Neuropsychologia 51 (2013) 358–371

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

Temporal event structure and timing in schizophrenia: Preserved binding in a longer ''now''

Brice Martin[d,e,f], Anne Giersch[g], Caroline Huron[a,b,c], Virginie van Wassenhove[a,b,c,*]

[a] INSERM, Cognitive Neuroimaging Unit, Bât 145 Point Courrier 156, 91191 Gif Yvette, France
[b] CEA, DSV/I2BM, NeuroSpin Center, F-91191 Gif/Yvette, France
[c] Univ Paris—Sud, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[d] Univ Paris—Diderot, F-75205 Paris, France
[e] Univ Paul Sabatier, 31062 Toulouse, France
[f] Centre de Réhabilitation, Service Hospitalo—Universitaire de Réhabilitation, CHS Le Vinatier, 98 rue Boileau, 69006 Lyon, France
[g] INSERM, U666, Centre Hospitalier Régional Universitaire de Strasbourg, Département de Psychiatrie, 67091 Strasbourg, France

ARTICLE INFO

Available online 16 July 2012

Keywords:
Multisensory integration
Audiovisual speech
Schizophrenia
Time perception
Fragmentation
Simultaneity
Psychophysics

ABSTRACT

Patients with schizophrenia experience a loss of temporal continuity or subjective fragmentation along the temporal dimension. Here, we develop the hypothesis that impaired temporal awareness results from a perturbed structuring of events in time—i.e., canonical neural dynamics. To address this, 26 patients and their matched controls took part in two psychophysical studies using desynchronized audiovisual speech. Two tasks were used and compared: first, an identification task testing for multisensory binding impairments in which participants reported what they heard while looking at a speaker's face; in a second task, we tested the perceived simultaneity of the same audiovisual speech stimuli. In both tasks, we used McGurk fusion and combination that are classic ecologically valid multisensory illusions. First, and contrary to previous reports, our results show that patients do not significantly differ from controls in their rate of illusory reports. Second, the illusory reports of patients in the identification task were more sensitive to audiovisual speech desynchronies than those of controls. Third, and surprisingly, patients considered audiovisual speech to be synchronized for longer delays than controls. As such, the temporal tolerance profile observed in a temporal judgement task was less of a predictor for sensory binding in schizophrenia than for that obtained in controls. We interpret our results as an impairment of temporal event structuring in schizophrenia which does not specifically affect sensory binding operations but rather, the explicit access to timing information associated here with audiovisual speech processing. Our findings are discussed in the context of curent neurophysiological frameworks for the binding and the structuring of sensory events in time.

1. Introduction

A core distinction in cognitive neurosciences is the dissociation between automatic processes and attention-driven processes that implicate higher-order operations such as ''top-down'' control (i.e., distinction between implicit or explicit processes, respectively). The set of automatic operations implicated in the temporal organization of information is called ''temporal event-structure'' (Zacks & Tversky, 2001) and necessitates the segmentation of temporal units of information (Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Functionally, these temporal units are time segments or temporal windows of various duration within which information is integrated in the brain (Theunissen & Miller, 1995; van Wassenhove, 2009; Wittmann, 2011). Neurophysiologically, temporal windows are the natural outcome of synaptic delays at the neuronal level or neural oscillations at the population level (for review see: Wang, 2010; Buzsáki, 2006, 2010). The automatic and implicit temporal segmentation thus provides the building blocks for more abstract levels of representations and has crucial implications for the qualitative and phenomenological aspect of conscious experience. However, it is unclear whether implicit and explicit temporal event structuring share similar functional properties or rely on entirely different neural mechanisms. This distinction is crucial for patients with schizophrenia: schizophrenia is typically characterized by a loss of experiential continuity, consisting of the subjective fragmentation of the experienced world, including its temporal dimension, and this, we argue, could be accounted for by impaired temporal event-structuring.

* Corresponding author at: INSERM, Cognitive Neuroimaging Unit, Bât 145 Point Courrier 156, 91191 Gif Yvette, France. Tel.: +33 16 908 1667.
E-mail addresses: Virginie.van-Wassenhove@cea.fr, Virginie.van.Wassenhove@gmail.com (V. van Wassenhove).

0028-3932/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.07.002

Several psychiatrists consider the experienced loss of continuity in the sense of time as a key factor in the pathophysiology of schizophrenia (Andreasen, 1999; Minkowski, 1933); what does this precisely entail? Although self-reports ought to be taken with caution, we cite one case illustrating alterations that have been clinically described (Fuchs, 2007; Kimura, 1994; Minkowski, 1933; Vogeley & Kupke, 2007), *"Time splits up and doesn't run forward anymore. These arise uncountable disparate now, now, now, all crazy and without rule or order"* (quoted in Kimura, 1994). Other similar reports can be found illustrating the need to integrate phenomenological reports with current cognitive neuroscientific approaches (Uhlhaas & Mishara, 2007).

In addition to clinical descriptions and self-reports, a number of studies have reported impairments of duration perception (Davalos, Kisley, & Freedman, 2005; Elvevåg et al., 2003; Volz et al., 2001) and a perturbed discrimination of simultaneous vs. synchronous events (Foucher, Lacambre, Pham, Giersch, & Elliott, 2007; Giersch et al., 2009; Schmidt, McFarland, Ahmed, McDonald, & Elliott, 2011). The latter studies show that for patients to become aware of the asynchrony between two sensory events, these events have to be separated by longer delays than for controls. The range of temporal delays that lies below the asynchrony detection threshold constitutes the actual temporal window of integration; within that window, events are considered to be simultaneous. Hence, the enlarged temporal window observed in patients suggests that they are binding or integrating events for a longer time or "in excess" compared to controls. These enlarged temporal windows are observed when explicitly accessing time information (i.e., when patients are asked to report the temporal characteristics of stimuli) and may be at the core of the general inability in organizing events in time.

Besides these explicit temporal impairments, recent results also suggest that patients with schizophrenia are sensitive to desynchronies at an implicit level: it has notably been shown that patients' responses are influenced by short and unconscious asynchronies (Giersch et al., 2009; Lalanne, van Assche, & Giersch, 2012, submitted). Sensitivity to short asynchronies does not tell us how different events are integrated in time, especially at an implicit level. However, the "unity assumption" in multisensory research posits that events are most likely to bind if they are perceived as belonging to a unique underlying cause: in other words, events perceived to be simultaneous should be more likely to bind together (Vatakis & Spence, 2007; Welch & Warren, 1980). For instance, in a populated room, the auditory utterance and the movements of a speaker's face that perceived to be in-sync are more likely to bind together in a single stream of speech. In schizophrenia, impaired audiovisual (AV) integration has previously been reported (de Gelder, Vroomen, Annen, Masthof, & Hodiamont, 2003; de Gelder et al., 2005; Ross et al., 2007) but impairments are not uniform (Pearl et al., 2009; Surguladze et al., 2001) and speech-specific (de Gelder et al., 2003).

Taken all together then, patients with schizophrenia would show less integration despite an enlarged temporal window of integration. This is clearly inconsistent: enlarged temporal windows should be associated with more, and not less, integration. Here, we thus aim at disentangling this conundrum by testing the possible dissociation between implicit and explicit temporal processing and by defining which specific impairments lead to the time distortions experienced by patients with schizophrenia.

For this, we focused on the possible consequences of temporal event-structure impairment in the perceptual binding of ecologically relevant stimuli such as AV speech—which bear obvious daily life relevance. We predicted that such a temporal-event structure deficit would affect the known temporal constraints of AV speech integration and that the subjective temporal estimation of these constraints would be perturbed. The former hypothesis can be addressed using an identification task (ID) in which participants report their perception of AV speech stimuli implicating the integration of visual and auditory information: this is equivalent to measuring the implicit timing of perceptual binding operations. The latter hypothesis can be tested using a simultaneity judgment task (SIM) in which participants report their perceived simultaneity of auditory and visual components of speech events: this assesses the explicit access to the encoding of temporal information. Using these approaches concomitantly (e.g., Conrey & Pisoni, 2006; van Wassenhove, Grant, & Poeppel, 2007) empirically addresses a tricky theoretical issue at the core of temporal perception research: namely, can we experimentally dissociate the temporal content of a representation (explicit time encoding) from the temporal characteristics of a representation (implicit time) (Dennett & Kinsbourne, 1992; van Wassenhove, 2009)?

Well known ecologically relevant illusions necessitating the binding of information across auditory and visual sensory modalities are the McGurk effects (McGurk & MacDonald, 1976). In McGurk/illusory fusion, dubbing an auditory "ba" (A_b) onto a visual place of articulation "ga" (V_g) leads to the illusory fused percept "da"; in McGurk illusion/combination, dubbing an auditory "ga" (A_g) onto a visual place of articulation "ba" (V_b) leads to the illusory combination percept "bga". Fusion is used as an index of automatic *AV* speech integration (Sams et al., 1991; van Wassenhove, Grant, & Poeppel, 2005) because it leads to a unique perceptual outcome that is nothing like any of the original sensory inputs (i.e., neither "ga" nor "ba"). Combination has been much less studied: unlike fusion, the resulting percept is not unique but the product of co-articulated AV speech information (such as "bga"). Fusion and combination stimuli were specifically chosen for the identification task to provide an insight on the binding mechanisms of speech: since auditory and visual speech stimuli and perceptual reports differ from each other, an index of multisensory integration is clearly obtained when desynchronizing the auditory and visual speech stimuli. AV speech integration has been shown to tolerate asynchronies in the order of 200 to 300 ms in healthy population (Conrey & Pisoni, 2006; Munhall, Gribble, Sacco, & Ward, 1996; Maier, Di Luca, & Noppeney, 2011; van Wassenhove et al., 2007). These temporal windows reflect precise neurophysiological correlates that have recently been described within a predictive coding framework for AV speech processing (Arnal, Morillon, Kell, & Giraud, 2009; Arnal, Wyart, & Giraud, 2011; van Wassenhove et al., 2005) and are in line with temporal units necessary for speech parsing (Poeppel, 2003; Giraud & Poeppel, 2012). Thus, AV speech makes an ideal ecological test for our question.

In healthy participants, no major differences were observed when comparing the temporal windows obtained in an ID or a SIM task (Conrey & Pisoni, 2006; van Wassenhove et al., 2007): the temporal properties of AV speech integration appear to reflect directly the temporal information available for the conscious perception of AV speech simultaneity. As previously emphasized, this is in marked contrast with what is currently observed in patients with schizophrenia. Patients appear to have a deficit in integrating multisensory information whereas their explicit impairments would have predicted excessive integration. The limit of the current literature in schizophrenia is that explicit and implicit judgments have not been directly compared using multisensory information. This study fills this gap by directly comparing patients' AV speech integration and simultaneity ratings on the same stimuli and in two tasks. First, we proceeded with assessing AV speech integration in two groups of patients with schizophrenia using illusory McGurk fusion and combination. We then tested whether patients showed an enlarged tolerance to AV desynchrony when identifying the illusions—namely, do AV speech illusions tolerate more asynchrony in patients than in controls (ID task, implicit timing)? Third, we used