UNIVERSITY OF HELSINKI
FACULTY OF BEHAVIOURAL SCIENCES
STUDIES IN PSYCHOLOGY 103: 2014

INTERACTIONS OF THE PROCESSING OF LETTERS AND SPEECH SOUNDS AS REFLECTED BY EVENT-RELATED BRAIN POTENTIALS

MARIA MITTAG

Interactions of the processing of letters and speech sounds as reflected by event-related brain potentials

Maria Mittag

Cognitive Brain Research Unit
Cognitive Science
Institute of Behavioural Sciences
University of Helsinki, Finland

Academic dissertation to be publicly discussed,
by due permission of the Faculty of Behavioural Sciences
at the University of Helsinki in Auditorium XII, Fabianinkatu 33,
on the 18th of June 2014, at 12 o'clock noon.

University of Helsinki
Institute of Behavioural Sciences
Studies in Psychology 103: 2014

Supervisors:

Professor Teija Kujala, Ph.D.
Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Finland
Cicero Learning, University of Helsinki, Finland

Professor Kimmo Alho, Ph.D.
Division of Cognitive Psychology and Neuropsychology, Institute of Behavioural Sciences, University of Helsinki, Finland
Helsinki Collegium for Advanced Studies, University of Helsinki, Finland

Dr. Rika Takegata
Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Finland

Reviewers:

Professor Emerita Patricia Michie, Ph.D.
Functional Neuroimaging Laboratory, University of Newcastle, Australia

Dr. Piia Astikainen
Department of Psychology, University of Jyväskylä, Finland

Opponent:

Professor Paavo Leppänen, Ph.D.
Department of Psychology, University of Jyväskylä, Finland

Cover illustration: Mikko Eerola

ISSN 1798-842X
ISSN-L 1798-842X
ISBN 978-952-10-9965-6 (pbk.)
ISBN 978-952-10-9966-3 (PDF)
http://www.ethesis.helsinki.fi
Unigrafia
Helsinki 2014

CONTENTS

ABSTRACT

The processing of audiovisual information is ubiquitous in our daily life. As such, understanding the cortical correlates of audiovisual processing and its interactions offers a promise of practical interventions in many real-life settings. Reading, as one example, relies on the formation of artificial audiovisual associations and requires adaptions from brain mechanisms in order to process and integrate these connections effortlessly. In dyslexia, reading problems are associated with a failure in forming those associations, and neural changes and improvements of reading skills in children with dyslexia were reported after interventions ameliorated those processes. The present thesis investigates the neural networks associated with speech sound processing and discrimination when accompanied by printed text. In all studies, a high-density EEG system was utilized, enabling the examination of spatio-temporal dynamics of audiovisual processing in adult fluent readers and in readers with dyslexia.

In fluent adult readers, change-related responses to consonant and pitch changes were greater when presented with printed text than with scrambled images, suggesting that letters modulate speech sound discrimination at an early cortical processing stage. This integration was sensitive to precise temporal alignment between the sounds and printed text, as it broke down when a time delay between the sounds and print was introduced. In contrast to fluent readers, adult readers with dyslexia showed a general attenuated discrimination of speech sounds when presented with print. Their neural responses for speech sounds presented with print did not differ from those presented with scrambled images. Our results, therefore, suggest that audiovisual processing is generally impaired in dyslexia, and support the notion that letter representations are poorer in readers with dyslexia than fluent readers. In addition, audiovisual processing was delayed in readers with dyslexia, suggesting a deficit in concurrent processing of multiple sensory cues. The studies of this thesis also show that

attention to one of the modalities is needed for the audiovisual integration to occur, and, moreover, that audiovisual attention boosts the integration. Furthermore, our results reveal that, in addition to attention, the phonological content of the task modulates letter-speech sound processing.

The studies presented in the present thesis confirmed, with a more controlled methodology, that letters modulate speech sound discrimination at an early neural level. The present results illuminate the way these processes are impaired in dyslexia, and, further, that audiovisual attention is most beneficial for such an integration to occur. To conclude, the studies at hand have shed novel light on the basic and aberrant mechanisms of letter-speech sound processing, and can be used, for instance, in training programs to promote accurate mapping of letters and speech sounds, and, consequently, reading skills in individuals with dyslexia.

TIIVISTELMÄ

Tiedon audiovisuaalinen käsittely liittyy jokapäiväiseen toimintaamme. Näkö- ja kuulotiedon aivomekanismien ymmärtämisen pohjalta voidaan muun muassa kehittää erilaisia interventioita. Esimerkiksi lukemisen edellytyksenä on, että aivot käsittelevät tehokkaasti äänteiden ja kirjainten välisiä audiovisuaalisia yhteyksiä. Lukihäiriöisillä lukivaikeuksien taustalla voi olla audiovisuaalisten yhteyksien muodostamisen ongelmia ja audiovisuaalisten interventioiden onkin osoitettu lapsilla tehostaneen sekä hermoston tiedonkäsittelyä että lukutaitoa. Tässä väitöskirjassa selvitetään puheäänteiden ja samaan aikaan esitetyn tekstin yhtäaikaisen tiedonkäsittelyn hermostollista perustaa. Tutkimuksissa käytetään monikanavaista elektroenkefalografiaa (EEG), joka mahdollistaa aivojen audiovisuaalisen tiedonkäsittelyn tutkimisen sekä sujuvasti lukevilla että lukihäiriöisillä henkilöillä.

Väitöskirjan tulokset osoittavat, että sujuvasti lukevilla aikuisilla konsonantin ja äänenkorkeuden muutosten hermostollinen tiedonkäsittely voimistui kun ne esitettiin kirjoitetun tekstin yhteydessä verrattuna siihen, että ne esitettiin merkityksettömien symbolien yhteydessä. Tulos viittaa siihen, että kirjainten näkeminen muokkaa puheäänteiden hermostollista tiedonkäsittelyä jo hyvin varhaisessa tiedonkäsittelyn vaiheessa. Tutkimuksissa havaittu hermostollisen tiedonkäsittelyn voimistuminen edellytti kuitenkin, että puheäänet ja teksti esitettiin samanaikaisesti sillä hermostollinen tiedonkäsittely ei voimistunut kun puheäänen ja tekstin esittämisen välillä oli ajallinen viive. Sujuvasti lukeviin aikuisiin verrattuna lukihäiriöisillä samanaikaisesti esitettyjen puheäänten ja tekstin hermostollinen tiedonkäsittely oli vaimeampaa eikä siihen vaikuttanut se, että teksti oli korvattu merkityksettömillä symboleilla. Tulosten mukaan lukihäiriöisillä audiovisuaalinen tiedonkäsittely on kauttaaltaan heikentynyttä ja löydökset tukevat ajatusta siitä, että kirjainten

hermostolliset edustukset ovat heikommin muodostuneita lukihäiriöisillä kuin sujuvasti lukevilla. Lisäksi lukihäiriöisillä audiovisuaalinen tiedonkäsittely oli ajallisesti viivästynyttä, mikä viittaa ongelmiin useiden aistimusten samanaikaisessa tiedonkäsittelyssä. Väitöskirjatutkimukset osoittivat myös tarkkaavaisuuden kohdistamisen ääniin tai tekstiin olevan tarpeellista, jotta tiedon integrointia aivoissa tapahtuisi. Tämä integraatioprosessi voimistui tarkkaavaisuuden kohdistuessa molempien aistien ärsykkeisiin. Tutkimusten mukaan tarkkaavuuden lisäksi myös tehtävässä käytetyn fonologisen aineksen sisällöllä oli vaikutusta kirjainten ja puheäänten audiovisuaaliseen tiedonkäsittelyyn.

Kokonaisuudessaan väitöskirjatutkimukset osoittavat, että kirjaimet vaikuttavat puheäänteiden hermostolliseen tiedonkäsittelyyn jo hyvin varhaisessa tiedonkäsittelyn vaiheessa. Tulokset antavat uutta tietoa siitä, miksi nämä tiedonkäsittelyn prosessit ovat heikentyneet lukihäiriössä ja kuvaavat, kuinka audiovisuaalinen tarkkaavuus edesauttaa kirjainten ja puheäänten yhdistämistä aivoissa. Väitöskirjan tulokset laajentavat tietämystämme puheäänten ja kirjainten samanaikaisen tiedonkäsittelyn taustalla olevista aivomekanismeista ja tuloksia voidaan hyödyntää esimerkiksi interventiotutkimuksissa, joilla pyritään sujuvoittamaan lukemista lukihäiriöisillä tehostamalla kirjainten ja puheäänten yhdistämistä aivoissa.

ACKNOWLEDGEMENTS

First and foremost, I am deeply indebted to my primary supervisor, Professor Teija Kujala, for her never-ending support and trust in my abilities as a scientist. She is my science mom, and has given me indispensable guidance and structure throughout my doctoral studies. Her seminal work on speech perception, disorders, and plasticity has been a constant source of inspiration, and her encouragement throughout the years was crucial to the development of my research. Not only has she been a role model for my research, she also provided me with emotional care and wisdom. She sets an example for a strong leading woman in science with passion, warmth, and hard work.

My sincere thanks also go to Professor Kimmo Alho, who stepped in as my supervisor halfway through my PhD studies, and has encouraged me with his knowledge and eager mind for science. His enthusiasm, persistence, sharp mind, and continuous drive for improvement were essential to my success in the submission of my thesis.

I also wish to thank my supervisor Dr. Rika Takegata, who introduced me to EEG analysis during my master's thesis. Her creative and brave scientific spirit left me frequently in awe. I am very thankful for her confidence in my scientific abilities from early on.

I am greatly indebted to Professor Risto Näätänen, who opened the door for me to come to Finland in the first place, when he accepted my application to work at CBRU. He supported me during difficult financial times by giving me a research assistant's job. Because of his groundbreaking research and his kind nature, I am where I am today.

I owe special thanks to co-authors: Docent Marja Laasonen, Docent Teemu Rinne, Mrs. Emma Salo, Mr. Tatu Huovilainen, and Ms. Paula Thesleff for a pleasant collaboration. I also wish to thank Ms. Marja Junnonaho and Ms. Piiu Lehmus for their immeasurable administrative support and friendly motherly manner. I am indebted to my co-authors Mr. Tommi Makkonen and Mr. Miika Leminen for the technical support and for their clever input in solving programming issues. Furthermore, I wish to thank Professor Mari Tervaniemi and Professor Minna Huotilainen for opening their homes to CBRU members and friends, for creating a nurturing family atmosphere, and for their advice in matters above and beyond career. In addition, I also thank Docent Kaisa Tiippana for fruitful discussions on multisensory processing, Dr. Jari Lipsanen for statistical support, and Docent Petri Paavilainen for his guidance in teaching.

CBRU has given me many long-lasting friendships such as my two beloved “boys”. I am extremely grateful to have had the privilege of working with Dr. Tuomas Teinonen, also known as my personal secretary, a bright, young scientist who is now in the hands of the medical world, and Dr. Eino Partanen, another bright, young scientist and former office-mate who was able to cope with my irritating nature and spread pony love throughout CBRU and in my heart, brohoof! Our precious friendships formed in Helsinki will continue to grow.

My deepest thanks to my international colleagues and friends who are very dear to my heart: Ms. Marina Klyuchko, Dr. Brigitte Bogert, Dr. Alexis Bosseler, Mrs. Tiziana Quarto, Mr. Carlos Silva Pereira, Mr. Ben Gold, Dr. Caroline Jacquier, and my invaluable and smart co-author and friend, Mrs. Karina Inauri. Thank you all for discovering and sharing Finnish life and culture with me, and for being such great spirits. I owe special thanks to Dr. Dries Froyen for inviting me to Maastricht University, into his home, and for stimulating discussions on our common research topic.

I also wish to thank Ms. Lilli Kimppa, Ms. Hanna Poikonen, Ms. Soila Kuuluvainen, Ms. Anna Rämä, Mrs. Saila Seppänen, Mr. Valtteri Wikström, Dr. Eva Istók, Dr. Veerle Simoens, Dr. Satu Pakarinen, Mr. Roope Heikkilä, Dr. Heike Althén, Dr. Riikka Lovio, Mrs. Riikka Lindström, Docent Elvira Brattico, Ms. Henna Markkanen, Dr. Alina Leminen, Ms. Ansku Nieminen, Dr. Juha Silvanto, Dr. Sari Ylinen, Mrs. Ritva Torppa, and all my other colleagues at the CBRU for their support and friendship during the various phases of my thesis.

The thesis work was supported financially by the University of Helsinki, the Centre for International Mobility, Oskar Öflunds Stiftelse, the National Doctoral Programme of Psychology, and the Academy of Finland. I am grateful to the two reviewers of this dissertation, Professor Patricia Michie and Dr. Piia Astikainen. I also thank Professor Paavo Leppänen for agreeing to act as my opponent during the public examination of this work.

I also thank Professor Patricia Kuhl for accepting me as a post-doctoral researcher, and for being so kind to allow me the time to complete my thesis work. I would also like to thank my new colleagues for their warm-hearted welcome to the I-LABS community. Especially, I wish to thank Ms. Mihwa Kim for the ‘U.S.S. So What boat’ and beautiful turtle moments and Dr. Ping Mamiya and Dr. Kambiz Tavabi for their encouragements and support.

I owe deep thanks to Ms. Annika Forstén for being such a wonderful friend with her thoughtful and gentle spirit, a positive attitude to life, and an extremely comfortable couch. I most warmly wish to thank Ms. Sirke Nieminen for her insightful discussions on life and for being my support pillow. My warmest thanks go out to the “coconut pan cooking group” including Dr. Tuomas Teinonen, Ms. Elina Aho, Mrs. Siiri Laari, Ms. Satu Pihlaja, and Dr. Anna Wilschut. Thank you for providing me with pescetarian gluten-free lactose-free delicacies at your homes. I owe special thanks to Mr. Mikko Eerola for specially providing me with a cover illustration for this thesis, and for bringing carrots and bunnies into my life. I warmly thank Ms. Anna-Stina Wiklund, who inspired me to make healthy life choices during my doctoral time. I also thank Ms. Sarah Stephan and Dr. Stefan Wahlen for their friendship and for exploring Finnish language and Finnish pubs with me. I give my deepest thanks to Ms. Taina Heimo. Thank you for being my first friend in Helsinki and for sharing with me St. Petersburg and Frankfurt experiences, and for introducing me to sauna culture in Finland.

From Frankfurt, I thank the Goethe group for welcoming me warmly, for providing me with a free desk whenever I needed it, and for sharing conference experiences with me. From Dresden, I thank my friends Ms. Anke Gaebel, Dr. Antje Gerner, Ms. Claudia Anders, and Dr. Patricia Garrido Vásquez-Schmidt for always being great cheerleaders.

My deep gratitude goes to Dr. Ann-Mari Estlander for providing me with the best guidance and care I could wish for in becoming a butterfly. I express my warmest thanks to my Finnish family, Ms. Michaela Björklund and Mrs. Carita Björklund, for sharing impeccable moments of comfort and constant reminders of my inner strength ~ your Pearl. I am also grateful to Mr. Jasper van den Bosch and his family for their support during various phases of this thesis. I dearly thank my childhood friend and my beautiful Brummer, Ms. Julia Garten, for sharing yet another life experience with me and for being such a joy. I am also indebted to Mr. Michael Ranft, not only a great “squirrel” programmer and insightful friend, but also my rock in Finnish weathers. I also wish to thank my precious baby kitten, Maui Nervzwerg, for her contribution of typing a few letters and numbers into my thesis document, which I did appreciate, however, unfortunately had to delete.

I owe my deepest gratitude to my brother, Dr. Marcus Mittag, for believing in my strengths and being the best role model a sister can have. We both learned how to fly.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

Study I	**Mittag, M.**, Takegata, R. & Kujala, T., 2011. The effects of visual material and temporal synchrony on the processing of letters and speech sounds. *Experimental Brain Research*, 211, 287-298.
Study II	**Mittag, M.**, Thesleff, P., Laasonen, M., & Kujala, T., 2013. The neurophysiological basis of the integration of written and heard syllables in dyslexic adults. *Clinical Neurophysiology*, 124, 315-326.
Study III	**Mittag, M.**, Alho, K., Takegata, R., Makkonen, T. & Kujala, T., 2013. Audiovisual attention boosts letter-speech sound integration. *Psychophysiology*, 50, 1034-1044.
Study IV	**Mittag, M.**, Inauri, K., Huovilainen, T., Leminen, M., Salo, E., Rinne, T., Kujala, T., & Alho, K. 2013. Attention effects on the processing of task-relevant and task-irrelevant speech sounds and letters. *Frontiers in Neuroscience*, 7, article 231.

ABBREVIATIONS

AEP	Auditory event-related potentials
ANOVA	Analysis of variance
EEG	Electroencephalogram
EOG	Electrooculogram
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
GBR	Gamma band response
IR	Incongruency response
MEG	Magnetoencephalography
MMN	Mismatch negativity
MMR	Mismatch response
Nd	Negative difference
Pd	Positive difference
PN	Processing negativity
PSP	Postsynaptic potentials
PT	Planum temporale
RP	Rejection positivity
SOA	Stimulus-onset asynchrony
STG	Superior temporal gyrus
aSTP	Anterior superior temporal plane
STS	Superior temporal sulcus

1 INTRODUCTION

The crosstalk between auditory and visual information is ubiquitous in our daily lives, whereby the brain integrates information from both senses into a coherent percept (B. E. Stein & Mededith, 1993). In order to focus on relevant information, the brain has the ability to suppress irrelevant information in one modality when it is unrelated to relevant information in another modality (Hillyard, Mangun, Woldorff, & Luck, 1995). However, in some processes like reading, audiovisual integration is required; one must effortlessly map familiar speech sounds to artificial visual symbols (Ehri, 2005). Whereas around 90% of people learn to read without problems after adequate instruction, 5 to 17% of children show difficulties learning how to read that cannot be explained by cognitive deficits, sensory deficits, or by lack of adequate reading instruction or motivation (e.g., American Psychiatric Association, 1994). In children with developmental dyslexia, reading problems are linked to unsuccessful mapping of graphemes associated with phonemes (Snowling, 1980). However, despite the relevance of audiovisual processing and its interactions in reading, there is little knowledge on the neural mechanisms underlying this initial mapping process leading to reading acquisition.

The present **Studies I, II, and IV** give insight into neural networks underlying interactions of auditory and visual linguistic processing. As these interactions are important for reading, we compared them in fluent readers to readers with dyslexia in **Study II**. In **Studies III and IV**, we investigated how attention affects the processing of letters and speech sounds.

1.1 Perception and neural basis of letter-speech sound integration

Behavioural studies have demonstrated that printed text can modulate auditory speech processing (Frost & Kampf, 1993; Frost, Repp, & Katz, 1988; Massaro, Cohen, &

Thompson, 1988). In the study of Frost and colleagues (1988), participants were instructed to detect noise-masked speech that was presented in synchrony with matching or non-matching print, or alone. The result of a strong response bias to identify masked speech in the matching print condition suggests that print modulates auditory speech because participants had to generate speech representations from print to perform the task. Evidence for generation of auditory representations by letters was also given in the study of Dijkstra and colleagues (1989), in which reaction times were faster to speech sounds when congruent, in contrast to incongruent letters were presented prior, synchronously to, or after the sounds. Furthermore, Massaro (1998) investigated whether a well-known sensory fusion between auditory and visual inputs, the so-called McGurk effect (McGurk & MacDonald, 1976), is limited only to visual speech, or if printed text similarly influences speech perception. Various studies suggested that this phenomenon is unique to speech since seeing articulatory movements provides complementary information for speech comprehension (Sams et al., 1991; Tuomainen, Andersen, Tiippana, & Sams, 2005). In the study of Massaro (1998), seven spoken consonants on the /bi/ and /di/ continuum were presented with either the letter “B”, the letter “D” or, as a control, with visual speech of /bi/ or /di/ presented in the same trial. The participants were instructed to report the letters they heard. At ambiguous auditory levels within the auditory continuum, letters facilitated auditory stimulus perception to the same extent as visual speech streams did.

Using magnetoencephalography (MEG), the time-course of letter-speech sound mapping was determined by recording magnetic brain responses to different Finnish consonants or vowel speech sounds and letters when presented alone or in matching or non-matching combinations (Raij, Uutela, & Hari, 2000). Activations were elicited at 60−120 ms after stimulus onset in sensory-specific areas and around 225 ms in the left superior temporal sulcus (STS), indicating feed-forward projections to multisensory convergence areas.

Evidence for an interaction of auditory and visual responses was found at 280 ms in the right temporo-occipito-parietal junction and differential interaction effects for matching and non-matching letter-speech sound pairs were observed at 380–540 ms in the STS. In addition, changes in cortical oscillations to congruent and incongruent grapheme-phoneme connections were also investigated (Herdman et al., 2006). Congruent pairs evoked 2–10 Hz activation in the left auditory cortex, followed by smaller 2–16 Hz activation bilaterally in the visual cortex, indicating that congruent letter input can modify cortical activity in the left auditory cortex. This was also substantiated by shorter response times to congruent letter-speech sound pairs than to incongruent pairs, the results being consistent with previous behavioral results (Dijkstra et al., 1989).

Functional magnetic resonance imaging (fMRI) studies investigated the neuro-anatomical structures underlying letter-speech sound integration in greater detail by manipulating semantic congruency and bottom-up information processing such as temporal accuracy between letters and speech sounds (van Atteveldt, Formisano, Blomert, & Goebel, 2007; van Atteveldt, Formisano, Goebel, & Blomert, 2004). In the study of van Atteveldt et al.’s study (2004), participants were presented with unimodal single letters, speech sounds, bimodal congruent, or incongruent letter-speech sound pairs and were asked to passively view and/or listen to these stimuli. Activations to congruent and incongruent letter-speech sound pairs were stronger than responses to speech sounds or letters alone in the STS and superior temporal gyrus (STG). In addition, low-level auditory cortex regions, specifically Heschl's sulcus extending to the planum temporale (PT), showed enhanced responses to congruent pairs, but suppressed responses to incongruent pairs. A follow-up study introduced a time delay of 150 ms and 300 ms between letters and speech sounds and replicated the results of enhanced activation for congruent and incongruent letter-speech sound pairs in the STS/STG (van Atteveldt, Formisano, Blomert, et al., 2007). In addition, the results showed that this enhanced activation was unaffected by the time delay, since larger responses were still observed when letters were asynchronously presented to the speech sounds. However, this was not the case for the PT and the anterior superior temporal plane (aSTP), which showed enhanced responses to letter-speech sound pairs only when synchronously presented. It was concluded that the STS serves as an integration site for letters and speech sounds over a wide temporal range, followed by feedback to regions of the auditory cortex only if letters and speech sounds are in accurate temporal alignment (for a review, see van Atteveldt, Roebroeck, & Goebel, 2009).

The question of whether task-irrelevant congruent or incongruent letters influence auditory cortex activation was examined in the study of Blau and colleagues (2008). Speech sounds were presented together with congruent or incongruent letters degraded at different levels and the task was to identify the speech sounds. Even though the visual information was not needed in the task, a congruency effect was found in the auditory cortex and in the fusiform gyrus of the visual cortex for speech sounds paired with letters with low amount of visual noise suggesting that letters and speech sounds are automatically linked in literate adults.

Furthermore, the influence of different top-down demands on letter-speech sound perception was manipulated by active versus passive tasks (van Atteveldt, Formisano, Goebel, & Blomert, 2007). Participants were presented with congruent or incongruent letter-speech sound pairs with the task either to actively judge whether the letters were congruent or incongruent with the speech sounds they heard or to passively listen to and view the bimodal stimuli. During the passive task, congruent letter-speech sound pairs elicited enhanced responses while incongruent pairs suppressed responses as compared to speech sound presentation alone in the auditory association cortex, a result consistent with earlier findings during passive designs (van Atteveldt, Formisano, Blomert, et al., 2007; van Atteveldt et al., 2004). The congruency effect observed in the passive condition, however, vanished during the active matching task. This was associated with enhanced responses in several frontal and parietal areas and increased activity in the auditory cortex for incongruent pairs relative to congruent pairs. It was suggested that responses in the auditory cortex to congruent versus incongruent letter-speech sound pairs are dependent on the demands of the task and, further, that attentive processing changes the neural substrate of congruency processing.

1.2 Auditory event-related potentials

1.2.1 Event-related potentials (ERPs)

Event-related potentials (ERPs) have recently become an attractive tool to investigate the neural time course underlying letter-speech sound integration in fluent readers and readers with dyslexia (Froyen, Bonte, van Atteveldt, & Blomert, 2009; Froyen, van Atteveldt, & Blomert, 2010; Froyen, van Atteveldt, Bonte, & Blomert, 2008; Froyen, Willems, & Blomert, 2011). ERPs are voltage fluctuations time-locked to perceptual, cognitive, or motor events (Picton et al., 2000). These potentials can be non-invasively measured with electrodes attached to the human scalp and extracted with signal averaging and filtering techniques. ERPs provide accurate information on the timing of neural activity due to their high millisecond temporal resolution (Picton et al., 2000; Picton, Lins, & Scherg, 1995).

ERPs are summated extracellular products of excitatory postsynaptic potentials (PSPs) originating during neurotransmission, i.e., the binding of neurotransmitters to postsynaptic receptors elicits short-term changes to the flow of ions across postsynaptic cell membranes (Luck, 2005). Thus, the electroencephalogram (EEG) measures instantaneous neural activity from summated PSPs of large numbers of similarly oriented and synchronized neurons

(Luck, 2005). Almost the entire EEG signal comes from cortical pyramidal cells oriented perpendicular to the cortex (Luck, 2005).

1.2.2 Auditory ERPs

Auditory evoked potentials (AEPs) allow investigating the neural mechanisms underlying the processing and discrimination of speech sounds and their modulation by letters with high temporal accuracy. In the present studies, long latency AEPs were recorded that are commonly classified as exogenous or endogenous responses depending on whether they reflect transient physical stimulus characteristics or cognitive processes, respectively (Näätänen, 1992; Picton et al., 1995; Sutton, Braren, Zubin, & John, 1965). Long latency AEPs occur between 50 to 300 ms after stimulus onset and are referred to as the P1-N1-P2 complex, usually originating from several spatially distinct neural sources (e.g., Näätänen & Picton, 1987). The P1 response with a positive polarity over central scalp areas is evoked between 55 to 80 ms with its maximum at the vertex and originates from the lateral portion of Heschl's gyrus which belongs to the secondary auditory areas (Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994). The P1 is followed by the N1 response, with its negative polarity usually peaking around 90 to 110 ms from stimulus onset and with multiple generators in the primary and secondary auditory cortex (Näätänen & Picton, 1987).

1.2.3 Change-related ERPs reflecting letter-speech sound integration

The present **Studies I-III** investigated processing of changes in speech sounds, as reflected by the N2 ERP response, and modulation of this processing by letters. The auditory

N2 response associated with deviant processing consists of two components (Näätänen, Simpson, & Loveless, 1982): the mismatch negativity (MMN) and the N2b.

1.2.3.1 The mismatch negativity (MMN)

The MMN reflects pre-attentive cortical stages of auditory discrimination and is usually elicited when a sound violates the memory trace formed by regularity in the preceding sounds (Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is elicited by any change in the auditory stimulation that exceeds a certain threshold that roughly corresponds to the behavioural discrimination threshold (Näätänen et al., 2007). The MMN usually peaks at 100 to 250 ms after deviance onset with maximum scalp distribution over frontal areas (Garrido, Kilner, Stephan, & Friston, 2009; Sams, Paavilainen, Alho, & Näätänen, 1985). The MMN reflects both simple representations of physical stimulus features of preceding sounds, such as pitch, and complex representations of more abstract auditory rules or regularities (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). With increasing magnitude of the stimulus deviation, the MMN latency shortens and amplitude increases until it reaches a plateau (Kujala & Näätänen, 2010). Additive effects on the MMN amplitude are observed when the deviant differs from the standard in two or more attributes (Näätänen & Alho, 1997; Näätänen et al., 2007; Takegata, Paavilainen, Näätänen, & Winkler, 1999).

The MMN gets contribution from several cerebral sources (for reviews, see Kujala, Tervaniemi, & Schröger, 2007; Näätänen et al., 2007) reflecting various stages in early cognition. The major subcomponent of the MMN originates from the bilateral supratemporal auditory cortices and is evidently related to pre-attentive auditory change detection (Alho, 1995). Another subcomponent is generated in the frontal lobes, predominantly in the right hemisphere, and is presumably associated with involuntary attention switching to a deviant auditory event (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000; Yago, Escera, Alho, & Giard, 2001). Additional MMN generators have been reported in subcortical areas (Csépe, 1995) and in the parietal lobe (Lavikainen, Huotilainen, Pekkonen, Ilmoniemi, & Näätänen, 1994; Levänen, Ahonen, Hari, McEvoy, & Sams, 1996).

The MMN can also be used to study how speech sounds are represented by neural traces in the brain. For instance, it was shown that MMN amplitude is stronger for a typical vowel category change in the native language than for an unfamiliar vowel category change in an unfamiliar language (Näätänen et al., 1997). The native-language memory traces were suggested to develop between 6 and 12 months in infants (Cheour et al., 1998; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). In addition, the MMN amplitude enhances for foreign-language phonemes after learning to master that language (Dehaene-Lambertz, Dupoux, & Gout, 2000; Winkler et al., 1999). In adults, the MMN for native-language phoneme changes is predominantly generated in the left hemisphere (Näätänen et al., 1997; Pulvermüller et al., 2001; Shtyrov, Kujala, Palva, Ilmoniemi, & Näätänen, 2000), whereas the MMN for acoustic changes is stronger in the right hemisphere than in the left hemisphere (Giard et al., 1995; Paavilainen, Alho, Reinikainen, Sams, & Näätänen, 1991).

The MMN is traditionally recorded with the oddball paradigm in which repetitive standard sounds and occasional rare (e.g., $p = 0.1$) deviant sounds are presented. A main disadvantage of the oddball paradigm is the small percentage of deviants recorded in one sequence which makes recording times long (Kujala et al., 2007). As vigilance affects signal-to-noise ratio, the so-called multi-feature paradigm (originally called "Optimum 1 paradigm"; Näätänen et al., 2004) was developed to diminish recording times and introduce different types of deviants in one recording sequence. In this paradigm, deviant stimuli alternate with the standard stimuli (50%) and the rationale is that each deviant functions as a standard because the deviant strengthens the memory trace of the standard with the features they have in common (Kujala et al., 2007). MMN responses to frequency, duration, intensity, and location changes and sounds including a small gap recorded with the multi-feature paradigm were similar or even slightly larger in amplitude as those obtained with the oddball paradigm (Näätänen et al., 2004; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007). Also, similar results between the two paradigms were obtained for speech sounds including semi-synthetic consonant-vowel syllables with vowel, duration, consonant, frequency, and intensity changes (Pakarinen et al., 2009). Therefore, the multi-feature paradigm is an attractive tool for recording an extensive profile of auditory discrimination abilities in a short recording time.

1.2.3.2 The N2b

When sound sequences are attended to or the deviant stimuli are especially intrusive, the MMN elicited by deviant sounds within a sequence of standard sounds can partially be overlapped by the N2b (Näätänen & Gaillard, 1983; Näätänen et al., 1982). The N2b is elicited later than the MMN at around 200 to 250 ms from sound onset (for reviews, see Folstein & Van Petten, 2008; Näätänen, Kujala, & Winkler, 2011). The N2b's maximum shows more posterior distribution on the scalp than that of the N1 and the MMN. Also, the N1 and MMN show a polarity reversal at the mastoids, which the N2b does not show.

The N2b indexes a more conscious processing level than the MMN and was suggested to reflect a complementing process of the deviance detection system in case more automatic mechanisms do not sufficiently contribute to deviance detection (for reviews, see Folstein & Van Petten, 2008; Näätänen et al., 2011). For instance, the N2b was larger to task-relevant frequency modulations occurring later than 400 ms after sound onset as compared to frequency modulations at 100, 200, or 300 ms after sound onset indicating that further mechanisms as reflected by the N2b are needed to process the temporal position of the deviant (Grimm & Schröger, 2005). The N2b is usually followed by the P3a component, but it can also occur alone when the discrimination of the features is unsuccessful (Folstein & Van Petten, 2008). Vice versa, the P3a can be elicited by deviant auditory events without the N2b in ignore conditions when deviants are intrusive and catch attention (Escera, Alho, Winkler, & Näätänen, 1998). Thus, research suggests that separate cortical generators underlie the MMN and the N2b (Näätänen & Gaillard, 1983; Ritter & Ruchkin, 1992; Sams, Hämälainen, et al., 1985; Sams, Paavilainen, et al., 1985).

1.3 The MMN and N2b in dyslexia

The MMN and N2b can be used for probing impairments of the subsequent pre-attentive and attentive stages of auditory processing (for review, see Näätänen et al., 2012). MMNs were attenuated in several clinical conditions; usually reflecting diminished behavioural discrimination accuracy (Javitt, Grochowski, Shelley, & Ritter, 1998; Matthews, Todd, Budd, Cooper, & Michie, 2007; Rabinowicz, Silipo, Goldman, & Javitt, 2000). The MMN obtained with the multi-feature paradigm (Näätänen et al., 2004) is useful for establishing an extensive profile of the patient's auditory discrimination skills and also serves as an index for treatment efficacy (e.g., Lovio, Halttunen, Lyytinen, Näätänen, & Kujala, 2012).

Dyslexia is associated with several problems in perceptual processing and attention, which can be probed with ERPs. According to the leading theory, dyslexia results from a linguistic processing deficit, that is, impairments in translating the linguistic input into a phonological code despite accurate auditory perception (Mody, Studdert-Kennedy, & Brady, 1997; Ramus, 2003). Alternative theories have linked developmental dyslexia to various impairments in processing and integrating sensory information (Kujala et al., 2001; Laasonen, Tomma-

Halme, Lahti-Nuuttila, Service, & Virsu, 2000; Ramus et al., 2003; Snowling, 1981, 2000; Vellutino, Fletcher, Snowling, & Scanlon, 2004), or to a more basic auditory processing deficit in perceiving short or rapidly varying sounds (Farmer & Klein, 1995; Tallal, Miller, & Fitch, 1993). Furthermore, it has been postulated that dyslexia results from a neurodevelopmental abnormality of the magnocellular system (the magnocellular model, Galaburda, Menard, & Rosen, 1994; J. Stein & Walsh, 1997). The attentional sluggishness hypothesis (Hari & Renvall, 2001), in turn, proposes that individuals with dyslexia have a prolonged temporal window for processing input chunks that leads to deficits in processing rapid stimulus sequences.

The MMN, and N2b to a lesser extent, have been used to probe deficits in discriminating speech and non-speech sounds in dyslexia. Abnormal auditory processing has even been shown in infants at risk for dyslexia (e.g., Lovio, Näätänen, & Kujala, 2010; van Zuijen et al., 2012). In adults, MMN amplitudes were attenuated for frequency changes in individuals with dyslexia (Baldeweg, Richardson, Watkins, Foale, & Gruzelier, 1999; Kujala, Belitz, Tervaniemi, & Näätänen, 2003; Renvall & Hari, 2003), an impairment that was more prominent in the left hemisphere (Kujala et al., 2003; Renvall & Hari, 2003). In contrast, the MMN amplitude for intensity changes did not differ between readers with dyslexia and fluent readers (Kujala, Lovio, Lepisto, Laasonen, & Näätänen, 2006) and there was even an MMN amplitude enhancement to location changes in readers with dyslexia (Kujala, Lovio, et al., 2006). Some studies reported an aberrant MMN for duration changes in dyslexia (Corbera, Escera, & Artigas, 2006; Huttunen, Halonen, Kaartinen, & Lyytinen, 2007; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1999), whereas other studies showed no MMN amplitude difference between fluent readers and readers with dyslexia (Baldeweg et al., 1999; Kujala, Halmetoja, et al., 2006). Furthermore, MMNs were attenuated for temporal changes in tone patterns in dyslexia (Kujala et al., 2000; van Zuijen et al., 2012).

The MMN enables investigation of deficits in the speech system as it reflects the neural mechanisms associated with speech sound discrimination (Kuuluvainen et al., 2014; Näätänen et al., 1997). MMN amplitudes were attenuated to consonant changes (Lachmann, Berti, Kujala, & Schröger, 2005; Lovio et al., 2010; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998; Sharma et al., 2006) and to vowel changes in children at risk for dyslexia (Lovio et al., 2010). In adult readers with dyslexia, however, MMNs for vowel changes were not different as opposed to fluent readers (Froyen et al., 2011). The discrepancies in these results may be explained by differences in the ages of the participants (children versus adults), by differences in the magnitudes of the stimulus changes, or by different subtypes of dyslexia. For instance, attenuated MMNs were reported in readers with dyslexia who were impaired in reading high frequency words but not in those who were impaired in non-word reading (Lachmann et al., 2005).

Discrimination abilities at different processing levels in dyslexia were also probed with the MMN and the N2b. For instance, duration changes embedded within pseudowords (200 ms deviation of 100 ms long standard vowel) or complex sounds showed no differences in MMN amplitudes between readers with dyslexia and fluent readers (Kujala, Halmetoja, et al., 2006). However, readers with dyslexia had difficulties in detecting duration contrasts attentively as reflected in the lack of N2b responses and poor accuracy in identifying the deviant stimulus segment. These results suggest that even easily discriminable changes eliciting normal MMNs in individuals with dyslexia are difficult to detect when they are embedded in complex word-like stimuli. This aberrant detection process is neurally reflected in the N2b following the MMN.

While the studies reported above suggest an association between auditory processing deficits and dyslexia, follow-up and intervention studies provide more compelling evidence on possible causal factors underlying dyslexia. For example, an inherited risk for dyslexia as reflected by the MMN is evident even in infancy (Leppänen et al., 2010; Leppänen, Pihko, Eklund, & Lyytinen, 1999; Leppänen et al., 2002; Pihko et al., 1999). Follow-up studies have also shown that MMN to e.g., phoneme or rise-time changes predicts later reading deficits at school (Maurer et al., 2009; Maurer, Bucher, Brem, & Brandeis, 2003; Plakas, van Zuijen, van Leeuwen, Thomson, & van der Leij, 2013; van Zuijen et al., 2012). Furthermore, intervention studies showed beneficial effects on reading skills in dyslexia (Temple et al., 2003). For instance, auditory training improved reading skills and enhanced activation of left temporo-parietal cortex and left inferior frontal gyrus in 8–12-year-olds with dyslexia (Kujala et al., 2001; Lovio et al., 2012; Temple et al., 2003). Also, in 7-year-olds with dyslexia, enhanced MMNs for tone-order reversals and improved reading skills were found after non-linguistic audiovisual training (Kujala et al., 2001). Even a brief 3-hour training supporting the connections between letters and speech sounds was found to improve pre-reading skills and to enhance the MMNs to speech sound changes in 6-year-olds at risk for dyslexia (Lovio et al., 2012).

1.4 Letter- speech sound integration investigated with the MMN

1.4.1 The MMN as a probe for audiovisual integration

The MMN can be used to probe audiovisual integration by assessing how the activity in the auditory cortex is affected by visual material. The MMN is, for instance, elicited when a visual deviance induces an illusory perception of an auditory change. MMNm was elicited in the auditory cortex by presenting videotaped face articulating stimuli of non-matching audiovisual deviant syllables (visual /ka/ synchronously with acoustic /pa/) which were perceived as /ta/ (the McGurk effect, McGurk & MacDonald, 1976) among matching audiovisual standard syllables (visual /pa/ synchronously with acoustic /pa/) (Sams et al., 1991). The MMN is also sensitive to the ventriloquist illusion, i.e., a perceptual bias of underestimating the spatial separation of simultaneously presented visual and auditory stimuli (Colin, Radeau, Soquet, Dachy, & Deltenre, 2002).

Furthermore, it was shown that the transient memory system as reflected by the MMN encodes not only single features of bimodal events, but also their conjunctions, regardless of whether there was an illusionary set up or not (Besle, Fort, & Giard, 2005; Bidet-Caulet et al., 2007). In the study of Besle and colleagues (2005), audiovisual standards (tone + ellipse) were presented with occasional changes in the tone frequency of the audiovisual pairs (A′V), or in the orientation of the ellipse (AV′), or in both (A′V′). The participant's task was to respond to changes in a fixation cross in the middle of the screen. The unimodal deviants (A′V, AV′) elicited sensory-specific MMNs and the audiovisual deviants (A′V′) elicited auditory (at frontocentral sites) and visual MMNs (at occipital sites). The visual MMN (V′), which was recorded as a control in a visual-only experiment (ellipse changes without the tones), differed from the visual MMNs in the audiovisual sequences (AV′) indicating that information from both senses interact before the MMN process.

Brain processes associated with predicting rules and regularities in one modality with the information given in the other modality can be probed with the incongruency response (IR), a negative-polarity MMN-like brain response (Widmann, Kujala, Tervaniemi, Kujala, & Schröger, 2004). For example, the IR was elicited at around 100 ms to sounds incongruent with a visual pattern whereas no such response was observed to sounds congruent with a visual pattern (Widmann et al., 2004). This response was associated with a mismatch between the visually induced prediction and the auditory sensory information.

1.4.2 The MMN and letter-speech sound integration

In a pioneering ERP study, letter-speech sound integration was probed with the MMN (Froyen et al., 2008). An auditory-only condition with a deviant speech sound /o/ and standard speech sound /a/ was compared to an audiovisual condition in which a written letter ‘a’ was simultaneously presented with each speech sound used in the auditory-only condition. The participant's task was to watch a silent movie in the auditory-only condition and to press a button to a target color picture in the audiovisual condition. The MMN amplitude was larger in the audiovisual condition than the auditory-only condition. The authors argued that the enhancement was due to a double deviation, that is, the deviant speech sound /o/ violated the neural memory trace formed by the standard speech sound /a/ as well as the neural memory trace formed by the standard letter ‘a’. It was concluded that letters interacted with the sounds before the MMN process indicating that letter-speech sound integration is an early and automatic process (Froyen et al., 2008). In addition, letters were either synchronously presented with the speech sounds or they preceded the sound onset by 100 ms or 200 ms. The MMN amplitude linearly decreased with temporal asynchrony between letters and speech sounds, to the extent that the MMN amplitude was not significantly different between the 100-ms time delay condition and the auditory-only condition. It was concluded that temporal synchrony between letters and speech sounds is needed for integration to occur.

In a follow-up study with school children, the MMN process emerged only after several years of reading education (Froyen et al., 2009). After one year of reading instruction children showed a full mastery of letter knowledge; however, they did not show an effect of letters on speech sound discrimination within the MMN time window. Advanced readers after four years of reading instruction, on the other hand, showed an MMN but only when letters were presented 200 ms before the speech sounds. In addition, there was a late effect at

650 ms after stimulus onset in both beginner and advanced readers for synchronously presented letters and speech sounds. It was concluded that the mapping of letters with sounds was not yet automated in beginner readers, whereas in advanced readers there was some evidence of automatic integration due to the early effect in the asynchronous condition (Froyen et al., 2009). This was interpreted to indicate that the development from mere mapping to automatic integration of letters and speech sounds takes years of reading experience (Blomert, 2011; Blomert & Froyen, 2010; Froyen et al., 2009).

Neural correlates underlying letter-speech sound integration were also explored in children with dyslexia by means of the MMN (Froyen et al., 2011). In the study of Froyen and colleagues (2011), the results of the advanced readers (Froyen et al., 2009) were compared with responses in age-matched readers with dyslexia who behaviourally showed a full mastery of letters after four years of reading experience. Vowel changes elicited an MMN in children with dyslexia, which was comparable with that in controls (Froyen et al., 2009; Froyen et al., 2008) suggesting that vowel discrimination works equally well in readers with dyslexia and fluent readers. However, whereas advanced readers showed larger MMNs in the asynchronous audiovisual condition than in the auditory-only condition (Froyen et al., 2009); no difference in MMN amplitude in children with dyslexia was found between those conditions. The results suggested a deficiency in the automatic modulation of letters of early speech sound processing in children with dyslexia. Furthermore, the late negativity found in advanced readers for the synchronous audiovisual condition (Froyen et al., 2009) was not observed in readers with dyslexia. The late negativity, however, was found in the asynchronous condition in the children with dyslexia, indicating that their neural processes in the integration of letters with speech sounds is less matured than in their age-matched controls.

The role of speech sounds on letter processing, in turn, was investigated with the visual mismatch negativity (vMMN), the visual analogue of the auditory MMN (Czigler, Balazs, & Pato, 2004; Maekawa et al., 2005; Tales, Newton, Troscianko, & Butler, 1999). No differences in vMMNs were found when letter deviants were presented alone or synchronously with speech sounds that corresponded to standard letters (Froyen et al., 2010). Whereas speech sound processing was modulated by the presentation of letters (Froyen et al., 2008), letter processing was not affected by concurrent presentation of speech sounds, suggesting an asymmetric relationship of letters and speech sounds in the mapping process.

There are several limitations in the studies of Froyen and colleagues (2009; 2010; 2008; 2011). Firstly, attention demands between the auditory and the audiovisual condition differed (Froyen et al., 2009; Froyen et al., 2008; Froyen et al., 2011). The participants viewed a silent movie in the auditory-only condition while they viewed letters in the audiovisual condition and responded to a target color picture. Therefore, the difference in ERPs to speech sounds caused by the differences in attention demands between the auditory and audiovisual conditions cannot be excluded from consideration. Furthermore, the enhanced MMN response in the audiovisual condition as compared to the auditory-only condition in the studies of Froyen and colleagues (2009; 2008; 2011) could alternatively reflect the sum of the ERPs to auditory and visual features per se (Giard & Peronnet, 1999) as opposed to genuine integration processes. Therefore, a control condition with non-speech visual stimuli would make it possible to study genuine integration of auditory and visual information.

1.5 Selective attention effects on speech sound processing

The ability to direct our attention selectively to particular sensory inputs enables us to process relevant stimuli further and to ignore irrelevant information (Pashler, 1997). The role of attention on the processing of letters and speech sounds can be examined with ERPs. Selective attention modulates ERPs and their magnetic counterparts elicited by simple tones and speech sounds within the first hundred milliseconds after stimulus onset (e.g., Hari et al., 1989; Hillyard, Hink, Schwent, & Picton, 1973; Näätänen, Gaillard, & Mäntysalo, 1978; Rif, Hari, Hämäläinen, & Sams, 1991; Teder, Kujala, & Näätänen, 1993; Woldorff et al., 1993).

Enhanced negatively-shifted ERPs are elicited by attended tones delivered in a rapid sequence to one ear compared to ERPs elicited by ignored tones delivered in a concurrent sequence to the other ear (Hillyard et al., 1973; Woldorff et al., 1993). These ERPs are composed of N1 and the processing negativity (PN). PN reflects cortical stimulus selection underlying a matching process between sensory information and an attentional trace, an actively formed and maintained neuronal representation of attended stimulus features (Alho, 1992; Michie, Bearpark, Crawford, & Glue, 1990; Näätänen, 1982, 1990, 1992; Näätänen et al., 1978; Näätänen & Michie, 1979). The early part of the negative difference (Nd) between the ERPs for attended and unattended tones has an auditory origin with its maximum at fronto-central sites whereas the late portion is more frontally distributed (Alho, 1987, 1992; Hansen & Hillyard, 1980; Michie et al., 1990). The early Nd to auditory stimuli was found to be distributed more posteriorly in an intermodal setting (selection of auditory stimuli among visual stimuli) than in an intramodal setting (selection of auditory stimuli among other auditory stimuli) indicating that auditory attention recruits slightly different brain networks during intermodal than intramodal contexts (Alho, 1992; Woods, Alho, & Algazi, 1992). Nds are also elicited by spoken syllables and words during selective listening tasks (Hansen, Dickstein, Berka, & Hillyard, 1983; Woods, Hillyard, & Hansen, 1984). For example, Woods and colleagues (1984) found enhanced negative ERPs over the left hemisphere at 50-1000 ms to speech probes (“but” and “a”) in the attended message delivered to one ear compared to ERPs to unattended tone probes at different speech-formant frequencies.

Unattended stimuli not matching the attentional trace elicit the so-called *rejection positivity* (RP) (Alho, 1992; Alho, Töttöla, Reinikainen, Sams, & Näätänen, 1987; Alho, Woods, & Algazi, 1994; Degerman, Rinne, Särkkä, Salmi, & Alho, 2008; Michie et al., 1990). Depending on the task, the RP usually lasts for more than 100 ms and may reflect active suppression of unattended sounds (Alho et al., 1987; Alho et al., 1994). Evidence for suppression of task-irrelevant speech stimuli comes also from a recent fMRI study in which participants selectively attended to independent streams of spoken syllables and written letters, and performed a simple task, a spatial task, or a phonological task (Salo, Rinne, Salonen, & Alho, 2013). Activity in the STS to unattended speech sounds was decreased during a visual phonological task as compared to non-phonological visual tasks (see also, Crottaz-Herbette, Anagnoson, & Menon, 2004). The suppression effects in the STS may indicate that suppression is needed during such a task because performance in the visual phonological task could have easily been distracted by the phonological content of task-irrelevant speech sounds.

2 AIMS OF THE STUDY

This thesis aimed at investigating interactions of cortical processing of letters and speech sounds with ERPs. A series of studies focused on the neural networks involved in the mapping of written and heard syllables (**Study I**), differences between the neural networks of fluent readers versus those with dyslexia (**Study II**), and attentional influences on the processing of letters and speech sounds (**Studies III and IV**).

Study I aimed at determining neural networks associated with the integration of written and heard syllables by using the MMN. To this end, MMNs were recorded to syllable sound changes in combination with either corresponding written syllables or scrambled images of the written syllables. Auditory stimuli included vowel and consonant changes, and changes in intensity, frequency, and vowel length. Visual stimuli were either presented synchronously with auditory stimuli or with a time delay. We expected that speech sound processing would be modulated differently by letters than by non-linguistic visual stimuli, and, further, that letter-speech sound integration would break down with a time delay.

The goal of **Study II** was to assess differences in the neural networks involved in mapping speech sounds with printed text in adult readers with dyslexia and fluent adult readers. We investigated integration of written and heard syllables in readers with dyslexia and fluent readers by using the design of *Study I.* We expected to find abnormal audiovisual syllable processing in the readers with dyslexia as reflected by diminished MMNs compared to fluent readers. Because previous studies reported longer integration times in readers with dyslexia than fluent readers, we also expected a sluggish integration in readers with dyslexia as indicated by delayed MMNs.

Study III aimed at investigating attention effects on the integration of written and spoken syllables. By utilizing a similar paradigm as in *Study I,* we determined the effect of attention on letter-speech sound integration. Attention was directed to 1) the auditory, 2) the visual, 3) both modalities (audiovisual), or 4) away from the stimuli (a mental counting condition). We expected to find an increased and/or earlier MMN/N2 response to speech sounds when presented synchronously with letters during audiovisual attention than during the other three conditions. This would imply that the mapping process of letters with speech sounds is facilitated by attending to both modalities.

With **Study IV**, our aim was to assess selective attention effects on cortical processing of speech sounds and letters. We presented syllables randomly to the left or right ear with a concurrent stream of consonant letters. The participants performed a phonological task or a non-phonological task in the auditory or visual domain, respectively. We expected to find an Nd to attended spoken syllables in relation to unattended spoken syllables as an indication of selective attention effects on speech. In addition, we also expected to find a visual Nd to attended letters during the visual than during the auditory tasks as an evidence of selective attention to letters. We also expected to find an RP in response to unattended spoken syllables delivered to one ear during attention to syllables presented to the other ear indicating that ignored spoken syllables were actively suppressed. In addition, we expected an RP to unattended spoken syllables during a visual phonological task in relation to a visual non-phonological task because suppression for speech stimuli is probably needed more during a linguistic visual task than a non-linguistic visual task.

3 METHODS

3.1 Participants

Participants were healthy adults with no reported neurological deficits or deficits in hearing or vision (**Study I, III, IV**). In **Study II**, adults with dyslexia were compared with an age-matched control group. All participants were monolingual Finnish speakers. Details of the participants for each study are reported in Table 1. The participants gave written informed consent prior to the experiment and received movie tickets, cultural vouchers, or monetary compensation for their participation.

Table 1 Number, gender, age, and test results (of WAIS III FIQ, Phonological processing, Reading) of the participants. The number does not include rejected participants.

	N	**Male/ Female ratio**	**Mean age in years (range)**	**WAIS III FIQ[a,d)]**	**Phonological Processing[b,d)]**	**Reading[c,d)]**
Study I	18	6/12	26.1 (19-31)	N/A	N/A	N/A
Study II	11 readers with dyslexia	5/6	26.3 (17-35)	115.45 (9.6)	5.9 (5.9)	-.35 (15.6)
	16 fluent readers	5/11	27.2 (19-34)	127.81 (20.5)	10.23 (10.3)	9.54 (2.5)
Study III	17	6/11	27.0 (22-43)	N/A	N/A	N/A
Study IV	26	11/15	25.0 (20–43)	N/A	N/A	N/A

[a)] The participants' full scale intelligence quotient (FIQ) was estimated with Wechsler Adult Intelligence Scale-third edition (WAIS-III) subtests vocabulary and matrix reasoning

[b)] Includes: phonological naming (RAS speed and accuracy (Wolf, 1986)), phonological memory (WAIS-III, subtest digit span forward length (Wechsler, 2005), and phonological awareness (Pig Latin (Nevala, Kairaluoma, Ahonen, Aro, & Holopainen, 2006))

[c)] Reading skills (reading speed and accuracy of each participant were assessed with reading a word and a pseudo word list aloud (Nevala et al., 2006)

[d)] Scores represent means and standard deviations.

Approval of **Studies I-III** was acquired from the Ethical Committee of the former Department of Psychology, University of Helsinki, and **Study IV** was approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences. In **Study II**, the performance criterion for adult readers with dyslexia was below -1 standard deviation in reading. The statistical analysis yielded poorer phonological processing and reading skills in readers with dyslexia than in fluent readers, whereas the groups were not significantly different in age or their FIQ.

3.2 Event-related potential measurements

3.2.1 Stimuli

In **Studies I** and **II**, auditory stimuli were Finnish consonant–vowel syllables /te:/ and /pi:/, the standard stimulus having a fundamental frequency (F0) of 101 Hz and a stimulus duration of 170 ms. The syllables were created with a Semisynthetic Speech Generation Method (Alku, Tiitinen, & Näätänen, 1999) from long isolated vowels /i:/ and /e:/ and short words /pe:ti/ and /pito/ uttered by a male Finnish speaker. From those words, the plosive /t/ and /p/ waveforms were extracted. Thereafter, the natural glottal excitation waveform was estimated from the vowel /e:/ and this signal was applied to the vowel tract models of the vowels /e:/ and /i:/, yielding semi-synthetic vowels. The plosive /t/ and /p/ waveforms were added to the beginning of the semi-synthetic vowels to obtain the syllables. In this manner, the spectrum of the consonant was kept the same, independent of which vowel followed it. The deviant syllables differed from the standard in the following parameters: consonant (/pe:/ or /ti:/, respectively); vowel (/ti:/ or /pe:/, respectively); vowel duration (-70 ms), frequency (±8% of F0, 93/109 Hz), and intensity (±6 dB). Corresponding to the auditory syllables, visual stimuli were either written syllables (‘‘tee’’ or ‘‘pii’’, respectively) or scrambled pictures of the written syllables. The target stimuli of the detection task were size (a uniform scaling of 130% was used) and color changes (from white to gray) of one of the three parts of the syllables and scrambled syllables, whereas the distractors included only size or color changes.

In **Study III**, a set of auditory stimuli was used, similar to those in **Studies I and II**, with the exception of the standard syllable /ke:/ in place of /te:/, and that only two deviants were included: consonant (/pe:/ or /ki:/, respectively) and frequency changes (±8% of F0, 93/109 Hz). Auditory target syllables had a duration of 200–280 ms, depending on the participant's individual threshold of 80% detection hit rate determined in a separate session. Corresponding to the spoken syllables, written standard syllables were “kee” and “pii”, and, as control standard stimuli, scrambled pictures of the written syllables were used. Visual deviants were consonant changes in the syllable (or the first part in the scrambled picture, respectively) and visual luminance-deviants (75% or 125% of the standard-stimulus contrast). Visual targets were 300 to 480 ms long, depending on the participant's individual threshold of 80% detection hit rate determined in a separate session.

In **Study IV**, auditory stimuli were eight meaningless consonant-vowel and vowel-consonant syllables, i.e., four starting with a vowel (/ah/, /ak/, /ap/, /at/) and four ending in a vowel (/ku/, /lu/, /mu/, /pu/) with the duration of 250 ms for each syllable. Visual stimuli were eight written consonants: four of the consonant names started with a vowel (L, M, R, S; for example, in English, the name of the letter “R” is pronounced like “are” and thus starts with a vowel) and four ended in a vowel (C, P, T, V; for example, in English, the name of the letter “T” is pronounced like “tea” and thus ends in a vowel). The fonts of the letters were gray: four of them being lighter (R, G, and B values each either 16, 32, 48 or 64) and four of them being darker than the background (R, G, and B values each either 192, 208, 224 or 240).

3.2.2 Experimental paradigms and conditions

In **Studies I** and **II**, the syllable sounds were presented in the multi-feature paradigm (identical to the 'Optimum-1'; Näätänen et al., 2004), wherein the standard alternates with 5 types of deviants (Fig. 1). In this paradigm, every other syllable sound is a standard (*p* = .5) and every other is one of the five deviants (*p* = .1, for each deviant), presented in a pseudo-randomized order, following the rule that the same type of deviant was is never repeated after the standard following it.

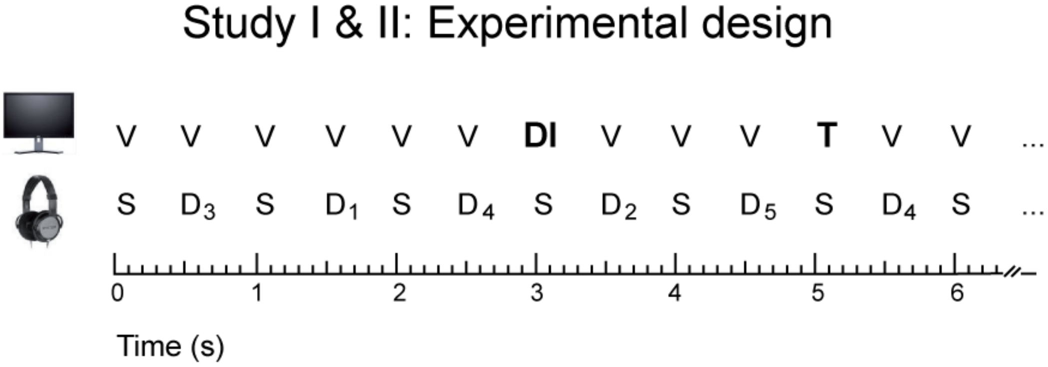


Figure 1. Schematic illustration of the experimental design of **Studies I** and **II**. Auditory stimuli were presented in the multi-feature paradigm including standard (S) and deviant (D_{1}-$_{5}$) syllable sounds (paradigm adapted from Näätänen et al. (2004)) together with corresponding written syllables or scrambled images of the written syllables (V = visual stimuli). The participants responded to visual targets (T) and ignored interspersed distractors (DI).

The experiment included four conditions, in all of which the stimuli were presented with a fixed stimulus-onset asynchrony (SOA) of 670 ms. In two conditions, the on- and offsets of spoken syllables were synchronized with either written syllables (synchronous syllable condition) or scrambled syllables (synchronous scrambled syllable condition). In the other two conditions, the written syllables (asynchronous syllable condition) or scrambled syllables (asynchronous scrambled syllable condition) always preceded the sounds by 200 ms. Participants responded when one part of the written or scrambled syllable changed in size and color (p = .025; targets) and ignored changes in stimuli in one of the following features, size or color (p = .0125 for each feature change; distractors). Participants were instructed to ignore the sounds and focus on the task.

Instead of the five auditory changes used in **Studies I** and **II**, only consonant and F0 changes were presented in **Study III**, since those changes were significantly modulated by synchronous visual letters in **Studies I** and **II**. In addition to auditory changes, consonant and luminance changes were used in the visual domain to keep the level of arousal between the auditory and visual sequences similar. Both auditory and visual stimuli were presented in an oddball sequence, wherein audiovisual standard pairs, synchronously presented spoken and written/scrambled syllables (*p* = .67), were randomly interspersed with deviants in either the visual or auditory domain (*p* = .07 for each deviant type). For target stimuli, duration changes were inserted in the sequences, the length of which was determined in a pre-experiment (participants' individual hit rate was set to 80%). **Study III** included four attentional conditions: auditory attention (A), visual attention (V), audiovisual attention (AV), and mental counting (MC) (Fig. 2).

Figure 2. Schematic illustration of the four attentional conditions: auditory attention (A), visual attention (V), audiovisual attention (AV), and mental counting (MC).

During A conditions, the participants responded whenever they detected a longer spoken syllable and ignored the visual stimuli. During V conditions, the participants responded when they perceived a longer duration visual stimulus and ignored the spoken syllables. During AV conditions, the participants responded when they detected a longer-duration auditory or visual stimulus. In MC conditions, the participants counted backwards mentally from 500 and responded after reaching multiples of ten (490, 480, 470, etc.), while fixating on the middle of the screen and ignoring all stimuli. In the A, V, and MC conditions, the probability of target stimuli was .05, whereas it was set to .25 during AV conditions to keep the overall target probability at .05.

In **Study IV**, independent sequences of auditory and visual stimuli were presented (Fig. 3). For each ear, syllable streams were randomly delivered with SOAs varying between 400 and 600 ms in 10 ms steps. Sequences included syllables spoken by a male voice and ending in a vowel (auditory "standards", *p* = .6), syllables spoken by a female voice and ending in a vowel (*p* = .2), and syllables spoken by a male voice and starting with a vowel (p = .2). Visual letter sequences were randomly delivered with SOAs varying between 400 and 1600 ms in 100 ms steps. Each letter sequence included letters written in lighter-than-background font and ending in a vowel (visual "standards", *p* = .6), letters written in darker-than-background font and their names ending in a vowel (*p* = .2), and letters written in lighter-than-background font and their names starting with a vowel (*p* = .2). Auditory syllables were delivered in a random order except that within each ear, a standard syllable ending in a vowel and spoken by a male voice was always presented after a voice-deviant or phonologically deviant syllable. A similar procedure was used for the visual stimuli i.e., a standard letter written in darker font and with its name ending in a vowel was always presented after a font-shade deviant or phonologically deviant letter. In each of the three auditory and visual stimulus categories, the four different voices/font shades and the four different syllables/letters occurred in a random order. The experiment included six conditions: Phonological and non-phonological left-ear conditions, wherein participants responded to syllables in the left ear starting with a vowel or syllables spoken by female voices, respectively; phonological and non-phonological right-ear conditions, wherein participants responded to syllables in the right ear starting with a vowel or syllables spoken by female voices, respectively; and phonological and non-phonological visual conditions, wherein participants responded to letters when the letter name began with a vowel or to letters written in darker fonts, respectively.

In all studies, the spoken syllables were delivered via headphones at an intensity of 50 dB above each subject's hearing threshold. Stimuli were delivered using Presentation 14.9.07.19.11 software (Neurobehavioral Systems, Inc., Albany, California, USA). The conditions occurred in a counterbalanced order between the participants.

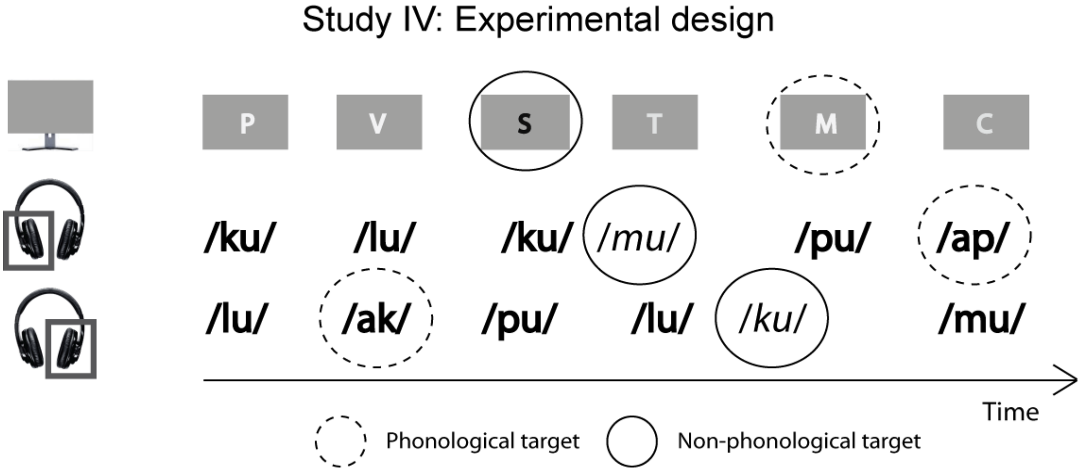


Figure 3. Schematic illustration of the experimental design of **Study IV**. Participants selectively attended to syllables delivered to the left ear or to the right ear and performed a phonological (syllables starting with a vowel, dashed circle) or non-phonological (female spoken syllables, printed in italic) task with the attended syllables. In separate conditions, they responded to visual phonological (letters with a name starting with a vowel, dashed circle) or non-phonological targets (letters darker than the background, continuous circle). Standard syllables were spoken by male voices and started with a consonant (printed in bold).

3.2.3 Data acquisition and analysis

In all Studies, the experiments were carried out in an acoustically and electrically shielded room with the EEG being recorded with 64 active scalp electrodes placed according to the international 10/20 layout (BioSemi ActiveTwo System and ActiView605-Lores, BioSemi

B.V., Amsterdam, the Netherlands). External electrodes were attached to the left and right mastoids and on the tip of the nose. Horizontal and vertical electro-oculogram (EOG) was recorded with electrodes placed near the outer canthus of each eye and with an electrode placed below the left eye. Table 2 shows further details of data acquisition and analysis. The ERPs were baseline corrected with respect to the mean voltage of the 100-ms pre-stimulus, filtered, and separately averaged for each stimulus type.

Table 2 Details of data acquisition and analysis

	Study I-III	**Study IV**
EEG recording bandpass	.1-100 Hz	DC-104 Hz
Sampling rate	256 Hz	512 Hz
Offline reference	Nose	Averaged mastoids
Filtering bandpass	1-25 Hz	.5-30 Hz
Epoch duration	-100-500 ms	-100-700 ms
Artefact rejection	±100 µV	±150 µV
Analysis Software	Matlab/ toolbox eeglab 1)	Besa 5.3 2)

1) 2009b, The Math-Works, Natick, MA./ (Delorme and Makeig 2004) (http://sccn.ucsd.edu/eeglab)
2) Besa Software GmbH, Gräfelfing, Germany

In **Studies I-III**, the change-related response to auditory deviant stimuli was quantified from grand-average difference waveforms by subtracting ERPs to the standard syllables from the ERPs to the corresponding deviant syllables. Mean amplitudes were measured at the FCz (**Study I & II**) and at Oz (**Study III**) as a mean voltage of a ±15 ms time window around peak latency of the difference waveform. In **Study III**, two earlier consecutive 30-ms latency windows immediately preceding the latency window aligned at Oz peak latency were additionally inspected.

In **Study IV**, only standard stimuli were analyzed due to the low number of reliable ERPs to deviant stimuli. Auditory attention effects to left-ear syllables were quantified by subtracting ERPs to left-ear syllables during right-ear phonological and non-phonological tasks from ERPs to right-ear syllables during left-ear phonological and non-phonological tasks (ERPs to right-ear syllables were analyzed accordingly). Attention effects to letters were quantified by subtracting ERPs to letters during auditory phonological tasks from ERPs to letters during the visual phonological task (ERPs during non-phonological tasks were analyzed accordingly). Suppression of speech during auditory tasks was studied by subtracting ERPs to auditory syllables during the visual non-phonological task from ERPs to the same stimuli during phonological and non-phonological tasks of the opposite ears. Suppression of speech during the visual phonological task was examined by subtracting ERPs to auditory syllables during the visual non-phonological task from ERPs to the same auditory syllables during the visual phonological task. The significance of difference-wave amplitudes was tested with t-tests over consecutive 50-ms or 100-ms averaged data points. Time windows in which t-tests exceeded .05 in most conditions were selected for further analysis.

In **Studies I** and **II**, peak latencies were identified from the difference waveforms by retrieving the most negative peak at the FCz electrode at 100–250 ms after the stimulus onset. However, a later window (150–300 ms) was used for the vowel-duration deviant since its stimulus change onset started later. For **Studies III** and **IV**, no individual peak latencies were analyzed because the peaks were often difficult to detect in individual ERP waveforms.

The significance of each response was assessed with t-tests against zero. Differences in amplitudes and latencies between the conditions, stimuli, and the groups were analyzed with the analysis of variance (ANOVA) for repeated measures over selected electrodes depending on the site of the effects. The Greenhouse–Geisser correction of degrees of freedom was applied wherever appropriate and post-hoc tests (Fisher's LSD tests for **Study I**, Bonferroni tests for **Studies II-IV**) were applied to determine the underlying patterns yielding interactions. In the Results section, only *p*-values less than .05 are reported unless otherwise explicitly stated.

4 RESULTS AND DISCUSSION

4.1 Letter-speech sound integration in fluent readers (Study I)

In **Study I**, we investigated the neural networks involved in the mapping of written with heard syllables. We found significantly larger change-related responses for the consonant and frequency deviants in heard syllables when they were presented with written syllables than with scrambled syllables (Fig. 4). In addition, time delay between heard and written material diminished the amplitudes for all deviants (Fig. 5). Participants responded faster when presented with written syllables than with scrambled syllables and when heard and written material was presented synchronously than asynchronously.

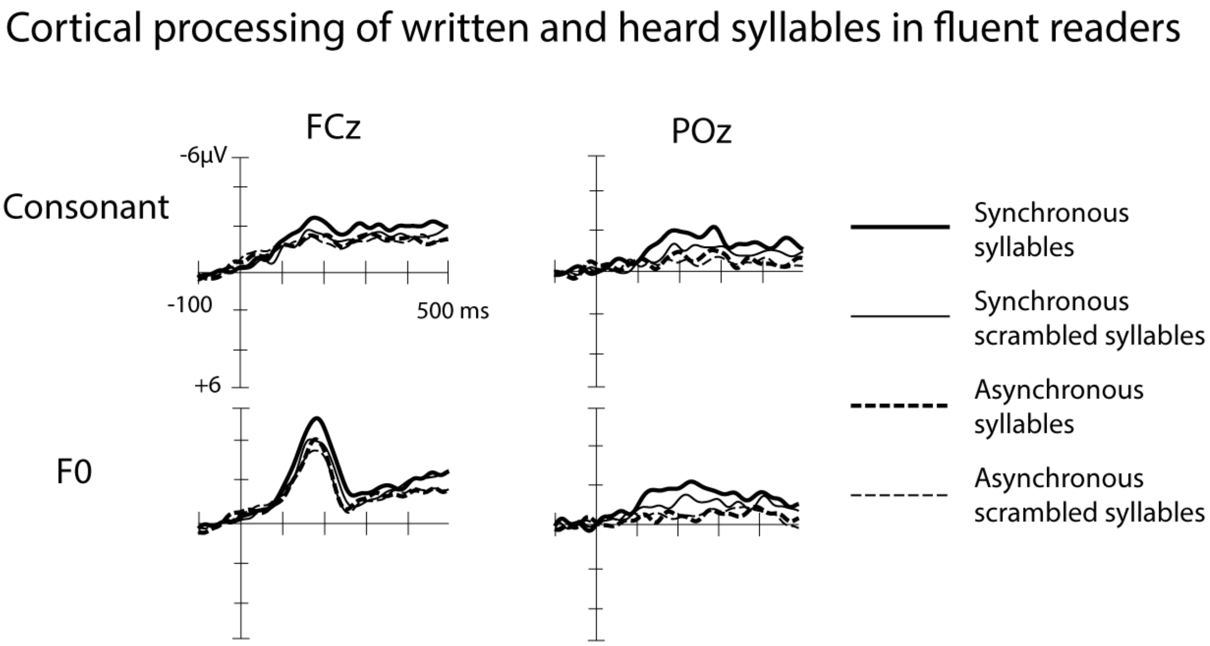


Figure 4. Deviant-minus-standard difference waves for ERPs elicited by consonant and F0 changes at the FCz and POz electrodes when syllable sounds were presented concurrently or with a time delay with written syllables or scrambled text.

Our results suggest that speech sound processing is modulated when the sounds are presented together with written syllables in relation to when they are presented together with non-linguistic visual stimuli, and further, that integration of written and heard syllables depends on their precise temporal alignment. In addition, the results show that a variety of parameters, relevant and irrelevant, for reading can be tested with our paradigm within one experiment. Our results are consistent with previous findings showing an early effect on the ERPs during letter-speech sound integration in adults, which was dependent on accurate temporal alignment of letters and speech sounds (Froyen et al., 2008).

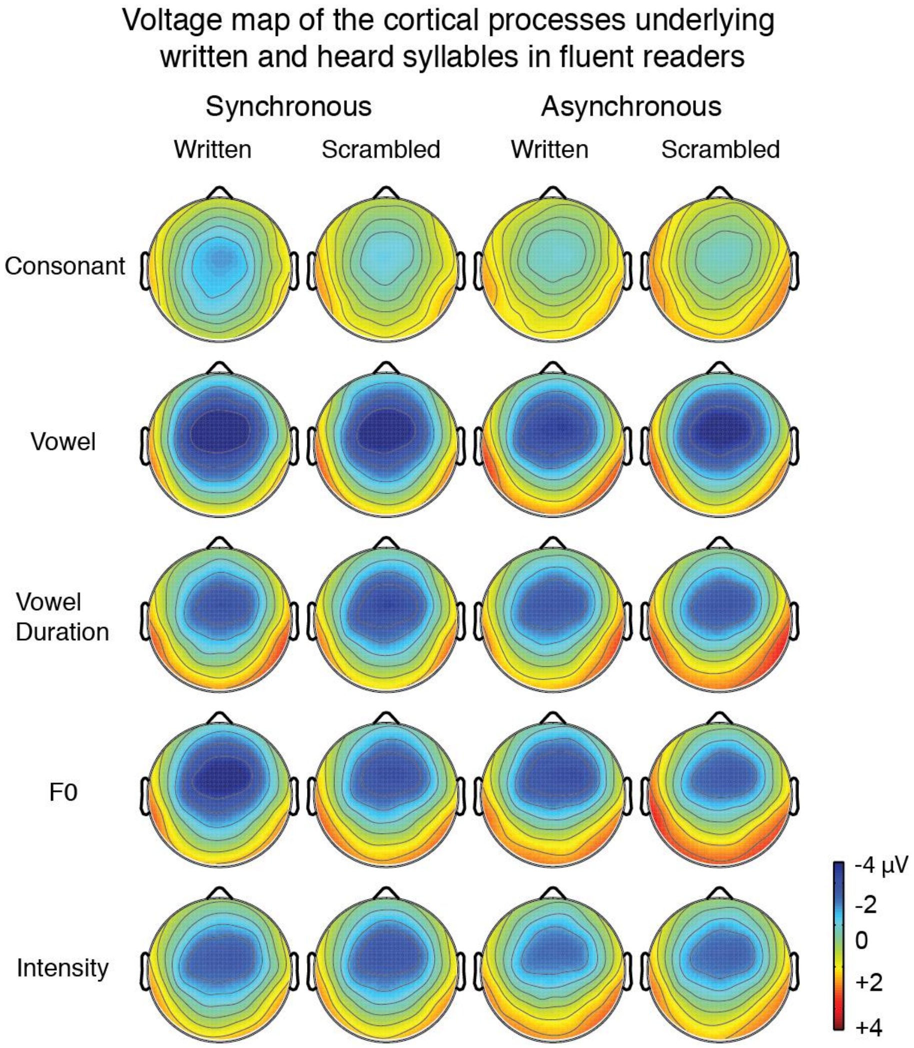


Figure 5. Voltage maps of the grand-average difference waveforms at the +/- 15 ms peak latency interval for all five deviant types in all conditions.

4.2 Letter-speech sound integration in readers with dyslexia (Study II)

Study II examined differences of the neural processes underlying letter-speech sound integration between fluent readers and readers with dyslexia. Fluent readers showed significantly larger N2 amplitudes to deviant syllables when presented synchronously with written syllables than with their scrambled images over the left hemisphere (Figs. 6 and 7). N2 amplitudes in fluent readers were also significantly larger over the left hemisphere than over the right hemisphere when auditory syllables were presented synchronously than asynchronously with written syllables. Additionally, the peak latency was significantly earlier during synchronous presentations than asynchronous presentations in fluent readers. Correspondingly, behavioral results showed faster reaction times in fluent readers when auditory and visual material was synchronously presented than asynchronously presented. Our results for fluent readers support the results of **Study I,** suggesting an early modulation of neural speech sound discrimination by printed text in fluent adult readers, which breaks down with a time delay between heard and written syllables.

Neither visual material nor time between written and heard syllables had an effect on the N2 amplitudes in readers with dyslexia. However, the N2 responses to frequency and consonant deviants peaked later in readers with dyslexia than in fluent readers when heard and written stimuli were presented synchronously.

These results suggest a deficit in speech sound discrimination when presented with written syllables in dyslexia, since, unlike fluent readers, readers with dyslexia showed no distinct effect of written text on speech sound discrimination as reflected by the N2 response. Furthermore, our results of no differences in N2 responses to auditory deviants when presented with written syllables than with symbols in dyslexia could also suggest a general problem in audiovisual processing since readers with dyslexia might, in general, need more resources and/or time to process synchronously different kinds of visual material with sounds. In addition, delayed responses during synchronous presentation of speech sounds and visual material in readers with dyslexia suggest that they, unlike fluent readers, do not profit from synchronous multimodal stimulus presentation.

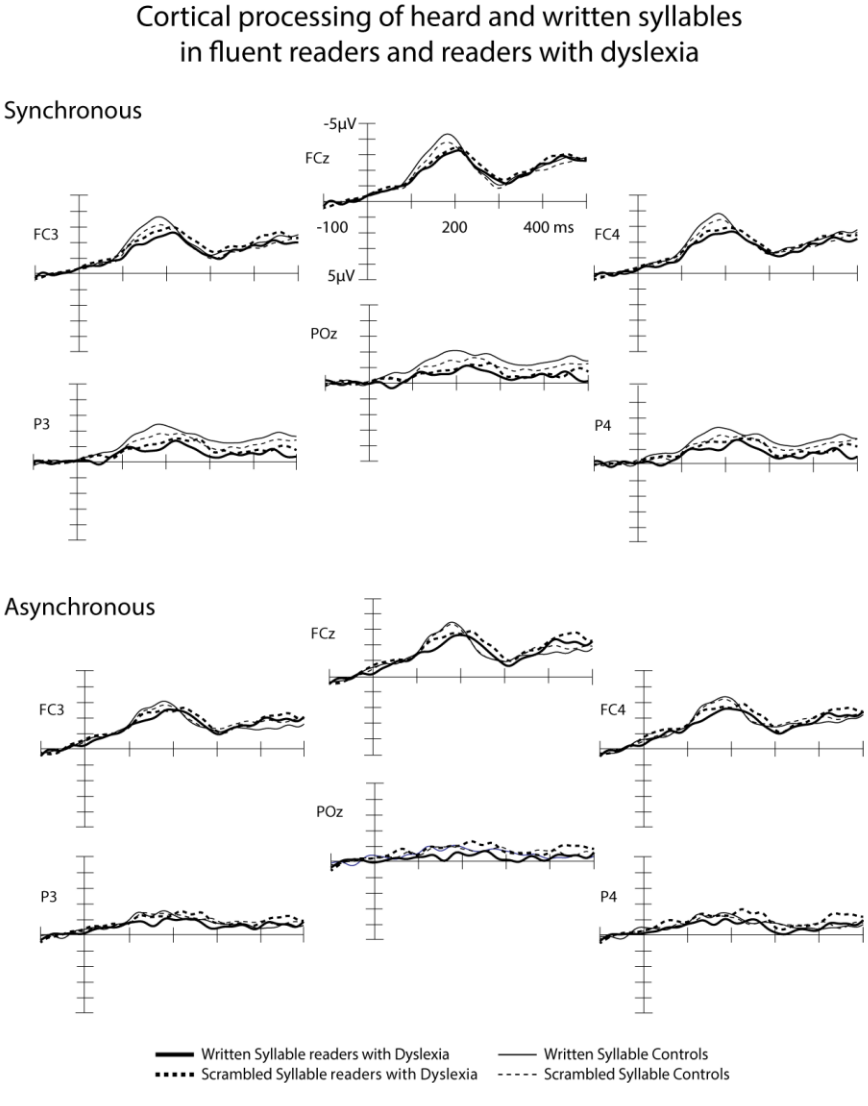


Figure 6. Grand-average difference waveforms averaged over the deviant types in readers with dyslexia and in fluent readers. The syllable sounds were presented either synchronously or asynchronously with written syllables or their scrambled versions.

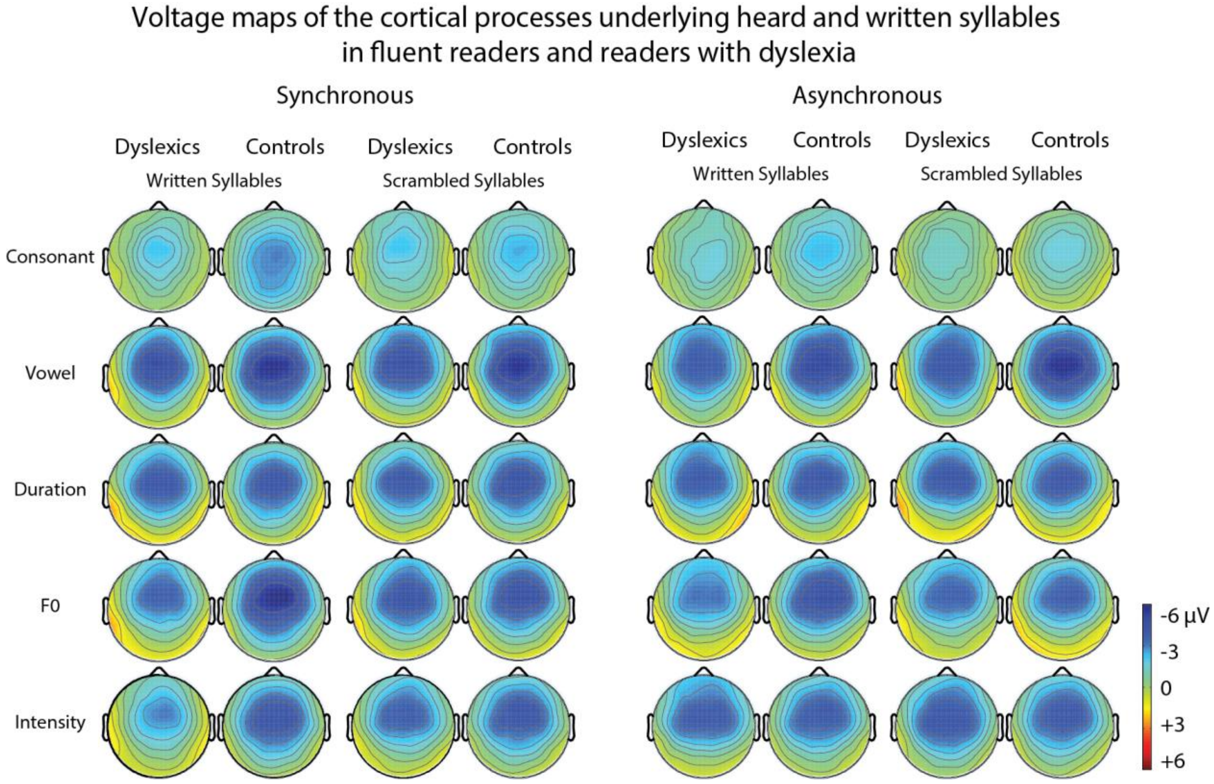


Figure 7. Voltage maps of the grand-average difference waveforms at the +/- 15 ms peak latency interval for all five deviant types in synchronous and asynchronous conditions for fluent readers and for readers with dyslexia.

4.3 Factors influencing letter-speech sound integration (Study III)

In **Study III**, we tested the effects of attention on letter-speech sound integration. We found larger negative responses to consonant changes accompanied by written text than to consonant changes accompanied by scrambled images. This effect occurred in the AV condition at ~140 ms (*first N2 time window*) and in the V condition later at ~200 ms (*third N2 time window*; Fig. 8). We found no such effect of visual material on spoken consonant changes in other conditions or for the F0 changes in any condition (Fig. 9).

The result of enhanced N2 responses to consonant changes accompanied by written syllables during visual attention is consistent with the results of **Studies I** and **II**, in which the participants responded to targets in the visual domain. These results suggest that speech sound discrimination is modulated by attended printed text.

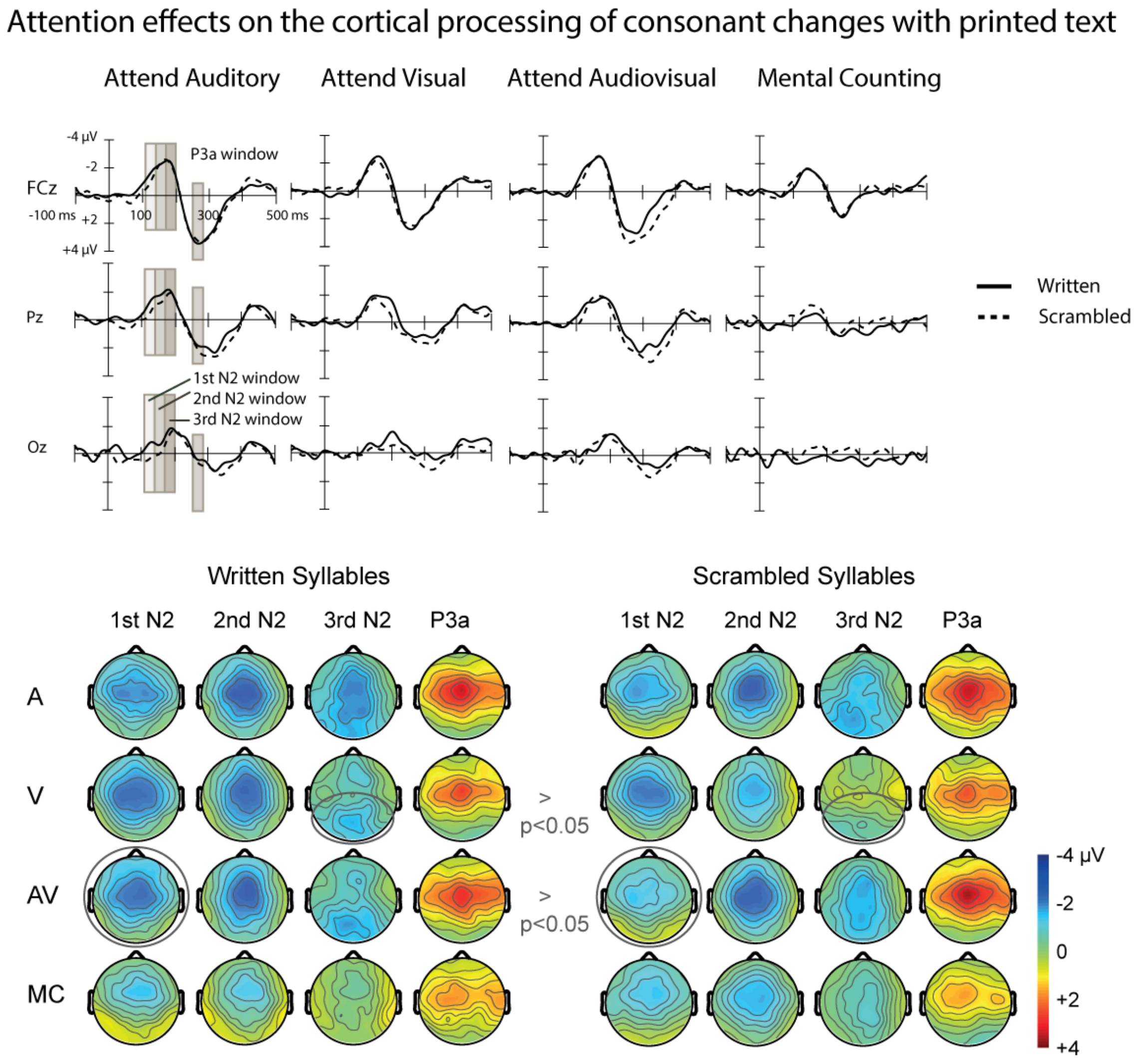


Figure 8. Difference waves and voltage maps for four 30-ms analyses time windows when consonant changes in spoken syllables were presented concurrently with written syllables or scrambled syllables during auditory attention (A), visual attention (V), audiovisual attention (AV), and mental counting (MC) conditions.

Furthermore, we found even earlier integration effects during audiovisual attention. At ~140 ms the responses to spoken consonant contrasts were more negative when accompanied by written text than when accompanied by scrambled images. This result is consistent with fMRI data showing stronger STS activation during attention to audiovisual features than during attention to a single modality (Degerman et al., 2007) and suggests that audiovisual attention boosts integration of written and heard syllables. This effect also is in agreement with our behavioral results yielding lower false alarm rates when processing written syllables during bimodal than unimodal attention.

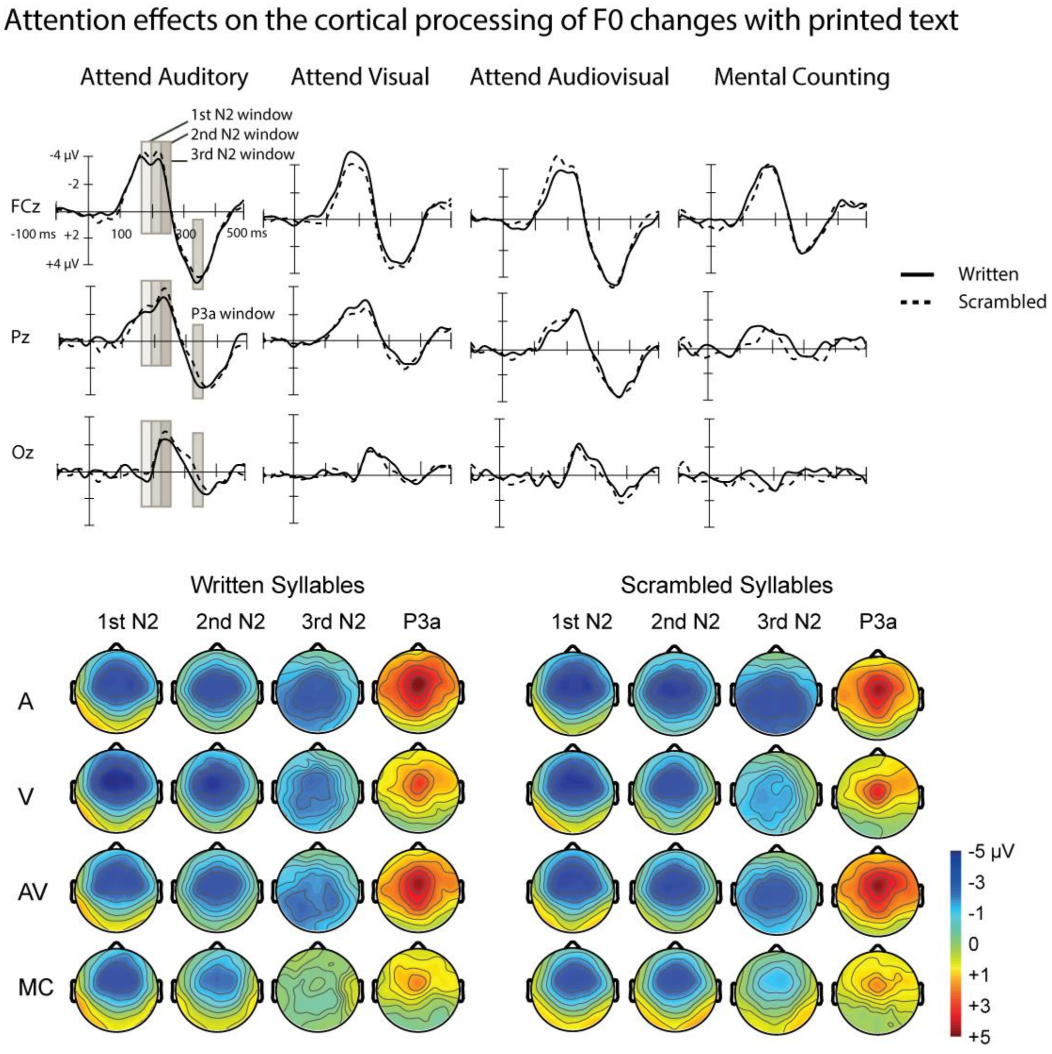


Figure 9. Difference waves and voltage maps for four 30-ms analyses time windows when F0 changes in spoken syllables were presented concurrently with written syllables or scrambled syllables during auditory attention (A), visual attention (V), audiovisual attention (AV), and mental counting (MC) conditions.

4.4 Selective attention effects on the processing of letters and sounds (Study IV)

In **Study IV**, we aimed at assessing selective attention effects on the cortical processing of speech sounds and letters while participants performed an auditory or visual phonological or non-phonological task. We found an early (150–200 ms) and late (300–700 ms) Nd between ERPs to attended and unattended spoken syllables during auditory selective attention, which was not dependent on whether the participants responded to phonological or non-phonological auditory targets (Fig. 10). Our results are consistent with earlier findings showing with tone stimuli that the early and late Nd reflect auditory attention effects (Alho et al., 1994; Degerman et al., 2008; Hansen & Hillyard, 1980).

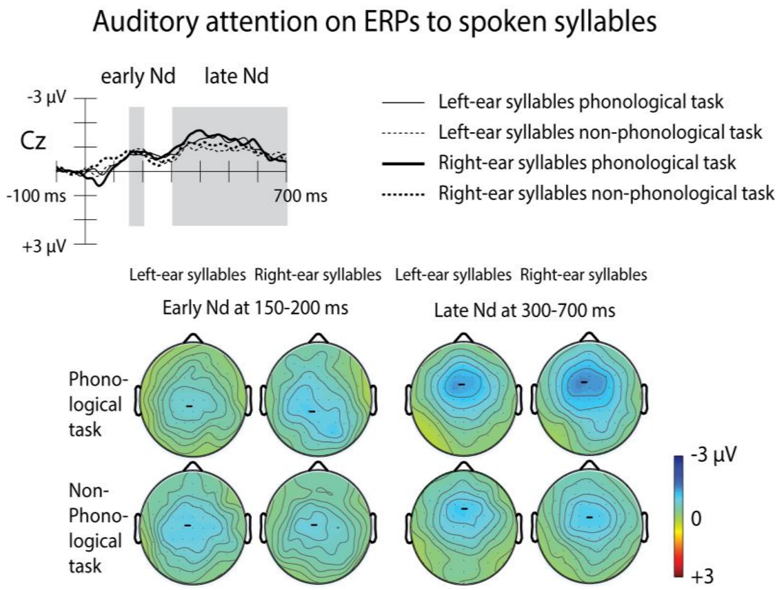


Figure 10. Top: Difference waves for ERPs to spoken syllables: ERPs to unattended syllable sounds were subtracted from ERPs to attended syllable sounds during auditory selective attention (early and late Nds are illustrated). Bottom: Voltage maps for the early and late Nd effects based on difference waves.

Consistent with previous results (e.g., Hillyard & Anllo-Vento, 1998; Salmi, Rinne, Degerman, & Alho, 2007), we also found an Nd (150–250 ms) and a positive difference (Pd; 300–500 ms) by subtracting visual ERPs to letters during auditory attention from visual ERPs during visual attention (Fig. 11). Both visual Nd and Pd were larger for the visual phonological than non-phonological tasks suggesting that phonological processing enhances selective attention.

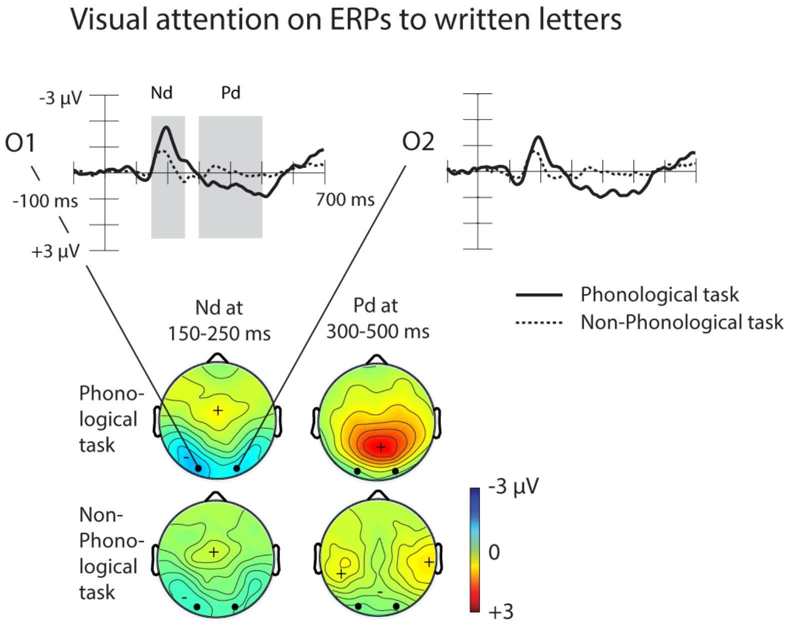
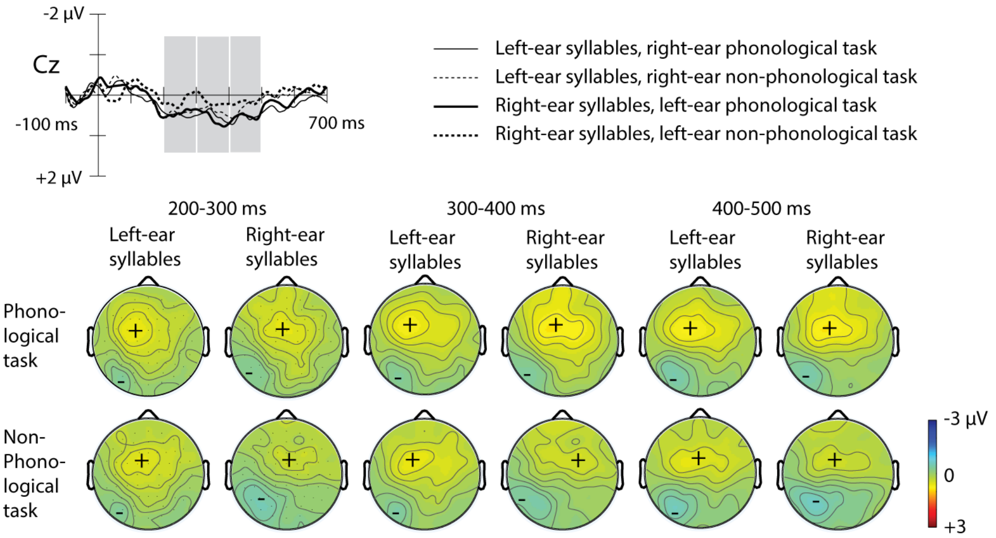


Figure 11. Top: Difference waves for ERPs to letters: visual ERPs during auditory tasks were subtracted from visual ERPs during visual tasks (Nd and Pd are illustrated). Bottom: Voltage maps for the Nd and Pd effects based on difference waves.

ERPs showed a positive displacement to spoken syllables when syllables in the other ear were attended in relation to a visual non-phonological task (Fig. 12a) suggesting active suppression of concurrent irrelevant speech during selective listening to speech sounds. At 200–300 ms, the RP was smaller over the left hemisphere for to-be-ignored right-ear than left-ear syllables, whereas no effect was found over the right hemisphere. Our results could be associated with right ear advantage (REA), that is, suppression is more difficult for the processing of right-ear than left-ear syllables (Alho et al., 2012; Kinsbourne, 1970; Takio, Koivisto, Laukka, & Hämälainen, 2011). At 300–400 ms, RP was larger during phonological than non-phonological tasks, an effect that became more prominent over the left than right hemisphere at 400–500 ms. Our results suggest generally delayed suppression of irrelevant speech during phonological processing of other speech stimuli and that this suppression is predominant in the language-dominant left hemisphere.

Suppression effects on ERPs to spoken syllables

a) Suppression effects during attention to the other ear



b) Suppression effects during visual phonological task

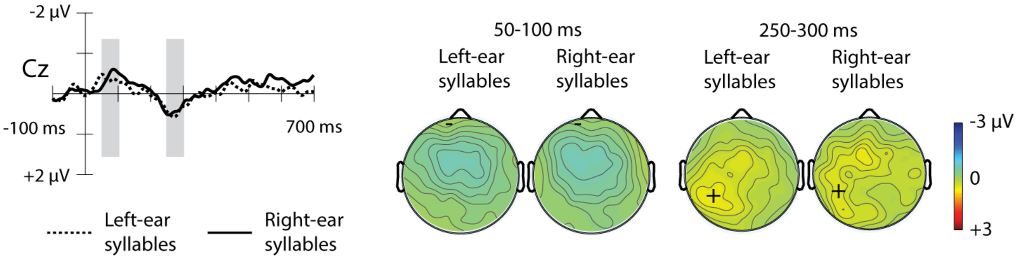


Figure 12. Difference waves and voltage maps for ERPs to unattended syllable sounds during a) auditory (200–300 ms, 300–400 ms, 400–500 ms) and b) visual tasks (50–100 ms, 250–300 ms).

We found stronger RPs to irrelevant speech sounds during the visual phonological task in relation to the visual non-phonological task over the left hemisphere, suggesting a stronger demand to ignore distracting irrelevant speech during a visual phonological task than non-phonological task (Fig. 12b). The RP was preceded by a stronger left-hemisphere dominant negativity at 50 to 100 ms to speech sounds during the visual phonological compared with non-phonological task indicating, possibly, that irrelevant speech automatically intrudes left-hemisphere speech processing systems when processing visual phonological information.

5 GENERAL DISCUSSION

5.1 Letter- speech sound integration

Speech sound discrimination is modulated by concurrent print as suggested by enhanced N2 to F0 and consonant contrasts when synchronously presented together with printed text than when presented together with non-linguistic icons in **Study I**. Converging evidence for early neural modulation of speech sound processing by printed text was reported in the study of Froyen and colleagues (2008) showing enhanced MMNs to vowel contrasts when presented with letters than when presented alone. However, the results in the study of Froyen and colleagues (2008) could also be explained by attention, since attentional demands between the conditions varied, and it was previously suggested that this could affect the MMN amplitude (Muller-Gass, Stelmack, & Campbell, 2006). In the study of Froyen and colleagues (2008), participants were instructed to detect a colored picture in the audiovisual condition whereas they watched a silent movie in the auditory-only condition. In our study, we compared ERPs to speech sound changes during the presentation of printed text with the presentation of printed scrambled text. The participants performed a similar visual detection task during both conditions. This way, we kept the attentional demands between the conditions the same. Therefore, we believe to have obtained genuine integration effects. We also can conclude that the phenomena observed in the study of Froyen and colleagues (2008) are valid, even though demands for attention varied between their conditions. Our results are also consistent with previous research showing a modulation of speech sound processing by printed text at a behavioral level (Dijkstra et al., 1989; Frost et al., 1988; Massaro, 1998) and with MEG and fMRI results reporting higher STS activation when congruent as opposed to incongruent letter-speech sound pairs were presented (Raij et al., 2000; van Atteveldt et al., 2004). Taken together, we conclude that letters influence speech sound processing and discrimination at an early cortical level.

The result of a modulatory effect of print on the F0 contrast was unexpected since F0 contrasts have no correspondence with written symbols in Finnish language. However, tonal features could influence word meaning in Finnish language (Vainio & Järvikivi, 2007). The F0 effect on speech in Finnish is supported by a study reporting larger MMNs elicited by F0 contrasts to speech than to non-speech stimuli, suggesting that pitch processing has a linguistic role in Finnish language (Sorokin, Alku, & Kujala, 2010). However, the modulation of the F0 contrast by print was not replicated in **Study III** and therefore, it might not be reliable.

Against our expectations, we found no effect of print on the vowel and the duration changes, which are both phonological cues for correct perception and production in Finnish language (e.g., Ylinen, Shestakova, Alku, & Huotilainen, 2005). The vowel change in **Study I** elicited the largest N2 amplitudes as compared to all other deviant types. Therefore, the responses elicited by the vowel contrast could reflect a ceiling effect, resulting from a large acoustical difference between the standard and the vowel deviant. The lack of a modulatory effect of printed text on the duration contrast, in turn, could be explained by the possibility that the sound duration differences used in the present thesis were not sufficiently typical for Finnish language. Whereas previous studies found larger MMNs for duration contrasts of 200 vs. 400 ms in a speech than in a non-speech condition, Sorokin and colleagues (2010), who used a smaller duration contrast (120 vs. 170 ms) as also **Study I** (100 vs. 170 ms), found no difference between duration changes in a speech than a non-speech condition. Therefore, future studies should test with different deviance magnitudes the effects of duration contrasts on letter-speech sound processing in quantity languages.

Evidence of early neural modulation of speech perception by printed text is noteworthy since reading is a cultural invention and connections between letters and sounds are artificial. Also, recent studies with Dutch speakers revealed that fluently reading children take years to automate letter-speech sound associations (Froyen et al., 2009) and children with dyslexia hardly exhibit evidence for an integration (Froyen et al., 2011). It was argued that the reason for such effortful learning resides in the arbitrary nature of linking phonological code to letters (Blomert, 2011; Blomert & Froyen, 2010), which is rather artificial as compared to audiovisual processing of speech. Studies showed that integrating letters with sounds does not resemble similar processes underlying the integration of more natural audiovisual objects, such as audiovisual speech (Calvert, 2001; van Atteveldt et al., 2004). For instance, whereas audiovisual speech recruits heteromodal integration sites for bidirectional feedback to visual and auditory cortices (Calvert, 2001), letter-speech sound integration exhibits feedback from the STS area only to the auditory cortex (van Atteveldt, Formisano, Blomert, et al., 2007; van Atteveldt et al., 2004). Also, whereas the time window for integrating audiovisual speech is relatively wide (Massaro, Cohen, & Smeele, 1996; van Wassenhove, Grant, & Poeppel, 2007), the same does not hold for letter-speech sound integration. The integration breaks down when letters and sounds are temporally misaligned, as also indicated by our results of attenuated ERPs to all sound contrasts when presented 200 ms later than letter onset. Our results are in agreement with previous ERP results of smaller MMN amplitudes for spoken vowel contrasts when presented 100 ms after letters, as opposed to synchronously presented with letters suggesting insufficient letter-speech sound integration (Froyen et al., 2008). Also, our results are consistent with fMRI data showing that STS only provides feedback to the auditory cortex when letters and speech sounds are in accurate temporal alignment (for a review, see van Atteveldt et al., 2009).

The functional organization of the adult brain to form new connections between orthography and phonology is influenced by literacy skills during childhood (Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998). Therefore, it is important to study the neural mechanisms of effortlessly learning letter-speech sound connections at a young age. It was shown, for instance, that literate children after 4 years of reading instruction, as opposed to fluent adult readers, do not exhibit MMN response to vowel contrasts when synchronously presented with written letters (Froyen et al., 2009). Studies illuminating the development of generating accurate letter and speech sound connections would promote improvements in dyslexia interventions since impaired letter-speech sound integration was proposed to be a core deficit in dyslexia (Snowling, 1980; for a review, see Vellutino et al., 2004).

5.2 Audiovisual deficit in dyslexia

Based on the present results **(Study II)**, letter-speech sound discrimination is impaired in dyslexia. Unlike in fluent readers, there was no distinct effect of text on the N2 to speech sound contrasts in readers with dyslexia. Our results are consistent with recent findings showing deficient letter-speech sound processing in children with dyslexia (Froyen et al., 2011). Whereas MMNs to vowel contrasts were larger in advanced readers when presented with letters than alone (Froyen et al., 2009), no differences in MMN amplitudes were found between those conditions in age-matched children with dyslexia, indicating that their neural processes underlying letter-speech sound integration is aberrant (Froyen et al., 2011). Our results extend those results by showing that a deficit in the neural networks underlying that mapping process is still present in adulthood, after years of reading experience. Consistent with our results, an fMRI study showed decreased activations to letter-speech sound pairs in the auditory cortex and in the STS in adults with dyslexia, whereas fluent readers exhibited facilitation of congruent and inhibition of incongruent letter-speech sound pairs in the auditory cortex (Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009).

Our results that show no differences in the ERPs in the letter-speech sound and symbol-speech sound conditions in readers with dyslexia might reflect a general problem in processing audiovisual information in dyslexia, as suggested by previous studies (Kujala et al., 2001; Widmann, Schröger, Tervaniemi, Pakarinen, & Kujala, 2012). Namely, in a non-linguistic symbol to sounds matching paradigm, the N2b response was late and reduced over the left hemisphere and the P3a was absent in children with dyslexia (Widmann et al., 2012). The results indicate sluggish and inaccurate mapping and discrimination of symbol-sound patterns in dyslexia. Furthermore, the same study showed a lack of early induced auditory gamma band response (GBR) to congruent symbol-sound pairs and an enhanced desynchronization in GBR to incongruent symbol-sound pairs in children with dyslexia than without, suggesting atypical brain mechanisms in children with dyslexia in matching audiovisual information. Consistent with results showing a general audiovisual deficit in dyslexia (e.g., Widmann et al., 2012), our results could imply that unlike fluent readers, readers with dyslexia treat letters as non-linguistic items, as previously also theorized by Pugh and colleagues (2001). This assumption is supported by behavioral studies showing no difference in reaction times to visual dots than letters in children with dyslexia indicating that they used the same strategies to process letters and dots. In contrast, control readers showed clear differences and hence, processing strategies for these stimuli (Lachmann & van Leeuwen, 2007).

Our results of sluggish responses during synchronous presentation of auditory syllable contrasts with visual material in readers with dyslexia suggest that their neural basis of processing multiple sensory cues presented at one time point is aberrant. This finding is consistent with dyslexia studies suggesting impairment in temporal information segregation, that is, the judgment of the temporal order of sequentially presented stimuli (De Martino, Espesser, Rey, & Habib, 2001; Hari & Renvall, 2001; Tallal, 1980). Deficits in the processing of temporally presented material were found in the auditory, visual, and tactile domain, indicating a general amodal temporal impairment in dyslexia (Laasonen, Service, & Virsu, 2002; Laasonen et al., 2000). Impairments in temporal processing in the auditory modality, especially at a timescale that is relevant for the processing of phonemes, contribute to phonological deficits in dyslexia (Snowling, 1981, 2000; Snowling & Stackhouse, 2006).

In **Studies I** and **II**, we used a newly developed multi-feature paradigm for studying letter-speech sound integration (Näätänen et al., 2004; Pakarinen et al., 2009; Pakarinen et al., 2007). This way, it was possible to investigate the influence of visual print on several speech sound contrasts, which are either phonologically relevant or irrelevant for the participants' mother tongue. Our results suggest that this approach is feasible to address audiovisual integration and its deficits. New versions of the auditory multi-feature paradigm were introduced (Pakarinen et al., 2013; Partanen, Vainio, Kujala, & Huotilainen, 2011; Sandmann et al., 2010; Shtyrov, Kimppa, Pulvermuller, & Kujala, 2011; Thönnessen et al., 2010), for example, by embedding phonologically relevant and irrelevant contrasts within word-like stimuli (Partanen et al., 2013). By incorporating these stimuli in audiovisual paradigms we could, for instance, investigate the neural networks underlying the mapping process of heard and written input at a letter, syllable, or word level. With this approach, it could be determined how the complexity of the context affects audiovisual mapping in dyslexia.

5.3 Top-down effects on letter-speech sound processing

Our results suggest that attention towards the stimuli is needed for an audiovisual integration to emerge. In the mental counting condition, attention was withdrawn from the stimuli and consequently, the letter-speech sound integration effect vanished. Froyen and colleagues (2008) previously suggested that such integration would occur automatically. However, based on their study, the role of attention in these processes could not be reliably determined since there were two conditions only. In one condition, the participants watched a silent movie during the auditory stimulation, whereas during an audiovisual condition they responded to a colored picture occasionally presented among the visual stimuli. By manipulating the direction of attention while keeping the stimulation identical in all conditions, we found that withdrawing the attention from all stimuli abolished the effect, which was present during attention to the visual and both modalities. Thus, it can be concluded that letter-speech sound integration is not an automatic process.

In **Study III**, we found an early negative response already at 140 ms to spoken consonant contrasts when presented together with printed text during audiovisual attention. This novel finding suggests that audiovisual attention boosts letter-speech sound integration. These results are in agreement with fMRI data showing that attention to both modalities is needed for audiovisual integration (Degerman et al., 2007). These current results as well as previous ones indicating audiovisual integration effects for letters with tone pips as early as at the latencies of the P50 (Talsma, Doty, & Woldorff, 2007) suggest that the integration of written and heard syllables takes place rather effortlessly if attention is directed to both modalities. Another reason for the earliness of this effect could be that the Finnish language has a rather transparent orthography. This issue could be investigated by comparing this effect between transparent and opaque languages, such as English. Recent research provides evidence for a different neural tuning for irregular letter-speech sound pairs in English as opposed to regular and therefore, transparent letter-speech sound pairs in Dutch (Holloway, van Atteveldt, Blomert, & Ansari, 2013). For example, facilitation of congruent and inhibition of incongruent letter-speech sound pairs was observed in the STS and STG in Dutch readers (van Atteveldt et al., 2004). However, the reversed pattern of a greater activity in response to incongruent relative to congruent letter-speech sound pairs in the STG was observed in English readers (Holloway et al., 2013).

In **Study IV**, selective-attention effects on letter and speech sound processing were determined with the results suggesting that the processing of letters and syllable sounds is not only affected by the direction of attention, but also by the phonological content of the task. Nd and Pd responses elicited by attended letters were enhanced during the visual phonological in relation to the visual non-phonological task suggesting that there was a higher demand for attention during the more difficult phonological task. In addition, RPs to unattended spoken syllables, presumably associated with active suppression of their processing, were larger during the visual phonological than visual non-phonological tasks over the language-dominant left hemisphere. These results are in line with previous fMRI results (Salo et al., 2013) and suggest a stronger need to ignore distracting irrelevant speech during the visual phonological than non-phonological task. Taken together, our results suggest that the phonological content of the task modulates basic neural mechanisms of selective attention on the processing of heard syllables and written letters.

5.4 Clinical Implications

Our results have possible implications for training programs specifically targeting to improve the accurate generation of letters-speech sound associations in children with dyslexia, since the deficit is thought to be a major impairment in developmental dyslexia (Snowling, 1981). Future studies could determine, by employing the current experimental design (**Study III**), the effects of attention on the integration of print and speech sounds in readers with dyslexia. Based on the results of this thesis for fluent readers (**Study III**), one could assume that children with dyslexia could possibly profit best from such training when attending to both modalities at the same time rather than when attending solely to one modality. However, it could also be that different subtypes of dyslexia, for instance individuals with dyslexia with a more pronounced visual deficit (Lachmann, Steinbrink, Schumacher, & van Leeuwen, 2009; Vellutino et al., 2004), might benefit more from unimodal visual attention than bimodal audiovisual attention.

Previous research showed that impaired oscillations correlate with aberrant audiovisual predictions in children with dyslexia (Widmann et al., 2012). Future studies on dyslexia could investigate to what extent oscillatory networks and their deficiencies correlate with the discrimination of speech sound contrasts when accompanied by print. Audiovisual prediction, as tested with non-linguistic material (Widmann et al., 2012), and the discrimination of speech sound-print pairs, as tested in **Study II** (Mittag et al., 2013), might require the synchronization of different oscillatory networks. Our results and possible implications could bring us one step closer to understanding the core deficit(s) in dyslexia and could consequently promote the development of adequate educational interventions.

5.5 Conclusions

The present thesis investigated cortical processing and discrimination of speech sound contrasts when accompanied by print. The results suggest that letters influence speech sound processing and discrimination at an early cortical level. In contrast to audiovisual speech, this integration is dependent on precise temporal alignment between letters and speech sounds. One reason for such narrow integration window for letter-speech sound pairs could be that connections between letters and sounds are artificial since reading is a cultural invention.

We found distinct effects of perceived letters on speech sound processing in fluent adult readers, but not in readers with dyslexia. This suggests deficient audiovisual mapping of letters with speech sounds in dyslexia. Also, audiovisual processing was delayed in readers with dyslexia suggesting sluggish processing of simultaneous multiple sensory inputs.

Whereas previous research suggested that letter-speech sound associations are processed automatically in the brain, our data show that attention to one of the modalities is needed for an integration to take place and furthermore, that audiovisual attention boosts the integration. In addition, our results suggest that selective attention effects on processing letters and speech sounds are modulated by the phonological content of the task.

Our studies confirmed, with a more controlled methodology than used before, that letters influence speech sound discrimination at an early neural level and that audiovisual attention is most beneficial for such integration to occur. Furthermore, our results illuminate how letter-speech sound integration is aberrant in dyslexia. In addition, our results show that letter-speech sound processing is not only dependent on attention but also on the content of the task. To conclude, our results shed novel light on the basic and aberrant mechanism of letter-speech sound processing and can be used, for instance, for designing training programs to improve audiovisual integration and consequently, reading skills in individuals with dyslexia.

6 REFERENCES

Alho, K. (1987). Mechanisms of selective listening reflected by event-related brain potentials in humans. (Vol. 46, pp. 1–86): Annales Academiae Scientiarum Fennicae. Dissertationes Humanarum Litterarum.

Alho, K. (1992). Selective attention in auditory processing as reflected by event-related brain potentials. *Psychophysiol, 29*(3), 247–263.

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear, 16*(1), 38–51.

Alho, K., Salonen, J., Rinne, T., Medvedev, S. V., Hugdahl, K., & Hämälainen, H. (2012). Attention-related modulation of auditory-cortex responses to speech sounds during dichotic listening. *Brain Res, 1442*, 47–54.

Alho, K., Töttöla, K., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalogr Clin Neurophysiol, 68*(6), 458–470.

Alho, K., Woods, D. L., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiol, 31*(5), 469–479.

Alku, P., Tiitinen, H., & Näätänen, R. (1999). A method for generating natural-sounding speech stimuli for cognitive brain research. *Clin Neurophysiol, 110*(8), 1329–1333.

American Psychiatric Association. (1994). *Diagnostic and statistical manual on mental disorders* (fourth ed.). Washington, DC: American Psychiatric Press.

Baldeweg, T., Richardson, A., Watkins, S., Foale, C., & Gruzelier, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Ann Neurol, 45*(4), 495–503.

Besle, J., Fort, A., & Giard, M. H. (2005). Is the auditory sensory memory sensitive to visual information? *Exp Brain Res, 166*(3–4), 337–344.

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P. E., Giard, M. H., & Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J Neurosci, 27*(35), 9252–9261.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Curr Biol, 19*(6), 503–508.

Blau, V., van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2008). Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *Eur J Neurosci, 28*(3), 500–509.

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage, 57*(3), 695–703.

Blomert, L., & Froyen, D. (2010). Multi-sensory learning and learning to read. *Int J Psychophysiol, 77*(3), 195–204.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb Cortex, 11*(12), 1110–1123.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121 (Pt 6)*, 1053–1063.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nat Neurosci, 1*(5), 351–353.

Colin, C., Radeau, M., Soquet, A., Dachy, B., & Deltenre, P. (2002). Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clin Neurophysiol, 113*(4), 507–518.

Corbera, S., Escera, C., & Artigas, J. (2006). Impaired duration mismatch negativity in developmental dyslexia. *Neuroreport, 17*(10), 1051–1055.

Crottaz-Herbette, S., Anagnoson, R. T., & Menon, V. (2004). Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. *Neuroimage, 21*(1), 340–351.

Csépe, V. (1995). On the origin and development of the mismatch negativity. *Ear Hear, 16*(1), 91–104.

Czigler, I., Balazs, L., & Pato, L. G. (2004). Visual change detection: event-related potentials are dependent on stimulus location in humans. *Neurosci Lett, 364*(3), 149–153. doi: 10.1016/j.neulet.2004.04.048

De Martino, S., Espesser, R., Rey, V., & Habib, M. (2001). The "temporal processing deficit" hypothesis in dyslexia: new experimental evidence. *Brain Cogn, 46*(1-2), 104–108.

Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I. P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *Neuroimage, 34*(4), 1683–1691.

Degerman, A., Rinne, T., Särkkä, A. K., Salmi, J., & Alho, K. (2008). Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. *Eur J Neurosci, 27*(12), 3329–3341.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *J Cogn Neurosci, 12*(4), 635–647.

Dijkstra, T., Schreuder, R., & Frauenfelder, U. H. (1989). Grapheme context effects on phonemic processing. *Lang Speech, 32*(2), 89–108.

Ehri, L. C. (2005). Development of sight word reading: phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: a handbook.* (pp. 135–154). Oxford: Blackwell Publishing.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *J Cogn Neurosci, 10*(5), 590–604.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: a review. *Psychol Bull Rev, 2*, 460–493.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152–170. doi: 10.1111/j.1469-8986.2007.00602.x

Frost, R., & Kampf, M. (1993). Phonetic recoding of phonologically ambiguous printed words. *J Exp Psychol Learn Mem Cogn, 19*(1), 23–33.

Frost, R., Repp, B. H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory and Language, 27*, 741–755.

Froyen, D., Bonte, M. L., van Atteveldt, N., & Blomert, L. (2009). The long road to automation: neurocognitive development of letter-speech sound processing. *J Cogn Neurosci, 21*(3), 567–580.

Froyen, D., van Atteveldt, N., & Blomert, L. (2010). Exploring the Role of Low Level Visual Processing in Letter-Speech Sound Integration: A Visual MMN Study. *Front Integr Neurosci, 4*, 9.

Froyen, D., van Atteveldt, N., Bonte, M., & Blomert, L. (2008). Cross-modal enhancement of the MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neurosci Lett, 430*(1), 23–28.

Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Dev Sci, 14*(4), 635–648.

Galaburda, A. M., Menard, M. T., & Rosen, G. D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proc Natl Acad Sci U S A, 91*(17), 8010–8013.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol, 120*(3), 453–463.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Seperate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *J Cogn Neurosci, 7*, 133–143.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci, 11*(5), 473–490.

Grimm, S., & Schröger, E. (2005). Pre-attentive and attentive processing of temporal and frequency characteristics within long sounds. *Brain Res Cogn Brain Res, 25*(3), 711–721.

Hansen, J. C., Dickstein, P. W., Berka, C., & Hillyard, S. A. (1983). Event-related potentials during selective attention to speech sounds. *Biol Psychol, 16*(3-4), 211–224.

Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalogr Clin Neurophysiol, 49*(3-4), 277–290.

Hari, R., Hämäläinen, M., Kaukoranta, E., Mäkelä, J., Joutsiniemi, S. L., & Tiihonen, J. (1989). Selective listening modifies activity of the human auditory cortex. *Exp Brain Res, 74*(3), 463–470.

Hari, R., & Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends Cogn Sci, 5*(12), 525–532.

Herdman, A. T., Fujioka, T., Chau, W., Ross, B., Pantev, C., & Picton, T. W. (2006). Cortical oscillations related to processing congruent and incongruent grapheme-phoneme pairs. *Neurosci Lett, 399*(1-2), 61–66.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proc Natl Acad Sci U S A, 95*(3), 781–787.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182*(4108), 177–180.

Hillyard, S. A., Mangun, G. R., Woldorff, M. G., & Luck, S. J. (1995). *Neural systems mediating selective attention. Handbook of cognitive neuroscience.* Cambridge, MA: MIT Press.

Holloway, I. D., van Atteveldt, N., Blomert, L., & Ansari, D. (2013). Orthographic Dependency in the Neural Correlates of Reading: Evidence from Audiovisual Integration in English Readers. *Cereb Cortex*.

Huttunen, T., Halonen, A., Kaartinen, J., & Lyytinen, H. (2007). Does mismatch negativity show differences in reading-disabled children compared to normal children and children with attention deficit? *Dev Neuropsychol, 31*(3), 453–470.

Javitt, D. C., Grochowski, S., Shelley, A. M., & Ritter, W. (1998). Impaired mismatch negativity (MMN) generation in schizophrenia as a function of stimulus deviance, probability, and interstimulus/interdeviant interval. *Electroencephalogr Clin Neurophysiol, 108*(2), 143–153.

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychol (Amst), 33*, 193–201.

Kujala, T., Belitz, S., Tervaniemi, M., & Näätänen, R. (2003). Auditory sensory memory disorder in dyslexic adults as indexed by the mismatch negativity. *Eur J Neurosci, 17*(6), 1323–1327.

Kujala, T., Halmetoja, J., Näätänen, R., Alku, P., Lyytinen, H., & Sussman, E. (2006). Speech- and sound-segmentation in dyslexia: evidence for a multiple-level cortical impairment. *Eur J Neurosci, 24*(8), 2420–2427.

Kujala, T., Karma, K., Ceponiene, R., Belitz, S., Turkkila, P., Tervaniemi, M., & Näätänen, R. (2001). Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children. *Proc Natl Acad Sci U S A, 98*(18), 10509–10514.

Kujala, T., Lovio, R., Lepisto, T., Laasonen, M., & Näätänen, R. (2006). Evaluation of multi-attribute auditory discrimination in dyslexia with the mismatch negativity. *Clin Neurophysiol, 117*(4), 885–893.

Kujala, T., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J., & Näätänen, R. (2000). Basic auditory dysfunction in dyslexia as demonstrated by brain activity measurements. *Psychophysiol, 37*(2), 262–266.

Kujala, T., & Näätänen, R. (2010). The adaptive brain: a neurophysiological perspective. *Prog Neurobiol, 91*(1), 55–67.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biol Psychol, 74*(1), 1–19.

Kuuluvainen, S., Nevalainen, P., Sorokin, A., Mittag, M., Partanen, E., Putkinen, V., . . . Kujala, T. (2014). The neural basis of sublexical speech and corresponding nonspeech processing: a combined EEG-MEG study. *Brain Lang, 130*, 19–32.

Laasonen, M., Service, E., & Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain Lang, 80*(3), 340–354.

Laasonen, M., Tomma-Halme, J., Lahti-Nuuttila, P., Service, E., & Virsu, V. (2000). Rate of information segregation in developmentally dyslexic children. *Brain Lang, 75*(1), 66–81.

Lachmann, T., Berti, S., Kujala, T., & Schröger, E. (2005). Diagnostic subgroups of developmental dyslexia have different deficits in neural processing of tones and phonemes. *Int J Psychophysiol, 56*(2), 105–120.

Lachmann, T., Steinbrink, C., Schumacher, B., & van Leeuwen, C. (2009). Different letter-processing strategies in diagnostic subgroups of developmental dyslexia also occur in a transparent orthography: reply to a commentary by Spinelli et al. *Cogn Neuropsychol, 26*(8), 759–768.

Lachmann, T., & van Leeuwen, C. (2007). Paradoxical enhancement of letter recognition in developmental dyslexia. *Dev Neuropsychol, 31*(1), 61–77.

Lavikainen, J., Huotilainen, M., Pekkonen, E., Ilmoniemi, R. J., & Näätänen, R. (1994). Auditory stimuli activate parietal brain regions: a whole-head MEG study. *Neuroreport, 6*(1), 182–184.

Leppänen, P. H., Hämälainen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., . . . Lyytinen, H. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex, 46*(10), 1362–1376.

Leppänen, P. H., Pihko, E., Eklund, K. M., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects. *Neuroreport, 10*(5), 969–973.

Leppänen, P. H., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Dev Neuropsychol, 22*(1), 407–422.

Levänen, S., Ahonen, A., Hari, R., McEvoy, L., & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cereb Cortex, 6*(2), 288–296.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol, 92*(3), 204–214.

Lovio, R., Halttunen, A., Lyytinen, H., Näätänen, R., & Kujala, T. (2012). Reading skill and neural processing accuracy improvement after a 3-hour intervention in preschoolers with difficulties in reading-related skills. *Brain Res, 1448*, 42–55.

Lovio, R., Näätänen, R., & Kujala, T. (2010). Abnormal pattern of cortical speech feature discrimination in 6-year-old children at risk for dyslexia. *Brain Res, 1335*, 53–62.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.

Maekawa, T., Goto, Y., Kinukawa, N., Taniwaki, T., Kanba, S., & Tobimatsu, S. (2005). Functional characterization of mismatch negativity to a visual stimulus. *Clin Neurophysiol, 116*(10), 2392–2402.

Massaro, D. W. (1998). *Illusions and issues in bimodal speech perception.* Paper presented at the Proceedings of Auditory Visual Speech Perception (AVSP'98) Terrigal-Sydney Australia.

Massaro, D. W., Cohen, M. C., & Thompson, L. A. (1988). Visible language in speech perception: lipreading and reading. *Visible Lang, 22*, 8–31.

Massaro, D. W., Cohen, M. M., & Smeele, P. M. (1996). Perception of asynchronous and conflicting visual and auditory speech. *J Acoust Soc Am, 100*(3), 1777–1786.

Matthews, N., Todd, J., Budd, T. W., Cooper, G., & Michie, P. T. (2007). Auditory lateralization in schizophrenia--mismatch negativity and behavioral evidence of a selective impairment in encoding interaural time cues. *Clin Neurophysiol, 118*(4), 833–844.

Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biol Psychiatry, 66*(4), 341-348. doi: 10.1016/j.biopsych.2009.02.031

Maurer, U., Bucher, K., Brem, S., & Brandeis, D. (2003). Altered responses to tone and phoneme mismatch in kindergartners at familial dyslexia risk. *Neuroreport, 14*(17), 2245–2250.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.

Michie, P. T., Bearpark, H. M., Crawford, J. M., & Glue, L. C. (1990). The nature of selective attention effects on auditory event-related potentials. *Biol Psychol, 30*(3), 219–250.

Mittag, M., Thesleff, P., Laasonen, M., & Kujala, T. (2013). The neurophysiological basis of the integration of written and heard syllables in dyslexic adults. *Clin Neurophysiol, 124*(2), 315–326.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: auditory processing or phonological coding? *J Exp Child Psychol, 64*(2), 199–231.

Muller-Gass, A., Stelmack, R. M., & Campbell, K. B. (2006). The effect of visual task difficulty and attentional direction on the detection of acoustic change as indexed by the Mismatch Negativity. *Brain Res, 1078*(1), 112–130.

Näätänen, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. *Psychol Bull*(92), 605–640.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral Brain Sciences, 13,* 201–288.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.

Näätänen, R., & Alho, K. (1997). Mismatch negativity--the measure for central sound representation accuracy. *Audiol Neurootol, 2*(5), 341–353.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol (Amst), 42*(4), 313–329.

Näätänen, R., & Gaillard, A. W. K. (1983). The orienting reflex and the N2 deflection of the event-related potential (ERP). In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event related potential research: endogenous components.* (pp. 119–141). Amsterdam: North Holland.

Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., & Ponton, C. (2012). The mismatch negativity (MMN)--a unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clin Neurophysiol, 123*(3), 424–458.

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology, 48*(1), 4–22.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*(6615), 432–434.

Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. *Biol Psychol, 8*(2), 81–136.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol, 118*(12), 2544–2590.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clin Neurophysiol, 115*(1), 140–144.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophys, 24*(4), 375–425.

Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biol Psychol, 14*(1-2), 53–98.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends Neurosci, 24*(5), 283–288.

Nevala, J., Kairaluoma, L., Ahonen, T., Aro, M., & Holopainen, L. (2006). Lukemis- ja kirjoittamistaitojen yksilötestistö nuorille ja aikuisille (Standardization version ed.). *Jyväskylä: Niilo Mäki Instituutti.*

Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities. *Electroencephalogr Clin Neurophysiol, 78*(6), 466–479.

Pakarinen, S., Lovio, R., Huotilainen, M., Alku, P., Näätänen, R., & Kujala, T. (2009). Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biol Psychol, 82*(3), 219–226.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clin Neurophysiol, 118*(1), 177–185.

Pakarinen, S., Teinonen, T., Shestakova, A., Kwon, M. S., Kujala, T., Hämälainen, H., . . . Huotilainen, M. (2013). Fast parametric evaluation of central speech-sound processing with mismatch negativity (MMN). *Int J Psychophysiol, 87*(1), 103–110

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proc Natl Acad Sci U S A, 110*(37), 15145–15150.

Partanen, E., Vainio, M., Kujala, T., & Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. *Psychophysiology, 48*(10), 1372–1380.

Pashler, H. E. (1997). *The Psychology of Attention*. Boston, M.A.: MIT Press.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., . . . Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology, 37*(2), 127–152.

Picton, T. W., Lins, O. G., & Scherg, M. (1995). The recording and analysis of event-related potentials. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology.* (Vol. 10, pp. 3–73). Amsterdam: Elsevier Science B.V.

Pihko, E., Leppänen, P. H., Eklund, K. M., Cheour, M., Guttorm, T. K., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: I. Age effects. *Neuroreport, 10*(5), 901–905.

Plakas, A., van Zuijen, T., van Leeuwen, T., Thomson, J. M., & van der Leij, A. (2013). Impaired non-speech auditory processing at a pre-reading age is a risk-factor for dyslexia but not a predictor: an ERP study. *Cortex, 49*(4), 1034–1045.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *J Commun Disord, 34*(6), 479–492.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Näätänen, R. (2001). Memory traces for words as revealed by the mismatch negativity. *Neuroimage, 14*(3), 607–616.

Rabinowicz, E. F., Silipo, G., Goldman, R., & Javitt, D. C. (2000). Auditory sensory dysfunction in schizophrenia: imprecision or distractibility? *Arch Gen Psychiatry, 57*(12), 1149–1155.

Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron, 28*(2), 617-625.

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Curr Opin Neurobiol, 13*(2), 212–218.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain, 126*(Pt 4), 841–865.

Renvall, H., & Hari, R. (2003). Diminished auditory mismatch fields in dyslexic adults. *Ann Neurol, 53*(5), 551–557.

Rif, J., Hari, R., Hämäläinen, M. S., & Sams, M. (1991). Auditory attention affects two different areas in the human supratemporal cortex. *Electroencephalogr Clin Neurophysiol, 79*(6), 464–472.

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage, 12*(1), 14–19.

Ritter, W., & Ruchkin, D. S. (1992). A review of event-related potential components discovered in the context of studying P3. *Ann N Y Acad Sci, 658*, 1–32.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev Sci, 8*(2), 162–172.

Salmi, J., Rinne, T., Degerman, A., & Alho, K. (2007). Orienting and maintenance of spatial attention in audition and vision: an event-related brain potential study. *Eur J Neurosci, 25*(12), 3725–3733.

Salo, E., Rinne, T., Salonen, O., & Alho, K. (2013). Brain activity during auditory and visual phonological, spatial and simple discrimination tasks. *Brain Res, 1496*, 55–69.

Sams, M., Aulanko, R., Hämälainen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neurosci Lett, 127*(1), 141–145.

Sams, M., Hämälainen, M., Antervo, A., Kaukoranta, E., Reinikainen, K., & Hari, R. (1985). Cerebral neuromagnetic responses evoked by short auditory stimuli. *Electroencephalogr Clin Neurophysiol, 61*(4), 254–266.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol, 62*(6), 437–448.

Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Meyer, M. (2010). Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clin Neurophysiol, 121*(12), 2070–2082.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998). Auditory processing and dyslexia: evidence for a specific speech processing deficit. *Neuroreport, 9*(2), 337–340.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1999). Pre-attentive processing of auditory patterns in dyslexic human subjects. *Neurosci Lett, 276*(1), 41–44.

Sharma, M., Purdy, S. C., Newall, P., Wheldall, K., Beaman, R., & Dillon, H. (2006). Electrophysiological and behavioral evidence of auditory processing deficits in children with reading disorder. *Clin Neurophysiol, 117*(5), 1130–1144.

Shtyrov, Y., Kimppa, L., Pulvermuller, F., & Kujala, T. (2011). Event-related potentials reflecting the frequency of unattended spoken words: a neuronal index of connection strength in lexical memory circuits? *Neuroimage, 55*(2), 658–668.

Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., & Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *Neuroimage, 12*(6), 657–663.

Snowling, M. J. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. *J Exp Child Psychol, 29*(2), 294–305.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychol Res, 43*(2), 219–234.

Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, UK: Blackwell Publishers Ltd.

Snowling, M. J., & Stackhouse, J. (2006). *Dyslexia, Speech and Language: A Practitioner's Handbook* (2nd ed.). London: Whurr Publishers Ltd.

Sorokin, A., Alku, P., & Kujala, T. (2010). Change and novelty detection in speech and non-speech sound streams. *Brain Res, 1327*, 77–90.

Stein, B. E., & Mededith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA MIT Press.

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends Neurosci, 20*(4), 147–152.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187–1188.

Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neurosci Lett, 266*(2), 109–112.

Takio, F., Koivisto, M., Laukka, S. J., & Hämälainen, H. (2011). Auditory rightward spatial bias varies as a function of age. *Dev Neuropsychol, 36*(3), 367–387.

Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual modality. *Neuroreport, 10*(16), 3363–3367.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang, 9*(2), 182–198.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann N Y Acad Sci, 682*, 27–47.

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex, 17*(3), 679–690.

Teder, W., Kujala, T., & Näätänen, R. (1993). Selection of speech messages in free-field listening. *Neuroreport, 5*(3), 307–309.

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proc Natl Acad Sci U S A, 100*(5), 2860–2865.

Thönnessen, H., Boers, F., Dammers, J., Chen, Y. H., Norra, C., & Mathiak, K. (2010). Early sensory encoding of affective prosody: neuromagnetic tomography of emotional category changes. *Neuroimage, 50*(1), 250–259.

Tuomainen, J., Andersen, T. S., Tiippana, K., & Sams, M. (2005). Audio-visual speech perception is special. *Cognition, 96*(1), B13–22.

Vainio, M., & Jarvikivi, J. (2007). Focus in production: tonal shape, intensity and word order. *J Acoust Soc Am, 121*(2), EL55–61.

van Atteveldt, N., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex, 17*(4), 962–974.

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron, 43*(2), 271–282.

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2007). Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage, 36*(4), 1345–1360.

van Atteveldt, N., Roebroeck, A., & Goebel, R. (2009). Interaction of speech and script in human auditory cortex: insights from neuro-imaging and effective connectivity. *Hear Res, 258*(1-2), 152–164. van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia, 45*(3), 598–607.

van Zuijen, T. L., Plakas, A., Maassen, B. A., Been, P., Maurits, N. M., Krikhaar, E., van der Leij, A. (2012). Temporal auditory processing at 17 months of age is associated with preliterate language comprehension and later word reading fluency: an ERP study. *Neurosci Lett, 528*(1), 31–35.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *J Child Psychol Psychiatry, 45*(1), 2–40.

Wechsler, D. (2005). *Wechsler Adult Intelligence Scale-Third edition: Manual*. Helsinki, Finland: Psykologien Kustannus OY.

Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schröger, E. (2004). From symbols to sounds: visual symbolic information activates sound representations. *Psychophys, 41*(5), 709–715.

Widmann, A., Schröger, E., Tervaniemi, M., Pakarinen, S., & Kujala, T. (2012). Mapping symbols to sounds: electrophysiological correlates of the impaired reading process in dyslexia. *Front Psychol, 3*, 60.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophys, 36*(5), 638–642.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci U S A, 90*(18), 8722–8726.

Wolf, M. (1986). Rapid alternating stimulus naming in the developmental dyslexias. *Brain Lang, 27*(2), 360–379.

Woods, D. L., Alho, K., & Algazi, A. (1992). Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalogr Clin Neurophysiol*, *82*(5), 341–355.

Woods, D. L., Hillyard, S. A., & Hansen, J. C. (1984). Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing. *J Exp Psychol Hum Percept Perform, 10*(6), 761–777.

Yago, E., Escera, C., Alho, K., & Giard, M. H. (2001). Cerebral mechanisms underlying orienting of attention towards auditory frequency changes. *Neuroreport*, *12*(11), 2583–2587.

Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Lang Speech*, *48*(3), 313–338.