Please cite this article in press as: Binder M. Neural correlates of audiovisual temporal processing – Comparison of temporal order and simultaneity judgments. Neuroscience (2015), http://dx.doi.org/10.1016/j.neuroscience.2015.05.011

Neuroscience xxx (2015) xxx–xxx

NEURAL CORRELATES OF AUDIOVISUAL TEMPORAL PROCESSING – COMPARISON OF TEMPORAL ORDER AND SIMULTANEITY JUDGMENTS

M. BINDER *

Psychophysiology Laboratory, Institute of Psychology, Jagiellonian University, Kraków, Poland

Abstract—Multisensory integration is one of the essential features of perception. Though the processing of spatial information is an important clue to understand its mechanisms, a complete knowledge cannot be achieved without taking into account the processing of temporal information. Simultaneity judgments (SJs) and temporal order judgments (TOJs) are the two most widely used procedures for explicit estimation of temporal relations between sensory stimuli. Behavioral studies suggest that both tasks recruit different sets of cognitive operations. On the other hand, empirical evidence related to their neuronal underpinnings is still scarce, especially with regard to multisensory stimulation. The aim of the current fMRI study was to explore neural correlates of both tasks using paradigm with audiovisual stimuli. Fifteen subjects performed TOJ and SJ tasks grouped in 18-second blocks. Subjects were asked to estimate onset synchrony or temporal order of onsets of non-semantic auditory and visual stimuli. Common areas of activation elicited by both tasks were found the bilateral fronto-parietal network, including regions whose activity can be also observed in tasks involving spatial selective attention. This can be regarded as an evidence for the hypothesis that tasks involving selection based on temporal information engage the similar regions as the attentional tasks based on spatial information. The direct contrast between the SJ task and the TOJ task did not reveal any regions showing stronger activity for SJ task than in TOJ task. The reverse contrast revealed a number of left hemisphere regions which were more active during the TOJ task than the SJ task. They were found in the prefrontal cortex, the parietal lobules (superior and inferior) and in the occipito-temporal regions. These results suggest that the TOJ task requires recruitment of additional cognitive operations in comparison to SJ task. They are probably associated with forming representations of stimuli as separate and temporally ordered sensory events.

Key words: fMRI, temporal order judgment, simultaneity judgment, temporal processing, attention.

INTRODUCTION

Understanding perceptual processing cannot be achieved without considering brain mechanisms of multisensory integration. An awake organism constantly receives a plethora of sensory signals coming from separate modalities, delivering information about different aspects of its environment. For organism to behave in an adaptive way this variety must be transformed into a consistent and yet dynamic representation of the surrounding world. Obviously spatial distribution of the sources of stimulation is an important cue for the successful integration (Spence et al., 2003; Zampini et al., 2003a, 2005), but another crucial factor is the temporal relation between multisensory events. The majority of the effects of multisensory integration involve temporal coincidence of its components (Keetels and Vroomen, 2012). However, these effects are not only constrained to the cases when there is an objective, physical coincidence of two (or more) stimuli from separate sensory channels. There is compelling evidence for a conjecture that multisensory integration should not be viewed as an effect of passive coincidence detection of signals arriving by separate sensory channels. It is a well-known fact that subjects perceive as simultaneous the pairs of stimuli that are not physically synchronous (Stevenson and Wallace, 2013). The hypothesis of 'temporal window of integration' provides a conceptual account of this phenomenon (van Wassenhove et al., 2007; Lewkowicz and Ghazanfar, 2009; Vroomen and Keetels, 2010; Colonius and Diederich, 2012). This notion denotes the temporal interval between multisensory stimuli during which multisensory integration may only occur. In many cases the brain can dynamically adjust the perceived temporal relations between stimuli arriving at different times, for example using the mechanism of temporal perceptual recalibration (Fujisaki et al., 2004; Vroomen et al., 2004).

Temporal order judgments (TOJs) and simultaneity judgments (SJs) are the two most widely used paradigms for the assessment of temporal perception,

*Address: Psychophysiology Laboratory, Institute of Psychology, Jagiellonian University, ul. Ingardena 6, 30-060 Kraków, Poland. Tel: +48-12-663-24-37; fax: +48-12-663-24-15.
E-mail address: marek.binder@uj.edu.pl

Abbreviations: FEF, frontal eye field; FLAME, FMRIB's local analysis of mixed effects; JND, just noticeable difference; PSS, point of subjective simultaneity; SC, superior colliculus; SJs, simultaneity judgments; SMA, supplementary motor area; SOA, stimulus onset asynchronies; STG, superior temporal gyrus; STS, superior temporal sulcus; TE, threshold estimation; TOJs, temporal order judgments; TPJ, temporo-parietal junction.

http://dx.doi.org/10.1016/j.neuroscience.2015.05.011

also in the field of multisensory integration (Keetels and Vroomen, 2012). Two parameters are usually derived from these measures. The point of subjective simultaneity (PSS) parameter provides an estimate of the interval between stimuli at which there is the highest probability of the perception of simultaneity. The 'just noticeable difference' (JND) variable reflects the subject's sensitivity to changes in intervals between the stimuli. The JND value (in milliseconds) denotes the minimal temporal interval at which the change between the perceived temporal relation stimuli can be observed.

During the TOJ procedure subjects are presented with pairs of stimuli with variable stimulus onset asynchronies (SOA), and after each presentation they are asked to make an explicit judgment about which of them was the first. In case of audiovisual pairs the subject has to select from two alternatives: 'sound-first' or 'flash-first'. The obtained psychometric function has a characteristic sigmoid profile and it is usually modeled by a cumulative Gaussian or logistic function (Keetels and Vroomen, 2012). The PSS value for TOJ task is taken at the cross-over point of the psychometric function, when there is an equal probability of 'sound-first' and 'flash-first' judgments, and subjects are maximally unsure about the temporal relation between the members of the audiovisual pair. The measure of sensitivity, JND, is calculated as a half of SOA difference between 25% and 75% points of the psychometric function. An alternative estimate of sensitivity is the psychometric function slope coefficient at the PSS value.

During the SJ procedure subjects are presented with the same kind of stimuli as during the TOJ procedure, but this time they are asked to judge whether the stimuli were perceived as simultaneous or not. In this case the psychometric function is usually modeled by the Gaussian function. There is however one important observation related to audiovisual stimuli: the resulting psychometric function may be asymmetric, being steeper for pairs with the leading auditory stimulus and shallower for pairs with the leading visual stimulus. This phenomenon suggests that in both cases subjects display different sensitivity to the temporal structure of stimulation (van Eijk et al., 2008; Alcalá-Quintana and García-Pérez, 2013). The PSS estimate for SJs is taken from the point of the psychometric function with the maximum probability of 'synchronous' response and JND is calculated as a mean SOA for 75% point of the psychometric curve (both for 'sound-first' and 'flash-first' pairs).

According to most observations, the PSS values for audiovisual pairs observed during TOJ and SJ procedures are usually positive, i.e. do not correspond to the point of objective simultaneity (at SOA = 0 ms). The positive value means that both stimuli are perceived as simultaneous when the visual stimulus leads the auditory stimulus (usually by the order of tens of milliseconds). This is probably caused by the different sensitivity of auditory and visual systems to temporal cues (such as temporal dynamics of intensity changes). So far this phenomenon has not been a subject of extensive research (but see van Eijk et al., 2010; Stevenson and Wallace, 2013).

Though both procedures are used to investigate processes of temporal integration, they often give inconsistent results. Estimates of the PSS values obtained with TOJ and SJ do not correlate with each other (reviewed by van Eijk et al., 2008). Moreover, as for the audiovisual SJ judgments the PSS values are usually positive, for the TOJ judgments the negative SOA values are reported in some studies (so stimuli are perceived as simultaneous when auditory stimulus leads visual stimulus). Van Eijk et al. (2008) directly compared PSS estimates for two types of audiovisual stimuli (flash-click pairs and bouncing ball with an impact sound) and three types of procedures: two-alternative SJ ('synchronous', 'asynchronous'), three-alternative SJ ('sound-first', 'synchronous', and 'flash-first'), and TOJ ('sound-first', 'flash-first'). PSS values for both SJ tasks were indeed correlated, but the authors did not observe any correlation between TOJ and any of the SJ tasks. More recently, a similar result was obtained by Love et al. (2013) in the study involving five types of audiovisual pairs. As in Van Eijk et al. (2008), Love et al. also observed negative PSS values for the TOJ tasks and consistent positive PSS values for the SJ tasks.

This result suggests that there could be essential differences in the composition of cognitive processes engaged in both tasks. According to Hirsh and Sherrick (1961) and Jaśkowski (1991), perceiving the temporal asynchrony is a necessary, though insufficient, condition for achieving an accurate judgment of temporal order. They suggest a two-stage architecture for temporal judgments. For example, Jaśkowski (1991) proposed a two-stage model consisting of two separate processing centers. The first stage, labeled 'the simultaneity center', works as a 'moment-gating' mechanism. Depending on the relative signal delays and the applied threshold it can generate two possible 'perceptual states': synchronous or asynchronous. On the second stage, 'the order center' decides on the temporal order of the stimuli, taking into account their relative latency differences and the perceptual state of the simultaneity center. In effect it can generate three possible states (for a pair consisting of A and B stimuli): 'order AB', 'order BA' or a 'uncertainty' – in this latter case the emitted response is random. Thus this model allows an outcome where stimuli are perceived as non-simultaneous but an adequate decision concerning their order cannot be made.

However, other authors (e.g. Sternberg and Knoll, 1973; Allan, 1975; García-Pérez and Alcalá-Quintana, 2012 for review) maintain that perception of asynchrony is both the necessary and the sufficient condition for an adequate TOJ. This is achieved by a ternary decision system operating on a ternary decision rule applied to the arrival-time difference between the two signals. Thus the dedicated decision system may generate three types of responses: 'order AB', 'order BA', 'synchronous'.

More recently, Zampini et al. (2003a) and Shore et al. (2005) emphasized the different character of both tasks, while not proposing the specific theoretical accounts explaining those differences. For example, Zampini et al. (2003a) suggested that essentially the SJ task requires multisensory binding, while the TOJ task is