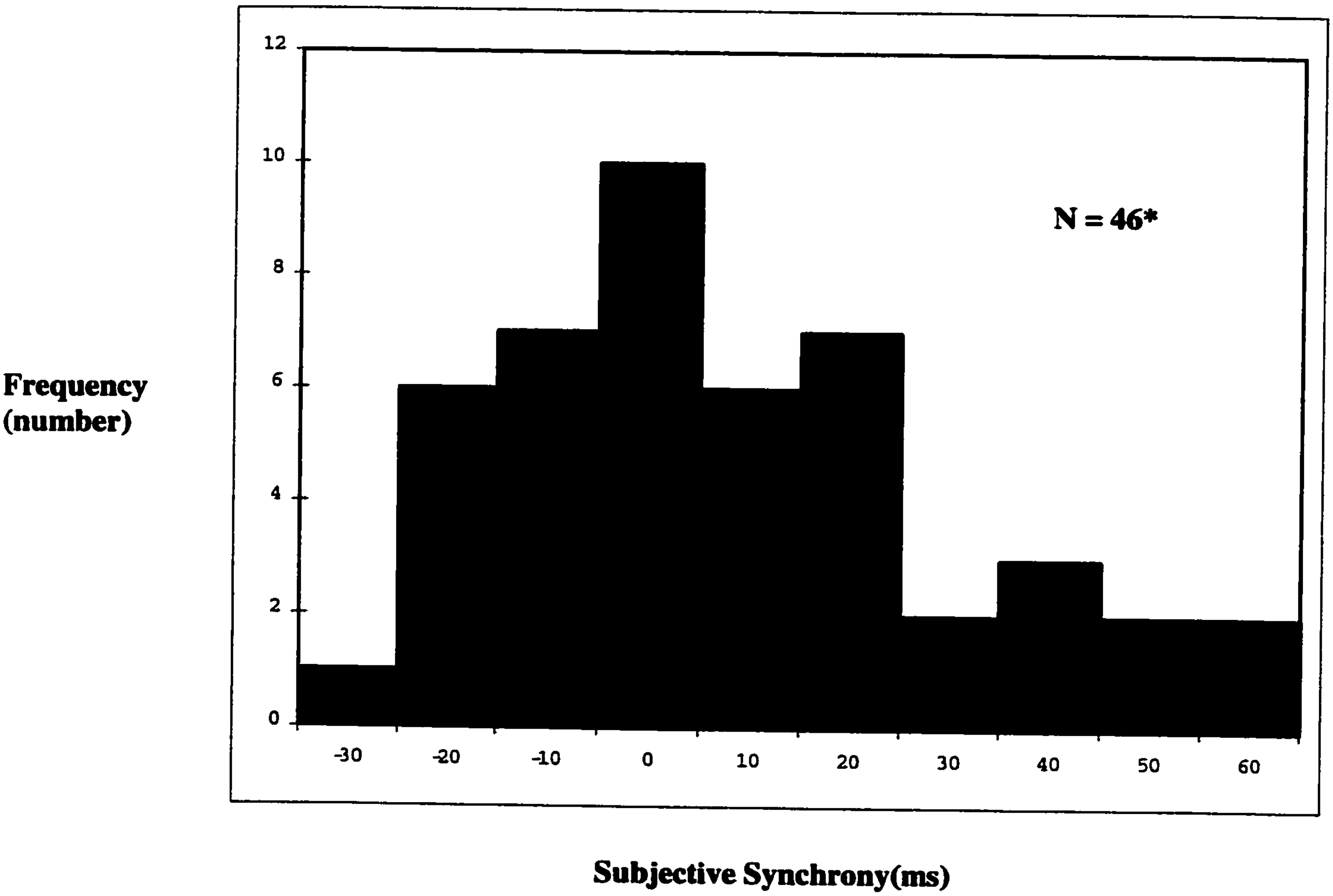
* Two subjects' (21 and 32) VA thresholds were not obtained as they judged the auditory following stimuli as in synchrony even at greatest delay.

Figure 10. Histogram of Synchrony Intervals



* Two subjects' (21 and 32) synchrony intervals were not obtained because their VA thresholds were not obtained.

Figure 11. Histogram of Subjective Synchronies



* Two subjects' (21 and 32) subjective synchronies were not obtained because their VA thresholds were not obtained.

The synchrony interval was calculated by subtracting the AV from the VA threshold. The synchrony interval is the range of delays in which subjects are uncertain of their Better Synchrony judgements. Table 3 indicates that the mean synchrony interval is 222.8ms (SD = 100.1), with a range of 66 to 498ms. Subjects' judgement uncertainty lay between 66 and 498ms of delay in both the AV and VA conditions. Figure 10 shows that synchrony intervals are positively skewed (sk =1.05).

The subjective synchrony — the centre of the synchrony interval — was determined by computing the average of the AV and VA thresholds. The mean value is 7.7ms (SD = 22.73), while the median value is 2ms (Figure 11). Many subjective synchronies are close to 0ms. Although there are more positive than negative subjective synchronies, the mean subjective synchrony was relatively smaller than the ones found in other studies.

As can be seen Figure 7, there were individual differences among subjects in this experiment. While the graph associated with subject 2's data resembled the theoretical graph in Figure 2, some subjects generated rather unique patterns of response. A number of subjects, for instance, 4 and 10 showed positive deviation of subjective synchrony, whereas subjects 11, 16, and a few others showed negative subjective synchrony. Many subjects including subject 2, had subjective synchrony at 0ms. The results for subjective synchrony do not replicate the findings of those previous studies that consistently reported positive deviation of subjective synchrony.

The unusual patterns of response for two of the subjects (subjects 21 and 32) are especially noteworthy. Unlike the others, their graphs do not fit the general U-shape. As the delay became larger in the AV condition, their rate of correct judgment on synchrony increased as well. On the other hand, in the VA condition their rate of judgment did not increase as much when the delay increased. Even at 300ms, the largest delay, the two subjects judged correctly only around 15 out of 25 trials. When the auditory followed the visual stimulus by 200ms, the two subjects perceived the stimulus more often as synchronous. Subject 21's rate of correct judgment increased somewhat after 150ms, but it was much lower than that of other subjects. Subject 32 did not detect VA asynchrony through to the 300ms delay. So, VA thresholds were not obtained for these two subjects, whose graphs indicated no perception of asynchrony, even at the largest VA delay. Only the AV threshold was measured for them.

The results for RT to auditory and visual stimuli are presented in Table 2. Visual stimulus, auditory stimulus, and visual and auditory stimuli together were presented 25 times for each variable, and the median values were obtained for each subject. Then the mean out of the median values for each variable was determined, as shown in Table 3. Most subjects responded faster to auditory stimuli (mean = 202; range: 116 to 350) than to visual stimuli (mean = 233; range; 160 to 336). The reaction was quickest when the two stimuli were concurrently provided. The important calculation for these data is the difference between the RT to the visual and to the auditory stimuli, as this would be expected to match the subjective synchrony if the Lewkowicz Model is correct. For all subjects the median RT to auditory stimulus was subtracted from that to the visual stimulus, as shown in Table 2. The RT differences are positive for the majority of subjects, but there are negative values for a few of the subjects, indicating that these subjects' RT to the visual stimuli was shorter than their RT to the auditory stimuli. The RT differences ranged from −17ms to 86ms. The overall RT differences were computed for correlation with subjective synchrony ($r = -.19$, $p > .2$). The result indicates no relation of RT difference to subjective synchrony. The scatterplot of the correlation confirms a lack of association between subjective synchrony and RT difference (Figure 12). Most of the datapoints fall below those predicted using the Lewkowicz Model. These experiments failed to confirm a positive relationship between RT difference and subjective synchrony. Correlation was computed to see if age factor played a role in subjective synchrony, and was found to be moderate ($r = .64$, $p > .2$). The younger the subjects were, their subjective synchronies were smaller, and the older the subjects were, their subjective synchronies were larger. However, as the majority of the subjects were in their mid-twenties and only a few were younger or older than them, this result needs replication with a sample size with larger age difference.

A secondary experiment in the overall Better Synchrony experiment involved 11subjects in 5 experimental sessions. These sessions examined whether subjective synchrony changes over time. The results are presented in Table 4. The mean AV threshold changes from −85ms to −82ms. The change of the mean VA threshold is bigger, from 98ms to 81ms. Subjective synchrony values also change from 7ms to 1ms on average. Standard deviation for subjective synchrony across five sessions changed from 24.9 to 22.1 to 13.3 to 20.8 to 20.8 to and to 7.2. The decreasing pattern of standard deviation resulted because big positive subjective synchronies got smaller in value and big negative subjective synchronies got smaller as well. Repeated measures test indicated that the changes in all the three variables; AV threshold, VA threshold, subjective synchrony were not significant. RT to the stimuli were also compared in repeated measure and found insignificant.

Figure 12. Relationship betwen Auditory-Visual Reaction Time Difference and Subjective Synchrony with Lewkowicz Model superimposed

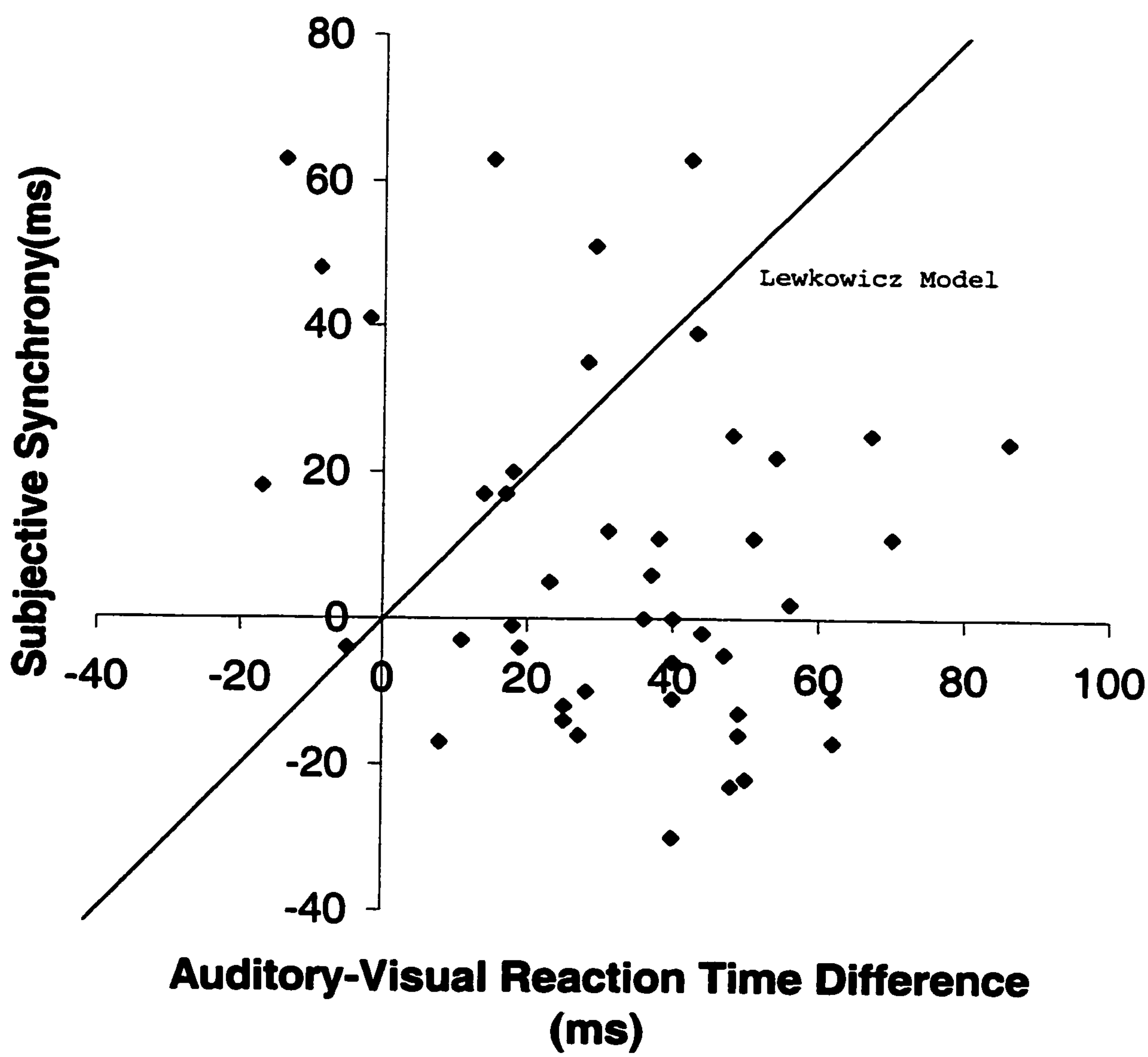


Table 4. AV thresholds, VA thresholds, subjective synchronies, synchrony intervals, and RT of eleven subjects who participated in five sessions of the experiment. Means are presented (All values are in ms).

	AV threshold					VA threshold					Subje-tive Syn-hrony					RT to Auditory Stimulus					RT to Visual Stimulus				
Subject	Session					Session					Session					Session					Session				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	-48	-56	-41	-71	-65	48	74	69	61	63	0	9	14	-5	-1	185	153	125	133	132	221	199	183	177	180
2	-70	-42	-71	-41	-73	63	88	38	36	70	-4	23	-17	-3	-2	181	164	163	190	179	176	195	200	188	192
3	-94	-48	-106	-74	-76	73	79	106	82	69	-11	16	0	4	-4	181	174	181	161	170	243	215	201	221	225
6	-79	-59	-81	-63	-72	74	77	72	84	81	-3	-5	-4	-1	-2	177	210	162	157	146	188	216	168	181	183
7	-115	-86	-98	-86	-59	111	76	117	113	70	-2	-5	10	13	5	116	115	119	110	115	160	154	163	149	150
9	-79	-59	-81	-63	-72	34	38	65	72	61	-23	-11	-8	5	-6	142	144	134	130	123	190	195	197	195	174
15	44	43	48	43	65	146	134	72	65	68	51	46	12	11	2	173	166	161	155	159	202	183	189	187	182
16	-98	-71	-88	-97	-67	77	48	69	42	34	-11	-12	-10	-28	-17	146	157	148	158	-11	226	182	188	178	177
31	-85	-89	-86	-81	-69	73	71	58	28	71	-6	-9	-14	-27	1	228	180	174	157	143	268	232	208	194	207
41	-98	-84	-89	-131	-84	176	138	70	82	106	39	27	-10	-25	11	186	178	195	187	178	229	222	215	217	214
42	-126	-122	-167	-134	-188	208	215	215	219	198	41	47	24	43	5	246	229	247	205	239	244	218	252	229	244
Mean	-85	-71	-87	-82	-82	98	94	86	80	81	7	11	0	1	1	182	169	165	158	158	213	203	197	192	193

DISCUSSION

Beginning with Exner (1875, cited in Hirsh & Sherrick, 1961), the literature on temporal perception has fairly consistently shown that thresholds differ, depending on the order of stimuli, such that the threshold is smaller when auditory stimulus precedes visual stimulus and the threshold is larger when auditory stimulus follows visual stimulus. Theoretically, when temporal-perception data are established, using either the In-Synchrony or Out-of-Synchrony or Better Synchrony methods, the subjective synchrony should match the physical synchrony of the stimuli, 0ms. Subjects should detect synchrony between two stimuli when in fact they occur simultaneously (Figures 5, Panel A). As the results show, in the present study and others, however, it has been found that the subjective synchrony is pushed past 0ms toward the VA condition.

The synchrony interval represents delays where the subjects cannot tell which stimuli pair occurred in better synchrony. All stimuli coming with delays within the synchrony interval seem synchronous to the subjects. The subjects in the present study all had 0ms in their synchrony intervals. However, the two delays that set the synchrony interval — that is, the two thresholds — differed in absolute value. When the auditory stimulus occurred first the subjects detected the asynchronies at smaller AV delays.

The aim of the present study was to investigate what systematically causes the temporal asymmetry of the thresholds between the AV and VA conditions, which leads to the positive deviation of subjective synchrony revealed in most studies. Lewkowicz (1996) attributed this positive deviation to the difference in arrival times of auditory and visual stimuli, when traveling to the brain. The Lewkowicz Model was tested in the present experiment. RT difference was calculated and correlated with subjective synchrony. The mean subjective synchrony was 7.7ms, and its median was 2ms. The difference in RT between auditory and visual stimuli was 31ms. The correlation between subjective synchrony and RT difference was – .29.

The expected tendency of close to 30–40ms of subjective synchrony was not observed. The results of this study failed to conform to the Lewkowicz Model and the main results of most other previous studies. The present study found most subjective synchronies close to zero, with some in the VA condition and a few in the AV condition. This suggests that the subjects in the present study were more sensitive than those of other studies to the temporal gap when the auditory followed the visual stimulus. On average, it took a delay of only 7.7ms for the subjects to detect asynchrony when the auditory stimulus followed. This was a very different result from those consistentiy found in most other studies, and the weak positive deviation in the present study indicates that the subjects did not tolerate greater temporal discrepancies when the auditory followed than when it preceded the visual stimulus. The subjects were not faster and more accurate in their judgments in the AV than in the VA condition.

Although Lewkowicz's general thinking that auditory and visual stimuli reach the CNS at different points in time should be correct, the present results lead us to believe that subjective synchrony is not due to difference in arrival times, or at least it is not predictable from RT difference. On the basis of Lewkowicz's claim that the consistent positive deviation of subjective synchrony was due to a difference in how quickly auditory and visual stimuli arrive at the comparator in the CNS, the present study developed the hypothesis that RT difference would be correlated with subjective synchrony, but this was not clearly apparent. In two ways, the data did not support the Lewkowicz Model. First, the subjective synchrony was not as big as 30–40ms. Many subjective synchronies were close to zero, and there were a few with negative values. Secondly, there was no correlation between the subjective synchrony and the difference in RT to the stimuli.

Nevertheless, most of the previous studies supported the Lewkowicz Model, with findings of fairly large positive deviations of subjective synchrony. Dixon and Spitz (1980) reported larger VA than AV thresholds and 63ms and 51ms of subjective synchrony for the voice and the hammer stimuli, respectively. In McGrath and Summerfield's (1985) experiment, using the Lissajou figure to mimic the opening and closing of lips, the subjective synchrony was 30ms, as the subjects judged the Lissajou movement and sound as one event when in fact the sound occurred 30ms after the movement. Bushara et al. (2001) also found 25ms of positive subjective synchrony, as a result of a larger VA threshold. They also measured the RT needed for subjects to respond to synchrony trials and reported that RT was faster when the auditory preceded the visual stimulus. The study of Hollier and Rimell (1998) also resulted in a positive deviation of 37ms, suggesting a greater tolerance of synchrony when the auditory stimulus follows and a greater sensitivity to asynchrony when the auditory stimulus leads. The present study's subjective synchrony was much smaller than those found in most others studies.

On the other hand, Smeele (1994) observed a strong negative subjective synchrony. In her study of bimodal speech perception, using the method of Order, the mean subjective synchrony was −105ms. This negative value indicates that the subjects tolerated the asynchronous speech stimuli with auditory stimulus leading. However, such a different result from those of other studies could be attributed to the method used. There are two possible explanations in terms of the method of Order. First, Sternberg and Knoll (1975) argued that in the method of Order subjects are required to use a different strategy for making judgments than in the methods of In-Synchrony or Out-of Synchrony and Better Synchrony. The method of Order specifically calls for attention to the difference in arrival times of the stimuli, and these other two methods require an integration of the stimuli. If this reasoning is correct, then the strong negative deviation of subjective synchrony in this particular experiment by Smeele should not be compared with the results of studies using the In-Synchrony or Out-of-Synchrony and Better Synchrony methods. Smeele concluded that judgment in a temporal order task can be different from judgment in a synchrony task; the perception of order is based on information that is different from, and to some extent independent of, the information used in the perception of asynchrony.

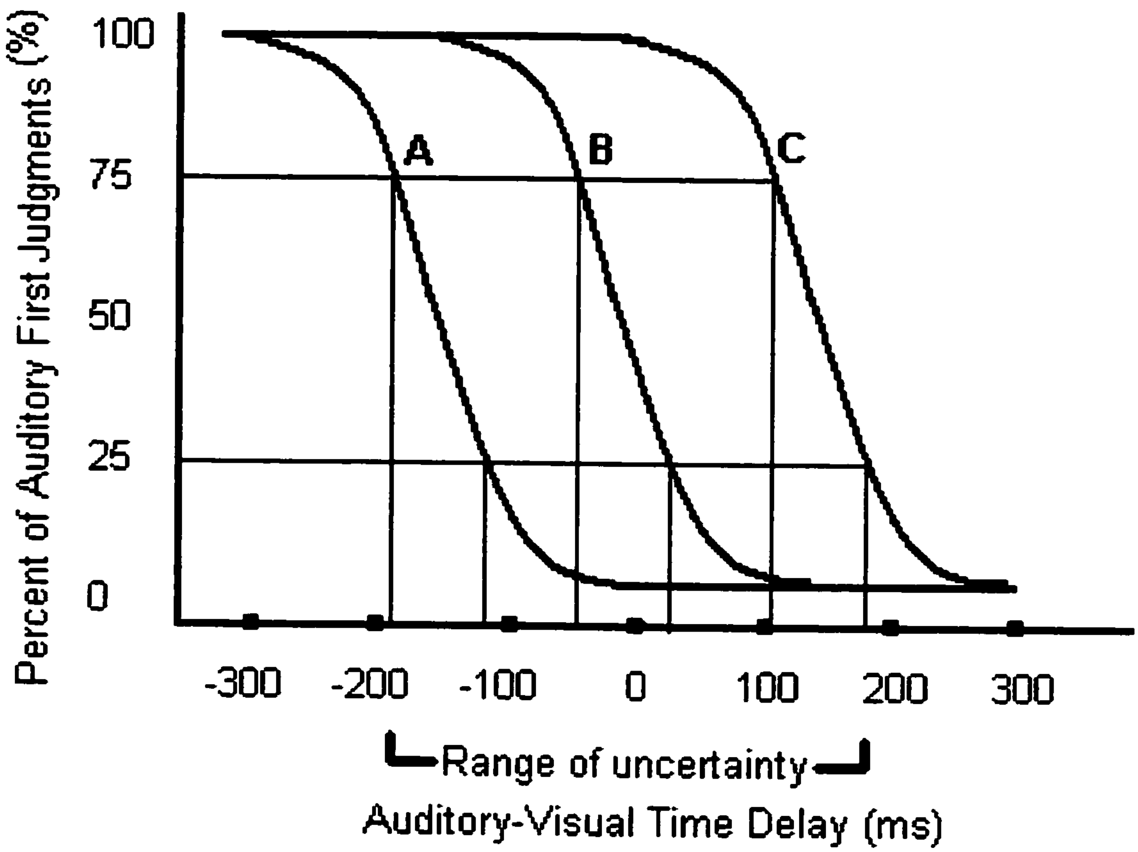
The subjects' task in the Order method might lead them to systematically judge in favor of a particular order of stimuli (Douglas, 2002). Theoretically the judgments of Auditory First should decrease as the auditory-visual delay approaches 0ms. The AV threshold would be located at 75% of the judgments in the AV condition, which is the dimension on which negative values are designated. The VA threshold obtained at 25% of judgments would have positive values in the VA condition. However, some subjects might voluntarily decide to affirm a particular order of the stimuli whenever they were uncertain of their judgments (Douglas, 2002).

For example, some subjects might decide to judge Visual First whenever they could not tell which stimulus occurred first. The resulting distribution graph would resemble that of Panel C in Figure15. In this case, the AV and VA thresholds — at 75% and 25%, respectively, of Auditory First judgments — would be measured both in the VA condition, yielding a AV threshold with a positive value. While the VA threshold would be valid, the AV threshold would be untrustworthy, as a result of the strategy adopted (Douglas, 2002).

On the other hand, when they could not tell which stimulus came first, they decide to answer that the auditory stimulus occurred first (instead of their guessing for each trial), then the distribution graph would appear like that of Panel A in Figure 13 (Douglas, 2002). Both the AV and VA thresholds would be found in the AV condition. Then the VA threshold would not be a valid value, because of the systematic strategy of judging Auditory First whenever in doubt. Thus, it would also have been possible for subjects in Smeele's (1994) experiment to decide to judge Auditory First whenever the order of stimuli seemed to them ambiguous.

The majority of studies reviewed earlier in the Introduction used the method In-Synchrony or Out-of-Synchrony. Table 1 shows the method that the studies used and the results. The method of In-Synchrony or Out-of-Synchrony appears to be the simplest and easiest compared in that it simply asks whether the pair of intermodal stimuli occurred together or separately. If subjects know that only one pair of stimuli will occur in synchrony at 0ms across the range of asynchronous delays, they are more likely to judge Out-of-Synchrony, even when they are not very sure of their judgment. This, then, falsely increases the percentage of Out-of-

Figure 13. Graph using Method of Order when subjects systematically adopt a strategy in favor of a particular order of stimuli presented: A - Graph with systematic "Visual First" strategy. Both AV (75%) and VA threshold (25%) are located in the AV condition; B - Theoretical graph, which is not endorsed by either strategy. AV threshold is located in AV condition and VA threshold is located in VA condition; C - Graph with systematic "Auditory First" strategy. Both AV and VA threshold are located in the VA condition (Douglas, 2002).



Synchrony responses. As a result, the method may not be suitable for measuring thresholds in the AV and VA conditions for the purpose of examining temporal asymmetry. Again, all of the studies except Smeele's (1994) that used this method observed positive deviation of subjective synchrony and temporal asymmetry of thresholds.

On the other hand, the method Better Synchrony requires subjects to choose between two pairs of stimuli, one of which is always in-synchrony, and the task here is to select the pair that seems more synchronous than the other. However, it is considered better as a psychophysical method than In-Synchrony or Out-of-Synchrony. As subjects in the latter method know that all stimulus pairs occur with delays except one pair with 0ms, they are more likely to make the Out-of-Synchrony judgment. That is, even when they cannot tell whether stimuli occur together or separately, they make a correct judgment of Out-of-Synchrony. In the Better Synchrony method, the subjects have to select the pair in better synchrony, and so they cannot make assumptions, and their judgments have a 50–50 chance of being correct.

However, not many studies have used this method to date. Among the studies reviewed in this paper, only McGrath and Summerfield (1985) made use of this method, having subjects indicate the target pair of stimuli against the comparison target in a special recognition task with Lissajou interference. They found the same asymmetrical thresholds, along with subjective synchrony in the VA condition. Based on McGrath and Summerfield's (1985) results, it seems that the method of Better Synchrony requires the same tasks and yields the same findings with respect to the direction of the subjective synchrony.

However, Matin and Bowen (1976) argued that different psychophysical methods might measure different aspects of time perception in investigating the perception of temporal asymmetry. Specifically, different methods may involve perceiving stimuli at different points of time in their occurrence, either at the onset or offset of stimuli. Detecting the delay between two stimuli can differ, depending on whether the delay was measured at the onset of the stimuli or at the offset. Moreover, latency, which is duration between physical onset of a stimulus and its perception, may add to the overall mechanisms of judgment of synchrony. Onset and offset latencies of stimuli produce could distort judgments of synchrony of stimuli (Matin & Bowen, 1976). The researchers suggested that to accurately determine synchrony it would be necessary to identify and measure the neural event that corresponds to the perception of the stimulus.

Regarding the judgmental situation related to the perception of onset and offset, Penner (1975) had a hypothesis that each subject selects his or her own criterion when judging onset and the offset of stimuli. Perception of stimuli by judging their onset or offset is not universal but may be different from one subject to another, and their choice of criteria can influence their judgments of synchrony. Penner's hypothesis was based on Robinson's (1974) model, which assumes that subjects still receive stimulus even when its physical source has ceased. Robinson argued that stimulus latency, or dragging sensation beyond their physical presence, can also affect the judgment of when the offset of a stimulus occurs. If Penner's suggestion holds true, the unexpected small positive deviation of subjective synchrony in the present study could be understood as a result of the subjects' having selected a different criterion for onset and offset from that chosen by subjects in other studies.

Although subjective synchrony did not correlate with RT and was small compared with the results of other studies, an asymmetrical pattern of the two thresholds was evident in the present study. The tendency of the AV threshold to be smaller and for the VA threshold to be larger, which is the characteristic of the temporal asymmetry of thresholds found in other studies, was replicated in the present data. Statistical results make it evident that the two thresholds differ in the absolute size. However, the question of what causes the temporal asymmetry remains if the arrival-time difference as in Lewkowicz Model is not the answer. The possible mechanisms underlying the positive deviation in temporal-synchrony perception are discussed in the literature.

Grondin (1996) discussed the mechanism of the temporal asymmetry. Grondin argued that visual stimuli have a longer latency, as they trail and fade more slowly than auditory stimuli, producing a more acute persistence after their occurrence. Because of these characteristics of the two stimuli, the internal representation of identical delays appears shorter in the VA than in the AV condition. This differential contributes to the positive deviation of subjective synchrony. On the other hand, when the auditory and visual stimuli occur simultaneously, the longer latency of the visual stimulus makes a subject perceive it as having occurred later than the auditory stimulus. This effect, in turn, makes the percentage of judgments of asynchrony to increase. When, on the other hand, the visual precedes auditory stimulus by a certain amount of time, the longer latency of the visual stimulus makes it seem to remain until the auditory stimulus appears. As a result, the stimuli are perceived as in synchrony and the percentage of judgments of asynchrony decreases, and subjective synchrony is pushed toward the VA condition.

Dixon and Spitz (1980) explained temporal asymmetry of thresholds in terms of perceptual learning. In their experiments using the hammer stimuli (image and sound of a hammer striking) and the voice stimuli (image and sound of someone speaking), they found two major results. One was that for each of the stimuli, the subjects tolerated longer delays when the sound followed than when it led. The researchers remarked that because sound travels more slowly than light, it takes longer for sound to travel from its source to the auditory receptors in the ears than for a visual stimulus to travel from its source to the visual receptors in the eyes. As a result, the brain may learn to tolerate temporal discrepancies between auditory and visual stimuli at their respective receptors so that it perceives the two stimuli as belonging to each other. Humans are accustomed to receiving visual earlier than auditory stimuli and perceiving them as occurring in synchrony. Conversely, when an experiment is set up to make an auditory stimulus occur and reach its receptor earlier than a visual stimulus, we become more alert to the unusual timing of sensory events. We detect asynchrony with smaller delays when an auditory stimulus leads and a visual stimulus follows. Our prolonged perceptual learning allows us to ignore or compensate for certain discrepancies between seeing and hearing a sound source (Dixon & Spitz, 1980). As a result of perceptual learning, the threshold is thus significantly smaller when the auditory precedes the visual stimulus.

The other result found by Dixon and Spitz was that the subjects perceived asynchrony with smaller delays in the case of the hammer than in that of the voice stimuli, both when its sound preceded and when it followed. In other words, the hammer stimuli had smaller thresholds both in the AV and VA conditions. While humans are daily exposed to voice stimulus in speech environments and therefore tolerant to a certain asynchrony between auditory and visual stimuli in speech, collisions between inanimate objects such as in the case of the hammer stimuli are experienced less frequently and require greater sensitivity to the sensory occurrence.

The perception of temporal synchrony between two distinct stimuli can therefore contain several sources of variability. Detecting the onset and offset of stimuli, as well as the method used in the experiment and its related behavioral tasks — along with general factors, such as attention and arousal — can all affect performance in temporal-synchrony perception. In explaining the mechanisms of the temporal asymmetry of thresholds, Lewkowicz (1996) did not discuss such variables but simply attributed its cause to the longer transduction time of visual stimuli. The present study concludes that the Lewkowicz Model (1996) is not complex enough to account for all possible variances of temporal-synchrony perception and calls for the development of a more thorough or new model to account for more of the possible mechanisms of synchrony perception.

On the other hand, the results of studies might be partly due to sampling fluctuation. These studies reviewed here did not have large enough sample sizes to generalize from the results. The present study had a sample size larger than those of any of the other studies, and did not find the subjective synchrony close to 30–40ms. It could be construed that in reality subjective synchrony is more or less close to 0ms for most of people and only a small portion of them show subjective synchrony around 30–40ms. Large enough sample size and sampling should be guaranteed in experiments in order to ensure valid and replicable results and allow generalization.

Nevertheless, the results of the present study might be partly due to the use of RT to reflect the arrival time of stimuli. On the basis of Lewkowicz's proposal that the positive subjective synchrony results from the difference in arrival times of the stimuli, we measured RT difference as a reflection of the difference between arrival times of the stimuli. Although we obtained 31ms of RT differences from our data, they did not correlate with subjective synchrony. It could be that subjective synchrony actually is related to arrival-time differences between auditory and visual stimuli, as Lewkowicz proposed, but RT differences do not manifest real arrival-time differences between the two stimuli. If so, then RT differences could only serve as a complementary indicator and a better indicator might be found to measure arrival-time differences between stimuli.

Another concern in the present study is that the subjects were not wearing the headphones on their ears, as the headphones were instead placed beside the apparatus on the table. The subjects were seated about 40cm away from the apparatus and the headphones. The difference between receiving auditory stimuli from the more distant source than from headphones closer to the ears, as in other studies, was considered in the experiment. Transduction should be considered in relation to the physical distance travelled by auditory and visual stimuli. The travelling time from the source of the stimulus to the sensory receptors in the brain is different, depending on the type of stimulus. Auditory stimuli travel more slowly to the auditory receptors in the ears than visual signals to the visual receptors in the eyes. Therefore, the distance between the source of the stimuli and the subject in the task should be carefully set up, and the expected additional time for the auditory stimulus should be subtracted from the thresholds. But this additional time was not subtracted from the thresholds in the present study, and this may have contributed to tranduction difference between auditory and visual stimuli.

We did attempt to calculate the effect of the extra 40cm of distance travelled by the auditory stimulus, using a rate of 340meter/seconds for the speed that sound travels. It was computed that the sound from the headphone placed 40cm away from the subjects took 1.178ms more than it would have had the earphones been worn on the subjects' ears. This small amount of time, however, does not seem to explain such a big difference in subjective synchrony as that of 22–32ms between the present study's results and the Lewkowicz Model.

If the industries using psychoacoustic systems want auditory and visual stimuli appear in synchrony in their products, they should ensure that the two stimuli come close enough in time. Based on the results of the present study, where the minimum thresholds were –22ms and 34ms, these industries should aim at making their products of which acoustic system meet the thresholds to make sure that even individuals with high sensitivity to asynchrony do not detect it.

Synchronization has increasingly been an issue in numerous areas of applied science, such as in the telecommunications, digital-media, and audio industries (Miner & Caudell, 1998). Along with having a role in the development of high technology in communications, audiovisual synchronization has become a crucial factor in developing virtual environments (VE) (Miner & Caudell, 1998). Creating an overall realistic and compelling virtual experience depends especially on the presentation of multisensory information (Miner & Caudell, 1998). Quantification of the audiovisual synchronization in thresholds is the key to successfully creating a VE experience (Miner & Caudell, 1998).

CONCLUSION

In conclusion, it may be said that the positive deviation of subjective synchrony occurs in perception of temporal synchrony, as found in the present study, although not as big as in other studies. Humans are less sensitive to asynchrony when auditory follows visual stimulus. However, the cause of this effect is still a matter of further debate. According to the present findings, RT difference is not a predictor of the positive deviation of subjective synchrony. As discussed, the perception of synchrony of stimuli of diverse modalities involves several factors, such as detection at onset and offset of stimuli, the experimental method used, and sampling. Further research on synchrony perception should examine these variables to determine the causes of the positive deviation of subjective synchrony. Findings in this area would have a wide range of application in today's world of telecommunications.

References

Bushara, Khalafalla, O., Grafman, Jordan, & Hallett, Mark (2001). Neural correlations of auditory-visual stimulus onset asynchrony detection. *The Journal of Neuroscience*, 21(1), 300-304.

Dixon, Norman F. & Spitz, Lydia (1998). The detection of auditory visual desynchrony. *Perception*, 9, 719-721.

Douglas, Brian (2002). A comparison of four methods used to measure thresholds of temporal asynchrony detection for audio-visual events. Undergraduate thesis: University of Ottawa.

Grondin, S. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics*, 58(3), 424-433.

Hirsh, Ira J. & Sherrick, Carl E. (1961). Perceived order in different sense modalities. *Journal of Psychology*, 62, 423-432.

Hollier, M.P. & Rimell, A.N. (1998). An experimental investigation into multi-modal synchronisation sensitivity for perceptual model development. In *705th Convention of the Audio Engineering Society*, San Francisco, California. Reprint No. 4790.

Kohlrausch, A. & van de Par, S. (1999). Auditory-visual interaction: From fundamental Research in cognitive psychology to (possible) applications. In *IS & T1 SPIE Conference on Human Vision and Electronic Imaging IV*. San Jose, California. SPIE Vol. 3644.

Lewkowicz, David J. (1996). Perception of auditory-visual temporal synchrony in human infants. *Journal of Experimental Psychology Human Perception and Performance*, 22, 1094-1106.

Matin, L. & Bowen, Richard W. (1976). Measuring the duration of perception. *Perception &*

Psychophysics, 20(1), 66-76.

McGrath, M. & Summerfield, Q. (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *Journal of Acoustic Society of America*, 77(2), 678-685.

Miner, Nadine & Caudell, Thomas (1998). Computational requirements and synchronization issues for virtual acoustic displays. Presence, 7(4), 396-409.

Penner, M.J. (1976). The effect of marker variability on the discrimination of temporal intervals. *Perception & Psychophysics*, 19, 466-469.

Robinson, C.E. (1974). Simple form of the auditory running-average hypothesis: Application to the temporal summation of loudness and to the delayed perception of offset of brief stimuli. *Journal of the Acoustical Society of America*, 55, 645-649.

Smeele, P. M. T. (1994). *Perceiving Speech: Integrating auditory and visual speech.* Delft: MIT Press.

Sternberg, S. & Knoll, R.L. (1975). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and Performance IV*. New York: Academic Press.

Wesley, Michelle (2002). The effect of unbalanced time-dalays on audio-visual judgments of synchrony. Undergraduate thesis: University of Ottawa.