Acta Psychologica xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Acta Psychologica

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

Duration perception in crossmodally-defined intervals

Katja M. Mayer [a,b], Massimiliano Di Luca [a,c,*], Marc O. Ernst [a,d]

[a] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[b] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[c] University of Birmingham, Birmingham, United Kingdom
[d] University of Bielefeld, Bielefeld, Germany

ARTICLE INFO

Article history:
Received 8 January 2013
Received in revised form 29 May 2013
Accepted 10 July 2013
Available online xxxx

PsycINFO classification:
2320 Sensory Perception
2221 Sensory & Motor Testing

Keywords:
Crossmodal intervals
Perceived duration
Sensory latency
Linear regression model

ABSTRACT

How humans perform duration judgments with multisensory stimuli is an ongoing debate. Here, we investigated how sub-second duration judgments are achieved by asking participants to compare the duration of a continuous sound to the duration of an empty interval in which onset and offset were marked by signals of different modalities using all combinations of visual, auditory and tactile stimuli. The pattern of perceived durations across five stimulus durations (ranging from 100 ms to 900 ms) follows the Vierordt Law. Furthermore, intervals with a sound as onset (audio-visual, audio-tactile) are perceived longer than intervals with a sound as offset. No modality ordering effect is found for visualtactile intervals. To infer whether a single modality-independent or multiple modality-dependent time-keeping mechanisms exist we tested whether perceived duration follows a summative or a multiplicative distortion pattern by fitting a model to all modality combinations and durations. The results confirm that perceived duration depends on sensory latency (summative distortion). Instead, we did not find evidence for multiplicative distortions. The results of the model and the behavioural data support the concept of a single time-keeping mechanism that allows for judgments of durations marked by multisensory stimuli.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Humans are capable of perceiving durations of less than a second with high precision (e.g., Grondin & Rousseau, 1991). One common concept of the underlying time-keeping system enabling duration judgments is the pacemaker–accumulator model (see Church & Broadbent, 1990, for a review). The model postulates a pacemaker that generates pulses that are sent to an accumulator at a certain frequency. Emitted pulses reach the accumulator through a switch mechanism. Time is inferred from the number of pulses registered by the accumulator. It is an ongoing debate whether the human time-keeping system might consist of a single pacemaker–accumulator mechanisms or whether multiple pacemakers and accumulators might exist that are used depending on the tasks at hand (see Grondin, 2010, for example). For different time scales, for example, such as milliseconds, seconds to hours, or circadian cycles, previous research suggests that several internal clocks exist that differ from each other in their time-keeping properties (see Buhusi & Meck, 2009; Ivry & Schlerf, 2008). Whether even more specialized time-keeping mechanisms such as modality specific mechanisms, for example, exist is still under debate (e.g., Gamache & Grondin, 2010; Ulrich, Nitschke, & Rammsayer, 2006). In everyday life, time-keeping tasks can involve simultaneity or duration judgments of signals of multiple sensory modalities. Here, we investigate if judging durations of crossmodally-defined intervals is achieved by a single modality-independent mechanism (i.e., a unique pacemaker–accumulator) or whether multiple modality-specific mechanisms exist.

A large body of literature has shown that time perception depends on the modality of the signals involved in a task at hand. In simultaneity judgments, for example, it was found that when presenting crossmodal signals such as brief light flashes, beeps, or tactile stimuli in physical simultaneity observers perceive them as occurring sequentially (Poliakoff, Shore, Lowe, & Spence, 2006; Zampini, Brown, Shore, Maravita, Röder and Spence, 2005; Zampini, Guest, Shore and Spence, 2005). Simultaneity of auditory and visual stimuli is commonly perceived when the visual stimulus precedes the auditory stimulus by about 20 to 30 ms (Zampini, Guest, et al., 2005), simultaneity of auditory and tactile stimuli is commonly perceived when the tactile stimulus precedes the auditory stimulus between 1.1 ms and 13.4 ms (Zampini, Brown, et al., 2005), and simultaneity between visual and tactile stimuli is commonly achieved when visual stimuli are presented 40 ms before tactile ones (Poliakoff et al., 2006).

The psychophysical results are generally consistent with the differences in the sensory latency of event-related potentials for each of the modality pairs. Allison, Matsumiya, Goff, and Goff (1977) found that the latency of visually evoked potentials was around 130 ms (referred to as VP130). The latency of auditory evoked potentials was 90 ms (referred to as AP90) and the latency of somatosensory evoked potentials was 100 ms (referred to as SP100). Moreover, single-cell recordings

* Corresponding author at: School of Psychology, Hills Building 2.04, Edgbaston, Birmingham, B15 2TT, United Kingdom. Tel.: +44 121 414 5526; fax: +44 121 414 4897.
E-mail address: m.diluca@bham.ac.uk (M. Di Luca).

0001-6918/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2013.07.009

in guinea-pigs confirmed that auditory signals had shorter transmission times than visual signals (King & Palmer, 1985).

Sensory latencies were found to affect duration judgments (Grondin & Rousseau, 1991; Rousseau, Poirier, & Lemyre, 1983; Grondin, Ivry, Franz, Perreault, & Metthè, 1996). Grondin and Rousseau (1991) found that when a brief beep marked the onset of an interval and a brief light flash marked the offset, the interval was perceived to be longer than when the onset was marked by a light flash and the offset by a beep. Other studies investigating duration judgments reported that simple biases due to perceptual latencies are not sufficient to explain all modality-dependent effects (Ulrich et al., 2006; Wearden, Edwards, Fakhri, & Percival, 1998). Wearden et al. (1998), for example, demonstrated that filled intervals defined by lights were perceived to be shorter than filled intervals defined by sounds. Such findings suggest differences in the internal processing of signals of different modalities defining duration that go beyond the effect of signal latency.

Following previous work (e.g., Grondin & Rousseau, 1991), the current study uses the biases in temporal perception caused by processing input of different sensory modalities to specify whether there is one modality-independent mechanism or whether multiple mechanisms are involved in processing crossmodal durations. Sensory latencies lead to summative distortion of time. Summative distortion of time means that perceptual latency causes a constant delay between the activation of the receptors by the stimulus and the percept. In the framework of the pacemaker–accumulator model, effects of constant distortions on perceived duration are explained by different latencies in activating and/or deactivating the switch. Latency in the activation of the switch (signal marking the beginning of the interval) causes fewer pulses to be accumulated, while latency in deactivation of the switch (signal marking the end) causes more pulses to be accumulated. If, however, the time-keeping system consists of multiple clocks, the frequency of the accumulated pulses depends on the underlying characteristics of the activated pacemaker. This means that a different clock is activated depending on the modalities defining onset and offset of a duration. In this case, each modality would have its own timing mechanism in form of a modality-specific pacemaker generating pulses at a modality-specific frequency. Both the single clock model and the multiple clocks model (e.g., Buhusi & Meck, 2009; Ulrich et al., 2006; Wearden et al., 1998) were able to account for a large range of phenomena that occur in human time perception.

Here, we directly test which of the two models describes duration perception in crossmodally-defined intervals. We presented participants with audiovisual, audiotactile, and visualtactile intervals and asked them to compare the intervals to probe durations (see Grondin & Rousseau, 1991). To test whether perceived duration follows a summative or a multiplicative pattern, we fitted a linear regression model to the observed data across five stimulus durations. Summative and multiplicative distortions make differential predictions on the properties of the regression line. Namely, summative distortion predicts a constant bias across all interval durations leading to an identical slope when the order of onset and offset of an interval is reversed. Multiplicative distortion, in contrast, predicts that the bias increases linearly with the interval duration because according to the pacemaker–accumulator model the longer the switch is activated the larger is the difference in the accumulated pulses between pacemakers with high pulse emission frequency and pacemakers with low pulse emission frequency (see Killeen & Taylor, 2000). It should be noted that the effect of multiplicative distortions is additional to the one of latencies in the activation/deactivation of the switch.

2. Method

2.1. Participants

The study received ethical approval by the ethics committee of the University of Tübingen. Thirty-six participants volunteered for the experiment (16 male, 20 female, age range: 19 to 33 years). They were recruited from the subject database of the Max Planck Institute for Biological Cybernetics, they were all naive to the purpose of the experiment and they gave written informed consent before taking part in the experiment. Participants had normal or corrected to normal vision and did not report any somatosensory or auditory deficit. They were randomly assigned to an experimental condition and no participant was tested in more than one condition.

2.2. Apparatus

Stimuli were presented using a custom-built device designed to generate co-located sound, vibration, and light with high temporal accuracy (for a picture, see Di Luca, Machulla, & Ernst, 2009). Two vertically aligned speakers with a center-to-center distance of 7.5 cm and a 2.5 cm radius produced the auditory stimuli. A vibration device (electro-magnetic shaker, Monacor Bass Rocker BR25) was situated between the speakers. It was mounted on a damping mass to produce tactile stimulation without audible noise. A LED array was mounted on top of the vibration device, serving as the vibrating surface as well as the light source (7×5 red LEDs, 1.6 cm $\times$ 1.3 cm). A multichannel sound card (M-audio 1010LT) was used to generate the stimuli. Sounds were 1000 Hz signals (61 dB SPL), lights were 145 Hz signals (93 cd/m^2), and vibrations were 120 Hz signals. The temporal accuracy of the stimuli generated by the device was verified by using an oscilloscope before the experiment.

Participants sat approximately 50 cm from the device in a dimly lit, sound-attenuated room. Noises from the computer fans were measured as approximately 30 dB SPL. In the conditions involving tactile stimulation participants were instructed to place their left index finger onto the LED array of the device and to maintain fixation on this location throughout the entire experiment. In the audiovisual experiment participants were asked to keep their gaze on the vibrating surface but they did not place the finger there.

2.3. Procedure

The paradigm was a two-interval forced-choice (2IFC) discrimination task. On each trial, two intervals were presented with an inter-stimulus interval (ISI) of 1000 ms. Within one trial, we used two different types of intervals, one empty "standard" interval and one filled "probe" interval. The empty interval was marked by two signals of different modalities, each with a duration of 20 ms. All signals were linearly ramped ($\pm$ 5 ms). Five stimulus onset asynchronies (SOA) were used for the empty interval (100, 300, 500, 700, and 900 ms). The filled interval was a continuous sound that lasted for 30%, 60%, 80%, 100%, 120%, 140%, and 170% of the duration of the empty interval. We used the combination of empty and filled intervals within a trial to control for perceptual time perception biases such as temporal shrinking (Nakajima, ten Hoopen, & van der Wilk, 1991 cited in Nakajima et al., 2004; Sasaki, Suetomi, Nakajima, & ten Hoopen, 2002), temporal stretching (Sasaki et al., 2010) or temporal ventriloquism (Morein-Zamir, 2003) which were likely to occur if we had asked participants to compare two empty intervals. Temporal shrinking refers to a perceptual bias that occurs when a train of three or four brief signals are presented. If three signals are presented in succession the duration of the interval marked by the second and third signal is underestimated. Sasaki et al. (2002) reported that temporal shrinking of the last interval occurs as well when four signals mark three intervals. If we had used two empty intervals in our 2IFC task participants could have interpreted that as a train of four signals. Therefore, the duration of the interval presented second might have been underestimated due to temporal shrinking. Temporal stretching (Sasaki et al., 2010) refers to a perceptual bias that occurs when two filled intervals are presented. It was found that the duration of the second interval was significantly overestimated. Though only reported in the auditory modality so far we were concerned that

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009

temporal shrinking and temporal stretching might occur if we ask participants to compare two empty or two filled intervals. Therefore, we chose a paradigm consisting of both an empty and a filled interval on a single trial. Temporal ventriloquism (Morein-Zamir, 2003) refers to a perceptual bias that occurs when auditory and visual signals are presented in succession. Morein-Zamir (2003) asked participants to judge which of two light flashes was presented first. Sensitivity decreased significantly when task-irrelevant clicks were presented before the first and after the second light flash. This result suggests that crossmodal signals can capture each other and therefore might affect their perceived occurrence in time. Therefore, we chose a paradigm with an empty and a filled interval to minimize the chance of temporal ventriloquism across the ISI.

Participants judged whether the first or the second interval on a trial was shorter. They were told that the duration of the empty interval was defined from the onset of first marker to the onset of the second marker. They entered their responses pressing one of two keys on a standard keyboard (left and right arrow) with the right hand. If participants missed one interval they had the possibility to repeat the trial one time by pressing the up arrow key. Three groups of participants were tested. One group was tested with audiovisual intervals, another group was tested with audiotactile intervals, and the third group was tested with visualtactile intervals. For each group, the experimental design was a $2 \times 2 \times 5$ factorial design with the within-subjects factors order and SOA and the between-subjects factor probe position. The factor order specified which of the two modalities that defined the interval was chosen as onset on a given trial (e.g., in the audiotactile group, participants were presented with both kinds of intervals: audio-tactile with auditory onset and tactile offset and tactile-audio with tactile onset and auditory offset). For the factor probe position, each group was split into two subgroups. One subgroup was presented with the filled "probe" interval before the empty "standard" interval to control for possible time-order effects (see Allan & Kristoff, 1974, for a review).

Each combination of the order of the markers (e.g., audio-visual vs. visual-audio), SOA of the empty interval (100, 300, 500, 700, 900 ms), and duration of the filled interval (30%, 60%, 80%, 100%, 120%, 140%, 170%) was presented six times to each participant. In total, the experiment consisted of a single 420 trials block with a break in the middle.

Responses were fitted with a cumulative Gaussian using the psignifit toolbox (Wichmann & Hill, 2001). From these fits we determined the Points of Subjective Equality (PSEs), which indicate the duration of the filled interval that is perceived to be equal in duration to the empty interval. From the fits we also determined the Just Noticeable Differences (JNDs), which indicate performance change from 50% "shorter" responses to 75% "shorter" responses in a two-interval discrimination judgment. Additionally, we tested whether our results confirmed with Weber's Law. To obtain Weber Fractions, we divided the JNDs by the PSEs rather than using the physical duration of the intervals for every participant, modality combination, and SOA.

3. Results

The PSEs, JNDs, and Weber Fractions were analyzed separately for each modality combination in 2 (order of markers defining the standard interval) × 2 (probe first vs. probe second on a trial) × 5 (SOA of the standard interval) mixed design ANOVAs with the within-subjects factors being order and SOA and the between-subjects factor being probe position. Since we found no significant effect of probe position nor interactions with this factor, we report a 2 (order) × 5 (SOA) repeated measures ANOVA for each modality combination.

3.1. PSE

The PSEs are plotted in Fig. 1, left column. Fig. 1 shows that participants tended to overestimate the shorter SOAs in the study and to underestimate the longer SOAs which is consistent with the range effect reported by von Vierordt (1868). For audiovisual intervals, there are significant main effects of order ($F(1,11) = 11.0$; $p = .007$; $\eta_p^2 = .50$) and SOA ($F(4, 44) = 38.3$; $p < .001$; $\eta_p^2 = .78$). On average, audiovisual intervals are perceived to be 46.2 ms (SEM = 13.9 ms) longer than visual-audio intervals. Similarly, audio-tactile intervals are perceived to be 49.0 ms (SEM = 11.8 ms) longer than tactile-audio intervals across SOAs as reflected in a main effect of order ($F(1, 11) = 17.1$; $p = .002$; $\eta_p^2 = .61$). The main effect of SOA is also significant ($F(4, 44) = 55.5$; $p < .001$; $\eta_p^2 = .84$). For visualtactile intervals, the main effect of SOA is significant ($F(4, 44) = 34.3$; $p < .001$; $\eta_p^2 = .76$). However, there is no significant difference between the two orders of the modalities (M = 17.1 ms; SEM = 12.0 ms; $p > .18$). For all modality combinations, the interaction of SOA and order did not have a significant effect on the PSEs ($ps > .22$). The lack of interactions between SOA and order is at odds with multiplicative distortions of perceived duration. To confirm this, we calculated the differences between the PSEs of the two orders at every SOA for every modality combination. If there were multiplicative distortions in the perceived durations we expect to find linear trends in the differences of the PSEs between the two modality orders. We submitted the differences between the modality orders to 1 × 5 (SOA) repeated measures ANOVAs for each modality combination. Neither the main effect of SOA ($ps > .22$) nor the linear trend ($ps > .09$) are found to be significant for any of the modality combinations.

3.2. JND

The JNDs are plotted in Fig. 1, middle column. For audiovisual intervals, there are significant main effects of order ($F(1, 11) = 13.0$; $p = .004$; $\eta_p^2 = .54$) and SOA ($F(4, 44) = 13.0$; $p < .001$; $\eta_p^2 = .54$). The main effect of order indicates that JNDs are smaller (better performance) for visual-audio intervals compared to audio-visual. The main effect of SOA indicates that JNDs increase with interval duration. For audiotactile intervals, there is only a significant effect of SOA indicating that the JND increases with interval duration ($F(4, 44) = 24.6$; $p < .001$; $\eta_p^2 = .69$; any other effects $ps > .64$). For visualtactile intervals we find a significant effect of order ($F(1, 11) = 5.6$; $p = .037$; $\eta_p^2 = .34$) indicating smaller JNDs for visual-tactile intervals in comparison to tactile-visual intervals. As for the other modality combinations, JNDs increase with SOA ($F(4, 44) = 13.4$; $p < .001$; $\eta_p^2 = .55$). There were no interactions of order and SOA for any of the modality combinations ($ps > .11$).

3.3. Weber fraction

The Weber Fractions obtained by dividing JND by PSE values are plotted in Fig. 1, right column. For audiovisual intervals, there is a main effect of SOA ($F(4, 44) = 6.97$; $p < .001$, $\eta_p^2 = .39$). Neither the main effect of order nor the interaction of order and SOA are significant ($ps > .40$). No significant effects are found for audiotactile and visualtactile intervals (all $ps > .07$, Greenhouse-Geisser corrected).

4. Linear fit

Our psychophysical results indicate that perceived duration of empty intervals depends on the modalities of the markers. For audiovisual and audiotactile intervals we find that perceived duration is longer when the onset of the duration is marked by the auditory signal (i.e., audio-visual and audio-tactile intervals are perceive longer than visual-audio and tactile-audio, respectively). For visualtactile signals there is no difference in perceived duration depending on the order of the markers. The aim of the present study was to infer whether a single or multiple time-keeping mechanisms enable duration judgments by testing for summative or multiplicative distortions of perceived duration. Interactions of SOA and order of the modalities indicate multiplicative distortions of perceived duration. In the ANOVAs, however, there were no interactions of SOA and order of the modalities for any of the modality combinations.

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009

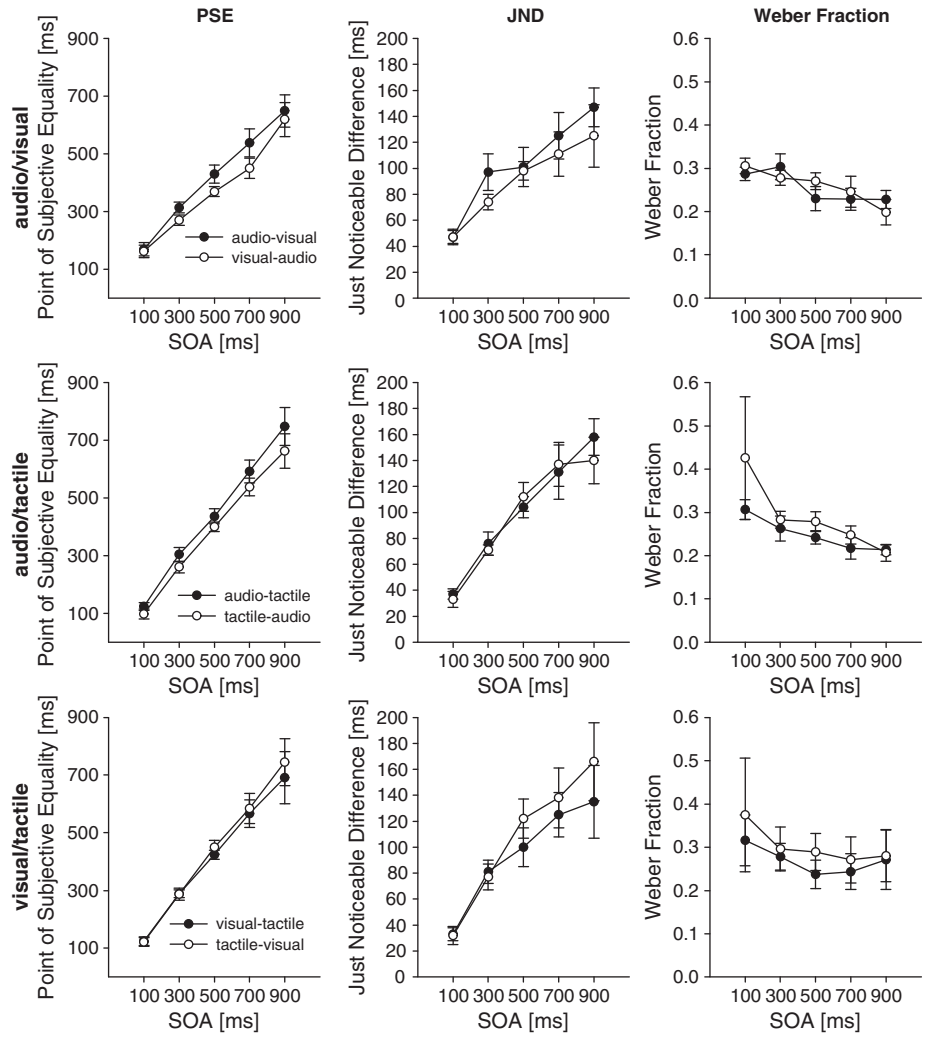

Fig. 1. Left column: Points of Subjective Equality (PSEs) defined as the auditory "probe" interval that matches the crossmodally-defined "standard" empty interval as a function of the duration of the standard (stimulus onset asynchrony, SOA, between the two signals defining the standard). The rows represent the three modality combinations of the standard empty interval. Middle column: Just Noticeable Differences (JNDs) with respect to the duration of the auditory probe interval that leads to a performance change from 50% "shorter" responses to 75% "shorter" responses in a two-interval discrimination judgment. Right column: Weber Fractions as a function of SOA and order of the modalities.

To ensure that the missing interactions are not due to lack of statistical power we fitted a linear model to the observed data.

To test whether the results found behaviourally indicate a summative and/or a multiplicative pattern of temporal distortion, we calculate the contribution of each of the three modalities as follows. If we consider perceived duration of an interval marked by two stimuli x and y as being distorted *only* by the perceptual latency (i.e., summative distortion of time), the perceived duration of the interval xy can be expressed as

$$P_{xy} = SOA - SL_x + SL_y, \tag{1}$$

in which case SL_x and SL_y denote the time required from the onset of the stimulus to the moment at which it is perceived (perceptual latency). Analogously, the duration of the interval with stimuli in reverse order (yx) would be

$$P_{yx} = SOA - SL_y + SL_x. \tag{2}$$

Since we are interested in the *difference* between latencies of signals in different sensory modalities, we can estimate the relative latency and simplify this formula further. To do this, we define the latency of a signal in modality i to be $SL_i = BL + L_i$ where BL represents the base latency across modalities defined as $BL = (SL_A + SL_T + SL_V)/3$ and where L_i is the relative latency from this value and as such it can be greater or less than zero. Therefore, Eqs. (1) and (2) can be expressed as:

$$P_{xy} = SOA - L_x + L_y \text{ and } P_{yx} = SOA - L_y + L_x. \tag{3}$$

The effect of latency L on perceived duration P is independent of the actual duration of the interval as its influence is only present at the beginning and at the end of the interval.

To model multiplicative distortion (i.e., distortion due to the pacemaker frequency rather than the switch) we have to consider the change in perceived duration across the SOAs. This can be done by modifying Eq. (3):

$$P_{xy} = SC_x SOA - L_x + L_y \text{ and } P_{yx} = SC_y SOA - L_y + L_x. \tag{4}$$

SC_x and SC_y represent the multiplicative factors by which the perceived duration P depends on the physical duration SOA. The indexes of SC (x and y) are used because we assume that the distortion at the perceptual level depends only on the type of stimulus marking the beginning (onset) of the interval because the order of the two modalities defining the empty interval is still unknown at the beginning of a trial. The alternative possibility (multiplicative distortions due to the offset marker) has also been tested but discarded, as it does not lead to a

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009

unified pattern of distortions across participants. Thus, if there were multiple time-keeping mechanisms, it must be the mechanism of the onset signal that is used for duration judgments. Deviations of perceived duration from veridical duration are indicated by values of $SC \neq 1$. If $SC > 1$ the perceived duration of the interval will increase at a higher rate than the physical duration. If, in contrast, $SC < 1$ perceived duration is going to be shorter than veridical as simulated duration increases. In Eq. (4), we can also use the relative term $C = SC\text{-}1$ instead of SC to express the multiplicative distortion explicitly. Here we make the assumption that the common denominator in clock speed is veridical, but given our formulation of the problem such assumption does not need to be met. Eq. (4) then becomes:

$$P_{xy} = (1 + C_x)SOA - L_x + L_y \text{ and } P_{yx} = \left(1 + C_y\right)SOA - L_y + L_x. \quad (5)$$

Values of C larger than zero indicate an overestimation of duration. If we compare perceived duration of empty intervals across the two orders of crossmodal markers, we can compute the difference of the values obtained in Eq. (5) for both orders and obtain:

$$P_{xy,SOA} - P_{yx,SOA} = (1 + C_x)SOA - L_x + L_y - \left[\left(1 + C_y\right)SOA - L_y + L_x\right]. \quad (6)$$

Simplifying, Eq. (6) becomes:

$$P_{xy,SOA} - P_{yx,SOA} = \left(C_x - C_y\right)SOA - 2L_x + 2L_y.$$

This formula can now be used to fit a straight line described by

$$DPD_{xyyx,SOA} = mSOA + q \quad (7)$$

with DPD being the *difference in perceived duration* at each SOA obtained from each participant conceiving multiplicative and summative distortions (in case of summative distortions: $m = 0$). In this way, it is possible to obtain an estimate of the intercept q that corresponds to $q = 2^*(L_y\text{-}L_x)$ and, in addition, the slope m that corresponds to $m = (C_x\text{-}C_y)$.

Using the values of q and m obtained for each of the three modality pairs based on measured data we can write a system of equations that can be solved to find the three values of L for the summative model (one common C value) and successively the values of C for the multiplicative model as described in Appendix A. Results are shown in Fig. 2 and the means and 95 percentiles are also reported in Table 1. Regarding the values of L, Table 1 suggests that L of the auditory modality is not included in the 95 percentiles of the distribution of visual and tactile modality Ls. Therefore, the model indicates that the latency of the auditory modality is shorter than the latency of the other two modalities (see also Fig. 2). Regarding the values of C for all three modalities, the 95 percentiles include zero (as well as the mean value of the C for the other two modalities) and therefore do not provide a significant evidence for differences in the scaling of the SOA. Thus, the values of C do not support the concept of multiplicative distortion of time.

In summary, the lack of interactions between SOA and order of the modalities in the PSEs, the lack of linear trends for the differences between the two modality orders in the PSEs and the lack of significant differences among the C values are at odds with the predictions for multiplicative distortions of time.

5. Discussion

Our study investigated whether a modality-independent or multiple modality-specific time-keeping mechanisms underlie duration judgments of crossmodally-defined intervals. We used all combinations of auditory, visual, and tactile stimuli to define empty intervals. Across all modality pairs, we find a trend to overestimate short SOAs and to underestimate long SOAs which is consistent with the range effect reported by von Vierordt (1868). Furthermore, consistent with previous studies (Grondin & Rousseau, 1991; Zampini, Brown, et al., 2005; Zampini, Guest, et al., 2005), we find that intervals with an auditory onset are perceived to be longer than intervals with an auditory offset. For visualtactile intervals there is no difference in perceived duration depending on whether the visual or the tactile signal marks the onset of the interval. Additionally, we fitted a linear regression model to the PSEs measured across interval durations to test whether distortions in perceived duration are summative (i.e., independent of SOA) or multiplicative (i.e., increasing with SOA). Our finding is consistent with summative distortion of perceived duration, implying that the underlying mechanism of perceived duration is based on a single modality-independent clock.

In general, the psychophysical results of the present study are in line with the findings of previous work (Grondin & Rousseau, 1991; Zampini, Brown, et al., 2005; Zampini, Guest, et al., 2005). When testing for ordering effects of the modalities defining an interval, we find that intervals with an auditory onset are perceived to be longer than intervals with an auditory offset (i.e., audio-visual intervals are perceived longer than visual-audio and audio-tactile intervals longer than tactile-audio). This result is consistent with the shorter perceptual latency of auditory signals compared to visual and tactile signals (Grondin & Rousseau, 1991). The magnitude of the stimulus order effect is comparable with the magnitude of the effects reported in previous behavioural studies (Grondin & Rousseau, 1991; Zampini, Brown, et al., 2005; Zampini, Guest, et al., 2005), human neuroimaging findings (Allison et al., 1977), and neurophysiological research (King & Palmer, 1985).

No differences between the ordering of modalities are found for visualtactile intervals suggesting that there are similar perceptual latencies for light flashes and tactile stimuli. This finding deviates from the results of Poliakoff et al. (2006), which show that for a

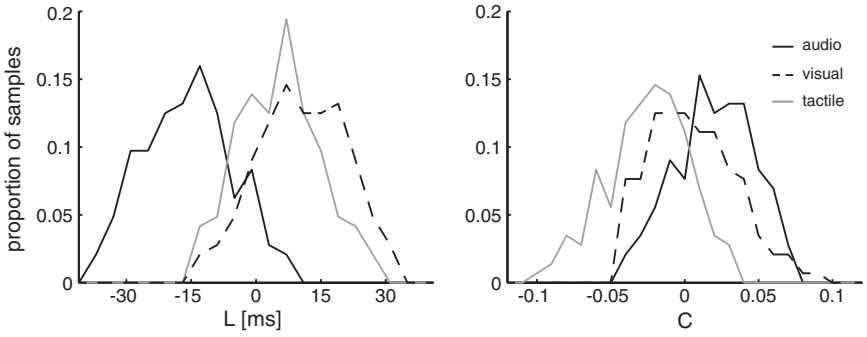

Fig. 2. Results of the bootstrapping procedure (see Appendix A). Plotted are the distribution of latency L (left panel) and multiplicative distortion C (right panel).

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009

Table 1
Values of the latency L and the multiplicative distortion C calculated from bootstrapping Eq. (7) (see Appendix A).

	$L_{audio\ [ms]}$	$L_{tactile\ [ms]}$	$L_{visual\ [ms]}$	$C_{auditory\ [\ ]}$	$C_{tactile\ [\ ]}$	$C_{visual\ [\ ]}$
CI low	−33.8	−10.8	−12.9	−0.0866	−0.1373	−0.1269
CI high	6.5	24.2	26.7	0.1493	0.0770	0.1174
Mean	−13.8	6.5	7.3	0.0292	−0.0247	−0.0045

Note: CI: 95% confidence interval.

population of young adults a visual stimulus had to be presented 40 ms before the tactile stimulus for the signals to be perceived as occurring simultaneously. Poliakoff et al. reported that the different perceptual latencies they found for visual and tactile signals are not constant across lifetime but vary with age. Our group of participants was similar in age to the sample of young adults in Poliakoff et al. and age differences therefore cannot account for the diverging results found in the two studies. The age dependency in the study of Poliakoff et al., however, suggests that the simultaneity percepts of visual and tactile signals can depend on factors other than sensory latencies per se. Possibly, temporal judgments of visual and tactile signals are also sensitive to the specific devices used to present the stimuli. In the study of Poliakoff et al., for example, tactile stimuli were presented to both thumb and index finger of both hands whereas in our study only the left index finger was stimulated. It might therefore be that the device used to present the stimuli caused the diverging results between the study of Poliakoff et al. and the current study.

Concerning the pattern of JNDs and Weber Fraction values, we found that the JNDs strongly depend on the perceived duration of the interval, rather than on the physical duration. The effect of order of the modalities on the JNDs for audiovisual and visualtactile intervals disappears when Weber Fractions are considered by dividing JNDs by PSEs. This suggests that the ordering effect on the JNDs is due to the summative distortions of perceived duration (as captured by the PSE) due to the modalities defining the interval. This result limits the possibility of a direct modality influence on sensitivity as measured by JND alone. Only for audiovisual intervals the Weber Fraction decreases with interval duration, suggesting that audiovisual duration judgments might be inconsistent with Weber's Law.

Because in our experiment we asked participants to compare crossmodally-defined empty intervals to filled auditory intervals, we have been able to limit the influence of possible biases such as temporal shrinking (Nakajima et al., 1991 cited in Nakajima et al., 2004; Sasaki et al., 2002), temporal stretching (Sasaki et al., 2010), and temporal ventriloquism effects (Morein-Zamir, 2003). The psychophysical findings on duration perception of crossmodally-defined intervals in our study are overall consistent with previous work (i.e., Grondin & Rousseau, 1991). Together with previous research our results confirm that the effects of sensory latencies on perceived duration cannot be explained by response biases caused by the experimental paradigm.

To investigate whether a modality-independent or several modality-specific time-keeping mechanisms enabled duration judgments in our experiment, we modeled PSE values for all modality combinations and durations. The fit of our linear regression model is consistent with the predictions made by a single time-keeping mechanism. Lack of multiplicative distortions could either be due to the presence of a single mechanism enabling duration judgments, or to the presence of multiple mechanisms with similar pacemaker frequencies. It is frequently suggested that stimulus intensity (among other factors) can modify the pacemaker frequency (e.g., Treisman, 1963), thus the latter possibility could be simply due to the particular stimulus level chosen for the auditory, visual, and tactile stimuli used in our study. This explanation is unlikely as if several modality-dependent mechanisms with similar pulse frequencies had been involved in the duration judgments, then no differences in the JNDs would have occurred. Note that JNDs are inversely related to pulse frequency (the higher the pulse frequency the lower the JND). The pattern of the PSEs, the JNDs, and the modeling results are consistent with the concept of a single time-keeping mechanism.

This conclusion does not imply that the human time-keeping system *as a whole* consists of a single time-keeping mechanism. It is well established that multiple clocks underlie time keeping at different time scales (see Buhusi & Meck, 2009, for example) and that even for sub-second intervals not all time perception phenomena can be explained by a single internal clock (Gamache & Grondin, 2010; Ulrich et al., 2006). The current study, however, shows that for empty crossmodal intervals in the milliseconds-to-second range that are compared to a filled interval, the human time-keeping system relies on a mechanism leading to consistent distortions across sensory modalities. This could be a pacemaker in a pacemaker–accumulator mechanism whose frequency does not depend on the modality marking the onset of the empty interval.

Other studies investigating duration judgments with multisensory stimuli reported results that imply multiple modality-specific timing mechanisms when processing sub-second durations (Gamache & Grondin, 2010; Grondin & Rousseau, 1991). The diverging results of previous work and the present study might originate from the different paradigms used. Crucially, the current study used crossmodal intervals (one interval defined by two modalities) while Gamache et al., for example, used bimodal intervals (one auditory and one visual interval that were presented simultaneously). In our paradigm, participants were forced to coordinate timing across input signals of different modalities and types of stimuli (filled vs. empty). Transferal processes of timing information between modalities are perceptually difficult (Lapid, Ulrich, & Rammsayer, 2009) and the most efficient way to solve the task in our study might be by using a single (modality-independent) mechanism. For the task of Gamache et al., in contrast, the most efficient strategy would be using two independent mechanisms that process the signals in parallel. One way of reconciling the diverging findings of whether timing mechanisms are modality-specific or modality-independent would be to assume the existence of both modality-specific and modality-independent mechanisms. Under this assumption, specific task demands would be the critical factor that decides whether a modality-specific or a modality-independent mechanism is selected for judging durations. Future research is necessary to determine whether such task demands can switch the distortion pattern found in perceived duration studies.

6. Conclusion

Our study shows that crossmodally-defined empty intervals induce a pattern of distortion in perceived duration that deviates from the physical interval duration. Intervals with an auditory onset are perceived to be longer than intervals with an auditory offset. No ordering effect is found for visualtactile intervals. Our results suggest a pattern of summative distortions in perceived duration, without significant multiplicative distortions. This is consistent with a single modality-independent time-keeping mechanism that enables duration judgments across all modality combinations.

Acknowledgments

We would like to thank Prof. Rolf Ulrich for his invaluable comments on the project and Tonja K Machulla for inspiring suggestions.

Appendix A

First, we test for differences in latencies between auditory, visual and tactile signals to test for summative distortions of perceived duration. We can formulate Eq. (7) (Section 4) in matrix form and set a single pacemaker frequency by forcing the m term to be 0. This allows us to

create a system of equations for the summative model in which each of the three types of empty intervals in the experiment corresponds to one row. In this way we obtain:

$$\begin{bmatrix} DPD_{\text{audio}-\text{visual}/\text{visual}-\text{audio}} \\ DPD_{\text{audio}-\text{tactile}/\text{tactile}-\text{audio}} \\ DPD_{\text{visual}-\text{tactile}/\text{tactile}-\text{visual}} \end{bmatrix} = 2 \cdot \begin{bmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} L_A \\ L_T \\ L_V \end{bmatrix} \tag{A.1}$$

or

$$d = 2\ M\ l.$$

Eq. A.1 gives the possibility to estimate the latencies of each sensory signal contained in l by solving the system of equations and substituting the vector d with the values of the *DPDs* averaged across the all the tested SOAs. The solution requires computing the inverse of M and expressing the system of equations as $l = d/M/2$. Since there is no general solution to the system of equations, it is not possible to find the inverse of M (it is a singular matrix). However it is possible to compute the Moore-Penrose pseudoinverse of M, approximating the solution to

$$pinv(M) = \begin{bmatrix} -0.33 & -0.33 & 0.00 \\ 0.00 & 0.33 & 0.33 \\ 0.33 & 0.00 & -0.33 \end{bmatrix}. \tag{A.2}$$

In this way the value of the vector l can be found approximately by using the following equation:

$$l = d \cdot pinv(M)/2. \tag{A.3}$$

If we substitute d with the average values of *DPD* (see Eq. (7)) obtained across observers with $d = [46, 49, -17]'$ we obtain the values $L_A = -14$ ms, $L_T = 7$ ms, and $L_V = 7$ ms (Table 1). These values give an estimate of the difference in sensory latency from the average base latency *BL* (see Eq. (2)) with the auditory channel having the shortest latency.

To obtain an estimate of the variance of the latency parameters, we applied Eq. (7) to all possible combinations of the signals defining the intervals obtained from each participant (the vector d contained the average *DPDs* for each of the 1728 combinations across 12 participants). The distribution of values indicates that auditory latency is significantly shorter than both visual and tactile latency (2.5–97.5 percentiles of the audio latency do not include the mean values of the distributions of visual and tactile latency values) and that tactile latency is not different from visual latency.

Second, we tested for multiplicative distortions (i.e. differences between the modality-specific pacemaker frequencies). Multiplicative distortions would support the existence of multiple time-keeping mechanisms. In a similar way as done with Eq. A.1, we want to express Eq. (7) using matrices. This time, however, we only consider the steepness of the regression lines fitted to the *DPDs* across SOAs because the steepness indicates the pacemaker frequency of each modality. After accounting for differences in latencies and rearranging the terms, Eq. (7) becomes:

$$DPD_{xyyx,SOA} - \left(-2L_x + 2L_y\right) = \left(C_x - C_y\right)SOA \tag{A.4}$$

Note that by taking the difference between the two terms of C we discount any common multiplicative distortion across modalities x and y. Using Eq. A.4 it is possible to subtract the difference in latency (see above) from the *DPDs*. By performing an individual fit with a linear regression line passing through the origin it is possible to obtain an estimate of the slope m that corresponds to $m = (C_x - C_y)$. From the values of m obtained for each of the three pairs of stimuli with all combinations of subject data we can write a system of equations as done before.

$$\begin{bmatrix} m_{\text{audio}-\text{visual}/\text{visual}-\text{audio}} \\ m_{\text{audio}-\text{tactile}/\text{tactile}-\text{audio}} \\ m_{\text{visual}-\text{tactile}/\text{tactile}-\text{visual}} \end{bmatrix} = -M \cdot \begin{bmatrix} C_A \\ C_T \\ C_V \end{bmatrix}. \tag{A.5}$$

Again, we solve the system of equations using the pseudoinverse of M by applying the following equation:

$$c = -m \cdot pinv(M) \tag{A.6}$$

Values based on averages across observers are $C_A = 0.0292$, $C_T = -0.0247$, and $C_V = -0.0045$ (see Table 1). Because the values of C are obtained from the differential terms m, they give an estimate of the deviation of each modality's pacemaker frequency from a base pacemaker frequency across modalities that discounts all common multiplicative distortions. Our findings indicate that for intervals with auditory onset the perceived duration increments less than veridically. Intervals with visual or tactile onsets instead are perceived to be incrementing more in duration than veridically. Critically, the 95 percentile range of each distribution includes both the 0 value and the means of the other two distributions. This suggests that pacemaker frequencies do not differ across modalities (for the values and their bootstrapped distributions see Fig. 2 and Table 1).

References

Allan, L. G., & Kristoff, A.B.. (1974). Psychophysical theories of duration discrimination. *Perception & Psychophysics*, *16*(1), 26–34. http://dx.doi.org/10.3758/Bf03203244.

Allison, T., Matsumiya, Y., Goff, G. D., & Goff, W. R. (1977). The scalp topogrphy of human visual evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *42*(2), 185–197.

Buhusi, C. V., & Meck, W. H. (2009). Relativity theory and time perception: single or multiple clocks? *PLoS One*, *4*(7), e6268. http://dx.doi.org/10.1371/journal.pone.0006268.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*(1–2), 55.

Di Luca, M., Machulla, T., & Ernst, M.O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency. *Journal of Vision*, *9*(12:7), 1–16.

Gamache, P. L., & Grondin, S. (2010). Sensory-specific clock components and memory mechanisms: Investigation with parallel timing. *European Journal of Neuroscience*, *31*(10), 1908–1914. http://dx.doi.org/10.1111/j.1460-9568.2010.07197.x.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582. http://dx.doi.org/10.3758/App.72.3.561.

Grondin, Simon, Ivry, Richard B., Franz, E., Perreault, L., & Metthè, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics*, *58*(3), 424–433.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics*, *49*(3), 245–256.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*(7), 273–280. http://dx.doi.org/10.1016/j.tics.2008.04.002.

Killeen, P. R., & Taylor, T. J. (2000). How the propagation of error through stochastic counters affects time discrimination and other psychophysical judgments. *Psychological Review*, *107*(3), 430–459. http://dx.doi.org/10.1037/0033-295x.107.3.430.

King, A. J., & Palmer, A.R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*(3), 492–500.

Lapid, E., Ulrich, R., & Rammsayer, T. (2009). Perceptual learning in auditory temporal discrimination: No evidence for a cross-modal transfer to the visual modality. *Psychonomic Bulletin & Review*, *16*(2), 382–389. http://dx.doi.org/10.3758/Pbr.16.2.382.

Morein-Zamir, S. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154–163. http://dx.doi.org/10.1016/S0926-6410(03)00089-2.

Nakajima, Y., ten Hoopen, G., Sasaki, T., Yamamoto, K., Kadota, M., Simons, M., et al. (2004). Time-shrinking: The process of unilateral temporal assimilation. *Perception*, *33*(9), 1061–1079. http://dx.doi.org/10.1068/p5061.

Nakajima, Y., ten Hoopen, G., & van der Wilk, R. (1991). A new illusion of time perception. *Music Perception*, *8*, 431–448.

Poliakoff, E., Shore, D. I., Lowe, C., & Spence, C. (2006). Visualtactile temporal order judgments in ageing. *Neuroscience Letters*, *396*(3), 207–211.

Rousseau, R., Poirier, J., & Lemyre, L. (1983). Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics*, *34*(6), 541–548.

Sasaki, T., Nakajima, Y., ten Hoopen, G., van Buuringen, E., Massier, B., Kojo, T., et al. (2010). Time stretching: Illusory lengthening of filled auditory durations. *Attention, Perception, & Psychophysics*, *72*(5), 1404–1421. http://dx.doi.org/10.3758/APP.72.5.1404.

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009

ARTICLE IN PRESS

Sasaki, T., Suetomi, D., Nakajima, Y., & ten Hoopen, G. (2002). Time-shrinking, its propagation, and Gestalt principles. *Perception & Psychophysics*, *64*, 919–931. http://dx.doi.org/10.3758/Bf03196796.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the 'internal clock'. *Psychological Monographs*, *77*(1–31).

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Crossmodal temporal discrimination: Assessing the predictions of a general pacemakercounter model. *Perception & Psychophysics*, *68*(7), 1140–1152.

von Vierordt, K. (1868). *Der Zeitsinn: nach Versuchen*. Tübingen: Laupp.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights". *The Quarterly Journal of Experimental Psychology. B*, *51*, 25.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313.

Zampini, M., Brown, T., Shore, D., Maravita, A., Röder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica*, *118*, 277–291.

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio–visual simultaneity judgments. *Perception & Psychophysics*, *67*(3), 531.

Please cite this article as: Mayer, K.M., et al., Duration perception in crossmodally-defined intervals, *Acta Psychologica* (2013), http://dx.doi.org/10.1016/j.actpsy.2013.07.009