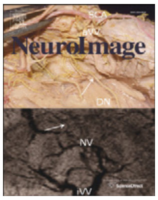
NeuroImage 101 (2014) 404–415

Contents lists available at ScienceDirect

NeuroImage

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



Interregional alpha-band synchrony supports temporal cross-modal integration

Joram van Driel [a,b,*], Tomas Knapen [c], Daniel M. van Es [c], Michael X Cohen [a]

[a] *Department of Psychology, University of Amsterdam, The Netherlands*
[b] *The Amsterdam Brain and Cognition Center, University of Amsterdam, The Netherlands*
[c] *Department of Cognitive Psychology, VU University Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Accepted 14 July 2014
Available online 17 July 2014

Keywords:
Cross-modal integration
Time perception
Alpha-band
Functional connectivity
Phase synchrony

ABSTRACT

In a continuously changing environment, time is a key property that tells us whether information from the different senses belongs together. Yet, little is known about how the brain integrates temporal information across sensory modalities. Using high-density EEG combined with a novel psychometric timing task in which human subjects evaluated durations of audiovisual stimuli, we show that the strength of alpha-band (8–12 Hz) phase synchrony between localizer-defined auditory and visual regions depended on cross-modal attention: during encoding of a constant 500 ms standard interval, audiovisual alpha synchrony decreased when subjects attended audition while ignoring vision, compared to when they attended both modalities. In addition, alpha connectivity during a variable target interval predicted the degree to which auditory stimulus duration biased time estimation while attending vision. This cross-modal interference effect was estimated using a hierarchical Bayesian model of a psychometric function that also provided an estimate of each individual's tendency to exhibit attention lapses. This lapse rate, in turn, was predicted by single-trial estimates of the stability of interregional alpha synchrony: when attending to both modalities, trials with greater stability in patterns of connectivity were characterized by reduced contamination by lapses. Together, these results provide new insights into a functional role of the coupling of alpha phase dynamics between sensory cortices in integrating cross-modal information over time.

Introduction

When the sound coming from a television gets out-of-sync with what is visually displayed, you immediately feel that something is wrong. Multimodal processing is ubiquitous in perception: different senses provide us with complementary evidence about external events, which can aid our responses to these events (McDonald et al., 2000; Yang et al., 2013), or can result in perceptual illusions (Alais and Burr, 2004; McGurk and MacDonald, 1976). Over the past several decades, neuroscience of multisensory processing has shifted from a strict hierarchical view of unisensory signals converging onto higher supramodal areas (Meredith and Stein, 1983; Stein and Stanford, 2008), to a growing consensus that cross-modal integration can occur even at early stages of sensory processing (Foxe et al., 2000; Ghazanfar and Chandrasekaran, 2007; Ghazanfar and Schroeder, 2006; Giard and Peronnet, 1999; Kayser and Logothetis, 2007; Martuzzi et al., 2007; Molholm et al., 2002). However, how these early-stage interactive processes are neurophysiologically organized remains a topic of active exploration (Klemen and Chambers, 2012; Sarko et al., 2013).

One proposed mechanism of neural interaction is "binding through coherence" (Fries, 2005; Varela et al., 2001; Ward, 2003; Womelsdorf et al., 2007), or the idea that effective windows of cortico-cortical communication may arise by transiently synchronized electrophysiological oscillations of the involved neural populations. Evidence is accumulating that this principle may apply to the integration of multisensory information as well (Doesburg et al., 2008; Hummel and Gerloff, 2005; Sarko et al., 2013; Senkowski et al., 2008; von Stein et al., 1999). For example, a stimulus of one modality can modulate the processing of a concurrently delivered stimulus of another modality, through the phase resetting of ongoing oscillatory activity in the corresponding primary sensory region (Diederich et al., 2012; Kayser et al., 2008; Lakatos et al., 2007). This results in increased cortical excitability, and thus improved behavioral performance towards the bimodal stimulus. Attention seems to determine which modality "controls" the phase-resetting (Lakatos et al., 2009). Given the tight link between alpha-band (8–12 Hz) activity and attentional processing (Jensen and Mazaheri, 2010; Klimesch et al., 2007), we hypothesized that during cross-modal attention, alpha phase synchrony may be an important mediator of large-scale communication between distant sensory regions (Hummel and Gerloff, 2005; Palva and Palva, 2011). Although several frequency bands have been implicated in multisensory integration (Senkowski et al., 2008), it has been proposed that phase

---

* Corresponding author at: Department of Psychology, The University of Amsterdam, 1018 XA Amsterdam, The Netherlands.
E-mail address: j.vandriel@uva.nl (J. van Driel).

http://dx.doi.org/10.1016/j.neuroimage.2014.07.022
1053-8119/

synchronization in the alpha-band may be especially important for coordinating functional integration between cortical regions (Doesburg et al., 2009) of relatively longer inter-areal distances (Palva and Palva, 2007; Palva et al., 2005). This active role of alpha activity has been shown in a variety of cognitive and perceptual tasks, such as spatial attention (Doesburg et al., 2009), working memory (Palva et al., 2005), object recognition (Bar et al., 2006), and error-processing (van Driel et al., 2012), and may thus represent a general mechanism of coherent network functioning. Importantly, increases in interregional alpha-band phase synchrony can co-occur with local decreases in alpha-band power (Palva and Palva, 2007, 2011), where the latter may reflect attention-related "active inhibition" of task-irrelevant areas (Jensen and Mazaheri, 2010). In this study, we were in particular interested in the role of interregional alpha phase dynamics during cross-sensory integration.

Most studies on multisensory processing use brief, momentary stimuli, and focus on the spatial domain (Driver and Spence, 1998, 2000; Macaluso and Driver, 2005), or on judgments of successiveness versus simultaneity (Jaekl and Harris, 2007; Keetels and Vroomen, 2007). However, multisensory events in a continuous environment are more likely to be prolonged (Ghazanfar and Chandrasekaran, 2007), and are not necessarily linked to one spatial location; in these more naturalistic situations, correlated temporal *durations* can provide key evidence for integration. Moreover, it is especially interesting to study cross-modal integration of elapsed time, because the perception of auditory duration is superior to that of visual duration. This is in contrast to the more frequently investigated spatial domain, in which visual spatial localization is superior to auditory spatial localization (Burr et al., 2009; Fendrich and Corballis, 2001; Pick et al., 1969).

The purpose of the present study was to investigate the potential neural mechanisms of multimodal integration via duration perception. We here report novel evidence that inter-regional phase synchrony in the alpha-band supports multimodal duration judgments in humans. Through time–frequency decomposition of high-density EEG activity, we found that alpha synchrony was modulated by cross-modal attention and correlated with subject-specific Bayesian estimated parameters of distractor interference and lapsing.

Materials and methods

Subjects

Nineteen subjects (age range 18–29, M = 22.4; 13 females) from the University of Amsterdam community participated in this study in exchange for €14 or course credits. All subjects signed an informed consent form before participation and reported to have normal or corrected-to-normal vision, and normal hearing. The study was approved by the local ethics committee; all procedures complied with relevant laws and institutional guidelines. Data of one subject were excluded from further analyses due to a programming error during data collection. Thus, the final dataset existed of data taken from eighteen subjects (12 females).

Bimodal duration discrimination task

Subjects performed a psychophysical duration discrimination task. In each trial they were presented with a "standard" 500 ms stimulus followed by a target stimulus whose duration was always shorter or longer than the standard (determined on a trial-to-trial basis by an adaptive staircase procedure; see below). After the target, they indicated with a button press whether they perceived its duration as shorter (left thumb) or longer (right thumb) than the duration of the standard. Both standard and target were composed of a concurrently presented tone (500 Hz sine wave with a 5 ms ramp-up/down envelope, played by speakers left and right from the screen) and a red LED (fixed at the center of a computer screen). During the standard, the auditory and visual stimuli had the same onset and offset times (i.e. always a perfectly simultaneous 500 ms combined audiovisual stimulus).

For the target intervals, stimuli had equal onset times. Offset times, as well as instruction on attention, were manipulated in three different conditions (see Fig. 1a). In the Audiovisual condition, the auditory and visual stimuli had the same offset time, and subjects were instructed to attend to both modalities, and to regard the two modalities as belonging to one coherent stimulus. In this condition, both the standard and target intervals thus consisted of a perfectly simultaneous auditory and visual stimulus. In two distractor conditions we manipulated the cross-modal offset of target intervals, thereby introducing "temporal conflict". That is, the target started simultaneously in both modalities, but the distractor modality ended earlier or later. We reasoned that the simultaneous onset initially triggers cross-sensory integration (as in the Audiovisual condition), but the unequal offset time results in a bias in responding to the duration of one or the other modality. Specifically, in the Attend Auditory condition, the visual stimulus had an earlier or later offset than the auditory stimulus, and subjects were instructed to attend to the auditory (target) and ignore the visual (distractor) modality. In the Attend Visual condition this was the other way around: subjects were instructed to attend to the visual target and ignore the auditory distractor stimulus that had an earlier or later offset time. For both Attend Auditory and Attend Visual conditions, distractor stimuli had a duration of 50% (short distractor) or 150% (long distractor) of the target interval, which was counterbalanced across short and long targets. Thus, for example, a target shorter than the standard could be accompanied by a distractor that was shorter or longer than this target. Similarly, a target longer than the standard could be accompanied by a distractor that was shorter or longer than this target. We hypothesized that, given the ubiquitous influence of cross-modal integration, this manipulation would bias the responses of the subject towards the duration of the distractor, which could be correct (e.g. responding "long" to a long target accompanied by a longer distractor), or incorrect (e.g. responding "short" to a long target accompanied by a shorter distractor).

Each trial started with a standard stimulus of 500 ms, followed by a 1000 ms inter-stimulus-interval, after which the target was presented. After target-offset, subjects were required to respond within 1500 ms. Trials ended upon responding, or when the 1500 ms response window had passed in which case feedback on response speed ("Respond faster!") was presented for 500 ms at the center of the screen. A response (or response-feedback) was followed by a 1500 ms inter-trial interval.

The duration of the target was titrated trial-by-trial by means of an adaptive staircase procedure. Within each block, two staircases of 36 trials were randomly interleaved: one with durations shorter than the standard (starting at 398 ms, with a minimum of 333 ms), and one with durations longer than the standard (starting at 654 ms, with a maximum of 750 ms). Titration followed a 2-up-1-down rule, such that after two consecutive correct responses to the same target, its duration approached the standard with a particular step size (0.15 times the difference between the target duration and 500 ms standard duration), and after one error, the target duration moved away from the standard with the same step size. With a 2-up-1-down staircase procedure, subjects will converge to ~71% accuracy level, i.e. around the just-noticeable difference (Leek, 2001). We took the proportional distance-to-standard step size approach to be able to account for Weber's law, or the scalar property of variation in interval timing (Buhusi and Meck, 2005; Gibbon, 1977; Grondin, 2010). This law predicts that, for example, a 500–750 ms difference is harder to perceive than a 500–250 ms difference. With our approach, step sizes were always greater for the long compared to the short staircases, and decreased proportionally as the staircase approached the standard duration. We chose to use a constant 500 ms standard to be able to use this titration procedure; this decision was based on extensive piloting. Further, we were in particular interested in the subsecond interval range, as this is thought to trigger automatic, rather than cognitively controlled (suprasecond)