Acta Psychologica 130 (2009) 225–234

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Asymmetric cross-modal effects in time perception

Kuan-Ming Chen, Su-Ling Yeh *

Department of Psychology, National Taiwan University, No. 1, Roosevelt Rd. Sec. 4, Taipei 10617, Taiwan

ARTICLE INFO

Article history:
Received 11 June 2008
Received in revised form 16 November 2008
Accepted 25 December 2008
Available online 4 February 2009

PsycINFO classification:
2300

Keywords:
Cross-modality
Scalar timing
Subjective time expansion
Time perception

ABSTRACT

It is common to judge the duration of an audiovisual event, and yet it remains controversial how the judgment of duration is affected by signals from other modalities. We used an oddball paradigm to examine the effect of sound on the judgment of visual duration and that of a visual object on the judgment of an auditory duration. In a series of standards and oddballs, the participants compared the duration of the oddballs to that of the standards. Results showed asymmetric cross-modal effects, supporting the auditory dominance hypothesis: a sound extends the perceived visual duration, whereas a visual object has no effect on perceived auditory duration. The possible mechanisms (pacemaker or mode switch) proposed in the Scalar Expectancy Theory [Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Annals of the New York Academy of Sciences: Vol. 423. Timing and time perception* (pp. 52–77). New York: New York Academy of Sciences] were examined using different standard durations. We conclude that sound increases the perceived visual duration by accelerating the pulse rate in the visual pacemaker.

1. Introduction

Studies of time perception have well established that subjective judgment of duration is not always faithful to the veridical interval that is timed (Efron, 1970a,1970b), and is affected by several factors. For example, the judged duration is longer than the veridical duration for dynamic stimuli (Goldstone & Lhamon, 1974; Kanai, Paffen, Hogendoorn, & Verstraten, 2006; Lhamon & Goldstone, 1974), larger magnitudes (Ono & Kawahara, 2007; Xuan, Zhang, He, & Chen, 2007), high aroused negative emotions (Angrilli, Cherubini, Pavese, & Mantredini, 1997), emotions requiring urgent action (Droit-Volet & Meck, 2007), impulsive personality (Wittmann & Paulus, 2008), and lower amounts of concurrently processed information (Hicks et al., 1977; Hicks, Miller, & Kinsbourne, 1976; Zakay, 1993). Moreover, the method of time estimation (e.g. production, reproduction, comparative time judgment, and verbal estimation) is also a critical factor influencing the subjective judgment of durations (Bueti, Walsh, Frith, & Rees, 2008; Hicks et al., 1976; Ulbrich, Churan, Fink, & Wittmann, 2007; Zakay, 1993; Zakay & Block, 1997).

Human timing ability has been mainly explained by the clock stage of the Scalar Expectancy Theory (Droit-Volet & Meck, 2007; Gibbon, 1977; Gibbon, Church, & Meck, 1984; van Wassenhove, Buonomano, Shimojo, & Shams, 2008; Wearden, Edwards, Fakhri, & Percival, 1998; Wittmann & Paulus, 2008), which consists of a pacemaker, a mode switch, and an accumulator. The pacemaker emits pulses at a certain rate, the mode switch controls the gating of pulses, and the accumulator stores the number of pulses. While facing a duration judgment, the mode switch closes at the onset of the duration that is timed. Its closure allows the pulses emitted from the pacemaker to flow into the accumulator and be collected. At the offset of this duration, the mode switch opens again so that the pulses are no longer accumulated. The number of pulses collected in the accumulator in the given duration represents the subjective duration, with a linear relationship: the more the accumulated pulses, the longer the judged duration.

In a prospective duration judgment, the judged duration can be influenced by different mechanisms in the clock stage, such as the rate of the pacemaker's pulses and/or the operation latency of the mode switch. These two mechanisms are independent but not necessarily exclusive; they can work serially to cause fluctuation in the pulses that are collected in the accumulator (Angrilli et al., 1997; Burle & Casini, 2001; Droit-Volet & Meck, 2007; Gibbon et al., 1984; Wearden et al., 1998; Wittmann & Paulus, 2008). When the pulse rate is accelerated, or the latency difference between the closing and opening of the mode switch is larger, or both, the perceived duration tends to be longer than the veridical duration, a phenomenon known as subjective time expansion or subjective time dilation (Kanai & Watanabe, 2006; Kanai et al., 2006; Tse, Intriligator, Rivest, & Cavanagh, 2004).

Tse et al. (2004) used the oddball paradigm to study subjective time expansion. In this paradigm, the participants conduct a prospective duration task in which they know a priori that they will

---

* Corresponding author. Tel.: +886 2 33663097; fax: +886 2 23629909.
E-mail address: suling@ntu.edu.tw (S.-L. Yeh).

0001-6918/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2008.12.008

judge the duration of a stimulus, so that the attention is oriented toward the temporal dimension of the stimuli (Block & Zakay, 1997; Hicks et al., 1976; Zakay, 1993). The duration of an oddball is compared to that of a standard, when one oddball is inserted among a train of standards in a serial presentation. Tse et al. (2004) observed that subjective time expansion occurred when the veridical duration was longer than 120 msec, in either visual or auditory modality. This is interpreted in terms of attentional orienting due to the unpredictability of the oddball: the oddball attracted more attention relative to the high-probability standards, thus its perceived duration was also relatively lengthened.

In their study Tse et al. (2004) focused mainly upon the perceived duration of stimuli presented in a single sensory modality. However, cross-modal effects have been well illustrated on different percepts (e.g. Chen & Yeh, 2008; MacDonald & McGurk, 1978; McGurk & MacDonald, 1976; Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Shams, Kamitani, & Shimojo, 2000; Vroomen & de Gelder, 2000; Watanabe & Shimojo, 1998). Time perception is no exception; it has been shown that temporal processing in one modality may be affected by the signals from other modalities (Goldstone, Boardman, & Lhamon, 1959; Guttman, Gilroy, & Blake, 2005; van Wassenhove et al., 2008; Walker & Scott, 1981). In one of these earlier contributions regarding intersensory temporal processing, Walker and Scott (1981) observed that perceived visual duration was affected by sound, and thus concluded that the perceived duration of audiovisual stimuli was resolved in favor of auditory information, because audition dominates vision for temporal processing (Welch, DuttonHurt, & Warren, 1986; Welch & Warren, 1980).

After Walker and Scott's (1981) proposition that audition was dominant in intersensory temporal processing, van Wassenhove et al. (2008) observed a contrary result favoring visual dominance. They modified the oddball paradigm in order to remove the unpredictability of the oddballs. In their study, an oddball was always presented at the fourth position among four standards in each trial, and the modality of stimuli was manipulated as a factor. They found, contrary to Walker and Scott's study, that perceived auditory duration was altered by visual input but that the reverse was not true.

Rather than the comparative time judgment adopted by van Wassenhove and her colleagues, Walker and Scott required the participants to reproduce their perceived duration by pressing a key after viewing the uni- or bi-modal stimuli. As the method of time estimation is also a factor influencing the perceived duration (Bueti et al., 2008; Hicks et al., 1976; Ulbrich et al., 2007; Zakay, 1993; Zakay & Block, 1997), the difference between their results may not reflect the cross-modal effect on time perception per se, but rather an effect that can be attributed to the different methods used for time estimation. To resolve this issue, we adopted the same comparative time judgment as adopted in van Wassenhove et al. (2008), and tested whether the opposite results observed between the two studies resulted from the methods used, or from other, overlooked, possibilities.

Regarding cross-modal time perception, three possible outcomes are predicted for the duration judgments of audiovisual stimuli. Firstly, either an auditory or a visual dominance should be observed, as observed by Walker and Scott (1981) or by van Wassenhove et al. (2008), respectively. The auditory dominance hypothesis predicts that adding a sound affects the subjective visual duration, but adding a visual object has little or limited effect on subjective auditory duration, since audition dominates vision for temporal processing (Welch & Warren, 1980; Welch et al., 1986). Based on the results of van Wassenhove et al. (2008), in contrast, visual dominance is also a possible outcome even though "The pattern of results … is difficult to reconcile with our current understanding of duration perception and classic model of multisensory integration" (van Wassenhove et al., 2008, p. 10).

Secondly, just as the attentional distraction model (Hicks et al., 1976, 1977) predicts, concurrent information should orient attention away from the temporal dimension of the stimulus, and a signal from another modality may act as a distractor. Therefore, it is subjective time contraction rather than expansion that is expected, due to the reduced attentional resource toward the temporal dimension of the stimulus. Moreover, Dalton and Spence (2007) observed that while irrelevant auditory singletons can capture attention in a sequential visual search task, irrelevant visual singletons cannot similarly affect a sequential auditory search task, rendering an asymmetric cross-modal effect on temporal attentional capture. Accordingly, subjective time contraction could affect either visual or auditory duration judgment for bi-modal stimuli, depending on which signal and modality capture attention most effectively.

Lastly, Tse et al. (2004) predicted that the degree of subjective time expansion would increase with the oddness of an oddball, and one of the methods of increasing the oddness is to increase the saliency of the stimulus. Bi-modal stimuli can be more salient than uni-modal stimuli if the signals from both modalities interact, and indeed, it has been shown that audiovisual stimuli have higher visibility than visual stimuli alone (e.g. Bolognini, Frassinetti, Serino, & Ladavas, 2005; Chen & Yeh, 2008; McDonald, Teder-Salejarvi, & Hillyard, 2000; Sheth & Shimojo, 2004). For that reason, the oddness hypothesis predicts a higher degree of subjective time expansion in the duration judgments of bi-modal stimuli, regardless of whether the stimulus is auditory or visual.

In this study, we used the oddball paradigm (Tse et al., 2004) to re-examine subjective time judgment of both uni- and bi-modal stimuli, using the duration judgment of the uni-modal stimuli as our baseline.[1] Our aim was twofold: to resolve the discrepancy in the results obtained by van Wassenhove et al. (2008) and Walker and Scott (1981), and to address the happening of the cross-modal effect by applying the internal clock stage of the Scalar Expectancy Theory (Gibbon, 1977; Gibbon et al., 1984).

2. Experiment 1

In this experiment, the judged duration of bi-modal stimuli was compared to that of uni-modal stimuli, and the difference between them was defined as the cross-modal effect. We examined the cross-modal effect of (1) sound on visual duration judgment and (2) visual objects on auditory duration judgment.

A factorial design with two treatments (task: visual or auditory × modality: uni- or bi-modal) was used in this experiment, with the task as a between-subject factor in Experiment 1A and a within-subject factor in Experiment 1B. Each participant was instructed to compare the duration of the oddball to that of the standards according to either the visual modality (i.e., visual task) or the auditory modality (i.e., auditory task). The judged duration was obtained for both the uni- and the bi-modal conditions in each task and the comparison between them revealed the influence of sound on visual duration judgment and vice versa.

2.1. Method

2.1.1. Participants

In Experiment 1A, 52 undergraduates at National Taiwan University participated in this study in exchange for course credit. Half were assigned to the visual task and the other half to the auditory task. In Experiment 1B, another group of 16 paid volunteers

---

[1] We did not attempt to include all possible combinations of audiovisual stimuli: as a result the present study applies only to the condition in which the onset and offset of audiovisual stimuli correspond to each other.

participated in both the visual and the auditory tasks. All of them had normal or corrected-to-normal vision and hearing, and were naïve as to the purpose of the experiment.

2.1.2. Stimuli and apparatus

Stimuli were controlled by an ASUS D360 Pentium 4 PC and presented on a ViewSonic Graphics Series G90f+ 21-inch CRT monitor with a refresh rate of 85 Hz. The visual standards were black and the visual oddballs were red. Both were disks with a radius of 1.06° visual angle, presented on a white background and centered on the screen. The auditory standard and oddball were 440 and 1000 Hz sinusoidal tones, respectively. Auditory tones were presented binaurally with a stereo headphone, and the intensity was 55 dB SPL as measured at the participant's ear. In both cases, the oddballs and standards were static disks or steady tones without any dynamic change in size, in order to avoid the effect of size difference or dynamic property on subjective time experience (see Section 5).

The uni- and bi-modal conditions were presented in two blocks (see Fig. 1). In both conditions, the duration of the standards was set to a constant duration of 1059 msec. The duration of the oddballs was varied among nine test durations: 753, 835, 906, 976, 1059, 1129, 1200, 1282, and 1353 msec. The oddball appeared six times at each of the nine test durations. These 54 appearances of oddballs were randomly distributed within a train of 540 standards (with 7–12 standards inserted between oddballs), and therefore the probability of the oddball's appearance was 9.09%. All the stimuli were separated by one of the three interstimulus intervals (ISIs): 953, 1059, and 1153 msec, selected at random. This arrangement of irregular ISIs was to prevent the participants from making responses according to a constant rhythm.

A 1000 Hz sinusoidal tone was added to the visual task to create the bi-modal condition, with the oddball and its accompanying tone having the same duration (i.e., the onset and offset of each were identical). In the bi-modal condition, the participants were told to respond to the duration of the visual stimuli and to ignore the tone.

In the case of the auditory task, a red disk was added to create the bi-modal condition, with the oddball and its accompanying disk having the same duration. The participants responded to the duration of the auditory stimuli and ignored the red disk in this condition while maintaining their fixation on the center of the screen.

2.1.3. Procedure

In Experiment 1A, each participant completed only the visual or auditory task, with a counterbalanced order between uni- and bi-modal conditions. In Experiment 1B, participants completed both tasks, and the sequence of tasks and conditions followed the Latin Square arrangement across participants.

The participants wore an earphone and sat in front of the screen at a viewing distance of 50 cm, with their heads stabilized by a chin rest. They were instructed to fixate on the center of the screen. A fixation point was not used because the differing sizes of the fixation point and disks might cause apparent motion; however, the appearance of visual disks can be helpful in locating the center of the screen.

The participants were told that the duration of standards would remain constant and the duration of oddballs would vary randomly. They had to compare the duration of an oddball to the duration of the standards, and make their judgment after the disappearance of the oddball and before the appearance of the third standard following it. They were encouraged to use the duration of standards both before and after the oddball to minimize memory load while making the duration judgment (Tse et al., 2004), and they were informed that reaction time was not recorded. If the participants did not respond before the appearance of the third standard, the test duration of that oddball was re-inserted at random in the train of standards.

The participants pressed the key "/" with their right index finger when the duration of the oddball appeared longer than that of the standard and "z" with their left index finger when it appeared shorter. Each condition was continued until six responses at each of the nine test durations were successfully recorded.

2.2. Results and discussion

The data were fitted by a Weibull function (Tse et al., 2004) and the 50% point of the fitted curve was taken as the point of subjective equality (PSE) for each participant. The group-averaged

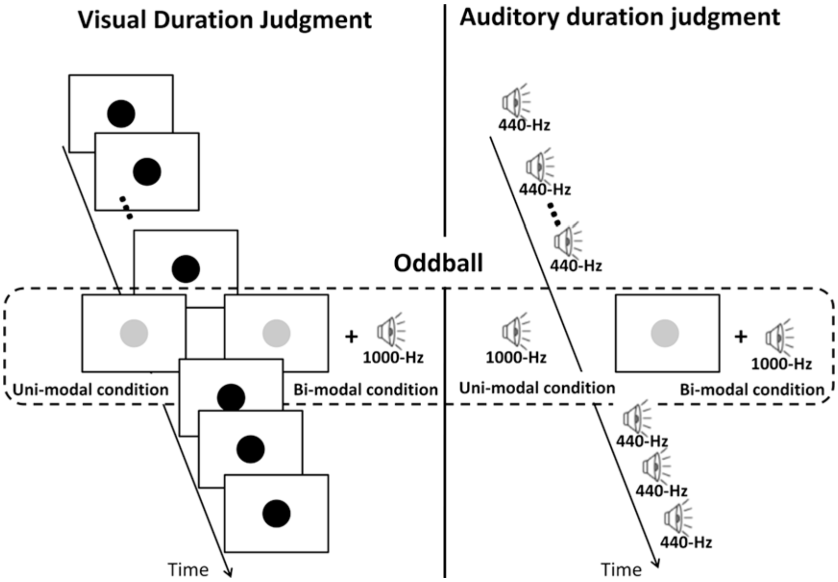

Fig. 1. Sequence of events for different duration judgment tasks and different modalities. The black disks and 440 Hz speakers symbolize the standards. The gray disks (red in experiments) and 1000 Hz speakers represent the oddballs. Before each oddball there were four to nine standards and after each oddball there were at least three standards, rendering 7–12 standards between two oddballs. No inter-trial interval existed in the experiment. The participants had to respond after the disappearance of the oddball and before the appearance of the third standard. The figures are not plotted to scale.

psychometric curves for the uni- and bi-modal conditions in each task of Experiments 1A and 1B are plotted in Fig. 2. We calculated the expansion ratio by dividing the standard duration by the PSE, with a value larger than 1 indicating subjective time expansion: the higher the expansion ratio, the larger the subjective time expansion.

In Experiment 1A, the expansion ratios (Fig. 3A) were submitted to a two-factor analysis of variance (ANOVA), with one between-subject factor (task: visual or auditory) and one within-subject factor (modality: uni- or bi-modal). The main effect of modality [$F(1, 50) = 24.496$, $MSE = 0.003$, $p < 0.001$] and the interaction of task × modality [$F(1,50) = 11.458$, $MSE = 0.003$, $p < 0.001$] were significant. The simple main effect confirmed that the expansion ratios in the bi-modal condition (the averaged expansion ratio = 1.092, $SE = 0.015$) were significantly higher than those in the uni-modal condition (the averaged expansion ratio = 1.007, $SE = 0.013$) for the visual task [$F(1,50) = 34.730$, $MSE = 0.003$, $p < 0.001$] but not for the auditory task [the averaged expansion ratio = 1.047, $SE = 0.015$ for the bi-modal condition and the averaged expansion ratio = 1.031, $SE = 0.015$ for the uni-modal condition, $F(1,50) = 1.224$, $MSE = 0.003$, $p = 0.274$], and the expansion ratios were significantly higher for the visual task than for the auditory task in the bi-modal condition [$F(1,100) = 4.680$, $MSE = 0.006$, $p < 0.05$] but not in the uni-modal condition [$F(1,100) = 1.317$, $MSE = 0.006$, $p = 0.255$].

In Experiment 1B, the expansion ratios (Fig. 3B) were submitted to a two-factor (visual/auditory task × uni-/bi-modality) repeated-measure ANOVA. The main effects of task [$F(1,15) = 19.359$, $MSE = 0.005$, $p < 0.001$], modality [$F(1,15) = 4.974$, $MSE = 0.002$, $p < 0.05$], and their interaction [$F(1,15) = 11.170$, $MSE = 0.002$, $p < 0.01$] were significant. The main effect of the task indicated that the expansion ratios for the auditory task were significantly higher than those for the visual task, when task was a within-subject factor. The simple main effect confirmed that while the expansion ratios in the bi-modal condition (the averaged expansion ratio = 1.067, $SE = 0.014$) were significantly higher than those in the uni-modal condition (the averaged expansion ratio = 1.000, $SE = 0.020$) for the visual task [$F(1,30) = 15.462$, $MSE = 0.002$, $p < 0.001$], the difference between the uni- (the averaged expansion ratio = 1.118, $SE = 0.020$) and bi-modal conditions (the averaged expansion ratio = 1.105, $SE = 0.021$) for the auditory task was not significant [$F(1,30) = 0.556$, $MSE = 0.002$, $p = 0.462$]. Moreover, the expansion ratios were significantly higher for the auditory task than for the visual task in the uni-modal condition [$F(1,30) = 30.439$, $MSE = 0.004$, $p < 0.001$], whereas the expansion ratios for the two tasks did not differ in the bi-modal condition [$F(1,30) = 3.134$, $MSE = 0.004$, $p = 0.086$].

The main effect of the task was significant: larger subjective time expansion was observed for the auditory task than for the visual task, but only in Experiment 1B (a within-subject design) and not in Experiment 1A (a between-subject design). It is well documented that the pulse rate is faster in the auditory pacemaker than in the visual pacemaker (e.g. Goldstone & Lhamon, 1974; Wearden et al., 1998), and therefore there is a modality difference while judging the same duration. It is possible that with a repeated-measures design the pacemaker would be less variable within a participant than across participants, and thus the phenomenon may merely reflect the modality difference since lights are judged shorter than sounds when they are actually equal in duration (Goldstone & Lhamon, 1974; Goldstone et al., 1959). However,

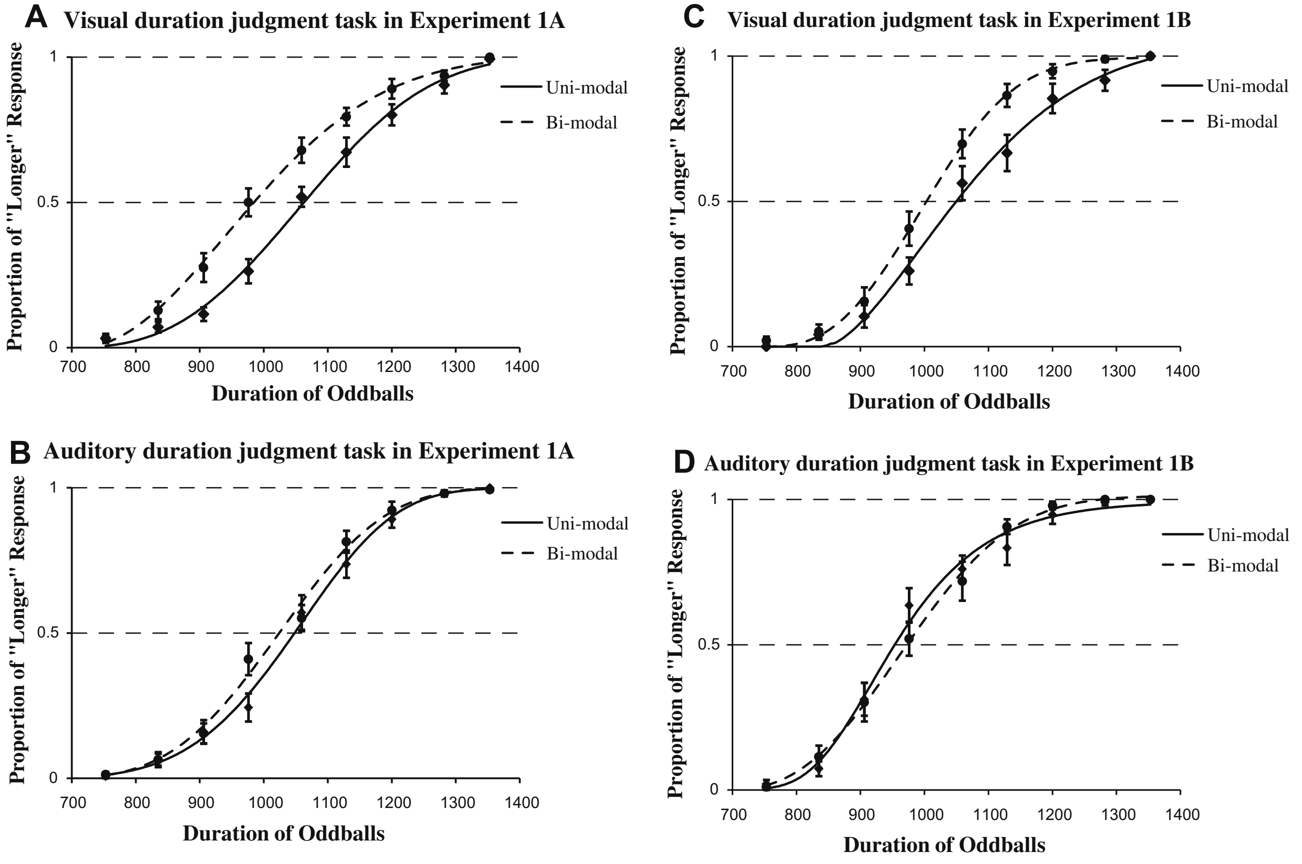

Fig. 2. The group-averaged psychometric curves, separated for the uni- and bi-modal conditions of Experiment 1A (task as a between-subject factor) and Experiment 1B (task as a within-subject factor).

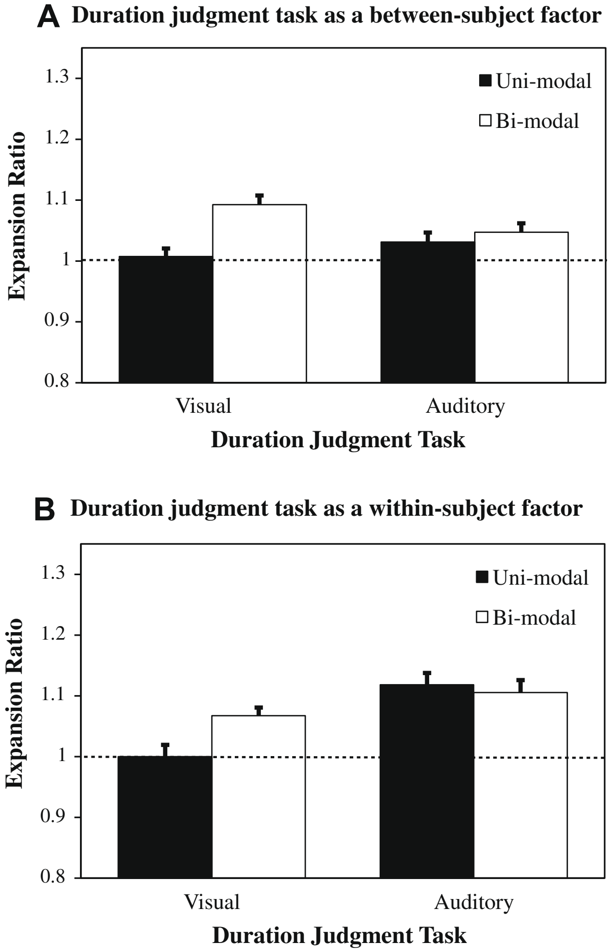

Fig. 3. The averaged expansion ratio for different duration judgment tasks and different modalities in Section 2. The dotted line indicates the 1.0 expansion ratio.

another possibility is that a uni-modal oddball is more salient when defined by a difference in pitch than by a difference in colour, and that difference in salience becomes more obvious with a within-subject design.

Our main interest was to compare the subjective time judgment in the bi-modal condition to that in the uni-modal condition for different tasks. In both experiments, a robust asymmetric result was observed, in that a concurrent sound influenced the visual duration judgment, causing it to expand, whereas a concurrent visual object did not influence the auditory duration. Moreover, in Experiment 1B, in the bi-modal condition, the subjective time expansion of the visual task was not different from that of the auditory task, which means that the added sound not only expanded the visual duration judgment but also equated it to the bi-modal condition of the auditory task with a repeated-measures design.

A possible reason that no difference was observed between the uni- and bi-modal conditions in the auditory task could have been the participants' failure to pay attention to the screen even though they were required to do so. In order to avoid the possibility that the participants simply did not see the visual pattern, we conducted a control experiment in which 16 participants performed only the auditory task, but with different manipulations in two blocks. In one manipulation, only the bi-modal oddballs (a red disk and a 1000 Hz tone) were included. In the other manipulation, catch trials (a black disk and a 1000 Hz tone) were randomly distributed, in addition to the bi-modal oddballs, within the train of standards and oddballs in order to force the participants to fixate their eyes on the screen. The procedure remained the same as the auditory task in Section 2, except that the participants were instructed to press the space bar for the catch trials. They had to rely on the visual difference (a black disk but not a red disk) in order to correctly respond to catch trials so that fixating on the screen was necessary. The sequence of these two manipulations was counterbalanced across participants. In both manipulations, the expansion ratios obtained in the control experiment did not differ [$F(1,15) = 2.553$, $MSE = 0.002$, $p = 0.140$]. We also compared the bi-modal condition of the auditory task in Experiment 1B to that in the control experiment and confirmed that they had the same expansion ratio [$F(1,30) = 0.720$, $MSE = 0.006$, $p = 0.403$]. Therefore, we cannot attribute the fact that no difference was observed between the uni- and bi-modal conditions in the auditory task in Section 2 to the participants' not seeing the visual changes in the bi-modal condition.

Regarding the three predictions described in Section 1, our results cannot be explained by either the attentional distraction model, which predicts that reduced attentional resources allotted to the temporal dimension of the bi-modal stimuli will cause subjective time contraction, or the oddness hypothesis, which predicts that in both tasks increased the saliency of the bi-modal stimuli, and thus the subjective time expansion, is to be expected. Our results support the auditory dominance hypothesis, which predicts that audition dominates vision in temporal processing (Welch & Warren, 1980; Welch et al., 1986). The added sound changed the visual duration judgment, and consequently more subjective time expansion was found than when the visual oddball was presented without sound. No difference was observed between the conditions for the auditory duration judgment regardless of the presence of a visual object. This suggests that the added sound dominated and altered the processing of visual temporal judgment but the reverse was not true, consistent with what Walker and Scott (1981) observed in their reproduction task.

To summarize, both Experiments 1A and 1B showed that there was a cross-modal effect of sound on visual duration judgment but a visual change had no effect on auditory duration judgment, and this pattern of asymmetric cross-modal effects supports the auditory dominance hypothesis in temporal processing.

3. Experiment 2

In Section 2, the observed cross-modal effects support the typical finding of auditory dominance for temporal processing. However, we wondered whether audition is so dominant that the visual change may not be necessary to induce the cross-modal effect of sound on visual duration judgment, or whether a contingent change of oddballs in both modalities is required.

To be more specific, the contingency defined here is the synchronous appearance of the audio and visual stimuli. We verified, in this experiment, whether the contingency between the visual and auditory presentations of the bi-modal oddball is necessary, in which case a lack of this contingency cannot demonstrate the cross-modal effect, or whether sound alone simply overrides the visual oddball in a winner-takes-all fashion, so that the dominant auditory modality is sufficient to capture the whole temporal processing.

The bi-modal condition of the visual task was modified in such a way that the 1000 Hz tone was paired with a black visual oddball and named as the bi-modal (black disk) condition. In this condition, the contingency of audiovisual stimuli did not exist and the expansion ratio was compared to that of the uni-modal condition. If mere exposure to the 1000 Hz tone without the visual change is sufficient to induce the cross-modal effect observed in Section 2, a difference in the expansion ratio between the uni-modal and bi-modal (black disk) conditions is to be expected. If, however, a contingency between sound and visual change is necessary to induce the cross-modal effect, then no difference should be observed between the uni-modal and bi-modal (black disk) conditions. The bi-modal condition as used in Section 2 was also included for comparison and named as the bi-modal (red disk) condition in this experiment.

3.1. Method

3.1.1. Participants

Fifteen undergraduates participated in this study in exchange for course credit. All were naïve about the purpose of the experiment and had normal or corrected-to-normal vision and hearing.

3.1.2. Stimuli and apparatus

The stimuli and apparatus were the same as those used in the visual task in Section 2. In addition to the uni- and bi-modal (red disk) conditions as used before, the bi-modal (black disk) condition in which a 1000 Hz sinusoidal tone was paired with a black disk was also included.

3.1.3. Procedure

Participants were required to compare the duration of an oddball to the duration of the standards. In the bi-modal (black disk) condition, they were told to identify the visual oddballs with the help of an auditory signal, since both the visual oddballs and standards were black disks, differing from each other only in duration. All other aspects were identical with the visual task in Section 2. The three conditions were conducted within participants, and they were presented in a Latin Square arrangement across participants.

3.2. Results and discussion

The PSEs and expansion ratios of uni- and bi-modal conditions were obtained for each participant. The expansion ratios were submitted to a one-factor [uni-modal, bi-modal (black disk), and bi-modal (red disk) conditions] repeated-measure ANOVA and were significant [$F(2,28) = 5.154$, $MSE = 0.004$, $p < .05$]. The Tukey HSD showed that while the expansion ratios between the bi-modal (black disk) condition (the averaged expansion ratio = 1.096, $SE = 0.027$) and the uni-modal condition (the averaged expansion ratio = 1.066, $SE = 0.022$) did not differ, the expansion ratio in the bi-modal (red disk) condition (the averaged expansion ratio = 1.139, $SE = 0.030$) was significantly higher than that in the uni-modal condition (Fig. 4). In other words, sound alone was not sufficient to induce the cross-modal effect, and contingency between the visual and auditory changes was necessary for sound to expand visual duration.

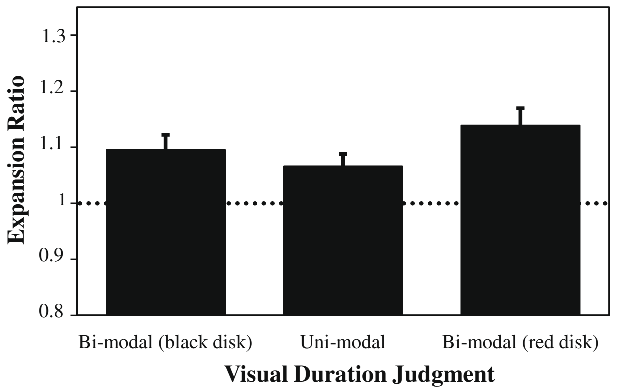

Fig. 4. The averaged expansion ratio for different modality conditions in Section 3. The dotted line indicates the 1.0 expansion ratio.

As auditory dominance over visual duration judgment was observed in Section 2, the result of Section 3 further suggests that the contingent change of the visual and auditory modalities was necessary to induce the cross-modal effect of sound on visual duration expansion, and the auditory change alone did not act in a winner-takes-all fashion. In other words, the observed cross-modal effect depends on the bi-modal interaction of audiovisual stimuli, rather than on the dominant modality alone.

4. Experiment 3

In Sections 2 and 3, the bi-modal interaction of audiovisual stimuli caused asymmetric cross-modal effects in time perception. In this experiment, we explored the mechanism responsible for the cross-modal effect of sound on visual duration expansion, with the help of the internal clock stage derived from the Scalar Expectancy Theory. As described previously, subjective time expansion is caused by the increased accumulation of temporal pulses during the event that is timed, which could be due to either (1) a higher pulse rate in the pacemaker or (2) a larger difference in latency between opening and closure of the mode switch (Angrilli et al., 1997; Block & Zakay, 1997; Burle & Casini, 2001; Droit-Volet & Meck, 2007; Gibbon, 1977; Gibbon & Church, 1984; Gibbon et al., 1984; Wearden et al., 1998; Wittmann & Paulus, 2008). The first possibility is similar to the phenomenon of auditory driving observed by Welch et al. (1986), in which the visual flicker would become perceptually synchronized with the auditory flutter. Also, sounds are judged to be longer than lights (Goldstone & Lhamon, 1974; Goldstone et al., 1959), because the pulse rate in the auditory modality runs faster than the one in the visual modality (Wearden et al., 1998). Therefore, it is feasible that the pulse rate of the visual pacemaker is driven faster by the concurrent sound, leading to the expansion of visual duration. If the pulse rate in the pacemaker is accelerated, the effect will be proportional to the durations that are timed (Gibbon & Church, 1984; Wearden et al., 1998).

The second possibility can be explained by the discrepancy of subjective simultaneity between the visual and auditory stimuli. It is well known that the perception of a sound onset precedes a simultaneous visual stimulus, so that the onset of the concurrent sound may serve as the marker for the audiovisual event (Fendrich & Corballis, 2001; Jaskowski, Jaroszyk, & Hojan-Jezierska, 1990; Kanai, Sheth, Verstraten, & Shimojo, 2007; Stone et al., 2001; Sugita & Suzuki, 2003; Zampini, Shore, & Spence, 2003). Accordingly, the closure of the mode switch may be associated with this sound onset marker. If the accumulation process is prolonged due to faster switch closure, an absolute horizontal placement of psychometric functions will appear, regardless of the duration that is judged (Gibbon & Church, 1984; Wearden et al., 1998).

We adopted the algorithm applied by Wearden et al. (1998) to the internal clock stage in order to differentiate the effects between the pacemaker and the mode switch. When facing a veridical duration (t) that is to be judged, the pulse rate of the pacemaker (r) and the total length of the mode switch closure ($t - lc + lo$, where lc denotes the latency of the mode switch to close and lo denotes the latency of the mode switch to open) work together to give rise to the number of pulses collected in the accumulator, and therefore the judged duration (y) is equal to $r(t - lc + lo)$. Based on the formula $y = rt + r(lo - lc)$, we can now use the veridical duration (t) to predict the judged duration (y). The first component rt is related to the pulse rate of the pacemaker and the veridical duration, and the second component is a multiple of the pulse rate and the difference of the latencies between the opening and closing of the mode switch.

If the added sound changes the pulse rate of the visual pacemaker to some extent, both components will also be changed to $\Delta rt + \Delta r(lo - lc)$ but $\Delta r(lo - lc)$ can be observed only when $(lo - lc)$ is greater than zero. Δrt can be inferred from the change in the slope of the function and $\Delta r(lo - lc)$ from the change in the intercept. Conversely, if the added sound influences the subjective time expansion through changing the latency of the mode switch, then a change will be observed in the intercept $r\Delta(lo - lc)$ only. If the sound casts its influence through both mechanisms, changes will be found in both the slope and the intercept.

Only the visual task was implemented in this experiment, and three durations were used for the standards (529, 1059, and 2106 msec): their judged durations were used to fit to the function $y = rt + r(lo - lc)$.

4.1. Method

4.1.1. Participants

Eighteen paid naïve participants, with normal or corrected-to-normal vision and hearing, took part in this experiment.

4.1.2. Stimuli and apparatus

All the stimuli and the apparatus were the same as those used in the uni- and bi-modal conditions of the visual task in Section 2, except for the three durations used for the standards (529, 1059, and 2106 msec) and their corresponding nine test durations. For the 1059 msec standards, the nine test durations of the oddballs were the same as those in Sections 2 and 3. For the 529 and 2106 msec standards, we directly adopted the nine test durations of oddballs used in Tse et al. (2004) study. The nine test durations for the 529 msec standards were 235, 282, 329, 376, 424, 471, 506, 553, and 600 msec. For the 2106 msec standards, the test durations were 1129, 1271, 1424, 1576, 1729, 1871, 2024, 2176, and 2329 msec.

4.1.3. Procedure

Eighteen participants took part in all six combinations of modality (uni- and bi-modal conditions) and durations of standards (529, 1059, and 2106 msec). The sequence of the six experimental blocks followed the Latin Square arrangement across participants.

4.2. Results and discussion

Fig. 5A shows the averaged expansion ratios of 18 participants in the uni- and bi-modal conditions with 529 msec standards (in the uni-modal condition the averaged expansion ratio = 1.072, *SE* = 0.023 and in the bi-modal condition the averaged expansion ratio = 1.252, *SE* = 0.037), 1059 msec standards (in the uni-modal condition the averaged expansion ratio = 1.039, *SE* = 0.019 and in the bi-modal condition the averaged expansion ratio = 1.136, *SE* = 0.026), and 2106 msec standards (in the uni-modal condition the averaged expansion ratio = 1.127, *SE* = 0.019 and in the bi-modal condition the averaged expansion ratio = 1.271, *SE* = 0.033). The expansion ratios were submitted to a two-factor repeated-measure ANOVA (529, 1059, or 2106 msec duration of the standards × uni- or bi-modality). The main effects of the duration of the standards [$F(2,34) = 13.376$, $MSE = 0.009$, $p < 0.001$] and modality [$F(1,17) = 32.586$, $MSE = 0.016$, $p < 0.001$] were significant. Their interaction, however, was not significant [$F(2, 34) = 2.753$, $MSE = 0.006$, $p = 0.08$].

A planned post -hoc analysis was conducted to verify whether the expansion ratios were larger in the bi-modal condition than those in the uni-modal condition, even though the interaction was not significant. The analysis confirmed significantly larger expansion ratios in the bi-modal conditions across all three durations of the standards [$F(1,51) = 31.623$, $MSE = 0.009$, $p < 0.001$ for the 529 msec duration of standard, $F(1,51) = 9.126$, $MSE = 0.009$, $p < 0.01$ for the 1059 msec duration of standard, and $F(1,51) = 20.121$, $MSE = 0.009$, $p < 0.001$ for the 2106 msec duration of standard], indicating that the effect of added sound on the visual duration expansion was significant regardless of the duration of the standards.

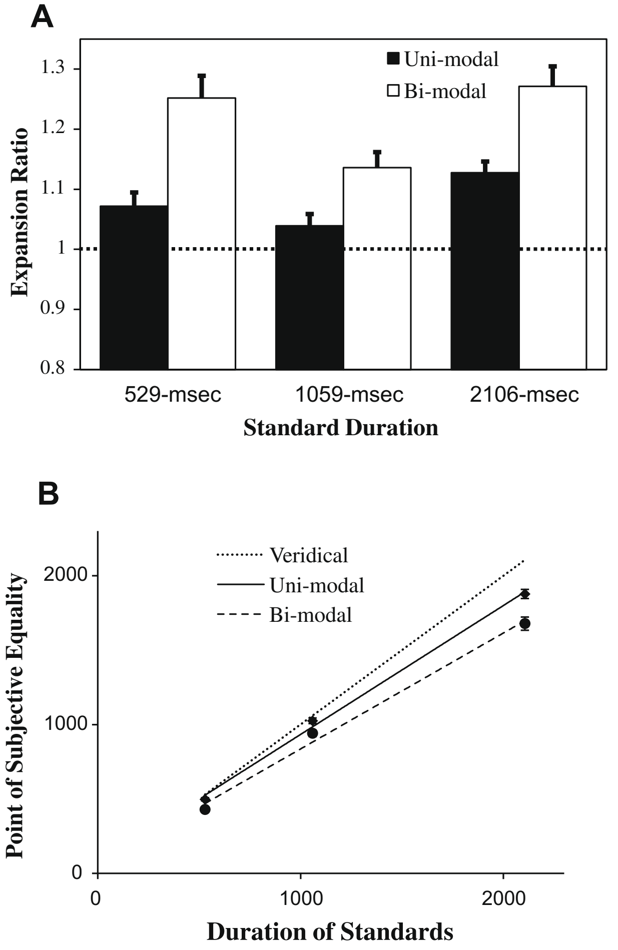

Fig. 5. (A) The averaged expansion ratio for different modalities at different standard durations in Section 4. (B) The averaged PSEs in the uni- and bi-modal conditions plotted against the duration of standards. The fine dotted line indicates the veridical duration.

The averaged PSEs obtained for the uni- and bi-modal conditions are plotted against the standard durations in Fig. 5B. The data showed an approximately linear relationship between the durations of the standards and their averaged PSEs. The ANOVA of the PSEs also supports this notion, showing significant differences in the main effects of the duration of the standards [$F(2,34) = 1497.812$, $MSE = 10527.043$, $p < 0.001$] and modality [$F(1,17) = 29.277$, $MSE = 12735.284$, $p < 0.001$], as well as their interaction [$F(2,34) = 7.047$, $MSE = 6504.661$, $p < 0.01$]. The interaction suggested that the magnitude of cross-modal effects increased with the duration of the standards: for the 529, 1059, and 2106 msec durations, the averaged PSEs were 497.747, 1025.161, and 1877.128 msec in the uni-modal conditions, respectively, and 428.927, 940.876, and 1677.696 msec in the bi-modal conditions). By performing a linear regression of the durations of the standards on the PSEs, the slopes r and intercepts $r(lo - lc)$ were obtained for each participant in the two modality conditions. The averaged slope was 0.866 ($SE = 0.017$) for the uni-modal condition and 0.780 ($SE = 0.028$) for the bi-modal condition. A t-test showed a significant difference between them [$t(17) = 2.965$, $p < 0.01$]. The differences in intercepts between uni- (the averaged intercept = 66.831, *SE* = 14.207) and bi-modal (the averaged intercept = 55.985, *SE* = 23.247) conditions were not significant [$t(17) = 0.419$, $p = 0.68$].

Obviously, the results suggested different slopes but identical intercepts between the two modality conditions. This means that the added sound accelerated the pulse rate of the pacemaker, whereas the operation latency of the mode switch did not change with modality. As an accelerated pulse rate caused more pulses to be collected in the accumulator within an identical veridical duration, more subjective time expansion was observed in the perceived visual duration when a sound was added.

5. General discussion

Using an oddball paradigm, we examined the effect of an added sound or visual object on the judgment of, respectively, visual or auditory duration. We found asymmetric cross-modal effects, in which an added sound expanded the perceived visual duration but an added visual object did not influence the judgment of auditory duration. This pattern of cross-modal effects supports the notion that our internal timer shares resources across modalities to some extent, with a constraint that audition dominates vision for temporal judgment.

In Section 3, the cross-modal effect has been further proved to result from the contingency between the visual and auditory changes, rather than from the sole contribution of sound alone. In Section 4, the cross-modal effect of sound on visual duration expansion was again robustly found, when a wider range of durations were used. Moreover, based on the internal clock stage derived from the Scalar Expectancy Theory (Block & Zakay, 1997; Droit-Volet & Meck, 2007), the effect of sound on visual duration is assumed to come from the accelerated pulse rate in the pacemaker, causing more pulses to be accumulated during the same interval that is timed, as compared to a condition involving no sound.

In the current study, a similar oddball paradigm was used, and the same comparative time judgment was applied as that applied in the study of van Wassenhove et al. (2008); however, we did not observe any cross-modal effects of the visual objects on auditory duration judgment. This might result from the lower salience of the red disk as an add-on object to the auditory oddball. However, as our results are consistent with those in the study of Walker and Scott (1981) and other studies (e.g. Goldstone et al., 1959; Guttman et al., 2005; Morein-Zamir et al., 2003; Welch & Warren, 1980; Welch et al., 1986), finding auditory dominance in intersensory temporal processing, the possible inequality of saliency as an add-on stimulus to the uni-modal oddballs cannot defeat our conclusion that duration judgments are made in favor of audition.

Our results clearly showed an asymmetric subjective time expansion for visual and auditory duration judgments. Since the attentional distraction model predicts subjective time contraction rather than expansion, and the oddball hypothesis predicts greater subjective time expansion for bi-modal than for uni-modal stimuli in both visual and auditory tasks, we find that neither of the two hypotheses is supported. The auditory dominance hypothesis better explains our asymmetric cross-modal effects. Further, our finding is consistent with that of Walker and Scott (1981) while using a different method of time estimation. This suggests that auditory dominance for temporal processing should stand equally between the time reproduction task and the comparative time judgment. Bueti et al. (2008) suggested that different brain circuits are recruited between temporal reproduction and perceptual representations of intervals, but that the activation of the basal ganglia and the cerebellum is common to different tasks. Therefore, it is reasonable to speculate that auditory dominance and bi-modal interaction of time perception may occur at relatively early stages of temporal processing.

Past studies have shown various forms of auditory dominance of temporal processing, in addition to the auditory driving (Welch et al., 1986) mentioned earlier. For example, Guttman et al. (2005) observed that the rhythmic stream of the to-be-ignored sounds interacted with the encoding of visual temporal structure: the incongruent sound interfered with a visual rhythm, whereas the congruent sound enhanced it. Some auditory dominance effects may be relevant to the time perception observed here. For example, Vroomen and de Gelder (2000) found a freezing phenomenon in which the detectability of a visual stimulus was enhanced by simultaneously presenting an abrupt tone in a perceptually organized sequence of tones. It is reasonable to infer from our study that their participants' freezing experience may have arisen from the effect of a concurrent oddball sound on the subjective time expansion of the visual target. Finally, Morein-Zamir et al. (2003) observed the phenomenon of temporal ventriloquism: a sound (versus no sound) trailing the second of two visual stimuli enhances judgment of their temporal order, which is accounted for by the sound pulling the second visual stimulus further away from the first visual stimulus in the temporal domain. According to our results, it is also possible that the temporal characteristic of the second visual stimulus was perceptually modified by the sound in a way similar to the subjective time expansion we observed here.

Using a comparative method similar to that of van Wassenhove et al. (2008), we observed results that were different from theirs, but were similar to those obtained by Walker and Scott (1981), who used a reproduction task instead. Therefore, the different results should not be attributed to the methods used. van Wassenhove et al. (2008) results depend heavily on the dynamic or size property of stimuli, but it is known that dynamic stimuli (Goldstone & Lhamon, 1974; Kanai et al., 2006; Lhamon & Goldstone, 1974) and stimuli with larger size (Ono & Kawahara, 2007; Xuan et al., 2007) per se can lengthen the perceived duration. In their study, the experimental conditions were separated into three types. In the loom condition, the visual and auditory oddballs were upward dynamic changes among steady standards (e.g. looming size or ascending tone). In the recede condition, the oddballs were shrinking disks or descending tones. In the reverse condition, the oddballs were steady stimuli among looming or ascending standards. They found subjective time expansion for visual, auditory, and bi-modal stimuli in the loom condition, but time compression for visual and bi-modal stimuli in the reverse condition. Based on the results that the auditory duration was expanded in the loom condition but compressed in the reverse condition by concurrent visual stimuli, they concluded that the concurrent visual stimuli modified auditory time perception. However, the fact that stimuli with dynamic attributes and/or larger size lead to time expansion may explain why they obtained time expansion in the loom condition and time compression in the reverse condition. We were careful to choose static stimuli, with no size difference between standards and oddballs, and thus avoided the effect of dynamic change or change in size on time expansion. Moreover, our finding of the asymmetric cross-modal effects was established on the basis of identical stimuli: the oddballs in the bi-modal conditions (i.e., a red disk with a 1000 Hz tone) were the same as those used for visual and auditory tasks. This was not the case in the study of van Wassenhove et al. (2008).

As the pacemaker and mode switch are the two proposed mechanisms of time perception (Angrilli et al., 1997; Block & Zakay, 1997; Burle & Casini, 2001; Droit-Volet & Meck, 2007; Gibbon, 1977; Gibbon & Church, 1984; Gibbon et al., 1984; Wearden et al., 1998; Wittmann & Paulus, 2008), we examined further which mechanism may have contributed to our cross-modal effects on time expansion. Our results in Section 4 suggested that the added sound accelerated the pulse rate of the pacemaker, causing more pulses to be collected in the accumulator within the same veridical duration. Since the bi-modal interaction was proved to be necessary and it has been shown that sound alone did not influence the visual duration expansion in Section 3, it is reasonable to infer that sound accelerated the pulse rate of the visual pacemaker, just as in the phenomenon of auditory driving (Welch et al., 1986). This also refutes the account that in the visual task participants may have used a slower pulse rate in the visual modality for the uni-modal oddballs (because it was the only modality to be used) and faster pulse rate in the auditory modality for the bi-modal (sound + red disk) oddballs (because both modalities were available, and sound was dominant and could take priority). Had the pacemakers of different modalities been used alternatively between the uni- and bi-modal oddballs, we would not have observed the same visual duration expansion in the bi-modal (black disk) condition, as compared to the uni-modal condition, in Section 3.

The perception of the sound onset is judged to occur before the visual stimulus when audiovisual events appear simultaneously (Fendrich & Corballis, 2001; Jaskowski et al., 1990; Kanai et al., 2007; Stone et al., 2001; Sugita & Suzuki, 2003; Zampini et al., 2003). It is thus possible that the closure of the mode switch could be associated with the marker of the sound onset, leading to more visual duration expansion. However, in Section 4 we did not obtain a significant change in the latency of the mode switch: thus, the discrepancy of subjective simultaneity between the visual and auditory stimuli may not play a critical role in inducing the cross-modal effect of sound on visual time expansion.

The uni-modal conditions used in our study are similar to those used in the study of Tse et al. (2004), but unlike them, across the three experiments we did not always observe subjective time expansion in the uni-modal conditions. This difference may come from the fact that in our study participants had to complete the uni- and bi-modal conditions in a single experiment, which might have caused fatigue in the participants or decreased salience of the uni-modal oddballs relative to the bi-modal oddballs. Even though we did not always observe subjective time expansion for uni-modal oddballs, the cross-modal effects of sound on visual duration judgment were consistently found, in all the bi-modal conditions of the visual task across all experiments and all durations of the standards.

In conclusion, we have demonstrated in this study that adding a sound expanded the subjective judgment of visual duration, but not vice versa. This asymmetric cross-modal effect was consistently found throughout the study using identical bi-modal stimuli, supporting the auditory dominance hypothesis. Further examination of the mechanism indicates that the added sound accelerates the pulse rate of the visual pacemaker, causing more pulses to be collected in the accumulator, thus extending the perceived visual duration.

Acknowledgements

This research was supported by grants from the National Science Council of Taiwan (NSC96-2752-H-002-008-PAE and NSC96-2413-H-002-009-MY3). We thank Ryota Kanai and an anonymous reviewer for their very constructive comments and suggestions.

References

Angrilli, A., Cherubini, P., Pavese, A., & Mantredini, S. (1997). The influence of affective factors on time perception. *Perception and Psychophysics, 59*(6), 972–982.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin and Review, 4*(2), 184–197.

Bolognini, N., Frassinetti, F., Serino, A., & Ladavas, E. (2005). “Acoustical vision” of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Experimental Brain Research, 160*(3), 273–282.

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience, 20*(2), 204–214.

Burle, B., & Casini, L. (2001). Dissociation between activation and attention effects in time estimation: Implications for internal clock models. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 195–205.

Chen, Y. C., & Yeh, S. L. (2008). Visual events modulated by sound in repetition blindness. *Psychonomic Bulletin and Review, 15*(2), 404–408.

Dalton, P., & Spence, C. (2007). Attentional capture in serial audiovisual search tasks. *Perception and Psychophysics, 69*(3), 422–438.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504–513.

Efron, R. (1970a). The minimum duration of a perception. *Neuropsychologia, 8*(1), 57–63.

Efron, R. (1970b). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia, 8*(1), 37–55.

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception and Psychophysics, 63*, 719–725.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279–325.

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, T. G. Bever & H. S. Terrace (Eds.), Animal cognition (pp. 465–488). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Annals of the New York Academy of Sciences: Vol. 423. Timing and time perception* (pp. 52–77). New York: New York Academy of Sciences.

Goldstone, S., Boardman, W. K., & Lhamon, W. T. (1959). Intersensory comparisons of temporal judgments. *Journal of Experimental Psychology, 57*(4), 243–248.

Goldstone, S., & Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: 1. Sounds are judged longer than lights. *Perceptual and Motor Skills, 39*(1), 63–82.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see. *Psychological Science, 16*(3), 228–235.

Hicks, R. E., Miller, G. W., Gaes, G., & Bierman, K. (1977). Concurrent processing demands and the experience of time-in-passing. *American Journal of Psychology, 90*, 431–446.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology, 89*, 719–730.

Jaskowski, P., Jaroszyk, F., & Hojan-Jezierska, D. (1990). Temporal-order judgments and reaction time for stimuli of different modalities. *Psychological Research, 52*(1), 35–38.

Kanai, R., Paffen, C. L., Hogendoorn, H., & Verstraten, F. A. J. (2006). Time dilation in dynamic visual display. *Journal of Vision, 6*(12), 1421–1430.

Kanai, R., Sheth, B. R., Verstraten, F. A. J., & Shimojo, S. (2007). Dynamic perceptual changes in audiovisual simultaneity. *PLoS ONE, 2*(12), e1253.

Kanai, R., & Watanabe, M. (2006). Visual onset expands subjective time. *Perception and Psychophysics, 68*(7), 1113–1123.

Lhamon, W. T., & Goldstone, S. (1974). Studies of auditory-visual differences in human time judgment: 2. More transmitted information with sounds than lights. *Perceptual and Motor Skills, 39*(1), 295–307.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception and Psychophysics, 24*(3), 253–257.

McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806), 906–908.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research, 17*(1), 154–163.

Ono, F., & Kawahara, J. (2007). The subjective size of visual stimuli affects the perceived duration of their presentation. *Perception and Psychophysics, 69*(6), 952–957.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature, 408*(6814), 788.

Sheth, B. R., & Shimojo, S. (2004). Sound-aided recovery from and persistence against visual filling-in. *Vision Research, 44*(16), 1907–1917.

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society B: Biological Sciences, 268*(1462), 31–38.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature, 421*(6926), 911.

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception and Psychophysics, 66*(7), 1171–1189.

Ulbrich, P., Churan, J., Fink, M., & Wittmann, M. (2007). Temporal reproduction: Further evidence for two processes. *Acta Psychologica, 125*(1), 51–65.

van Wassenhove, V., Buonomano, D. V., Shimojo, S., & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE, 3*(1), e1437.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1583–1590.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance, 7*(6), 1327–1339.

Watanabe, K., & Shimojo, S. (1998). Attentional modulation in perception of visual motion events. *Perception, 27*(9), 1041–1054.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology: Section B, 51*(2), 97–120.

Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception and Psychophysics, 39*(4), 294–300.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*(3), 638–667.

Wittmann, M., & Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences, 12*(1), 7–12.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision, 7*(10), 1–5.

Zakay, D. (1993). Time estimation methods – Do they influence prospective duration estimates? *Perception, 22*(1), 91–101.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science, 6*(1), 12–16.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research, 152*(2), 198–210.