Acta Psychologica 147 (2014) 2–9

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

A R T I C L E I N F O

Article history:
Received 8 January 2013
Received in revised form 29 May 2013
Accepted 10 July 2013
Available online 15 August 2013

PsycINFO classification:
2320 Sensory Perception
2221 Sensory & Motor Testing

Keywords:
Crossmodal intervals
Perceived duration
Sensory latency
Linear regression model

A B S T R A C T

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.

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 × 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 (± 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