Neuropsychologia 62 (2014) 245–261

Contents lists available at ScienceDirect

Neuropsychologia

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

The level of audiovisual print–speech integration deficits in dyslexia

Jens Kronschnabel [a,b], Silvia Brem [a], Urs Maurer [b,c,1], Daniel Brandeis [a,b,d,e,*,1]

[a] University Clinics of Child and Adolescent Psychiatry (UCCAP), University of Zurich, Zurich, Switzerland
[b] Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland
[c] Department of Psychology, University of Zurich, Zurich, Switzerland
[d] Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany
[e] Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland

ARTICLE INFO

Article history:
Received 3 November 2013
Received in revised form
28 March 2014
Accepted 22 July 2014
Available online 30 July 2014

Keywords:
Multisensory integration
Reading
Phonological deficit
Adolescence
Superior temporal
Planum temporale

ABSTRACT

The classical phonological deficit account of dyslexia is increasingly linked to impairments in grapho-phonological conversion, and to dysfunctions in superior temporal regions associated with audiovisual integration. The present study investigates mechanisms of audiovisual integration in typical and impaired readers at the critical developmental stage of adolescence.

Congruent and incongruent audiovisual as well as unimodal (visual only and auditory only) material was presented. Audiovisual presentations were single letters and three-letter (consonant–vowel–consonant) stimuli accompanied by matching or mismatching speech sounds. Three-letter stimuli exhibited fast phonetic transitions as in real-life language processing and reading. Congruency effects, i.e. different brain responses to congruent and incongruent stimuli were taken as an indicator of audiovisual integration at a phonetic level (grapho-phonological conversion). Comparisons of unimodal and audiovisual stimuli revealed basic, more sensory aspects of audiovisual integration. By means of these two criteria of audiovisual integration, the generalizability of audiovisual deficits in dyslexia was tested. Moreover, it was expected that the more naturalistic three-letter stimuli are superior to single letters in revealing group differences. Electrophysiological and hemodynamic (EEG and fMRI) data were acquired simultaneously in a simple target detection task. Applying the same statistical models to event-related EEG potentials and fMRI responses allowed comparing the effects detected by the two techniques at a descriptive level.

Group differences in congruency effects (congruent against incongruent) were observed in regions involved in grapho-phonological processing, including the left inferior frontal and angular gyri and the inferotemporal cortex. Importantly, such differences also emerged in superior temporal key regions. Three-letter stimuli revealed stronger group differences than single letters. No significant differences in basic measures of audiovisual integration emerged. Convergence of hemodynamic and electrophysiological signals appeared to be limited and mainly occurred for highly significant and large effects in visual cortices.

The findings suggest efficient superior temporal tuning to audiovisual congruency in controls. In impaired readers, however, grapho-phonological conversion is effortful and inefficient, although basic audiovisual mechanisms seem intact. This unprecedented demonstration of audiovisual deficits in adolescent dyslexics provides critical evidence that the phonological deficit might be explained by impaired audiovisual integration at a phonetic level, especially for naturalistic and word-like stimulation.

Abbreviations: AV, audiovisual; BOLD, blood oxygenation level dependent; CVC, consonant-vowel-consonant (string); CVCa, unimodal auditory CVC; CVCv, unimodal visual CVC; CVCcon, congruent CVC; CVCinc, incongruent CVC; DD, developmental dyslexia; ERP, event-related potential; GFP, global field power; M, mean; PT, planum temporale; ROI, region of interest; SD, standard deviation; SLcon, congruent single letter stimuli; SLinc, incongruent single letter stimuli; STG, superior temporal gyrus; STS, superior temporal sulcus; TANOVA, topographic analysis of variance

* Correspondence to: Department of Child and Adolescent Psychiatry, University of Zurich, Neumünsterallee 9/Fach, CH-8032 Zurich, Switzerland. Tel.: +41 43 499 2763; fax: +41 43 499 2604.
E-mail address: brandeis@kjpd.uzh.ch (D. Brandeis).

[1] Equal senior contribution.

http://dx.doi.org/10.1016/j.neuropsychologia.2014.07.024
0028-3932/

1. Introduction

Developmental Dyslexia (DD) is a learning disability of neurobiological origin with substantial genetic risk (Pennington & Olson, 2008; Schulte-Körne, Warnke, & Remschmidt, 2006). It is characterized by specific impairments in the acquisition of efficient reading and emerges despite normal intelligence, no obvious sensory deficits, and adequate instruction (Lyon, Shaywitz, & Shaywitz, 2003; WHO, 1992). Dyslexia is one of the most widespread developmental disorders, affecting around 5% of school-aged children in German speaking countries (Schulte-Körne, 2010; Schulte-Körne & Remschmidt, 2003). The most commonly accepted cause of DD is thought to be a phonological processing deficit (Bradley & Bryant, 1978; Goswami, 2000; Ramus, 2003; Shaywitz & Shaywitz, 2005; Snowling, 2000; Tree, 2008; Vellutino, Fletcher, Snowling, & Scanlon, 2004). It has been suggested that this deficit is characterized by impairments in converting print (graphemes) into corresponding sounds (phonemes; Snowling, 1980). This hypothesis is appealing, considering that such conversions are obviously required in order to acquire efficient reading (Bradley & Bryant, 1983; Ehri, 2005). It also receives strong support from intervention studies in dyslexic or pre-school children that demonstrate improvements in reading skills through audiovisual (AV) training programmes (Brem et al., 2010; Kujala et al., 2001; Lovio, Halttunen, Lyytinen, Näätänen, & Kujala, 2012; Törmänen & Takala, 2009). Given the importance of AV conversion for reading, it is surprising that neurobiological origins of DD have until lately mainly been investigated by means of unimodal paradigms, both in the fMRI and EEG literature. Pekkola et al. (2006) were the first to use bimodal AV stimulations in order to probe phonological deficits in dyslexic adults. They found increased activation to incongruent stimulation in motor speech regions (Broca, left premotor cortex), which was more pronounced in dyslexic compared to typical adult readers. This was interpreted as increased use of subvocal motor-articulatory strategies during AV speech processing.

More recently, reduced congruency effects in impaired compared to nonimpaired readers were reported for superior temporal sulci (STS) and auditory cortices (planum temporale, PT; Blau et al., 2010; Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009). Specifically, stronger activation to congruent than incongruent letter–speech sound pairings was observed in these regions for typically reading children (Blau et al., 2010) and adults (Blau et al., 2009) but not for their dyslexic peers. STS regions have repeatedly been associated with multisensory integration and it has been argued that a distinction between congruent and incongruent pairings can only emerge after the unisensory inputs have been integrated successfully (Blomert, 2011; van Atteveldt, Formisano, Goebel, & Blomert, 2007). Thus, Blau et al. (2009, 2010) reasoned that letter–sound integration is an emergent property of learning to read which develops inadequately in dyslexic readers, presumably as a result of lacking specialization at the neuroanatomical level. This has eventually been interpreted as a specific crossmodal binding deficit and as an impairment in the automated formation of unique grapho-phonological objects (Blomert, 2011). Note that Pekkola et al. (2006) had observed stronger activation to incongruent rather than congruent stimulation in the left STS of both groups. This difference to Blau et al. (2009, 2010) might be explained by attentional factors given that participants in Pekkola et al. (2006) but not in Blau et al. (2009, 2010) had to actively monitor congruency state (see van Atteveldt et al., 2007 for active versus passive processing). The significance of superior temporal regions for dyslexia is also supported by observations of reduced grey matter volume (Richlan, Kronbichler, & Wimmer, 2013; Welcome, Chiarello, Thompson, & Sowell, 2011).

According to a range of other criteria, the STS and adjacent gyral regions have repeatedly been related to multisensory integration: The "super-additivity" criterion, for instance, requires the response to crossmodally congruent stimuli to be stronger than the summed unisensory responses. In addition, the response to incongruent multisensory stimuli should be sub-additive compared to the summed unisensory stimulation, as has been shown for single cells of the superior colliculi in mammals (Kadunce, Vaughan, Wallace, Benedek, & Stein, 1997; Stein, 1998). These criteria are based on the general concept that sensory brain responses to stimulus elements that are processed independently should be additive. Accordingly, violations of additivity can indicate the presence of multisensory interactions, or of (non-additive) attentional or cognitive processes. Calvert, Campbell, and Brammer (2000) reported that BOLD signals in the ventral part of the human left STS also fulfilled both super- and sub-additivity criteria. Subsequent fMRI studies, however, could not replicate this rather strict conjunction of criteria (Hocking & Price, 2008, give an overview). Considering that the BOLD signal originates from large clusters of neurons containing not only multisensory integration cells, subsequent studies accepted more liberal criteria of AV integration for fMRI. van Atteveldt, Formisano, Goebel, and Blomert (2004) found super-additivity (without corresponding sub-additivity) in bilateral PT and Heschl's sulci, but nowhere in the STS. However, the STS of both hemispheres showed stronger responses to congruent stimulation than to the maximum of the unisensory responses. This "response enhancement" (Beauchamp, 2005) was met in several studies (Beauchamp, Lee, Argall, & Martin, 2004; van Atteveldt et al., 2007). Electrophysiological studies have also made use of the super-additivity criterion and supported the role of the STS in AV integration (e.g., Besle, Fort, Delpuech, & Giard, 2004; Klucharev, Möttönen, & Sams, 2003; Raij, Uutela, & Hari, 2000). Interestingly, all of these electrophysiological studies found sub- rather than super-additive effects for congruent AV stimulation.

In contrast to effects of congruency mode, criteria on the interplay of uni- and multimodal responses do not require grapho-phonological knowledge or conversion when probing AV integration. Comparisons of phonetically matching versus conflicting conditions therefore probe phonetic features of AV integration (Hocking & Price, 2008; Klucharev et al., 2003; Ojanen et al., 2005), whereas super-additivity effects probe general features of AV integration, including more basic and sensory aspects because phonological processing is no premise for the super-additivity criterion (e.g., Klucharev et al., 2003).

As mentioned above, reduced congruency effects have been reported for dyslexia (Blau et al., 2010, 2009). However, criteria of super-additivity have never been tested in dyslexia, although they could clarify whether reading impairments originate from a more general AV deficit. The present study therefore includes both congruency as well as super-additivity approaches of testing AV integration. Because previous studies had found reduced AV integration in DD during passive (Blau et al., 2010, 2009) but not active congruency matching (Pekkola et al., 2006), the design of Blau et al. (2010) was adopted to simulate AV integration during reading. Three further important aspects go beyond previous studies: First, by using consonant–vowel–consonant sequences (CVCs) in addition to single letter stimuli, we intended to go one step further towards realistic reading. CVCs are more speech-like than mere letters given their rapid acoustic changes (formant transitions) that are an immanent property of natural speech. A recent study used uni- and bimodal word and pseudoword stimuli and found reduced overall activation in the right STS of dyslexic adults, indicating less proficient engagement of circuits involved in AV processing (Kast, Bezzola, Jäncke, & Meyer, 2011). However, no contrasts to specifically test for AV integration were computed. The presently used CVCs were all without semantic content in order to avoid the engagement of confounding higher-order psycholinguistic processing. Of central interest was whether