

# The neural basis of speech sound discrimination from infancy to adulthood

**Eino Partanen**



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

Finnish Centre of Excellence in Interdisciplinary Music Research  
University of Jyväskylä  
Finland

Academic dissertation to be publicly discussed,  
by due permission of the Faculty of Behavioural Sciences  
in Auditorium 1 at the Institute of Behavioural Sciences, Siltavuorenpenger 1 A,  
on the 13<sup>th</sup> of November, 2013, at 12 o'clock noon

University of Helsinki  
Institute of Behavioural Sciences  
Studies in Psychology 97: 2013

Supervisors: Research Professor Minna Huutilainen, PhD  
Cognitive Brain Research Unit  
Cognitive Science  
Institute of Behavioural Sciences  
University of Helsinki *and*  
Finnish Centre of Excellence in Interdisciplinary Music Research  
University of Jyväskylä  
Finland *and*  
Finnish Institute of Occupational Health  
Helsinki, Finland

Professor Teija Kujala, PhD  
Cognitive Brain Research Unit  
Cognitive Science  
Institute of Behavioural Sciences  
University of Helsinki *and*  
Cicero Learning  
University of Helsinki  
Finland

Reviewers: Professor Valerie L. Shafer, PhD  
Program in Speech-Language-Hearing Sciences  
Graduate School and University Center  
The City University of New York  
New York, USA

Docent Elina Pihko  
Brain Research Unit  
O.V. Lounasmaa Laboratory  
Aalto University School of Science  
Espoo, Finland

Opponent: Professor Thomas Jacobsen  
Experimental Psychology Unit  
Faculty of Humanities and Social Sciences  
Helmut Schmidt University  
University of the Federal Armed Forces  
Hamburg, Germany

ISSN-L 1798-842X  
ISSN 1798-842X  
ISBN 978-952-10-9420-0 (pbk.)  
ISBN 978-952-10-9421-7 (PDF)  
<http://www.ethesis.helsinki.fi>

Unigrafia  
Helsinki 2013

# Contents

Abstract .....	4
Tiivistelmä.....	5
Acknowledgements.....	6
List of original publications.....	8
Abbreviations.....	9
1 Introduction.....	10
1.1 Speech sounds, speech sound discrimination and its development .....	12
1.1.1 Speech sound discrimination in adults.....	13
1.1.2 The development of speech sound discrimination in infants and children...	13
1.1.3 Fetal learning as the basis of neonatal auditory discrimination .....	15
1.2 The mismatch negativity (MMN).....	16
1.2.1 MMN, language, and speech sound discrimination.....	18
1.2.2 MMN as an index of language-related cognitive processes .....	20
1.2.3 MMN recording paradigms .....	21
1.2.4 MMN in children .....	22
1.2.5 MMN in fetuses and infants .....	24
2 Aims of the study .....	27
2.1 The main aims of the thesis .....	27
2.2. Specific aims of the studies.....	27
3 Methods.....	29
3.1 Participants.....	29
3.2 Stimuli.....	30
3.3 Procedure .....	33
3.4 Prenatal stimulation.....	34
3.5 Experimental paradigms.....	34
3.6 Data analysis .....	35
4 Results.....	39
4.1 Multifeature word paradigm in studies of adults and children (Studies I and II)..	39
4.2 The multifeature paradigm in studies of infants (Studies III and IV) .....	42
4.3 The effects of prenatal stimulation on speech sound discrimination (Study IV)..	45
5 Discussion .....	49
5.1 The feasibility of the multifeature word paradigm.....	49
5.2 Neural speech sound processing in childhood and infancy .....	52
5.3 Fetal learning and its implications .....	54
5.4 Study limitations and future directions .....	56
5.5 Conclusions .....	57
6 References .....	59

## Abstract

Rapid processing of speech is facilitated by neural representations of native language phonemes. However, some disorders and developmental conditions, such as developmental dyslexia, can hamper the development of these neural memory traces, leading to language delays and poor academic achievement. While the early identification of such deficits is paramount so that interventions can be started as early as possible, there is currently no systematically used ecologically valid paradigm for the assessment of the neural basis of speech sound processing. Thus, this thesis investigates the feasibility of a mismatch negativity (MMN) paradigm presenting speech sounds embedded in pseudo words to probe the neural discrimination accuracy of sounds in more natural context. Furthermore, the applicability of this paradigm for infant research was determined in a study investigating the effects of additional prenatal speech stimulation on newborn sound processing. The results show, on group level, that while healthy adults neurally detect all speech sound changes used in the experiments, the 4–12-year-old children's and infants' neural processing of speech sounds varies with age. In particular, the preschool children seem to be more proficient in neurally detecting small pitch changes in word context than school-aged children. Furthermore, children's MMNs were found to be associated with improved verbal IQ 14–17 months later, while the positively-displaced MMN, the p-MMR, correlated with poorer performance IQ. The results on the effects of additional prenatal exposure to pitch-modulated speech sounds showed that the MMNs to pitch changes in pseudo words were enhanced after birth, indicating specific learning effects due to additional prenatal stimulation. Furthermore, these learning effects generalized to other types of speech sounds not included in the learning material. Taken together, these findings suggest that the paradigm presenting speech sound changes in word context is viable for probing the neural memory traces for speech sounds from infancy to adulthood. The possibility to assess neural speech sound discrimination broadly in a single EEG recording could be used, for example, to classify between different subtypes of dyslexia. Finally, the neural effects induced by fetal learning suggest that prenatal exposure to sounds, for example, to the voice of the mother, may predispose the infant to be neurally sensitive to such sounds by birth.

## Tiivistelmä

Puheen ymmärtäminen perustuu osin oman äidinkielen puheäänteille muodostuneisiin hermostollisiin muistijälkiin. Jotkin oireyhtymät tai kehitykselliset häiriöt, kuten lukihäiriö, voivat kuitenkin haitata näiden muistijälkien muodostumista. Tämä voi ilmetä esimerkiksi kielen kehityksen viivästyminä tai huonona koulumenestyksenä. Kuntoutuksen kannalta kielenkehityksen vaikeuksien varhainen tunnistaminen on tärkeää, mutta systemaattisesti käytettyä ja luonnollisen kaltaista ääniympäristöä jäljittelevää asetelmaa puheen piirteiden hermostollisen perustan tutkimiseen ei ole. Tässä tutkimuksessa selvitettiin, soveltuuko puheäänteiden muutoksia sanakontekstissa mismatch negativity (MMN) -vasteen avulla mittaava monipiirreasetelma äänten hermostollisten muistijälkien ja sikiöaikaisen kuuloaineuksen oppimisen hermostollisten vaikutusten tarkasteluun. Ryhmätason tulosten perusteella terveiden aikuisten aivot tunnistivat puheen piirteiden muutokset, mutta 4–12 -vuotiaiden lasten ja vauvojen aivot käsittelivät puheen piirteiden muutoksia eri tavoin. Esimerkiksi esikouluikäisten kuulojärjestelmä vaikutti käsittelevän sanakontekstissa esitettyjen äänen korkeuksien muutoksia tarkemmin kuin kouluikäisten. 4–12 -vuotiailla lapsilla suuremmat MMN-vasteet olivat myös yhteydessä parempaan kielelliseen päättelysuoriutumiseen, ja MMN:n positiivinen muoto, p-MMR, oli yhteydessä heikompaan näönvaraiseen päättelysuoriutumiseen 14–17 kuukautta myöhemmin. Tutkimus osoitti myös, että sikiöaikainen altistus puheäänten taajuusmuutoksille vahvisti vauvojen aivojen kykyä erotella puheen taajuusmuutoksia. Sikiöaikaisen altistuksen myötä vauvojen kuulojärjestelmä oppi myös erottamaan muutoksia, joille vauvat eivät olleet altistuneet raskauden aikana ja joita äänimateriaalille altistamattomat vauvat eivät kyenneet erottelemaan. Kokonaisuudessaan tulokset osoittivat, että väitöskirjan tutkimuksissa testattu MMN-asetelma soveltuu puheen piirteiden hermostollisten muistijälkien tutkimukseen ryhmätasolla vauvoista aikuisiin. Useiden erityyppisten puheäänten muutosten hermostollisen erottelukyvyn tutkiminen samanaikaisesti voi hyödyttää esimerkiksi erilaisten lukihäiriötyyppien erottelua. Havaitut sikiöaikaisen oppimisen hermostolliset vaikutukset viittaavat siihen, että sikiöaikainen altistus esimerkiksi äidin äänelle voi edesauttaa samankaltaisten äänien hermostollista käsittelyä syntymän jälkeen.

## Acknowledgements

During the process of preparing my PhD I have had the privilege of meeting many awe-inspiring people whose talent, skill and enthusiasm in their own fields have left me utterly breathless. Even more, these people have taken the time to sit with me over a cup of coffee to patiently answer my questions, helping me understand both my and their work in greater detail. Were it not for encounters with such people and altruistic sharing of expertise, this thesis would not be as it is today. For facilitating this, I must thank the supportive and enthusiastic environments which I have been privileged to work at: the Cognitive Brain Research Unit, the Finnish Center of Excellence in Interdisciplinary Music Research, and the Clinic of Neurology.

While it is impossible to list everyone deserving to be mentioned, I will do the best I can with the space given. First, I am immeasurably in debt to my supervisor, Research Prof. Minna Huotilainen, for hiring me, trusting me, and taking me under her wing. Her example has taught me all I know about what it means to be a scientist. My sincere thanks go to my second supervisor, Prof. Teija Kujala, who taught me how to write scientific texts – my PhD would not have been completed without her supervision. I am much obliged to Docent Kaisa Tiippa for giving me my very first opportunity to teach at the university level. Prof. Mari Tervaniemi has always been supportive, encouraging and eager to follow my progress, for which I can only offer my humble gratitude. I am also thankful to Docent Martti Vainio for his extensive knowledge in language research, Assistant Prof. Anke Sambeth for teaching me to understand infant ERPs, Auli Liitola for her work on infant ERPs, and Prof. Vineta Fellman for her long career and help in doing clinical ERP research in infants and young children. I am indebted to Prof. Risto Näätänen for both his coauthorship and lifetime work, without his influence the research groups I have worked in would probably not exist as they are today. Thanks to Piiu Lehmus, Markku Pöyöhnen, and Marja Junnonaho for being invaluable in all administrative issues. Finally, I convey my thanks to the reviewers of my thesis, Prof. Valerie Shafer and Docent Elina Pihko, for helpful comments, and to Prof. Thomas Jacobsen for acting as an opponent at the public defense of the dissertation.

Many others have also shared their time and friendship with me, both in and out of work. The TELK posse (Tuomas Teinonen, Lea Oja, and Kristiina Relander-Syrjänen)

have been there for many of the highs and lows of my life. Maria Mittag without a doubt deserves the lifetime award of adorableness and irritation (and a bunny). Dr. Teppo Särkämö is certainly a role model for all of us and his extensive knowledge is thoroughly stunning. Without Dr. Satu Pakarinen and her pioneering work on MMN paradigms this thesis would not have been possible. The Brain and Language group (Docent Minna Lehtonen, Soila Kuuluvainen, Lilli Kimppa, Dr. Alina Leminen, Riikka Lindström, Dr. Riikka Lovio, and Dr. Sari Ylinen, just to name a few) have shared with me not only their interest and enthusiasm in language research but many hilarious moments both in and out of work. The Brain and Music Team (Docent Elvira Brattico, Tanja Linnavalli, Vesa Putkinen, Johannes Pykäläinen, Anna Rämä, Katri Saarikivi, Martti Sutinen, Ritva Torppa, and Paula Virtala, among others) have always made me feel welcome and taught me immensely about music, science and teamwork. Thanks to Miika Leminen and Tommi Makkonen for being around to talk to about engineering, mathematics, and programming, and for absolutely hilarious company. I also offer my sincere gratitude to Tarja Ilkka for conducting the EEG recordings in Studies III and IV.

While finishing this thesis I also have had the opportunity of doing clinical work in the Clinic of Neurology under the leadership of Docent Marja Hietanen. Her talent in understanding both clinical neuropsychology and science are phenomenal and I have learned much during my work in her unit. I especially want to thank Elisa Iivonen, Riikka Niinikuru, and Anna-Maria Wikström for their effort, patience, and copious use of green ink when teaching me to understand clinical neuropsychology.

I am also grateful to the projektimökki group (Salla Laaksonen, Tuukka Lehtiniemi, Anni Mattila, and Katja Sibenberg) for their friendship and for creating an extremely interdisciplinary, inspirational, and champagne-rich environment in which to participate in. Thanks to my parents for raising me to appreciate both science and art, and to Ilona, Jari, and Joel Sepänkumpu for accepting me as a part of their extended family. Finally, I wish to mention the House of Ninja Finland (Anneli Kanninen, Virpi Kurki, Venla Vuorio, and others), the three dance groups I've had the pleasure to participate in, and so many others in the Finnish dance community for inspiring me to dance.

Dedicated to Raija Partanen (1948-2006), who passed away when I was taking my first baby steps in science. You are sorely missed.

## List of original publications

This thesis is based on the following original publications, referred to in the text by Roman numerals (I–IV).

- Study I** Partanen, E., Vainio, M., Kujala, T., Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. *Psychophysiology*, *48*, 1372–1380.
- Study II** Partanen, E., Torppa, R., Pykäläinen, J., Kujala, T., Huotilainen M. (2013). Children’s brain responses to sound changes in natural pseudo words in a multi-feature paradigm. *Clinical Neurophysiology*, *124*, 1132–1138.
- Study III** Partanen, E., Pakarinen, S., Kujala, T., Huotilainen, M. (2013). Infants’ brain responses for speech sound changes in fast multifeature MMN paradigm. *Clinical Neurophysiology*, *124*, 1578–1585.
- Study IV** Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences U S A*, *100*, 15145–15150.

The articles are reprinted with the permission of the copyright holders.



## Abbreviations

APGAR	newborn health assessment scale (Appearance, Pulse, Grimace, Activity, Respiration)
EEG	electroencephalograph, electroencephalography
EOG	electro-oculogram
GA	gestational age
ICA	independent component analysis
MEG	magnetoencephalograph, magnetoencephalography
MMN	mismatch negativity
MMR	mismatch response
NICU	neonatal intensive care unit
PC	principal component
PCA, tPCA	(temporal) principal component analysis
PIQ	Performance IQ
SD	standard deviation
SLI	specific language impairment
SSG	Semisynthetic Speech Generation
tPCA	temporal principal component analysis
VIQ	Verbal IQ
WISC-IV	Wechsler Intelligence Scale for Children – IV

# 1 Introduction

Fast and error-free human communication is facilitated by the automatic and effortless processing of individual speech sounds. Such automatic neural processes for speech sound discrimination can be seen immediately after birth, when the infant is sensitive to exaggerated pitch contours of infant-directed speech (Fernald & Kuhl, 1987). In infancy, these automatic processes have been suggested to facilitate infants' perception and segmentation of speech (Karzon, 1985; Nelson, Hirshpasek, Jusczyk, & Cassidy, 1989) and the formation of neural representations of native language speech sounds during the first year of life (Cheour et al., 1998; for a review, see Kuhl, 2004). However, if the automatic process of recognizing individual speech sounds is hampered, or if the neural representations of native language speech sounds do not develop normally, difficulties in communication and learning may follow. These difficulties, in turn, can have long-term consequences on, for example, individuals' academic achievement.

Indeed, many developmental conditions and disorders are associated with abnormal speech sound processing. These include, for example, dyslexia (Lovio, Näätänen, & Kujala, 2010), autism spectrum (Kujala et al., 2010), specific language impairment (SLI; Archibald & Joanisse, 2012), and children born preterm (Jansson-Verkasalo et al., 2003). While the specific prevalence of speech sound processing difficulties is difficult to assess due to comorbidity with other conditions and deficits, a recent meta-analysis suggested that the prevalence of language deficits in the population varies between 2.6–14.6% (excluding deficits secondary to another condition, such as autism), depending on the age group studied, the definition of the deficit, and geographical location (Law, Boyle, Harris, Harkness, & Nye, 2000). Furthermore, a 12-year follow-up study suggested that the prevalence of learning deficits in the US is approximately 7.6% (Boyle et al., 2011), or for SLI, approximately 7% (Tomblin et al., 1997). In many cases, the aforementioned conditions, associated also with deficient speech sound processing, may predispose a child to, for example, poor study skills (Mortimore & Crozier, 2006) and academic achievement (Catts, 1991; Klein, Hack, & Breslau, 1989; McCormack, McLeod, McAllister, & Harrison, 2009), or worse than baseline performance in psychometric testing (Flax, Realpe-Bonilla, Roesler, Choudhury, & Benasich, 2009; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998).

Due to high prevalence of speech sound processing deficits, testing the at-risk populations using structured and efficient early screening methods shows merit (Law, Boyle, Harris, Harkness, & Nye, 1998). Consequently, recent Finnish guidelines advocate for early screening (Ahonen, 2010; Valtonen et al., 2003; Valtonen et al., 2009). Such early assessment methods may be cost-effective, if language deficits can be diagnosed and interventions started already early in childhood. Conversely, the lack of structured and efficient early identification and intervention programs cause enormous costs for the society: for example, the lifetime costs associated with learning deficits were suggested to be approximately 400 million euros per birth cohort in Finland (Willberg, 2002).

Previously, several behavioral tests assessing the processing and recognition of speech sounds have been developed and shown associations between speech sound processing and, for example, judgments of the severity of the child's speech disorder (Preston & Edwards, 2010) and receptive language abilities (Vance, Rosen, & Coleman, 2009). However, behavioral methods evaluating speech sound processing are limited in scope, as the child needs to be of sufficient age to be able to understand the instructions and perform the test. In contrast, the brain's event-related potentials (ERP), mismatch negativity (MMN) in particular (see Section 1.2), can be used to assess speech sound discrimination even in infancy (Cheour et al., 1998; Leppänen et al., 2002; Weber, Hahne, Friedrich, & Friederici, 2004, for a review, see Leppänen et al., 2012). ERPs and the MMN have also been shown to be viable biomarkers for a plethora of deficits, such as dyslexia (Kujala, Lovio, Lepistö, Laasonen, & Näätänen, 2006a; Leppänen et al., 2002; for a review, see Kujala, 2007b), autism and Asperger's syndrome (Lepistö et al., 2005; for a review, see Kujala, Lepistö, & Näätänen, 2013), schizophrenia (Takahashi et al., 2012; for a review, see Urban, Kremlacek, & Libiger, 2007), and cognitive deficits associated with low birth weight (Mikkola et al., 2007). It has also been suggested that a systematic ERP paradigm assessing many facets of neural speech sounds processing could benefit the early evaluation of language skills and their deficits, especially in conjunction with parental reports and psychometric tests (Bishop, 2007; Gabrieli, 2009).

While ERP paradigms have already been developed for the assessment of auditory processing and its' deficits, they have studied the discrimination of speech sounds in isolation (Lovio et al., 2009; Pakarinen et al., 2009), or used acoustic changes in non-

speech sounds only (Pakarinen, Huotilainen, & Näätänen, 2010). In contrast, speech sound changes in natural speech occur in word context, not in isolation. Thus, studies assessing acoustic changes only, or changes in isolated syllables, may not give a full picture of the individuals speech sound discrimination accuracy. For example, the effects of preceding and following sounds in a sound stream have been shown to hamper the discrimination of sound features in dyslexics (Kujala, Belitz, Tervaniemi, & Näätänen, 2003). Therefore, investigating auditory processing in an ecologically valid word context using naturally produced speech sounds may yield more information on actual auditory capabilities of the individuals than studying it with simpler paradigms.

## **1.1 Speech sounds, speech sound discrimination and its development**

In all languages, sentences can be broken into words, which consist of one or more individual meaning-carrying units, morphemes. For example, the word '*joyful*' consists of two morphemes: the free morpheme or the root of the word, '*joy*', and the bound morpheme or the suffix, '*ful*' (Leminen, 2012). Morphemes, in turn, consist of individual phonemes, which can be characterized as smallest units in language that can change the meaning of a word (for example, the change of one phoneme in the words '*pin*' and '*pen*'). Finally, phonemes, morphemes, and words can differ in how they are uttered: speaker can emphasize a word or a part of the word by, for example, increased pitch or intensity. Such different ways of utterance are usually referred to as speech prosody or prosodic features of speech.

Throughout this thesis, the term '*speech sound change*' is used to refer to changes in acoustic features of individual phonemes or individual syllables. Speech sound changes can be roughly divided into prosodic and phonological (or phonetic) changes. While phonological changes in words often change the meaning of the word (for example, vowel identity changes in the words '*pin*' and '*pen*', or vowel duration changes in Finnish: /tu-li/ - fire, /tu:-li/ - wind), prosodic changes usually denote word stress, emotion, importance or turn taking (Suomi, Toivanen, & Ylitalo, 2008).

### **1.1.1 Speech sound discrimination in adults**

In healthy adults, discrimination of speech sounds is fast, effortless, and efficient (Pardo & Remez, 2006). With little effort, adults can even comprehend the contents of speech from sine wave replicas, where the replica is formed by representing the first three formants of an utterance with time-varying sinusoids (Remez, Rubin, Pisoni, & Carrell, 1981). This shows that adults can perceive native language speech sounds accurately even when the auditory input is extremely ambiguous or lacks many familiar features of speech, which is the case with sine wave replicas. However, while discrimination or categorization of phonemes is fast and accurate, detecting changes within a phoneme category is not (Lieberman, Harris, Hoffman, & Griffith, 1957). Furthermore, prosodic changes in speech sounds do not form similar categories as phonemes do, and thus processing of prosodic changes may differ from that of phonological changes.

While adults are very proficient in discriminating changes in their native language speech sounds, their capability for discriminating nonnative speech sound contrasts is often limited (Best & Strange, 1992). However, adults can be trained to discriminate nonnative speech sound contrasts (Winkler et al., 1999), but effects can be seen only after extensive training. For example, short training has been shown to only modestly improve the ability of native adult speakers of Greek to discriminate phoneme duration changes in words (Giannakopoulou et al., 2013), which is a nonnative speech sound contrast in Greek.

### **1.1.2 The development of speech sound discrimination in infants and children**

The present theories on the development of speech processing suggest that initially the human infant discriminates sounds in a universal fashion. This means that the infant can process a variety of speech sound contrasts, unlike adults whose auditory processing is highly specialized to discriminate changes in, for example, their own native language (Kuhl, 2004) or musical environment (Hannon & Trehub, 2005). This shift from universal to culture-specific processing is often referred to (native language) neural commitment (Kuhl et al., 2008) or musical enculturation (Hannon & Trainor, 2007). In short, the shift from universal to culture-specific processing is seen as an increased

proficiency in processing sounds prevalent in the auditory environment surrounding the infant while the ability to process unfamiliar contrasts is diminished (Kuhl et al., 2006).

The native language neural commitment, or the enculturation, hypothesis has been confirmed in several studies and recent evidence suggests that this process starts already prior to birth (Moon, Lagercrantz, & Kuhl, 2013). Comparing native and non-native consonant contrasts in American infants, Rivera-Gaxiola, Silva-Pereyra, and Kuhl (2005) showed that 7-month-old infants could neurally discriminate both the native and nonnative consonant contrasts, but 11-month-old infants could not. Similar results have been found in a study comparing the Japanese and United States' infants discrimination ability (Kuhl et al., 2006) or neural responsiveness (Y. Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005) to /r-l/-contrasts, and in Finnish and Estonian infants' neural responses to native and nonnative vowel contrasts (Cheour et al., 1998). Furthermore, these effects are not limited to speech sound discrimination only, as 6-month-old infants could discriminate rhythmic changes in both familiar and unfamiliar musical contexts but additional exposure or training was needed in order to discriminate the changes at the age of 12 months (Hannon & Trehub, 2005). It must be noted, however, that the diminished ability to process nonnative speech sound contrasts is not found in all studies or in all infants within a study (Polka, Colantonio, & Sundara, 2001; Rivera-Gaxiola et al., 2005).

Taken together, the evidence suggests that the shift from universal to culture-specific perception of sounds occurs during the first year of life, and that this process is characterized by diminished ability to discriminate nonnative contrasts while the ability to discriminate contrasts present in auditory environment improves. This process is probably due to both neural commitment and improved sensitivity to smaller sound changes during development. However, while infants can process both native and nonnative sound contrasts early in development, they are much less proficient in speech sound processing than adults. For example, some studies suggest that infants are able to detect speech sound changes in infant-directed speech but not in adult directed speech (Karzon, 1985; Nelson et al., 1989), suggesting that infants need additional cues to discriminate some speech sound differences. Consistent with this, 2-year-old children were shown to be able to accurately discriminate phoneme contrasts of large magnitude only (Holt & Lalonde, 2012). Thus, while infants and children are capable of

discriminating speech sound contrasts, some of the changes may have to be large in magnitude, or presented in context where additional prosodic cues facilitate speech sound discrimination (e.g., infant-directed speech).

### **1.1.3 Fetal learning as the basis of neonatal auditory discrimination**

During the last trimester of pregnancy at the onset of hearing, usually by the 27 weeks gestational age (GA) (Hepper & Shahidullah, 1994) in humans, the external auditory input starts to rapidly reorganize the auditory cortex (Chang & Merzenich, 2003), as reflected by the formation of new synapses (Kostović & Judas, 2010) and myelination of neuronal connections (Moore & Linthicum, 2007). Animal studies suggest that this reorganization is first characterized by the formation of tonotopically organized primary auditory cortex, facilitating fast and accurate pitch perception (L. I. Zhang, Bao, & Merzenich, 2001). Furthermore, animal studies also indicate that the formation of the tonotopic map can be influenced by the early acoustic environment, as additional exposure to certain tone frequencies expanded the cortical representations of those tone frequencies in rat pups (Zhang et al., 2001). In humans, the fetal reactions to external sounds have been confirmed by showing that a fetus can be startled by sounds, or that fetuses heart rate changes due to external stimulation. (Gerhardt & Abrams, 2000). The fetal discrimination of external sounds have also been demonstrated with magnetoencephalography (MEG; Draganova et al., 2005; Huotilainen et al., 2005). Taken together, these results suggest that humans can perceive and discriminate sounds in utero and thus may have some learning abilities even prior to birth.

Fetal learning effects have been suggested by several behavioral experiments after birth by assessing differences in fetal reactions (cardiac rhythm accelerations and decelerations, changes in pacifier sucking rate, head turning) to familiar versus novel sounds. It was suggested that fetuses recognize the native language of the mother (DeCasper & Fifer, 1980; Moon, Cooper, & Fifer, 1993), familiar melodies (Hepper, 1988), and changes in vowels (Lecanuet et al., 1987), habituate to environmental sounds (Ando & Hattori, 1970), and even to recognize the voice of their mother from the voices of unfamiliar women (Kisilevsky et al., 2003). Fetuses also appear to discriminate fine-tuned auditory information as infants react differently to native versus non-native

phonemes (Moon et al., 2013). Although fetal learning effects have been convincingly demonstrated in previous studies, it is not known how much prenatal exposure is required for fetal learning to occur. Previous studies suggest that fetuses can learn sounds that they are not exposed to during normal development at least after 6 weeks of daily exposure prior to birth (DeCasper et al., 1994; for a review, see Moon & Fifer, 2000).

However, fetal learning might be strongest to low frequency components of sounds: while the results from studies assessing the attenuation of external sounds coming into utero vary, the low frequency sounds are much less attenuated than high frequency sounds (Gerhardt & Abrams, 2000; Lecanuet & Schaal, 1996; Peters, Abrams, Gerhardt, & Griffiths, 1993; Spence & Freeman, 1996). Consistent with this, infants prefer the low-pass filtered mothers' voice, but not the maternal whispered voice which does not contain the low frequency components of speech (Spence & Freeman, 1996).

The behavioral evidence also suggests that fetuses can form long-term neural memory traces for sounds they have been exposed to during pregnancy and use them to generate specific learned behaviors. For example, the pitch contours in infants' cry is shaped by the language they were exposed to in utero (Mampe, Friederici, Christophe, & Wermke, 2009). Furthermore, fetal learning effects may last for several weeks since a recent study showed that 6-week-old infants could recognize changes in the melody they heard in utero, even with no additional exposure after birth (Granier-Deferre, Bassereau, Ribeiro, Jacquet, & Decasper, 2011). Thus, it has been suggested that fetal learning may play a role in preparing the fetus for the life after birth, promoting attachment (Moon & Fifer, 2000), cognitive abilities (Huotilainen, 2010), and language learning (Lecanuet, Granier-Deferre, Jacquet, & DeCasper, 2000).

## **1.2 The mismatch negativity (MMN)**

The MMN is a negative deflection 150–250 ms after the stimulus onset in the deviant-minus-standard difference waveform (Kujala & Näätänen, 2010). According to the memory trace hypothesis, the MMN reflects the auditory systems' automatic change-detection response, elicited when a discriminable change between the frequently presented 'standard' sound and the occasionally presented 'deviant' takes place in the



sound stream (Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is elicited not only by changes in the physical features of the sounds, but also by changes in more abstract sound features, such as regularity violations in sound sequences (Tervaniemi, Maury, & Näätänen, 1994; for a review, see Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001a), omissions of sounds from the sound stream (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997), or even changes in complex spectro-temporal rules (for reviews, see Näätänen et al., 2001a; Paavilainen, 2013). While the MMN was suggested to be somewhat modulated by attention (for a review, see Sussman, 2007), its' elicitation does not require attentive action from the participant (for reviews, see Kujala, Tervaniemi, & Schröger, 2007a; Näätänen et al., 2007). Thus, the MMN can be used to study auditory discrimination in conditions without many of the confounding task-related effects, such as motor potentials arising from participant pressing a button after detecting a change in the sound stream. Due to its benefits, the MMN has been widely used to assess the neural basis of auditory discrimination of many different participant groups (for reviews, see Näätänen et al., 2012; Näätänen et al., 2011a), including infants (Alho, Sainio, Sajaniemi, Reinikainen, & Näätänen, 1990; Cheour, Leppänen, & Kraus, 2000) and children (Cheour, Korpilahti, Martynova, & Lang, 2001), and also various clinical groups, such as dyslexics (Baldeweg, Richardson, Watkins, Foale, & Gruzelier, 1999; for reviews, see Kujala, 2007b; Schulte-Korne & Bruder, 2010), children with specific language impairment (Bishop, Hardiman, & Barry, 2010), patients with schizophrenia (Michie et al., 2000), and preterm infants (Fellman et al., 2004; Mikkola et al., 2007).

According to the recent theories, the MMN is elicited in a process comparing the features of the incoming sound are with the neural memory trace of the previously repeated sounds, or with the auditory systems' predictions of the future auditory events (Näätänen, Kujala, & Winkler, 2011b). If a change in the auditory regularity is detected by the auditory system, or if a change in the auditory environment does not fit the auditory systems' predictions, then the MMN is elicited. However, alternative theories have suggested that the MMN does not reflect a memory-based process, but it is seen in the deviant-minus-standard waveform only due to differences in the refractoriness of the neurons eliciting the N1 response (May & Tiitinen, 2010). However, the N1 hypothesis fails to explain the MMN generated by the omission of the stimuli from a sound stream

(Yabe et al., 1997) or the MMN generated by the violations of complex spectro-temporal rules (for a review, see Paavilainen, 2013), where stimulus specificity is not the basis of the MMN elicitation.

Nevertheless, the MMN and N1 components do temporally overlap and the deviant-minus-standard waveform may include contributions from both N1 and MMN responses (Kujala et al., 2007a), at least in adults. If the difference between the standard and the deviant is sufficiently large (32% in adults; Horváth et al., 2008), the MMN amplitude may be overestimated due to the N1 confound (see, e.g., Jacobsen & Schröger, 2001). However, even with the confounding N1 effects, the MMN is an attractive tool for cognitive and clinical neuroscience (Kujala et al., 2007a; Näätänen, Kujala, Kreegipuu, et al., 2011; Näätänen et al., 2007).

### **1.2.1 MMN, language, and speech sound discrimination**

While the MMN reflects the auditory discrimination process, the MMN can also be used to probe the learning-induced long-term memory traces (Kujala & Näätänen, 2010), such as those associated with language processing (Pulvermüller & Shtyrov, 2006). Language learning was shown to modulate the MMN amplitude (Cheour, Shestakova, Alku, Čeponienė, & Näätänen, 2002a; Näätänen et al., 1997; Shestakova, Huotilainen, Čeponienė, & Cheour, 2003; Winkler et al., 1999; Ylinen, Shestakova, Huotilainen, Alku, & Näätänen, 2006) and the MMN can be used to assess, for example, differences in the neural organization of the phonological processing between bilinguals and monolinguals (Tamminen, Peltola, Toivonen, Kujala, & Näätänen, 2013). The MMN to changes in native language speech sounds is further enhanced in amplitude when the changes are presented in word than nonword contexts, both in adults (Pulvermüller et al., 2001) and children (Korpilahti, Krause, Holopainen, & Lang, 2001), suggesting that neural memory representations for meaningful words are accessed by the system generating the MMN response. Furthermore, the MMN is associated with word frequency, as MMNs to more frequently used words have shorter MMN latencies (Alexandrov, Boricheva, Pulvermüller, & Shtyrov, 2011), and larger MMN amplitudes (Shtyrov, Kimppa, Pulvermüller, & Kujala, 2011) than infrequent ones.

The effects of the linguistic expertise on the MMN are not limited to the enhanced MMNs for native language phonemes or words only. MMN amplitudes are enhanced for sound feature changes which have a linguistic role in a language the participant is native or proficient in. Such enhanced MMNs have been found to, for example, duration changes in quantity languages such as Finnish (Tervaniemi et al., 2006), to pitch changes of lexical tones in tonal languages like Mandarin (Kaan, Wayland, Bao, & Barkley, 2007), and even to violations of co-articulatory patterns in language (Steinberg, Truckenbrodt, & Jacobsen, 2010). The sensitivity to linguistic cues of native language also enhances the MMNs to such cues or changes in foreign language context, even if they do not have a linguistic role in that language (Lipski, Escudero, & Benders, 2012).

Furthermore, the MMNs to linguistic and nonlinguistic stimuli seem to originate from at least partially different processes (Pulvermüller & Shtyrov, 2006). It was proposed that linguistically relevant stimuli activate long-term memory traces of the native language speech sounds formed during language learning, whereas the MMN to nonlinguistic stimuli may predominantly reflect the processing of such sounds in the auditory sensory memory only. These differences in processes are reflected in the MMN topography, as MMN to linguistically relevant contrasts were suggested to be more left-lateralized than the processing of similar nonspeech contrasts (Shtyrov, Kujala, Palva, Ilmoniemi, & Näätänen, 2000; Sorokin, Alku, & Kujala, 2010).

Thus, it seems plausible that the MMN to linguistic stimuli is elicited by a two-step process, where the activation of the long-term memory representations by native language speech sounds, phonemes, or words facilitates and enhances the change-detection process in the auditory sensory memory, which is reflected as enhanced MMN amplitude and shortened MMN latency for native language speech sounds (Näätänen, 2001b; Pulvermüller & Shtyrov, 2006; Winkler et al., 1999). However, the linguistic MMN may be confounded by the enhancement of obligatory ERP components due to linguistic expertise. For example, Alain, Campeanu, and Tremblay (2010) showed a reduction of the N1 and P2 response amplitudes when participants learned to discriminate between syllables differing in the voice onset time. However, Tremblay, Kraus, McGee, Ponton, and Otis (2001) showed an opposite pattern, an enhancement of N1 and P2 responses as participants learned to discriminate between consonant-vowel -

syllables differing only in the voice onset time. While the evidence for the influence of the linguistic experience on the changes of other ERP components is not entirely conclusive, it should be noted that they may contribute to the learning-induced changes observed in the MMN.

### **1.2.2 MMN as an index of language-related cognitive processes**

Because the MMN reflects both auditory discrimination (for a review, see Näätänen, Kujala, & Winkler, 2011) and formation of long-term memory traces of native language phonemes (see Chapter 1.2.1), some studies have investigated whether the MMN is associated with performance in psychometric tests of language. Such studies have shown that the MMN amplitudes are positively correlated with, for example, VIQ or verbal fluency in 5 year old children (Mikkola et al., 2007), or that the incidence of MMN in 5–7-year-old children is associated with improved performance in auditory memory span test (Bauer et al., 2009). Further studies have linked the MMN with performance in more specific language-related psychometric tests and found that enhanced MMNs are associated with, for example, improved pronunciation skills in 10–12-year-old children (Milovanov et al., 2009), or object naming abilities in prematurely born 4- and 6-year-old children (Jansson-Verkasalo et al., 2003; Jansson-Verkasalo et al., 2004). Because of the association of the MMN and language-related psychometric measures, the MMN could be a viable biomarker for various deficits characterized by substandard auditory processing capabilities (Kujala, 2007b), such as SLI (Weber, Hahne, Friedrich, & Friederici, 2005) or dyslexia (Bruder et al., 2011; Leppänen et al., 2010). This is important for it may allow the early identification of language-related deficits or predict possible language-related difficulties of participants who cannot be evaluated with traditional psychometric testing, such as infants (Leppänen et al., 2010).

Consequently, the MMN has been used to predict normal and deficient developmental trajectories. In the extensive Jyväskylä Longitudinal Study of Dyslexia the MMN, and other ERP components, were used to assess differences between normally reading children, children at-risk for dyslexia but who do not develop dyslexia, and dyslexics (for a review, see Leppänen et al., 2012). In similar vein, the MMN was used to assess the efficacy of interventions aimed to ameliorate reading skill

deficits in children. In such studies the intervention effects were evident both in reading skills and enhanced MMNs to tone pairs in 6–7-year-old children (Kujala et al., 2001; Lovio, Halttunen, Lyytinen, Näätänen, & Kujala, 2012), although not all studies have shown improvements in both reading skills and MMN after intervention (Huotilainen et al., 2011).

However, there are some concerns with respect to the viability of the abnormal MMN amplitudes or latencies as an indicator of language deficits and the findings on MMN in children with dyslexia or SLI have been described by some as “highly inconsistent” (Bishop, 2007). The concerns raised by Bishop (2007) do not challenge the premise that the MMN is a viable biomarker per se, but are more methodological in nature. Due to, for example, small sample sizes or methodological inconsistencies, differences in MMNs between clinical groups and healthy controls, or associations between the MMN amplitude and performance in psychometric testing have not always been found (Shafer, Morr, Datta, Kurtzberg, & Schwartz, 2005). Consequently, some work has already been conducted on improving the analysis methods for assessing the MMN, such as use of tPCA (Leppänen et al., 2010) or wavelet analysis (Isler et al., 2012) to provide more accurate MMN estimates.

### **1.2.3 MMN recording paradigms**

Traditionally, the MMN has been recorded with an “oddball” paradigm, in which changed sounds, deviants, are occasionally presented ( $p = 0.1$ – $0.2$ ) among the frequently presented sounds, standards ( $p = 0.8$ – $0.9$ ). This tends to result in long MMN experiments, especially when MMNs for several change types are recorded. The recording time is often the limiting factor especially in clinical neuroscience and in studies of infants or children.

To shorten the MMN recording time, Näätänen, Pakarinen, Rinne, and Takegata (2004) developed a faster multifeature paradigm (Optimum-1 in the original paper), where every other stimulus is a standard and every other is one of the several different deviant types. The rationale behind the paradigm is that each deviant differs from the standard in one feature only. Thus, the unchanged features of the deviant strengthen the memory traces for those features while the change in the deviating feature is detected by

the auditory system, eliciting an MMN response. Previous studies have suggested that the MMNs recorded in the multifeature and oddball paradigms are very similar for tones in adults (Näätänen et al., 2004) and speech sounds both in adults (Pakarinen et al., 2009) and children (Lovio et al., 2009). The multifeature paradigm with complex tones was shown to be feasible for studying even 2–3-year-old children (Putkinen, Niinikuru, Lipsanen, Tervaniemi, & Huotilainen, 2012), and in infants using MEG (Sambeth et al., 2009).

Thus far, several variants of the multifeature paradigm have been developed in addition to the original multifeature paradigm with pure tones (Näätänen et al., 2004). For example, Pakarinen et al (2009) used this paradigm with changes embedded in single semi-synthetic syllables (/te:/ and /pi:/), while Pakarinen et al. (2013) developed a variant of the multifeature paradigm with three magnitudes of change using semi-synthetic syllables (Pakarinen et al., 2013). Further studies have investigated the MMNs in paradigms with no standard stimulus (Pakarinen et al., 2010), used the multifeature paradigm to assess MMNs of different deviant magnitudes of many different change types in one session (Pakarinen et al., 2013; Putkinen et al., 2012), or even developed paradigm variants investigating music processing (Vuust et al., 2011).

While the multifeature MMN paradigm does not shorten the recording time of MMN for one change type only, the main advantage of the multifeature paradigm is the possibility to record extensive auditory discrimination profiles for several change types simultaneously (Pakarinen et al., 2009). Such broad auditory discrimination profiles may be beneficial in, for example, classifying different types of dyslexics or disentangling general versus specific auditory deficits.

#### **1.2.4 MMN in children**

At least from the age of 4 years, an adult-like negative MMN is reliably elicited in children (Martin, Shafer, Morr, Kreuzer, & Kurtzberg, 2003; Shafer, Morr, Kreuzer, & Kurtzberg, 2000). However, children's MMN amplitudes may differ from those elicited by adults, even for identical stimuli. For example, Lovio et al. (2009) reported that according to visual analysis, 6-year-old children have smaller MMNs to acoustically identical speech sound changes than found for adults tested in the study of Pakarinen et

al. (2009). These results suggest that the MMN changes in amplitude during development and that these changes in the MMN amplitude might reflect speech sound processing proficiency between children and adults. Thus, MMNs might not be elicited by all sound changes or the MMNs may be very small in young children. Consistent with this, Putkinen et al. (2012) showed that an adult-like negative MMN seems to be reliably elicited in 2–3-year-old children by changes of larger magnitudes only, or by changes that are linguistically relevant for the participant (e.g., changes of vowel duration or silent gaps representing consonant duration changes in Finnish). While the MMN seems to change in amplitude during development, this effect is not found in all studies. For example, Kraus, Koch, McGee, Nicol, and Cunningham (1999) found no changes in the MMN amplitude to /da/-/ga/- or /ba/-/wa/-contrasts in synthetic syllables between the ages of 6 and 15 years.

Children's standard-minus-deviant waveform may also include other components than the adult-like MMN which may have different functional characteristics (Lee et al., 2012). For example, in addition to, or instead of an adult-like negative MMN, positive MMN-like responses (p-MMR) have been found in children and infants, (e.g. Ahmmed, Clarke, and Adams, 2008; Leppänen et al., 2004; Maurer, Bucher, Brem, and Brandeis, 2003; Shafer et al., 2000; Shafer, Yu, and Datta, 2010). While the interplay of negatively-displaced MMN and positive p-MMR is not entirely clear, it has been suggested that the positive p-MMR might reflect immature neural change-detection processes which mature into the adult-like negatively-displaced MMN during the development when children become more proficient in detecting the changes (Shafer et al., 2010). Consistent with this, Lee et al. (2012) showed that in Mandarin Chinese, deviants of large magnitude (lexical tones, vowels) elicited negatively displaced MMNs in 4–6-year-old preschoolers while deviants of smaller magnitude (initial consonant) elicited p-MMRs instead. Alternatively, the p-MMR was proposed to represent the recovery of the P100 response from refractoriness in the child standard response waveform (Shafer et al., 2010).

Thus, the MMN in children seems to differ somewhat from that of adults and there are several factors related to the maturation that underlie the MMN elicitation in children. First, previous studies have suggested that the duration of the sensory memory traces for tone stimuli, a prerequisite for MMN elicitation, are very short in 2–3-year-

old children (1–2 s; Glass, Sachse, & von Suchodoletz, 2008). While the duration of the sensory memory traces increases to 3–5 s by 6 years of age (Glass et al., 2008), they are still much shorter than in adults (up to 30 s; Winkler et al., 2002). Therefore, children's MMN elicitation may be more susceptible to changes in stimulus repetition rate than that of adults. Second, the test-retest reliability of the MMN in children may not be as good as in adults since MMN amplitudes in children have in some cases been suggested to change between two recording sessions (Uwer & von Suchodoletz, 2000). Taken together, the findings imply that the MMN in children may be smaller in amplitude than that of adults, require larger magnitudes of change to be elicited, include components not seen in the adult MMN waveform (e.g. the p-MMR), be more sensitive to differences between the experimental paradigms (e.g., to differences in SOA), and MMN test-retest reliability may be poorer in children than in adults.

### **1.2.5 MMN in fetuses and infants**

Previous studies have indicated that the MMN develops very early and MMN-like responses have been found already during the last trimester of pregnancy to pure tone contrasts in fetuses. For example, Huotilainen et al. (2005) found statistically significant responses in 12 out of 17 fetuses and Draganova et al. (2005), using similar stimuli, in 48%, or in 66% (Draganova, Eswaran, Murphy, Lowery, & Preissl, 2007) of the fetuses. Consistent with the findings in fetuses, MMN has been recorded from sleeping infants, first by Alho et al. (1990).

While the MMN develops early, the number of studies using the MMN to assess the neural basis of infant auditory perception is small and the studies have some disparity with regards to the infant MMN polarity. For example, both negative and positive MMN responses (with positive MMN often referred to as MMR; for the sake of clarity, the term MMN has been used throughout this thesis, regardless of response polarity) have been found in infants to tones (e.g., Čeponienė et al., 2000; Leppänen, Eklund, & Lyytinen, 1997). Furthermore, some studies report responses of both positive and negative polarity within the same group and using the same stimuli (e.g., tones: Cheour et al. (2002b); harmonic tones: Kushnerenko, Čeponienė, Balan, Fellman, & Näätänen (2002); vowel duration changes: Friedrich, Weber, & Friederici (2004)). In some



studies, the infants are analyzed in two separate groups depending on response polarity (e.g. Trainor et al., 2003; and Mueller, Friederici, & Mannel, 2012; 3 month old infants).

The different results found on the infant MMN polarity may be due to several factors. First, the MMN polarity in infants has been suggested to depend on the alertness (i.e., sleep stage) (Cheour et al., 2002b; Friederici, Friedrich, & Weber, 2002; Friedrich et al., 2004; Shafer, Yu, & Datta, 2011; Shafer, Yu, & Garrido-Nak, 2012), although this effect was not found in all studies (Martynova, Kirjavainen, & Cheour, 2003). Second, the MMN polarity was found to be correlated with heart rate variability and vagal tone, which are associated with the maturation of the nervous system, such as myelination (Leppänen et al., 2004). Third, the magnitude of the change affects the MMN polarity with negative MMNs being elicited by large and positive ones by small changes (tones: Morr, Shafer, Kreuzer, & Kurtzberg, 2002; French horn resonance changes: Vestergaard et al., 2009), (see, however, Kushnerenko et al., 2001). Fourth, the MMN polarity may partly depend on gender or possible gender differences in cortical maturation (Mueller, Friederici, & Mannel, 2012). Fifth, the MMN polarity may depend on the reference electrode used (e.g. ipsilateral mastoid: Pihko et al., 1999; Cz: Tew, Fujioka, He, & Trainor, 2009; mastoid average: Carral et al., 2005; Cz and re-referencing to an average electrode: He, Hotson, & Trainor 2009b). Sixth, also filtering parameters may affect the MMN polarity: if lower frequency components (0–1 or 0–3 Hz) are not filtered out, the MMNs may be predominantly positive and if the low frequency components are excluded, also negative MMN components appear in the data (He, Hotson, & Trainor, 2007; He et al., 2009b; Trainor et al., 2003; Weber et al., 2004). Seventh, fast presentation rate could diminish the neural refractoriness, which can be seen as a decreased MMN latency and amplitude, and may thus result in more negative waveforms (He et al., 2009b). Finally, individual differences in observed ERP polarities may not be limited to the MMN only. For example, Sambeth, Huotilainen, Kushnerenko, Fellman, and Pihko (2006) reported that one infant out of 11 elicited responses of the opposite polarity than the rest of the infants for all stimuli used in the experiment. This suggests that the MMN and ERP response polarity may not depend on the stimuli and paradigm only, but also on individual differences between infants.

Furthermore, the infant MMN responses seem to change during the first year of life, changing from positive MMNs to resemble adult-like negative MMNs (He et al., 2007; He, Hotson, & Trainor, 2009a; He et al., 2009b; Trainor et al., 2003; Weber et al., 2004), although different patterns of development have also been reported (Fellman et al., 2004; Kushnerenko et al., 2002). Trainor et al. (2003) suggested that early in development the positive MMN is generated by the slow and unsynchronized activity of layer IV neurons, representing thalamocortical input, while the negative MMN, appearing later in development, is generated by the mature function of neurons in deeper cortical areas. While the hypothesis that the infant MMN matures mainly from positive to negative in polarity during the first year of life seems to have been generally accepted (e.g. Mueller et al., 2012), the neuronal circuits associated with the changes in polarity are largely unexplored. However, the neural processes underlying the maturation of change detection during development cannot be solely explained by a shift from positive to negative MMN responses, because during development also other components than the MMN are seen in the deviant-minus-standard difference waveform (see Chapter 1.2.4). Some studies suggest that in 4–6-year-old children, negative adult-like MMNs are elicited only to changes which they proficiently discriminate and positive p-MMRs to changes which they may not accurately perceive (Lee et al., 2012).

Regardless of the plethora of possible factors associated with the infant MMN polarity, various approaches have been used to tackle the issue of some infants eliciting MMNs of opposite polarity to the majority of the group. Some studies reject the data of a minority of participants who showed an opposite polarity (Carral et al., 2005; Čeponienė et al., 2002; Sambeth et al., 2009), and other studies report infant data with positive and negative MMNs separately (Mueller et al., 2012).

## 2 Aims of the study

### 2.1 The main aims of the thesis

This thesis investigates neural discrimination of speech sound changes in word context from infancy to adulthood. In order to study speech sound discrimination in word context without the influence of the semantics of the word on the discrimination, the first aim of the thesis was to develop a multifeature MMN paradigm for studying changes in speech sounds in pseudo word context. The feasibility of the pseudo word paradigm for studies in adults and children was assessed in Studies I and II, respectively. For infants, the thesis assessed whether multifeature MMN paradigm is viable for investigating speech sound discrimination in easy (single syllables, Study III) and challenging (pseudo word context, Study IV) settings. The second aim of the thesis was to use the paradigm to explore whether infant sound discrimination skills are enhanced by prenatal auditory stimulation (Study IV).

### 2.2. Specific aims of the studies

**Study I** investigated whether MMNs are elicited in adults to changes of intensity, pitch, vowel duration, and vowel identity in the multifeature paradigm with the changes embedded in naturally produced pseudo words [tatata] (called the *multifeature word paradigm* from now on).

**Study II** assessed how children between the ages of 4 and 12 years process the changes of consonant duration, intensity, pitch, vowel duration, and vowel identity in pseudo word context using the multifeature word paradigm. In addition, it was determined whether the processing of speech sound changes, reflected by the MMN, predicted future performance in cognitive tests.

**Study III** explored the feasibility of the multifeature paradigm with single syllables for investigating infants' neural basis of discriminating changes of consonant identity, intensity, pitch, vowel duration, and vowel identity in single syllables. Furthermore, the study evaluated if the MMN responses to changes in speech sounds are similar between the traditional oddball and multifeature paradigms.

**Study IV** investigated how well infants neurally process the changes of intensity, pitch, vowel duration, and vowel identity in speech sounds in pseudo word context using the multifeature word paradigm, and whether this processing is enhanced by additional prenatal stimulation during the last trimester of pregnancy.

## 3 Methods

### 3.1 Participants

Studies I and II were carried out at the Cognitive Brain Research Unit at the University of Helsinki. The participants in Study I were healthy adults (N=18, 6 males) and reported having no neurological disorders or hearing deficits. The ages of the participants were between 20 and 35 years (mean 24 years). Thirty children participated in Study II. They were divided on basis of their age into preschoolers (children under the age of 7 years; from 4 years 1 month to 6 years 11 months, mean 5 years 4 months, N=15), and schoolchildren (children of 7 years of age and older; from 7 years 0 months to 12 years 9 months, mean 9 years 2 months, N=15). The parents of the children in Study II reported the children as having no neurological disorders, learning deficits or problems with hearing. The reading skills of the participants in Study II were determined with Lukilasse reading fluency test (Häyrinen, Serenius-Sirve, & Korkman, 1999) while the verbal IQ (VIQ) and performance IQ (PIQ) were assessed with block design and verbal comprehension subtests from the Wechsler Intelligence Scale for Children IV test (WISC-IV; Wechsler, 2004) 14 to 17 months after the study. Preschoolers VIQ and PIQ scores were 11.13 (SD 3.00) and 10.13 (SD 3.83), respectively, whereas schoolchildren's VIQ and PIQ scores were 11.47 (SD 2.85) and 13.01 (SD 2.28). The reading fluency test indicated that 6 preschoolers (for the 6 children, mean 10.17, SD 2.29), and all the schoolchildren (mean 11.00, SD 2.08) could read.

The infants in Studies III and IV were recruited from the internet discussion boards and the maternity ward of the Women's Hospital of the Hospital District of Helsinki and Uusimaa. For Studies III and IV, the hearing of the infants was tested with Evoked Oto-Acoustic Emissions (EOAE, ILO88 Dpi, Otodynamics Ltd., Hatfield, UK). All infants passed the test and were considered healthy by a neonatologist. Details of the participants in Studies III and IV are listed in Table 1.

**Table 1.** Participant details for Studies III and IV. Means in brackets.

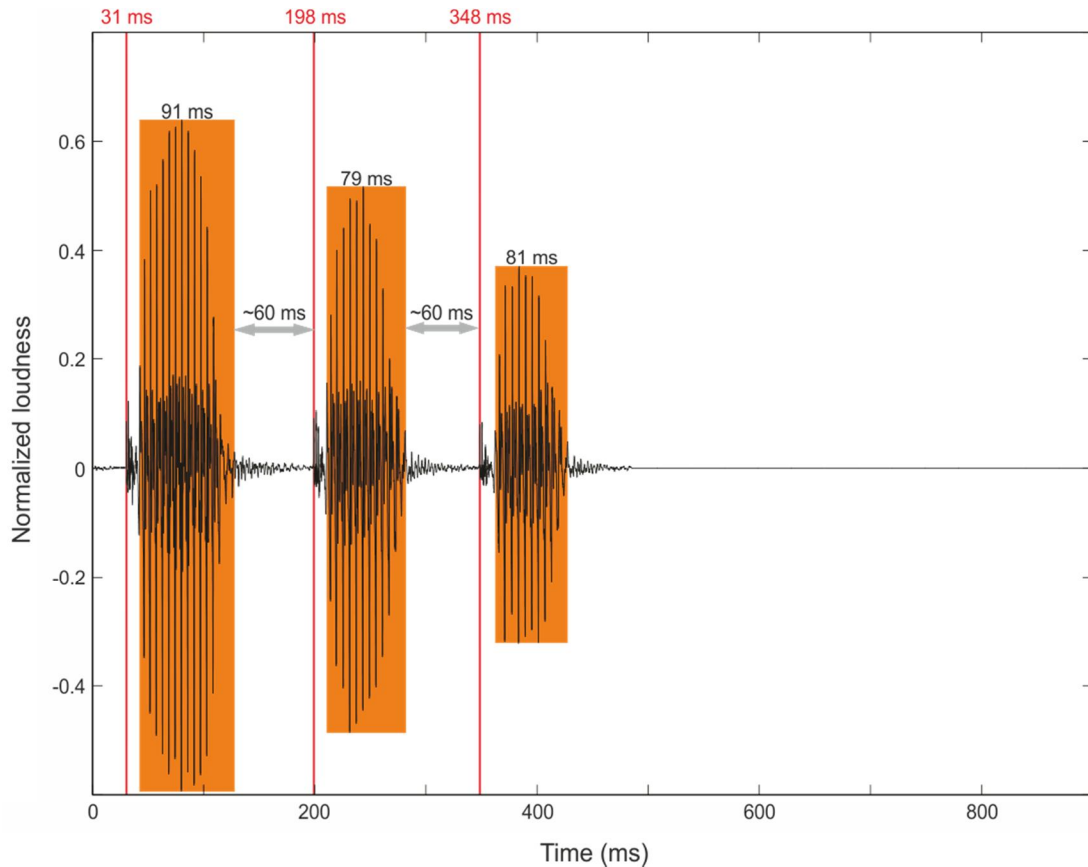
<b>Study</b>	<b>Age (days)</b>	<b>GA (weeks+days)</b>	<b>Weight</b>	<b>APGAR</b>
<b>III</b> (multifeature group, N=15)	0.5–3.5d (1.3d)	37+4–42+2 (39+5)	3260–4215 g (3710 g)	7–9 (9)
<b>III</b> (oddball group, N=13)	0.5–4d (1.9d)	38+4–42+1 (39+3)	3030–4150 g (3560 g)	9 (9)
<b>IV</b> (learning group, N=17)	1–27d (5.5d)	38+0–42+1 (39+6)	2880–4740 g (3650 g)	7–10 (9)
<b>IV</b> (control group, N=16)	1–7d (4.0d)	38+0–42+3 (40+2)	2485–4840 g (3590 g)	7–9 (8)

The adult participants in Study I gave their written consent to participate in the Study. Parents of Study II gave a written consent for their child to participate in the experiment while children gave their oral consent. The mother or both parents gave their oral consent for their infant to participate in Studies III and IV. Study I was approved by the Ethical Committee of the former Department of Psychology, University of Helsinki, Finland. Study II was approved by the Ethical Committee of the Hospital District of Helsinki and Uusimaa, Finland. Studies III and IV were approved both by the Ethical Committee of the Hospital District of Helsinki and Uusimaa, Finland and by the Ethical Committee of the former Department of Psychology, University of Helsinki, Finland.

### **3.2 Stimuli**

In Studies I, II, and IV, the multifeature word paradigm was used, where the standard stimulus was a non-synthetic trisyllabic pseudo word [tatata] produced by a native female speaker of Finnish (see Figure 1). As the stimulus was naturally produced, it included the natural variation present in Finnish words, namely the middle and final syllables being quieter and lower in pitch in comparison to the initial syllable. In addition, the vowel duration varied slightly between the syllables (from 79 to 91 ms). For the standard stimulus, the pitch of the initial syllable was 177 Hz, the middle syllable 169 Hz and the final syllable 167 Hz, the intensities of the middle and final syllables were 2 and 3 dB smaller than the intensity of the initial syllable, respectively,

and the vowel durations for the initial, middle, and final syllables were 91, 79, and 81 ms, respectively. In total, the standard pseudo word [tatata] was 480 ms in duration, with approximately 60 ms gaps between the syllables.



**Figure 1.** The waveform of the [tatata] pseudo word used in Studies I, II, and IV. The orange bars denote vowel durations, red lines denote the onset of syllables, and arrows between the syllables denote the approximate gaps between the syllables.

A total of four change types were used in Studies I and IV: intensity, pitch, vowel duration, and vowel identity. Intensity, pitch, and vowel duration changes were created by modifying the standard pseudo word [tatata] using Praat program (Boersma, 2001). The pseudo word containing the vowel identity change was produced separately by the aforementioned native female speaker of Finnish in order to capture all natural coarticulation effects and modified using Praat program to match the standard pseudo word in pitch, voice onsets, intensity, vowel length, and pauses between the syllables. In

Study II, another change type, consonant duration, was used in addition to the aforementioned 4 change types.

Table 2 lists the change types and magnitudes of the stimuli in the multifeature word paradigm used in Studies I, II, and IV. Differences in the vowels are apparent in F1 and F2 frequencies, which are 750 and 1450 Hz for vowel /a/ and 560 and 1240 for vowel /o/. Furthermore, the long vowels in Finnish are usually twice as long as short ones (Kukkonen, 1990), and thus the duration change used in Studies I, II and IV (approximately 100% increase in duration) is prototypical for Finnish. In Study I, the changes were presented in initial, middle, and final word syllables while in Studies II and IV changes were presented in the middle syllable only. For intensity and pitch changes in Studies I and IV, both increases and decreases of intensity and pitch were used while Study II included only pitch increments. Furthermore, two magnitudes of pitch changes were used (8% and 15% in Studies I and IV, 15% and 50% in Study II). In addition to the multifeature word paradigm, an oddball paradigm with pure tones (standard of 1000 Hz and a deviant of 1100 Hz) was used in Study IV.

**Table 2.** Change types of the multifeature word paradigm and their magnitudes used in Studies I, II, and IV. N/A indicates that the change type was not used in the experiment.

Study	Consonant duration	Intensity	Pitch	Vowel duration	Vowel identity
I	N/A	+6 dB, -6dB	+8%, -8%, +15%, ±15%	+80 ms	/a/-/o/
II	+100 ms	+6 dB, -6dB	+15%, +50%	+80 ms	/a/-/o/
IV	N/A	+6 dB, -6dB	+8%, -8%, +15%, ±15%	+80 ms	/a/-/o/

In Study III, 170 ms long semi-synthetic Finnish-language syllables /te:/ and /pi:/ were used. The pitch of the syllables was 101 Hz and the duration, 170 ms, is typical for a long vowel in Finnish (Kukkonen, 1990). The syllables were generated using the Semisynthetic Speech Generation method (SSG; Alku, Tiitinen, & Näätänen, 1999); identical stimuli were used previously by Pakarinen et al. (2009) and Lovio et al. (2009). Five types of sound changes were used: intensity ( $\pm 7$  dB), pitch ( $\pm 8\%$ ), consonant (/te:/ -> /pe:/ and /pi:/ -> /ti:/) or vowel identity (/te:/ -> /ti:/ and /pi:/ -> /pe:/) and vowel duration (-70 ms). Again, the change types can be classified as either



phonetic (intensity, pitch) or phonological (consonant identity, vowel duration, vowel identity).

### **3.3 Procedure**

In Studies I and II, the participants were seated in an electrically and acoustically shielded room watching a silent movie and instructed not to pay any attention to the sounds. In Studies III and IV, the infants were recorded in a silent room in the maternity ward of the Women's Hospital of the Hospital District of Helsinki and Uusimaa by a trained nurse who took notes on the sleep stages and activity of the infant.

The stimuli in Study I were presented via headphones at 60 dB (SPL) level. In Study II, the stimuli were presented through two high-quality loudspeakers in 45° angle at both sides of the participant, approximately at 1 meter distance from the participants' ear, at 60 dB SPL. In Studies III and IV stimuli were played from two loudspeakers placed on both sides of the head while the infants were lying in a crib or in an infants' care seat. The loudness of the stimuli in Studies III and IV was measured to be approximately 60 dB (SPL) at the location of the infants' head.

In Study I the EEG was recorded using the NeuroScan system and a Synamps amplifier (Compumedics NeuroScan, El Paso, TX) with 28 Ag/Cl electrodes with a 500-Hz sampling rate through a band pass filter with 0.1 Hz and 40 Hz as cutoff frequencies. In Study II, the EEG was recorded using a Biosemi Active2 system (BioSemi B.V., Amsterdam, The Netherlands) with a sampling rate of 512 Hz and band-pass filtering of 0.16–100 Hz. In Studies III and IV the EEG was recorded with the Neuroscan system with a Synamps amplifier using nine electrodes placed according to the international 10-20 system: F3, F4, C3, Cz, C4, P3, P4, T3, and T4. In Studies I and II, the EEG was referenced to the nose while Studies III and IV used an average of two mastoid electrodes as a common reference. In all studies, the EOG was recorded with two electrodes, one below and one at the outer corner of the right eye.

### **3.4 Prenatal stimulation**

During the last trimester of pregnancy in Study IV, the fetuses in the learning groups were exposed to some of the auditory stimuli used in the experiments. Starting from pregnancy week 29+0, the mothers were instructed to play the learning material at loud volume between five and seven times per week until birth and never during or after birth. The mothers kept diaries and documented how often they had played the material. On average, the mothers in Study IV played the material 50–71 times (mean 60).

The learning material consisted of two four-minute sequences, interspersed with children's music and other music chosen by the mother. The total length of the learning material was 15 minutes. In both sequences, the standard [tatata] stimulus of the multifeature word paradigm was presented 429 times ( $p = 0.7$ ), the pitch changes (+8% or -8%,  $p = 0.05$  for both; or +15% or -15%,  $p = 0.05$  for both) and vowel identity change 146 times ( $p = 0.1$ ). Both the pitch and vowel identity changes occurred in the middle syllable. During the prenatal exposure, the fetuses in the learning groups heard the standard [tatata] 21450–30459 times (mean 25740), each of the four pitch changes 7400–10508 times (mean 8880), and the vowel identity change 7300–10366 times (mean 8760).

### **3.5 Experimental paradigms**

In all studies, the MMNs were recorded using the multifeature paradigm (Näätänen et al., 2004) with standard and deviant stimuli alternating. Specifically, Studies I, II and IV used the multifeature word paradigm while Study III utilized the variation of the multifeature paradigm with speech sounds developed by Pakarinen et al. (2009). In addition, the MMNs to pitch and vowel identity changes elicited in the multifeature paradigm were compared to those elicited in a traditional oddball paradigm in Study III. SOA of one second was used in Studies I and IV, 900 ms in Study II, and 650 ms in Study III. The control condition with pure tone stimuli in Study IV had SOA of 800 ms. Probabilities and the number of repetitions for the stimuli in Studies I–IV are listed in Table 3.

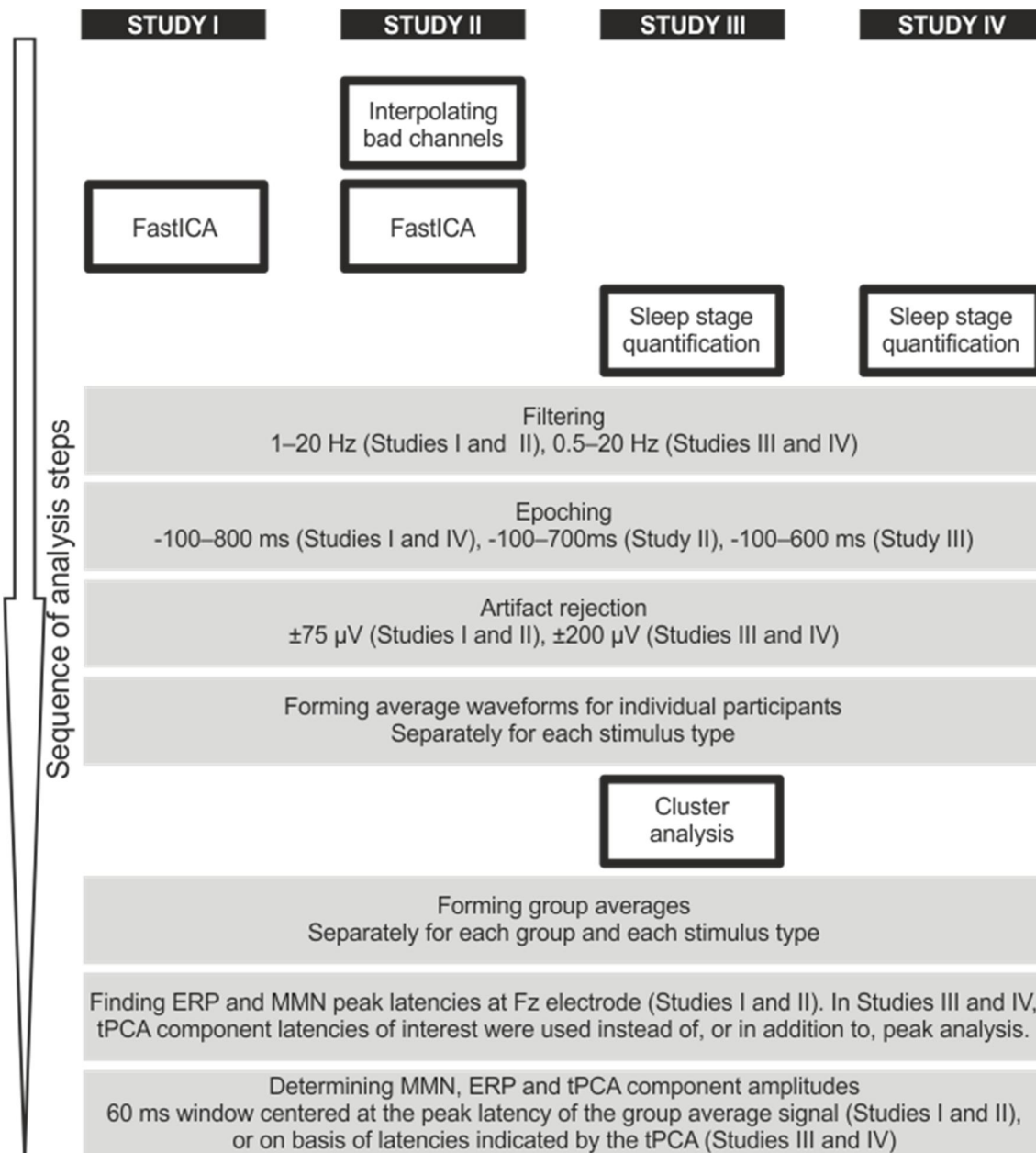
In Study I, all the change types were presented in three different word positions: initial, middle, and final. However, in Studies II and IV, all changes occurred in the middle syllable.

**Table 3.** Probabilities for the different stimuli (number of repetitions in brackets) for Studies I–IV. N/A indicates that the change type was not used in the experiment.

<b>Study</b>	<b>Consonant duration</b>	<b>Consonant identity</b>	<b>Intensity</b>	<b>Pitch</b>	<b>Vowel duration</b>	<b>Vowel identity</b>
<b>I</b> (for each word position)	N/A	N/A	5% (90) for each	5% (90) for each	10% (180)	10% (180)
<b>II</b>	10% (200)	N/A	10% (200)	5% (100) for both	10% (200)	10% (200)
<b>III</b> (multifeature group)	N/A	10% (300)	10% (300)	10% (300)	10% (300)	10% (300)
<b>III</b> (oddball group)	N/A	N/A	N/A	N/A	10% (135)	10% (135)
<b>IV</b> (main experiment)	N/A	N/A	5% (120) for each	5% (120) for each	10% (240)	10% (240)
<b>IV</b> (control condition with tones)	N/A	N/A	N/A	10% (60)	N/A	N/A

### 3.6 Data analysis

The offline data analysis sequence of Studies I–IV is shown in Figure 2. In Study II Matlab’s invdist-interpolation algorithm was used. Independent component analysis (ICA) was used to remove eyeblinks in Studies I and II. For ICA analysis in Studies I and II using FastICA algorithm (Hyvärinen & Oja, 2000), previously interpolated channels were excluded. Due to small number of channels, interpolation or ICA was not used in Studies III and IV.



**Figure 2.** Data analysis sequence in Studies I–IV. The steps were conducted from top to down.

Prior to filtering, the sleep stages of the infants in Studies III and IV were quantified. The active sleep (AS) stage is characterized by low-amplitude high-frequency activity while quiet sleep stage (QS) shows either high-amplitude low-frequency activity or trace alternants, defined by high- and low-voltage low-frequency waves alternating (e.g. Anders, 1971; Mirmiran, Maas, & Ariagno, 2003). For all infants, the amount of time spent in each sleep stage was determined by comparing the number of accepted epochs in active sleep to the total number of accepted epochs in both active and quiet sleep. The

data during which the infant was awake were discarded due extensive artefacts arising from, for example, movement and crying.

Filtering, epoching, and artefact rejection were conducted in similar manner in all studies (see Figure 2 for details; zero-phase band pass filter was utilized in all Studies), after which individual waveforms were separately created for each participant and stimulus type. Corresponding change types from /te:/ and /pi:/ stimulus blocks were averaged together in Study III, as were increases and decreases of intensity in Studies I–IV and increases and decreases of pitch in Studies I, III, and IV. The 8% and 15% pitch changes were averaged together in Study I. In study IV, pitch increments and decrements were also analyzed separately. In Study III, a subgroup of participants had predominantly negative MMNs whereas the majority of the group had positive MMNs. K-means cluster analysis, using the MMN amplitudes from seven electrodes (F3 F4, C3, Cz, C4, P3, P4) as variables, was conducted to assess whether this was a genuine effect. The cluster analysis divided the participants into two separate groups: one with predominantly positive MMNs and another with negative MMNs. The MMN amplitudes of these two groups were statistically significantly different from each other for all change types except for the vowel identity change in the *oddball group*. Specifically, in the *multifeature group*, for the consonant identity changes, the MMN amplitudes of positive and negative responders were statistically different from each other in six out of seven electrodes used in the analysis. For the other change types, statistically different MMNs were found as follows: intensity, 5/7 electrodes; pitch, 6/7 electrodes; vowel duration, 7/7 electrodes; vowel identity, 7/7 electrodes. In the *oddball group*, statistically different MMNs between positive and negative responders were found in 7/7 electrodes for the vowel duration changes and 0/7 electrodes for the vowel identity changes. For further analysis, the minority of the participants with negative MMNs was excluded from the statistical analysis but their data are shown in figures. In Studies III and IV, the responses from F3, F4, C3, Cz, and C4 electrodes were averaged together to reduce noise.

The ERP and MMN (and p-MMR in Study II) peak latency was determined from Fz electrode of the group-average waveform (Studies I and II) or from the average signal of the electrodes (Study IV). In Studies III and IV, temporal principal component analysis (tPCA) was used to determine the latencies of interest (see, e.g., Dien, 1998; Leppänen

et al., 2010) where variables were mean amplitudes in succeeding 10 ms windows and cases were ERP amplitudes from different electrodes, participants and stimuli, separately for both groups. Latencies, where the tPCA components had factor loadings of 0.8 or greater, were analyzed further. To determine if the response amplitudes were statistically significant, the mean amplitudes within a 60 ms window centered on the peak latency (or within the latency range indicated by the tPCA) were compared to zero using two-tailed t-tests.

In Study II, the incidence of MMN was also assessed at an individual level. For individual participants, the analysis was conducted similarly as on group level, except that the data from electrodes F3, Fz, F4, C3, Cz, and C4 were averaged together.

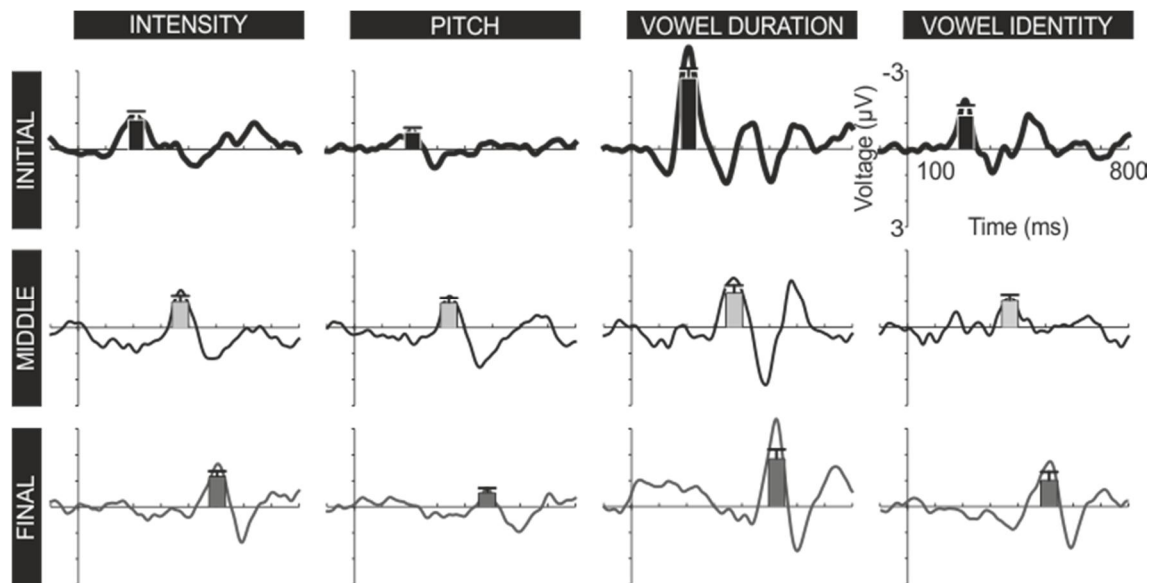
Analysis of variance (ANOVA) was used in Studies I–IV to investigate the effects of, for example, different paradigms, change types, deviation magnitudes, frontality, lateralization, and group on MMN (or p-MMR in Study II) amplitude or latency. Two-tailed t-tests, corrected for unequal variances where applicable, were used to study differences between groups in Study IV. Possible differences in background variables between the groups in Studies II–IV were determined with two-tailed t-tests, corrected for unequal variances when appropriate. Correlations between the MMN and ERP amplitudes or latencies on psychometric test scores (Study II) or amount of prenatal stimulation (Study IV) were analyzed with Pearson's correlation, Bonferroni-corrected for multiple comparisons in Study II. In Study II, the incidence of statistically significant MMN between the groups was compared using chi-square test ( $\chi^2$ ). Effect sizes (Partial eta squared ( $\eta^2$ ) for ANOVA, Cohen's  $d$  for t-tests, coefficients of determination,  $R^2$ , for correlations) are reported where applicable.

Greenhouse-Geisser correction was used for all ANOVAs where applicable, for t-tests the t-values were corrected for unequal variances where applicable, and all post-hoc tests were corrected using Bonferroni correction. For all Studies, the corrected p-values and the original degrees of freedom are reported. The Bonferroni-corrected p-values were calculated and reported as follows:  $p * n < \alpha$ , where  $p$  is the p-value obtained from the uncorrected test,  $n$  is the correction coefficient and  $\alpha$  is the alpha level.

## 4 Results

### 4.1 Multifeature word paradigm in studies of adults and children (Studies I and II)

In Study I, adults had statistically significant MMNs for all change types (intensity pitch, vowel duration, vowel identity) in all word positions (initial, middle, final;  $p < 0.01$  for all comparisons; see Figure 3). The vowel duration changes elicited larger MMNs than other change types in some word positions (stimulus  $\times$  word position,  $F_{6,11} = 3.04$ ,  $p < 0.021$ ,  $\eta^2 = 0.16$ ): the vowel duration change elicited larger MMNs than other change types in the initial ( $p < 0.001$ ) and final ( $p < 0.046$ ) position. Furthermore, MMNs were larger for the vowel duration changes in the initial position than in middle ( $p < 0.001$ ) or final ( $p < 0.012$ ) position. In addition, in the final position the intensity changes elicited larger MMNs than pitch changes ( $p < 0.011$ ).



**Figure 3.** MMN waveforms and amplitudes in Study I from electrode Fz. The bars denote the different word position (initial, middle, final), the error bars denote standard errors of the mean.

MMN distributions also varied across the scalp between the change types. In the front-back direction ( $F_{2,15} = 27.15$ ,  $p < 0.001$ ,  $\eta^2 = 0.63$ ), MMNs were larger in the frontal than parietal electrodes ( $p < 0.001$ ), and central than parietal regions ( $p < 0.002$ , except for the changes in the middle word position,  $p < 0.073$ ). Furthermore, the main

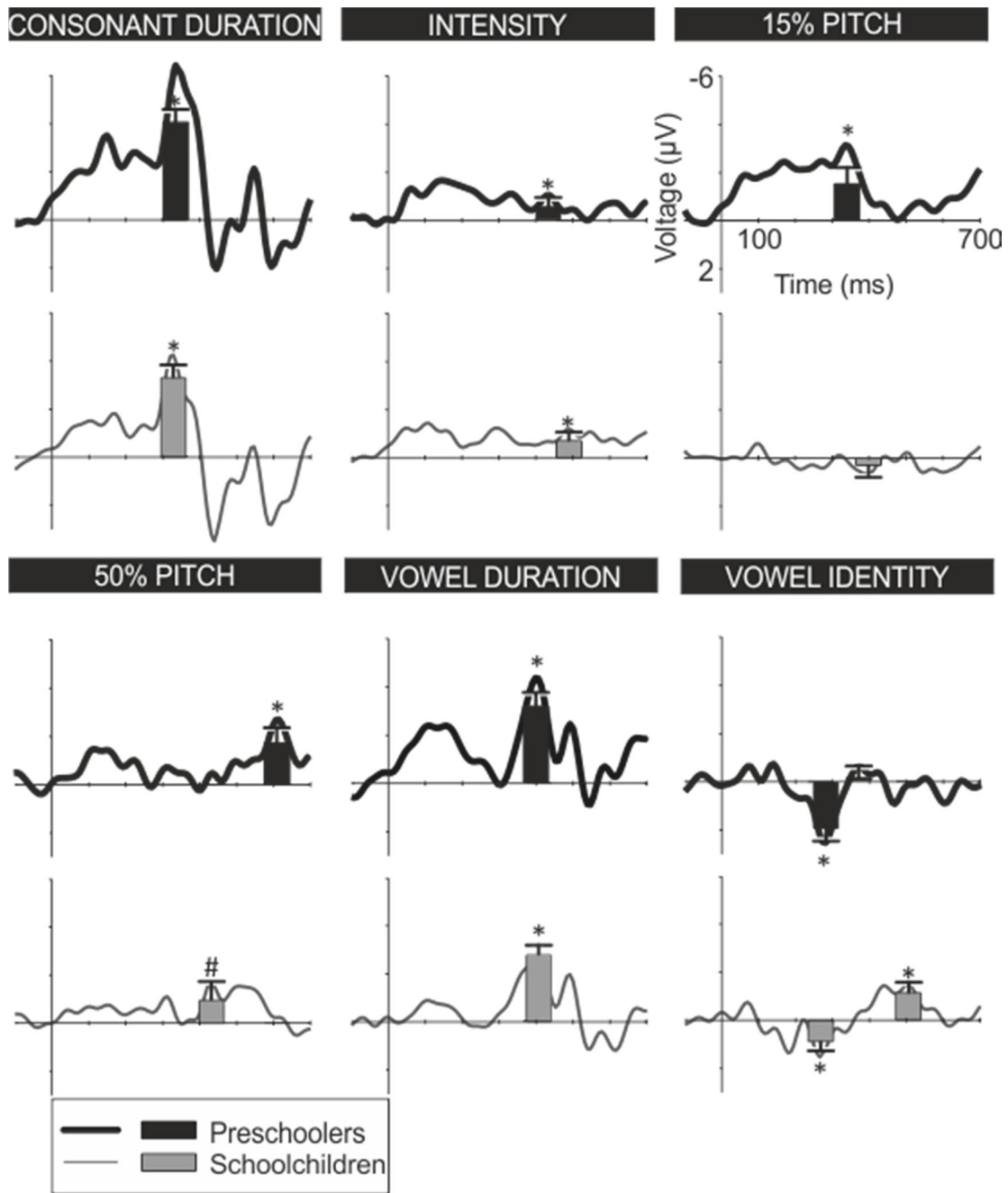
effect of lateralization ( $F_{2,15} = 5.55$ ,  $p < 0.009$ ,  $\eta^2 = 0.26$ ) showed that the responses to vowel identity changes were larger on the midline than on the right ( $p < 0.016$ ) and tended to be larger on the left than on the right scalp ( $p < 0.080$ ). MMNs to the vowel duration changes were larger on the midline than on the left scalp ( $p < 0.006$ ).

In Study II, preschoolers had statistically significant MMNs to consonant duration, intensity, 15% pitch, 50% pitch, and vowel duration changes, and schoolchildren to consonant duration, intensity, vowel duration, and vowel identity changes (see Figure 4). The p-MMR for the vowel identity change was statistically significant in both groups. Consonant and vowel duration changes elicited larger MMNs than other stimulus types (main effect of stimulus,  $F_{5,24} = 22.32$ ,  $p < 0.001$ ,  $\eta^2 = 0.44$ ;  $p < 0.001$  except for vowel duration versus 50% pitch change,  $p < 0.041$ ). Furthermore, the MMN amplitudes differed between the groups (group \* stimulus -interaction,  $F_{5,24} = 2.69$ ,  $p < 0.034$ ,  $\eta^2 = 0.09$ ): preschoolers had larger MMNs to 15% pitch changes than schoolchildren ( $F_{1,28} = 6.09$ ,  $p < 0.020$ ,  $\eta^2 = 0.18$ ) and schoolchildren tended to have larger MMNs to vowel identity changes than preschoolers ( $F_{1,28} = 2.95$ ,  $p < 0.097$ ,  $\eta^2 = 0.10$ ).

Tests on MMN topography showed that MMNs were larger in frontal than central electrodes ( $F_{1,28} = 29.15$ ,  $p < 0.005$ ,  $\eta^2 = 0.25$ ). MMNs were also larger at the Fz than F3 electrode (frontality \* laterality interaction,  $F_{2,27} = 3.39$ ,  $p < 0.043$ ,  $\eta^2 = 0.24$ ). MMN latency for the gap changes tended to become shorter with age ( $r = -0.49$ ,  $p < 0.063$ ,  $R^2 = 0.24$ ). Furthermore, larger MMNs for the intensity changes predicted larger VIQ scores 14-17 months after the EEG recording ( $r = -0.52$ ,  $p < 0.063$ ,  $R^2 = 0.27$ ).

In individual children, statistically significant or nearly significant MMNs were found in 7–80% of the cases (see Table 4). There were no group differences in the amount of statistically significant or nearly significant MMNs or p-MMRs ( $\chi^2$ -test,  $p > 0.283$  for all comparisons). However, the children who had statistically significant or near significant ( $p < 0.1$ ) p-MMRs to the vowel identity change had smaller PIQ scores 14–17 months after the EEG recording than children who had no statistically significant p-MMRs ( $t_{28} = 2.57$ ,  $p < 0.023$ ,  $d = 2.16$ ).





**Figure 4.** Preschooler and schoolchildren MMNs and p-MMRs for the change types of Study II from Fz electrode. Bars denote MMN amplitudes from the average of F3, Fz, F4, C3, Cz, and C4 electrodes. Error bars denote standard errors of the mean. \*:  $p < 0.05$ , #:  $p < 0.1$

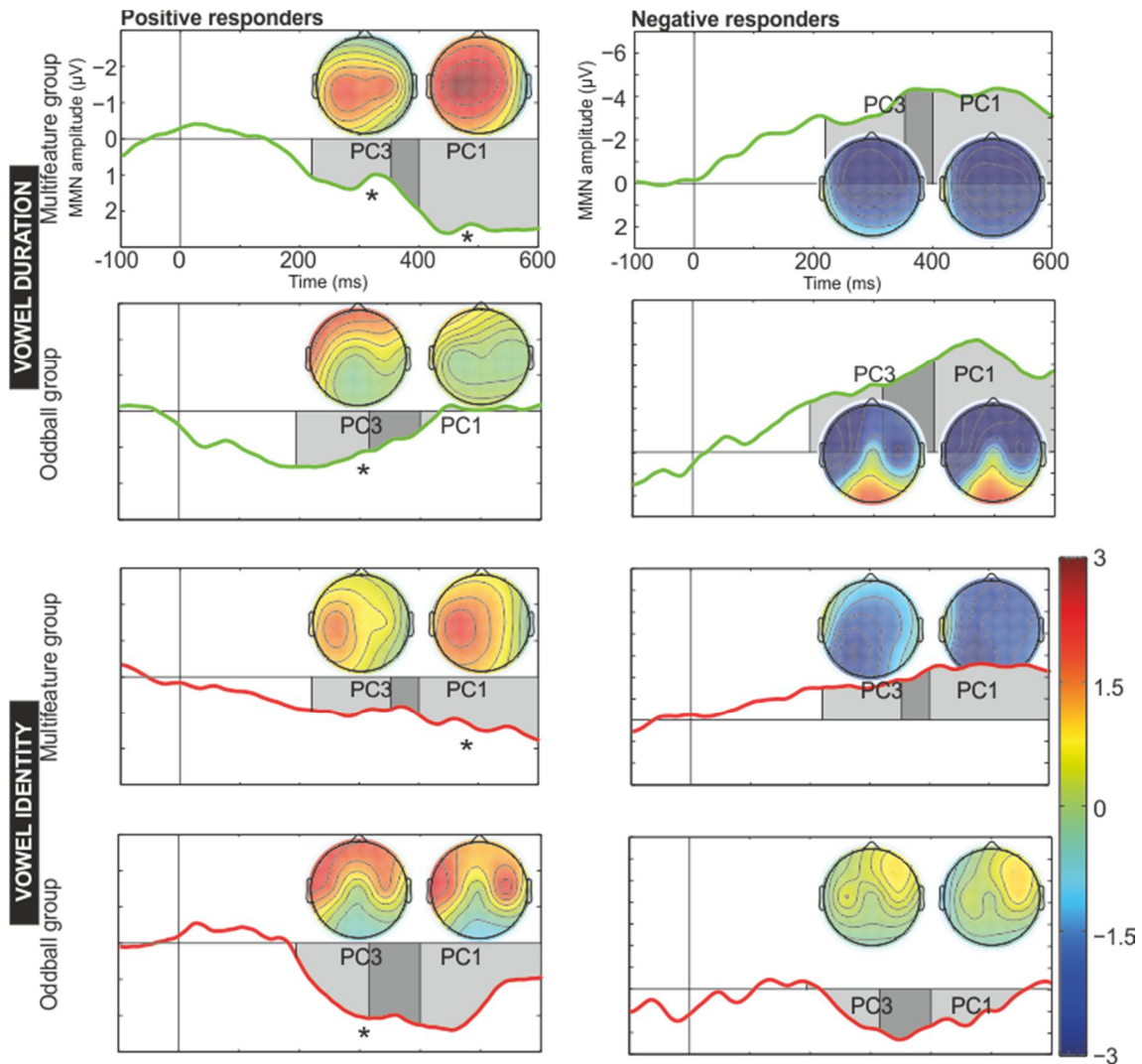
**Table 4.** The number of children with statistically significant MMNs and p-MMRs for each change type (the number of children with responses that tended to be significant ( $p < 0.1$ ) in brackets).

Group (N=15 for both groups)	Change type						
	Consonant duration	Intensity	15%	50%	Vowel duration	Vowel identity	
			pitch	pitch		MMN	p-MMR
Preschoolers	12(13)	1(3)	3(5)	7(8)	10(12)	2(2)	3(5)
Schoolchildren	11(12)	3(5)	2(3)	7(7)	9(10)	4(5)	2(4)

## 4.2 The multifeature paradigm in studies of infants (Studies III and IV)

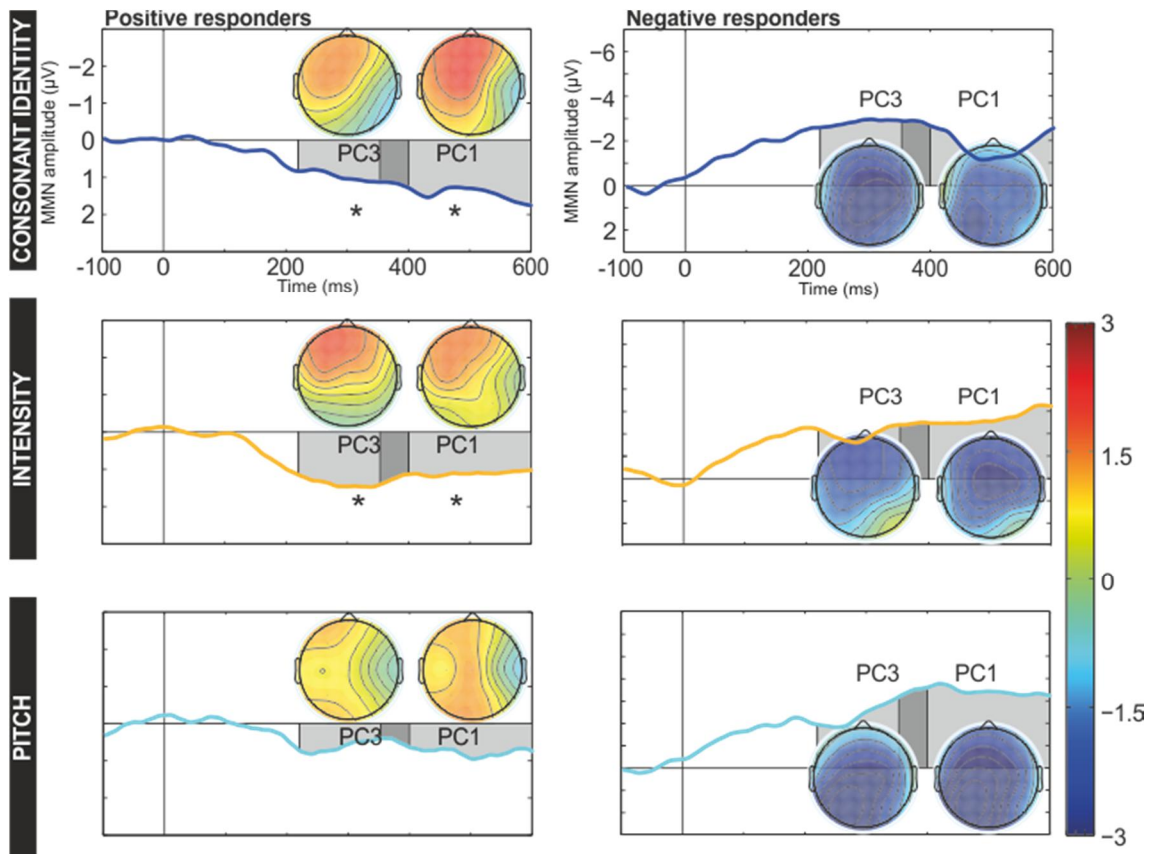
In Study III, the tPCA resulted in 3 principal components (PC's) with factor loadings above 0.8 after rotation and these 3 components accounted for 79.6% of the variance of the *multifeature* and 83.2% of the *oddball group* data. The aforementioned components were in the latency range of 350–600 (PC1; 360–600 ms for the *multifeature* and 320–600 ms for the *oddball group*), 50–200 (PC2) and 200–400 ms (PC3; 220–400 ms for the *multifeature* and 290–400 ms for the *oddball group*) after stimulus onset. As the latencies indicated by the tPCA components were similar between the paradigms, the average MMN amplitudes between the groups were compared with each other within the time windows indicated by those components. PC2 was not analyzed further as it was considered to reflect early processes not necessarily related to MMN.

Infants in the *multifeature group* had statistically significant MMNs for consonant identity, intensity, vowel duration, and vowel identity changes in single syllables while infants in the *oddball group* had a statistically significant MMN for both vowel identity and vowel duration changes. See Figures 5 and 6 for MMN waveforms and scalp distributions for both positive and negative responders in *multifeature* and *oddball groups*.



**Figure 5.** The MMNs for the vowel duration and vowel identity changes in *multifeature* and *oddball* groups. Light gray bars denote the time ranges of the principal components (PC; dark gray bar indicates an overlap between the two PCs). Statistically significant ( $p < 0.05$ ) MMNs are marked with asterisks (\*) for the positive responders only; the data of the negative responders was not analyzed further due to small number of participants (four in the *multifeature* and three in the *oddball* group).

Comparing the MMNs between the *multifeature* and *oddball* groups showed a group \* frontality interaction ( $F_{1,19} = 6.06$ ,  $p < 0.026$ ,  $\eta^2 = 0.24$ ), due to larger MMNs the *oddball* than *multifeature* group in the PC3 component latency range (1.50 vs 0.70  $\mu\text{V}$ ,  $F_{1,19} = 4.63$ ,  $p < 0.044$ ,  $\eta^2 = 0.20$ ). No other differences were found in MMN amplitudes between the groups.



**Figure 6.** The MMNs for the consonant identity, intensity and pitch changes in the *multifeature group*. Light gray bars denote the time ranges of the principal components (PC; dark gray bar indicates an overlap between the two PCs). Statistically significant MMNs ( $p < 0.05$ ) are marked with asterisks (\*) for the positive responders only; data of the negative responders was not analyzed further due to small number of participants (four for the multifeature and three for the oddball group).

Analysis on effects of background factors on MMN amplitudes showed that in infants in the *oddball group*, MMNs were the larger for the vowel duration changes the more time they spent in active sleep (electrode F4,  $r = 0.81$ ,  $p < 0.005$ ,  $R^2 = 0.66$ ). Furthermore, the positive responders in the *oddball group* spent more time in active sleep than positive responders in the *multifeature group* (19.6% versus 7.4%;  $t_{18} = 2.30$ ,  $p < 0.045$ ,  $d = 1.08$ ) and negative responders tended to have smaller birth weights than positive responders ( $t_{27} = 1.90$ ,  $p < 0.076$ ,  $d = 0.72$ ).

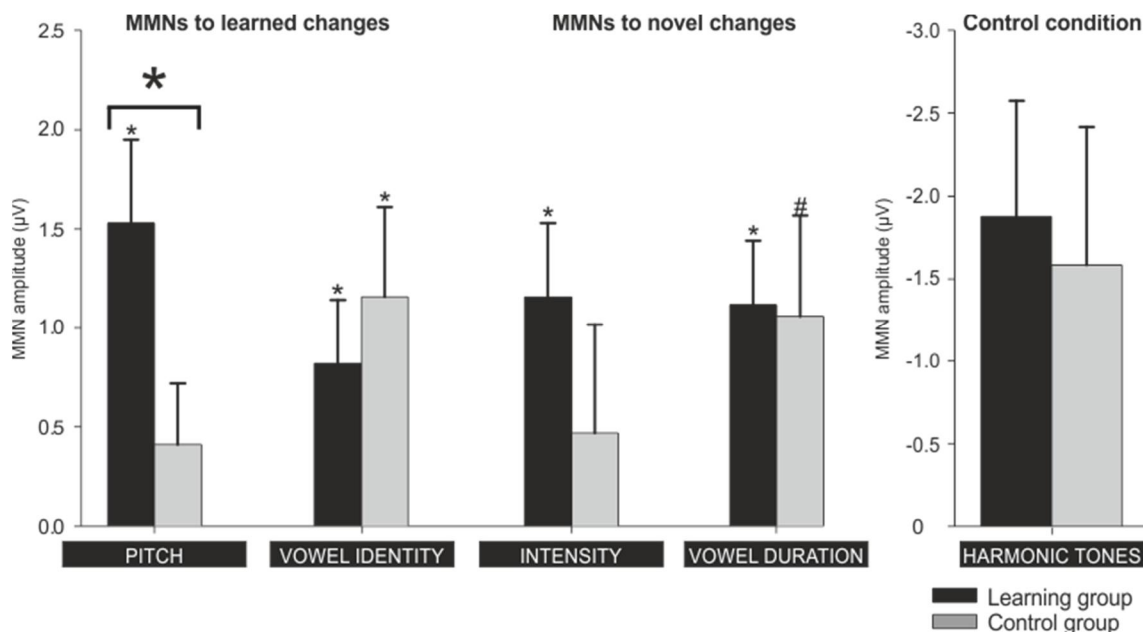
Study IV assessed the infant MMNs in a challenging pseudo word context. Infants not exposed to the stimuli in utero had statistically significant MMNs to the vowel identity change only ( $t_{15} = 2.56$ ,  $p < 0.021$ ) while the MMN to the vowel duration

change was near significant ( $t_{15} = 2.08$ ,  $p < 0.055$ ). MMNs to the intensity and pitch changes were not statistically significant. No differences in MMN amplitudes were found between the different change types. See Figure 7 for MMNs in Study IV.

### 4.3 The effects of prenatal stimulation on speech sound discrimination (Study IV)

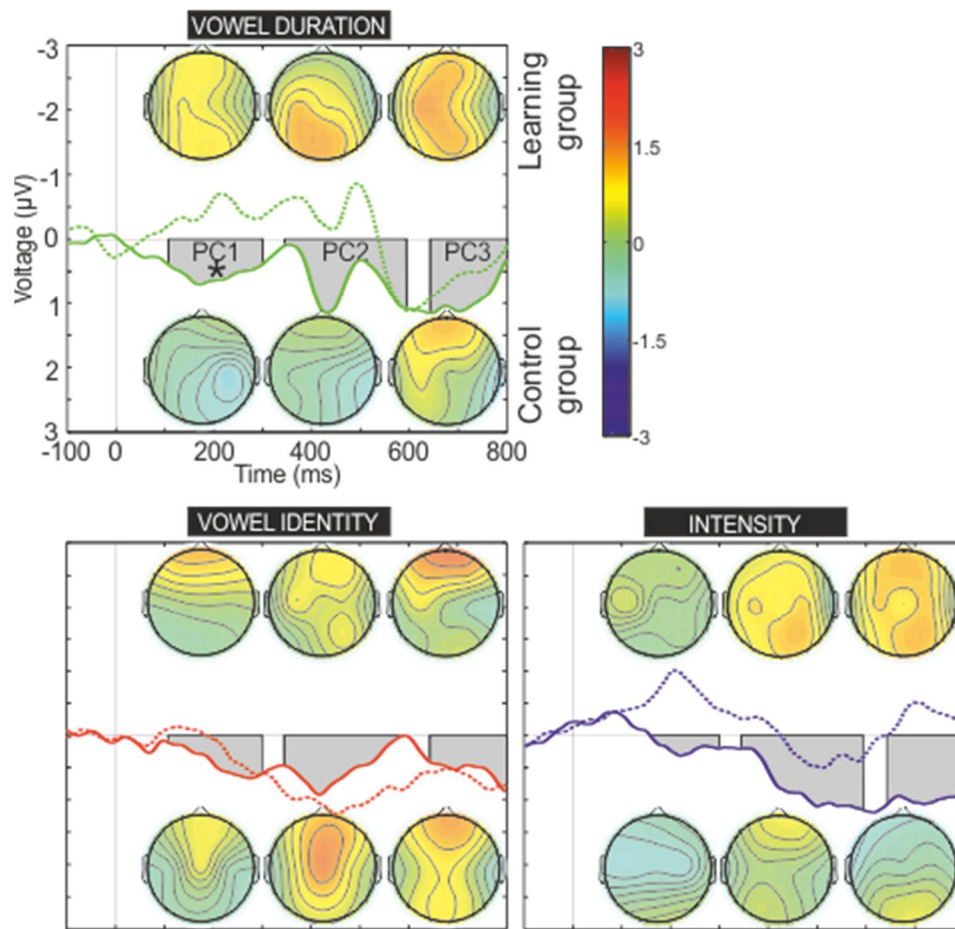
In Study IV, the effects of prenatal stimulation on the MMN were assessed using two methods. First, the MMN amplitudes were quantified as the mean amplitude in a 60-ms window centered at the largest peak in the grand-mean difference waveform. The second method utilized tPCA analysis.

Using the MMN peak amplitudes in 60-ms windows it was found that the *learning group* had statistically significant MMNs for all change types used in the experiment (see Figure 7). Furthermore, the MMN to pitch changes was stronger in the *learning group* infants that had been exposed to the stimuli prenatally, than in the naïve *control group* ( $t_{31} = 2.122$ ,  $p < 0.042$ ,  $d = 0.763$ ).

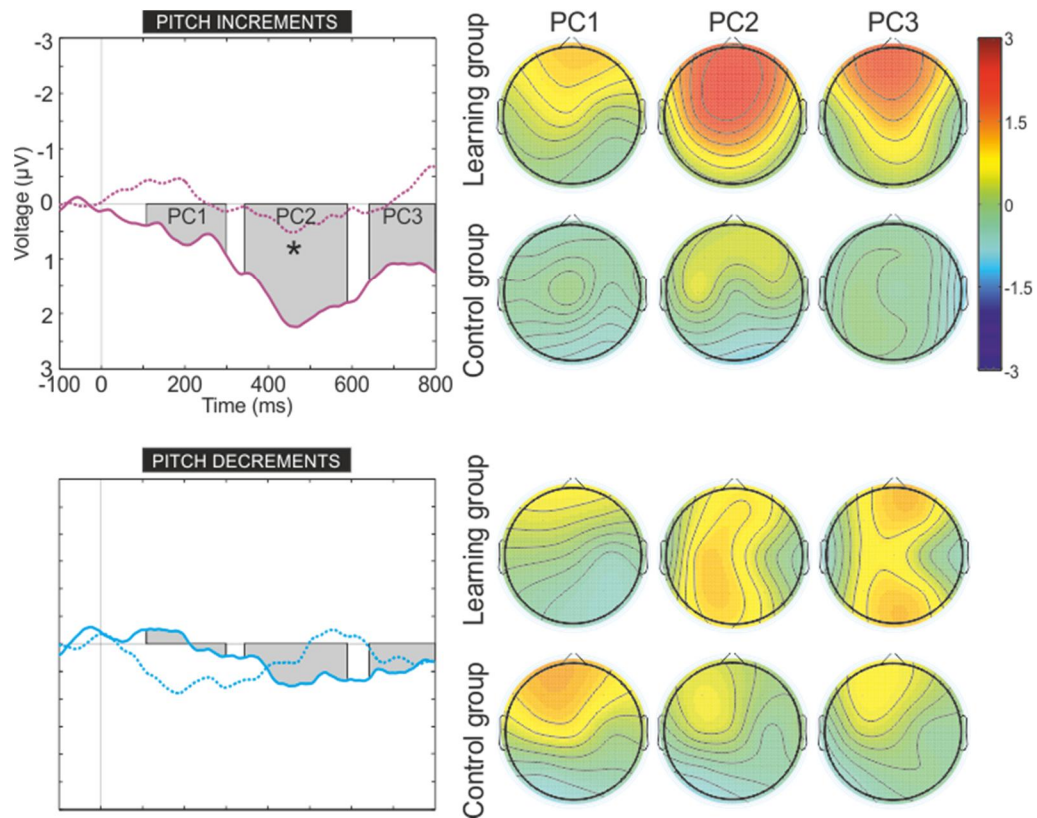


**Figure 7.** MMNs for learned and novel changes, and for the pure tones presented in the control condition in Study IV. The bars represent average MMN amplitude as the mean amplitude in a 60-ms window centered at the largest peak in the grand-mean difference waveform. Asterisks denote statistical significances (near significances ( $p < 0.1$ ) are marked with the number sign).

The effects of prenatal exposure were also assessed in greater detail using the tPCA analysis (see Figures 8 and 9). The analysis indicated that the infants in the *learning group* had larger responses to pitch changes in the 340–590 ms time range than the *control group* (stimuli \* component \* group interaction,  $F_{6,26} = 2.97$ ,  $p < 0.024$ ,  $\eta^2 = 0.41$ ). In a further analysis, it was found that this effect was due to *learning group* having larger responses to pitch increments than the *control group* (340–590 ms time range;  $F_{1,31} = 6.497$ ,  $p < 0.016$ ,  $\eta^2 = 0.17$ ), but not for pitch decrements. The amplitudes in the PC2 time range to pitch increases were also positively correlated with the amount of exposure to the prenatal stimulation (C4 electrode;  $r = 0.61$ ,  $p < 0.009$ ,  $R^2 = 0.37$ ). Furthermore, the learning effects generalized to other change types, not included in the learning material: the amplitudes in the PC1 time range for vowel duration changes were larger in the *learning* than *control group* (110–330 ms time range,  $F_{1,31} = 4.988$ ,  $p < 0.033$ ,  $\eta^2 = 0.14$ ). No statistically significant differences between the background factors (GA, APGAR, birth weight, age at measurement, sleep stages) were found between the *learning* and *control groups*. The MMNs to pitch changes in harmonic tones measured as the mean amplitude in a 60-ms window centered at the largest peak in the grand-mean difference waveform, equally unfamiliar for both groups, did not differ between the groups ( $F_{1,30} = 0.052$ ,  $p < 0.821$ ), suggesting no differences in the infants' basic auditory discrimination abilities.



**Figure 8.** The deviant-minus-standard waveforms and the principal components (PC) indicated by the tPCA for vowel duration, identity and intensity changes in Study IV. Dotted lines denote the *control group*, solid lines the *learning group*. Asterisks denote statistically significant differences between the groups.



**Figure 9.** The deviant-minus-standard waveforms and the principal components (PC) indicated by the tPCA for pitch increments and decrements in Study IV. Dotted lines denote control group, solid lines the learning group. Asterisks denote statistically significant differences between the groups.



## 5 Discussion

The first aim of the present thesis was to investigate the applicability of the multifeature MMN paradigm for studying the neural processing of speech sound changes in adults (Study I), children (Study II), and infants (Studies III and IV). Specifically, Studies I, II, and IV assessed the feasibility of the multifeature word paradigm to evaluate the neural processing of speech sound changes while Study III determined whether the multifeature MMN paradigm with single syllables is feasible for studying infants' neural discrimination of speech sounds. The second aim of the thesis was to apply the multifeature word paradigm to explore the neural correlates of fetal learning (Study IV).

The *first main finding* of the thesis was that the multifeature word paradigm is feasible for investigating neural speech sound discrimination from infancy to adulthood, although distinct MMN elicitation patterns were found in different age groups (see Chapter 5.1). The *second main finding* was that prenatal exposure to speech sounds enhances neural speech sound discrimination in infants, and that the amount of this exposure is positively correlated with the enhancement of ERPs (see Chapter 5.3). The following sections will discuss the topics relevant to this thesis in greater detail: 1) the feasibility of the multifeature word paradigm for assessing neural speech sound discrimination, 2) neural processing of speech sounds in infancy and childhood, as reflected by the multifeature word paradigm, and 3) fetal learning and its implications.

### 5.1 The feasibility of the multifeature word paradigm

In adult participants (Study I), all speech sound changes used in the experiment (vowel duration, pitch, intensity, and vowel identity) elicited statistically significant MMN responses for all word position (initial, middle, and final). In children (Study II) and infants (Study IV), for whom the changes were presented in the middle word position only, the pattern of MMN results was distinct. Both preschoolers and schoolchildren had MMNs to changes of vowel duration, consonant duration, and intensity. However, the MMN to the vowel identity change was not statistically significant in preschoolers

(children between 4 and 7 years of age) while the schoolchildren (7–12-year-old children) had no statistically significant MMNs to the 15% or 50% pitch increments. The p-MMR to vowel identity changes was statistically significant in both preschoolers and schoolchildren. In infants with no additional exposure to speech sounds in utero (the control group of Study IV), statistically significant MMNs were found to the vowel identity changes only while the MMNs to vowel duration changes tended to be statistically significant.

The consonant and vowel duration changes in children and vowel duration changes in adults elicited larger MMNs than the other speech sound changes. This is likely to be due to the Finnish participants' enhanced sensitivity to duration changes (Kirmse et al., 2008; Marie, Kujala, & Besson, 2012; Nenonen, Shestakova, Huotilainen, & Näätänen, 2003; Tervaniemi et al., 2006). However, the large magnitude of the duration changes might also have contributed to the MMN enhancement (Näätänen & Alho, 1997). Vowel duration increment also increases stimulus energy, which enhances the MMN amplitude (Takegata, Heikkilä, & Näätänen, 2011). Finally, duration changes are accompanied by the variation in subjective loudness, which may further enhance the MMN amplitude for duration changes (Todd & Michie, 2000).

However, the enhanced MMNs to duration changes may be partly due to the confounding effects of, for example, the sustained negative potential and the N1 response when the deviant-minus-standard waveform is formed from the responses to physically different sounds (Jacobsen & Schröger, 2003). The vowel and consonant duration changes used are approximately 100% longer than the vowel and consonant lengths in the standard pseudo word [tatata], and thus the MMNs to duration changes may be somewhat confounded by the differences in the N1 amplitudes for the standard and deviant stimuli. However, as the N1 attenuates due to rapid presentation of stimuli (Näätänen & Picton, 1987), it is likely that the N1 confound is the largest for the syllable in the initial word position, preceded by silence only, and diminishes for the middle and final word positions. Consistent with this, the results from Study I showed in adults that the MMN to duration changes was larger for the initial word positions than for middle or final positions. Thus it seems to be plausible that the larger MMNs to duration changes in the initial word position found in Study I result partly from the N1 confound.

Regardless of the possible N1 confound, the multifeature word paradigm has several benefits. First, preceding and following sounds may hamper the auditory processing in some developmental disorders like dyslexia (Kujala et al., 2003). Consequently, the multifeature word paradigm, presenting changes in word context, may be useful in assessing such effects. Second, this paradigm could provide a more broad profile of auditory discrimination abilities and deficits, which is one of most beneficial aspects of the multifeature paradigms (Bishop, 2007). Third, the sounds in the multifeature paradigm have constant variation, and thus represent normal speech better than the traditional oddball paradigm (Kujala et al., 2007a). Fourth, the multifeature word paradigm uses naturally produced sounds instead of synthetic speech sounds or tones, which may activate different neural networks than natural sounds (Blomert & Mitterer, 2004; Guttorm, Leppänen, Tolvanen, & Lyytinen, 2003; Vouloumanos & Werker, 2004). However, the disadvantage in the use of naturally produced speech sounds is that the changes in these sounds cannot be controlled for as accurately as changes in synthetic sounds.

Finally, the MMN and p-MMRs elicited in the multifeature word paradigm were associated with psychometric test scores in children. Specifically, the elicitation of MMN by intensity changes was associated with higher VIQ scores 14–17 months after the recording. Conversely, the elicitation of p-MMR was associated with lower PIQ scores. While it is not clear how neural speech sound processing, as reflected by the MMN, is associated with the performance in psychometric tests, the correlation between these measures promotes the usage of MMN in studies of children who are too young to be investigated with psychometric tests. While our results are consistent with previous studies (Bauer et al., 2009; Mikkola et al., 2007), the evidence for the associations between psychometric tests and MMN is still scarce. Thus, the results on the connections between the MMN and psychometric tests should be cautiously interpreted and further explored in future studies.

To conclude, the multifeature word paradigm appears to reliably elicit MMN for all change types and word positions in adults. In children, however, the MMN elicitation seems to depend on the children's age, and not all speech sound changes elicited statistically significant MMNs in infants and children. The findings on children and infant neural speech sound processing are discussed in detail in the next section.

## 5.2 Neural speech sound processing in childhood and infancy

In children (Study II), the MMN elicitation depended on age: while both preschoolers and schoolchildren had MMNs to vowel duration, consonant duration, and intensity changes, the preschoolers had no statistically significant MMN to vowel identity changes and schoolchildren had no MMNs to pitch increments. While previous studies have repeatedly shown statistically significant MMNs for vowel identity changes in children under 7 years of age (Čeponienė, Lepistö, Alku, Aro, & Näätänen, 2003; Lovio et al., 2009; Shafer et al., 2010), it is possible that the preceding and following syllables in the multifeature word paradigm challenged the auditory discrimination of young children to such an extent that no MMN was elicited at group level. Thus, while even children under 7 years of age can certainly discriminate the vowel identity changes, it may be that their auditory discrimination abilities have not yet sufficiently matured to elicit an MMN in word context. However, the p-MMR to vowel identity changes was statistically significant in both preschoolers and schoolchildren. Therefore, it seems to be plausible that children's processing of vowel identity changes in this challenging word paradigm relies at least in part on the immature change detection processes, as reflected by the p-MMR (Shafer et al., 2010).

The differences between preschoolers and schoolchildren in neural pitch change processing may be associated with the type of speech preschoolers and schoolchildren are exposed to in their daily lives. Infants and preschool children spend most of their time either with their parents or in daycare, where they are mostly exposed to parentese, infant directed speech, or to speech of other infants, which are characterized by higher pitch and larger pitch range than adult-directed speech (Lee, Potamianos, & Narayanan, 1999; Warren-Leubecker & Bohannon, 1984). This could influence the developing auditory system to focus predominantly on pitch cues in speech sound discrimination, even though young children's pitch discrimination abilities are still immature (Jensen & Neff, 1993). The statistically significant MMNs to pitch increments in pseudo word context in preschoolers might reflect the enhanced experience of, and orientation to, pitch variation in speech. In contrast, schoolchildren are more often exposed to adult-directed speech in formal school settings, facilitating in speech sound categorization the use of other auditory cues than exaggerated pitch contours. This increased exposure to

adult-directed speech in schoolchildren might then give rise to the adult-like MMN to vowel identity changes and diminish the MMN to pitch changes.

In infants, the changes in vowel identity and duration elicited statistically significant or nearly significant MMNs in the multifeature word paradigm (Study IV) while the changes in pitch and intensity did not. When similar changes were presented in isolated syllables to another group of infants (Study III), vowel duration, vowel identity, intensity, and consonant identity changes elicited statistically significant MMNs but pitch changes did not. Thus, infants seem to readily discriminate vowel identity, vowel duration, and consonant identity changes, but their auditory change-detection processes are not sufficiently matured to detect intensity changes in the challenging word context.

However, infants had no statistically significant MMNs to pitch changes presented either in isolation or in word context. This is surprising as infants should be particularly sensitive to pitch changes of parentese (Fernald & Kuhl, 1987). There are two possible explanations. First, the pitch changes may not have been sufficiently large to elicit an MMN in infants, as a previous study showed that in pure tones pitch changes of 5% did not reliably elicit a statistically significant MMN while those of 20% did (Novitski, Huotilainen, Tervaniemi, Näätänen, & Fellman, 2007). Second, the processing of pitch changes may be specific to the direction of change and infants may be predisposed to perceiving increases of pitch instead of decreases of pitch, as the use of increased pitch is a predominant feature in parentese and enhances speech sound discrimination in 2–3 month old infants (Karzon, 1985). While the differences in processing of pitch increments and decrements in infants could have been investigated further in Studies III and IV, in Study III increments and decrements of pitch and intensity were averaged together in order to control for the possible effects arising from the direction of change.

Taken together, the results suggest that the multifeature word paradigm reflects the effects of experience and neural development on speech sound discrimination from infancy to adulthood. This was evident in the distinct pattern of MMNs that differentiates infants from children, and preschoolers from schoolchildren.

### 5.3 Fetal learning and its implications

The results from Study IV indicated that the exposure to novel features in pseudo words before birth affects the neural responsiveness of infants in two ways. First, the MMNs to pitch increments in the middle position of the word were enhanced, and the amount of enhancement was associated with the amount of prenatal exposure. Second, the effects of the prenatal exposure generalized to other types of changes in the speech sounds, not included in the learning material. This was reflected by the statistically significant MMN to intensity changes in the *learning* but not in the *control group*. Furthermore, in the early time window, the PC1 for vowel duration changes indicated larger amplitudes for the *learning* than *control group*.

The enhanced MMNs to pitch changes indicate the development of neural commitment in fetuses due to additional prenatal exposure to speech sounds. Furthermore, the more the infants had received stimulation in utero, the stronger this neural commitment was. Since the infants were not exposed to the sounds for 1–27 days (average 5 days) after being born, the results suggest that neural memory traces developed in the fetal brain lasted for several days. However, also other types of fetal learning mechanisms have been suggested, such as habituation, which is characterized by diminished responsiveness to familiar stimuli. Such habituation effects were reported in the study of Ando and Hattori (1970), who found that infants whose mothers lived near an airport during pregnancy exhibited no startle reflex to aircraft noise. In addition, some studies on fetal learning of speech sounds also report diminished behavioral responses to familiar native language phonemes in comparison to unfamiliar nonnative phonemes (Moon et al., 2013), suggesting possible habituation effects to familiar stimuli. Alternatively, the diminished behavioral responses to familiar versus unfamiliar sounds after birth may be explained by attention as the infant might find stimuli interesting, thus reacting more strongly to new than familiar sounds. However, neither the habituation nor attention hypotheses can conclusively explain the results in Study IV as the fetuses were exposed to both the unchanged and changed pseudo words in utero, suggesting that both sounds were equally familiar to the infants. Furthermore, the enhanced MMNs to pitch increments in the learning group do not seem to be due to generally improved pitch discrimination abilities either, as no differences were found in

the MMNs to pitch increments of tones, equally unfamiliar for both *learning* and *control* groups in Study IV. Taken together, the enhanced MMNs in Study IV appear to reflect long-term memory traces developed in the fetal brain.

Regardless of the facilitating mechanism, the fetal exposure to structured sound environments seems to have positive effects. Studies in animal models have shown that structured sound environments during pregnancy are beneficial for cortical organization and synaptogenesis (Xu, Yu, Cai, Zhang, & Sun, 2009). In addition, the fetal learning effects may be long lasting as rat pups prenatally exposed to music exhibited better spatial learning abilities than rats unexposed to prenatal stimulation even 21 days after birth (Aoun, Jones, Shaw, & Bodner, 2005). Although additional prenatal stimuli resembling those that normally belong to the sound environment of the fetus seem to have positive effects for the development, novel, unstructured or abnormal sounds which the auditory system might process as noise may hinder neural development (Krueger, Horesh, & Crossland, 2012). For example, unstructured acoustic environments, such as noise, seem to have adverse effects both on adults (Kujala et al., 2004) and during development (Chang & Merzenich, 2003). Thus, fetal exposure to abnormal sound stimulation should be avoided at least until thorough long-term follow-up studies on the effects of such stimulation have been conducted.

The adverse effects of noise or unstructured acoustic environments may, in part, be associated with generally larger prevalence of language deficits in preterm children than in children born at term (Anderson & Doyle, 2008). Preterm infants are deprived of the positive effects of natural fetal learning and are mostly exposed to noisy sound environments in neonatal intensive care units (NICUs), possibly putting them at a developmental disadvantage in comparison to infants born at term (McMahon, Wintermark, & Lahav, 2012). These developmental disadvantages could possibly be ameliorated by exposing preterm infants in NICUs to sounds that infants at term are exposed to in utero, such as natural speech, maternal voice, and intrauterine sounds. Should additional exposure to sound material mitigate the possible negative consequences of preterm birth, even brief daily stimulation may be beneficial: the results of Study IV indicate that mere 15 minutes of daily exposure to speech sounds in utero enhances the neural responsiveness to speech sound changes after birth. However, due to possibly detrimental effects of abnormal sound stimulation early in development,

many of the early causal factors of auditory processing disorders need to be understood more thoroughly prior to developing interventions during the prenatal period.

## **5.4 Study limitations and future directions**

The studies of this thesis represent a developing field in neuroscience, utilized novel paradigms, and to an extent, novel analysis methods to assess the processing of speech sounds from infancy to adulthood. Therefore, there are certain limitations, which are discussed further below.

Although the multifeature word paradigm elicits MMNs in adults, 4–12-year-old children, and infants, it may not be sufficiently sensitive for investigating the speech sound discrimination in individual participants. This is relevant for studies where the development of speech sound discrimination of single individuals needs to be followed up. However, in the current multifeature paradigm only vowel and consonant duration changes elicited MMNs in at least 80% of the 4–12-year-old children. Furthermore, the MMN amplitudes of these children did not very well predict their performance in psychometric tests 14–17 months later. Although some associations between the MMN and future psychometric test performance were found in Study II, it is not apparent what the causal link between the MMN to intensity changes and children's verbal IQ is. Specifically, while the elicitation of p-MMR predicted poorer performance IQ in Study II, it is not clear why the effects were seen in the performance IQ only. Further limiting the conclusions on the interplay of MMN and performance in psychometric tests, the test battery in Study II was not broad enough to warrant far-reaching conclusions on the relationship of the MMN to psychometric test performance. For the MMN and the multifeature word paradigm to be viable for assessment or prediction of the onset of auditory processing deficits, the following issues should be investigated in future studies: 1) what are the optimal design and stimuli for reliable MMN elicitation in individual participants, 2) what are the data analysis methods most useful for assessing the MMN elicitation in individual participants, and 3) what are the associations between the MMN amplitude and performance in psychometric tests as determined with a more extensive battery of psychometric subtests.



The use of the MMN in infancy as a predictor for future outcome needs to be studied further since the MMNs for identical stimuli are in some infants positive and in others negative in polarity as shown by Study III (see also Sambeth et al., 2006). There may be a plethora of underlying effects which explain the variation in the infant MMR polarities reported by many studies (see Chapter 1.2.5). Further studies on this issue are relevant particularly if the MMN is used to assess the efficacy of interventions during the first years of life.

Finally, while the findings of Study IV show promise by suggesting that already the fetus is capable of learning fine-grained auditory information, future studies are needed to investigate the both beneficial and harmful long-term effects of increased sound stimulation on the fetus. For example, harmful sounds such as noise (Chang et al., 2003) or sounds disrupting the sleep/wake rhythms of the fetus (for a review, see Moon & Fifer, 2000) may have adverse effects. To avoid harmful effects to the fetus due to additional prenatal stimulation, future studies could teach the fetus by using sounds the fetus is exposed to during normal development, such as maternal voice, by, for example, having mother sing or speak very specific sounds or in a particular manner.

## **5.5 Conclusions**

In conclusion, the multifeature word paradigm is feasible for studies from infancy to adulthood and the neural responses in the multifeature word paradigm were associated with future performance in psychometric testing in children, suggesting that the paradigm could be feasible in assessing normal and deficient developmental trajectories on group level. In addition, the multifeature word paradigm seems to reflect changes in speech sounds processing during development, as reflected by the distinct pattern of MMNs in preschoolers and schoolchildren. A further benefit of the multifeature word paradigm over many other experimental paradigms is that it assesses stimulus discrimination in word context while many paradigms study the processing of speech sounds in isolation. This is important because in natural speech the preceding and following sounds may interfere with speech sound processing (Kujala et al., 2006b).

The results on neural correlates of fetal learning both confirm and extend the findings of previous behavioral studies on fetal learning (Moon & Fifer, 2000). While previous

studies have suggested that the fetus can learn to discriminate fine-grained auditory information (Moon et al., 2013), neural evidence for such learning effects has not previously been shown. The fetuses' ability to learn to discriminate fine-grained auditory information also suggests that auditory processing deficits might be alleviated with interventions for fetuses and infants with interventions for fetuses or infants. Taken as a whole, the findings of the present thesis hopefully encourage further development of the means for early identification and treatment of developmental disorders associated with speech sound processing deficits.

## 6 References

- Ahmed, A. U., Clarke, E. M., & Adams, C. (2008). Mismatch negativity and frequency representational width in children with specific language impairment. *Developmental Medicine and Child Neurology*, *50*, 938–944.
- Ahonen, T. (2010). Seulonta ja kielellisten erityisvaikeuksien varhainen toteaminen. Retrieved from <http://www.kaypahoito.fi/web/kh/suosituksset/naytaartikkeli/tunnus/nix01469>
- Alain, C., Campeanu, S., & Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *Journal of Cognitive Neuroscience*, *22*, 392–403.
- Alexandrov, A. A., Boricheva, D. O., Pulvermüller, F., & Shtyrov, Y. (2011). Strength of word-specific neural memory traces assessed electrophysiologically. *PLoS One*, *6*, e22999
- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology*, *77*, 151–155.
- Alku, P., Tiitinen, H., & Näätänen, R. (1999). A method for generating natural-sounding speech stimuli for cognitive brain research. *Clinical Neurophysiology*, *110*, 1329–1333.
- Anders, T. (1971). *A Manual of standardized terminology, techniques and criteria for scoring of states of sleep and wakefulness in newborn infants*. Los Angeles, California, United States: UCLA Brain Information Service, NINDS Neurological Information Network.
- Anderson, P. J., & Doyle, L. W. (2008). Cognitive and educational deficits in children born extremely preterm. *Seminars in Perinatology*, *32*, 51–58.
- Ando, Y., & Hattori, H. (1970). Effects of intense noise during fetal life upon postnatal adaptability (statistical study of reactions of babies to aircraft noise). *Journal of the Acoustical Society of America*, *47*, 1128–1130.
- Aoun, P., Jones, T., Shaw, G. L., & Bodner, M. (2005). Long-term enhancement of maze learning in mice via a generalized Mozart effect. *Neurological Research*, *27*, 791–796.
- Archibald, L. M., & Joanisse, M. F. (2012). Atypical neural responses to phonological detail in children with developmental language impairments. *Developmental Cognitive Neuroscience*, *2*, 139–151.
- Baldeweg, T., Richardson, A., Watkins, S., Foale, C., & Gruzelier, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Annals of Neurology*, *45*, 495–503.
- Bauer, P., Burger, M., Kummer, P., Lohscheller, J., Eysholdt, U., & Doellinger, M. (2009). Correlation between psychometric tests and mismatch negativity in preschool children. *Folia Phoniatrica Et Logopaedica*, *61*, 206–216.
- Best, C. T., & Strange, W. (1992). Effects of Phonological and Phonetic Factors on Cross-Language Perception of Approximants. *Journal of Phonetics*, *20*, 305–330.

- Bishop, D. V., Hardiman, M. J., & Barry, J. G. (2010). Lower-frequency event-related desynchronization: a signature of late mismatch responses to sounds, which is reduced or absent in children with specific language impairment. *Journal of Neuroscience*, *30*, 15578–15584.
- Bishop, D. V. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: Where are we, and where should we be going? *Psychological Bulletin*, *133*, 651–672.
- Blomert, L., & Mitterer, H. (2004). The fragile nature of the speech-perception deficit in dyslexia: natural vs synthetic speech. *Brain and Language*, *89*, 21–26.
- Boersma, P., & Weenink, T. (2001). Praat, a system for doing phonetics by computer. *Glott international*, *5*, 341–345.
- Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., . . . Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics*, *127*, 1034–1042.
- Bruder, J., Leppänen, P. H. T., Bartling, J., Csépe, V., Démonet, J. F., & Schulte-Körne, G. (2011). Children with dyslexia reveal abnormal native language representations: evidence from a study of mismatch negativity. *Psychophysiology*, *48*, 1107–1118.
- Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näätänen, R., & Escera, C. (2005). A kind of auditory 'primitive intelligence' already present at birth. *European Journal of Neuroscience*, *21*, 3201–3204.
- Catts, H. W. (1991). Early identification of dyslexia - evidence from a follow-up-study of speech-language impaired children. *Annals of Dyslexia*, *41*, 163–177.
- Čeponienė, R., Lepistö, T., Alku, P., Aro, H., & Näätänen, R. (2003). Event-related potential indices of auditory vowel processing in 3-year-old children. *Clinical Neurophysiology*, *114*, 652–661.
- Čeponienė, R., Kushnerenko, E., Fellman, V., Renlund, M., Suominen, K., & Näätänen, R. (2002). Event-related potential features indexing central auditory discrimination by newborns. *Cognitive Brain Research*, *13*, 101–113.
- Čeponienė, R., Hukki, J., Cheour, M., Haapanen, M. L., Koskinen, M., Alho, K., & Näätänen, R. (2000). Dysfunction of the auditory cortex persists in infants with certain cleft types. *Developmental Medicine and Child Neurology*, *42*(4), 258–265.
- Chang, E. F., & Merzenich, M. M. (2003). Environmental noise retards auditory cortical development. *Science*, *300*, 498–502.
- Cheour, M., Shestakova, A., Alku, P., Čeponienė, R., & Näätänen, R. (2002a). Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native speech sounds within two months. *Neuroscience Letters*, *325*, 187–190.
- Cheour, M., Čeponienė, R., Leppänen, P., Alho, K., Kujala, T., Renlund, M., . . . Näätänen, R. (2002b). The auditory sensory memory trace decays rapidly in newborns. *Scandinavian Journal of Psychology*, *43*, 33–39.

- Cheour, M., Korpilahti, P., Martynova, O., & Lang, A. H. (2001). Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. *Audiology & Neuro-Otology*, *6*, 2–11.
- Cheour, M., Leppänen, P. H., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology*, *111*, 4–16.
- Cheour, M., Čeponienė, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*, 351–353.
- DeCasper, A. J., Lecanuet, J-P., Busnel, M-C., Granier-Deferre, C. & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, *17*, 159-164.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, *208*, 1174–1176.
- Dien, J. (1998). Addressing misallocation of variance in principal components analysis of event-related potentials. *Brain Topography*, *11*, 43–55.
- Draganova, R., Eswaran, H., Murphy, P., Lowery, C., & Preissl, H. (2007). Serial magnetoencephalographic study of fetal and newborn auditory discriminative evoked responses. *Early Human Development*, *83*, 199–207.
- Draganova, R., Eswaran, H., Murphy, P., Huotilainen, M., Lowery, C., & Preissl, H. (2005). Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *NeuroImage*, *28*, 354–361.
- Fellman, V., Kushnerenko, E., Mikkola, K., Čeponienė, R., Leipälä, J., & Näätänen, R. (2004). Atypical auditory event-related potentials in preterm infants during the first year of life: a possible sign of cognitive dysfunction? *Pediatric Research*, *56*, 291–297
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development*, *10*, 279–293.
- Flax, J. F., Realpe-Bonilla, T., Roesler, C., Choudhury, N., & Benasich, A. (2009). Using early standardized language measures to predict later language and early reading outcomes in children at high risk for language-learning impairments. *Journal of Learning Disabilities*, *42*, 61–75.
- Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport*, *13*, 1251–1254.
- Friedrich, M., Weber, C., & Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology*, *41*, 772–782.
- Gabrieli, J. D. E. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science*, *325*, 280–283.
- Gerhardt, K. J., & Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *Journal of Perinatology*, *20*, S21–30.

- Giannakopoulou, A., Uther, M. & Ylinen, S. (2013). Enhanced plasticity in spoken language acquisition for child learners: Evidence from phonetic training studies in child and adult learners of English. *Child Language Teaching and Therapy*, 29, 201–218.
- Glass, E., Sachse, S., & von Suchodoletz, W. (2008). Development of auditory sensory memory from 2 to 6 years: an MMN study. *Journal of Neural Transmission*, 115, 1221–1229.
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A. Y., & Decasper, A. J. (2011). A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PLoS One*, 6, e17304.
- Guttorm, T. K., Leppänen, P. H. T., Tolvanen, A., & Lyytinen, H. (2003). Event-related potentials in newborns with and without familial risk for dyslexia: principal component analysis reveals differences between the groups. *Journal of Neural Transmission*, 110, 1059–1074.
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, 11, 466–472
- Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences U S A*, 102, 12639–12643.
- He, C., Hotson, L., & Trainor, L. J. (2009a). Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. *European Journal of Neuroscience*, 29, 861–867.
- He, C., Hotson, L., & Trainor, L. J. (2009b). Maturation of cortical mismatch responses to occasional pitch change in early infancy: effects of presentation rate and magnitude of change. *Neuropsychologia*, 47, 218–229.
- He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, 19, 878–892.
- Hepper, P. G., & Shahidullah, B. S. (1994). Development of Fetal Hearing. *Archives of Disease in Childhood*, 71, F81–F87.
- Hepper, P. G. (1988). Fetal Soap Addiction. *Lancet*, 1, 1347–1348.
- Holt, R. F., & Lalonde, K. (2012). Assessing toddlers' speech-sound discrimination. *International Journal of Pediatric Otorhinolaryngology*, 76, 680–692.
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2008). MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, 45, 60–69
- Huotilainen, M., Lovio, R., Kujala, T., Tommiska, V., Karma, K., & Fellman, V. (2011). Could audiovisual training be used to improve cognition in extremely low birth weight children? *Acta Paediatrica*, 100, 1489–1494.
- Huotilainen, M. (2010). Building blocks of fetal cognition: emotion and language. *Infant and Child Development*, 19, 94–98.
- Huotilainen, M., Kujala, A., Hotakainen, M., Parkkonen, L., Taulu, S., Simola, J., . . . Näätänen, R. (2005). Short-term memory functions of the human fetus recorded with magnetoencephalography. *NeuroReport*, 16, 81–84.
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, 13, 411–430.

- Häyrynen, T., Serenius-Sirve, S., & Korkman, M. (1999). *Lukilasse*. Helsinki, Finland: Psykologien kustannus.
- Isler, J. R., Tarullo, A. R., Grieve, P. G., Housman, E., Kaku, M., Stark, R. I., & Fifer, W. P. (2012). Toward an electrocortical biomarker of cognition for newborn infants. *Developmental Science*, *15*, 260–271.
- Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, *114*, 1133–1143.
- Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, *38*, 723–727.
- Jansson-Verkasalo, E., Korpilahti, P., Jäntti, V., Valkama, M., Vainionpää, L., Alku, P., . . . Näätänen, R. (2004). Neurophysiologic correlates of deficient phonological representations and object naming in prematurely born children. *Clinical Neurophysiology*, *115*, 179–187.
- Jansson-Verkasalo, E., Čeponienė, R., Valkama, M., Vainionpää, L., Laitakari, K., Alku, P., . . . Näätänen, R. (2003). Deficient speech-sound processing, as shown by the electrophysiologic brain mismatch negativity response, and naming ability in prematurely born children. *Neuroscience Letters*, *348*, 5–8.
- Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory-discrimination in preschool-children. *Psychological Science*, *4*, 104–107.
- Kaan, E., Wayland, R., Bao, M. Z., & Barkley, C. M. (2007). Effects of native language and training on lexical tone perception: An event-related potential study. *Brain Research*, *1148*, 113–122.
- Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, *39*, 326–342.
- Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *International Journal of Psychophysiology*, *67*, 131–143.
- Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., . . . Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, *14*, 220–224.
- Klein, N. K., Hack, M., & Breslau, N. (1989). Children who were very low birth-weight - development and academic-achievement at 9 years of age. *Journal of Developmental and Behavioral Pediatrics*, *10*, 32–37.
- Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language*, *76*, 332–339.
- Kostović, I., & Judas, M. (2010). The development of the subplate and thalamocortical connections in the human foetal brain. *Acta Paediatrica*, *99*, 1119–1127.
- Kraus, N., Koch, D. B., McGee, T. J., Nicol, T. G., & Cunningham, J. (1999). Speech-sound discrimination in school-age children: psychophysical and neurophysiologic measures. *Journal of Speech, Language, and Hearing Research*, *42*, 1042–1060.

- Krueger, C., Horesh, E., & Crossland, B. A. (2012). Safe sound exposure in the fetus and preterm infant. *Jognn-Journal of Obstetric Gynecologic and Neonatal Nursing*, *41*, 166–170.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B-Biological Sciences*, *363*, 979–1000.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.
- Kujala, T., Lepistö, T., & Näätänen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, *37*, 697–704.
- Kujala, T., & Näätänen, R. (2010). The adaptive brain: a neurophysiological perspective. *Progress in Neurobiology*, *91*, 55–67.
- Kujala, T., Kuuluvainen, S., Saalasti, S., Jansson-Verkasalo, E., von Wendt, L., & Lepistö, T. (2010). Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. *Clinical Neurophysiology*, *121*, 1410–1419.
- Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, *74*, 1–19.
- Kujala, T. (2007). The role of early auditory discrimination deficits in language disorders. *Journal of Psychophysiology*, *21*, 239–250.
- Kujala, T., Lovio, R., Lepistö, T., Laasonen, M., & Näätänen, R. (2006). Evaluation of multi-attribute auditory discrimination in dyslexia with the mismatch negativity. *Clinical Neurophysiology*, *117*, 885–893.
- 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. *European Journal of Neuroscience*, *24*, 2420–2427.
- Kujala, T., Shtyrov, Y., Winkler, I., Saher, M., Tervaniemi, M., Sallinen, M., . . . Näätänen, R. (2004). Long-term exposure to noise impairs cortical sound processing and attention control. *Psychophysiology*, *41*, 875–881.
- Kujala, T., Belitz, S., Tervaniemi, M., & Näätänen, R. (2003). Auditory sensory memory disorder in dyslexic adults as indexed by the mismatch negativity. *European Journal of Neuroscience*, *17*, 1323–1327.
- Kujala, T., Karma, K., Čeponienė, R., Belitz, S., Turkkila, P., Tervaniemi, M., & Näätänen, R. (2001). Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children. *Proceedings of the National Academy of Sciences U S A*, *98*, 10509–10514.
- Kukkonen, P. (1990). *Patterns of phonological disturbances in adult aphasia*. Helsinki, Finland: Suomalaisen Kirjallisuuden Seura.



- Kushnerenko, E., Čeponienė, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: a longitudinal ERP study. *NeuroReport*, *13*, 1843–1848.
- Kushnerenko, E., Cheour, M., Čeponienė, R., Fellman, V., Renlund, M., Soinen, K., . . . Näätänen, R. (2001). Central auditory processing of durational changes in complex speech patterns by newborns: An event-related brain potential study. *Developmental Neuropsychology*, *19*, 83–97.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language & Communication*, *35*, 165–188.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for primary speech and language delay: a systematic review of the literature. *International Journal of Language & Communication*, *33*, 21–23.
- Lecanuet, J. P., Granier-Deferre, C., Jacquet, A. Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology*, *36*, 29–39.
- Lecanuet, J. P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics Gynecology and Reproductive Biology*, *68*, 1–23.
- Lecanuet, J. P., Granier-Deferre, C., Decasper, A. J., Maugeais, R., Andrieu, A. J., & Busnel, M. C. (1987). Fetal perception and discrimination of speech stimuli as demonstrated by cardiac reactivity, preliminary-results. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences*, *305*, 161–164.
- Lee, C. Y., Yen, H. L., Yeh, P. W., Lin, W. H., Cheng, Y. Y., Tzeng, Y. L. & Wu, H. C. (2012). Mismatch responses to lexical tone, initial consonant, and vowel in Mandarin-speaking preschoolers. *Neuropsychologia*, *50*, 3228–3239.
- Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, *105*, 1455–1468.
- Leminen, A. (2012). Neurocognitive processing of inflected and derived words. (Doctoral dissertation). *Studies in Cognitive Science 1*.
- Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, *1066*, 147–157.
- Leppänen, P. H. T., Hämäläinen, J. A., Guttorm, T. K., Eklund, K. M., Salminen, H., Tanskanen, A., . . . Lyytinen, H. (2012). Infant brain responses associated with reading-related skills before school and at school age. *Neurophysiologie Clinique-Clinical Neurophysiology*, *42*, 35–41.
- Leppänen, P. H. T., Hämäläinen, 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*, 1362–1376.

- Leppänen, P. H. T., Guttorm, T. K., Pihko, E., Takkinen, S., Eklund, K. M., & Lyytinen, H. (2004). Maturational effects on newborn ERPs measured in the mismatch negativity paradigm. *Experimental Neurology, 190*, S91–S101.
- Leppänen, P. H. T., 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. *Developmental Neuropsychology, 22*, 407–422.
- Leppänen, P. H. T., Eklund, K. M., & Lyytinen, H. (1997). Event-related brain potentials to change in rapidly presented acoustic stimuli in newborns. *Developmental Neuropsychology, 13*, 175–204.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.
- Lipski, S. C., Escudero, P., & Benders, T. (2012). Language experience modulates weighting of acoustic cues for vowel perception: an event-related potential study. *Psychophysiology, 49*, 638–650.
- 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 Research, 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 Research, 1335*, 53–62.
- Lovio, R., Pakarinen, S., Huotilainen, M., Alku, P., Silvennoinen, S., Näätänen, R., & Kujala, T. (2009). Auditory discrimination profiles of speech sound changes in 6-year-old children as determined with the multi-feature MMN paradigm. *Clinical Neurophysiology, 120*, 916–921.
- Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology, 19*, 1994–1997.
- Marie, C., Kujala, T., & Besson, M. (2012). Musical and linguistic expertise influence pre-attentive and attentive processing of non-speech sounds. *Cortex, 48*, 447–457.
- Martin, B. A., Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2003). Maturation of mismatch negativity: A scalp current density analysis. *Ear and Hearing, 24*, 463–471.
- Martynova, O., Kirjavainen, J., & Cheour, M. (2003). Mismatch negativity and late discriminative negativity in sleeping human newborns. *Neuroscience Letters, 340*, 75–78.
- Maurer, U., Bucher, K., Brem, S., & Brandeis, D. (2003). Development of the automatic mismatch response: from frontal positivity in kindergarten children to the mismatch negativity. *Clinical Neurophysiology, 114*, 808–817.
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology, 47*, 66–122.
- McCormack, J., McLeod, S., McAllister, L., & Harrison, L. J. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology, 11*, 155–170.
- McMahon, E., Wintermark, P., & Lahav, A. (2012). Auditory brain development in premature infants: the importance of early experience. *Annals of the New York Academy of Sciences, 1252*, 17–24.

- Mikkola, K., Kushnerenko, E., Partanen, E., Serenius-Sirve, S., Leipälä, J., Huotilainen, M., & Fellman, V. (2007). Auditory event-related potentials and cognitive function of preterm children at five years of age. *Clinical Neurophysiology*, *118*, 1494–1502.
- Milovanov, R., Huotilainen, M., Esquef, P. A. A., Alku, P., Välimäki, V., & Tervaniemi, M. (2009). The role of musical aptitude and language skills in preattentive duration processing in school-aged children. *Neuroscience Letters*, *460*, 161–165.
- Mirmiran, M., Maas, Y. G., & Ariagno, R. L. (2003). Development of fetal and neonatal sleep and circadian rhythms. *Sleep Medicine Reviews*, *7*, 321–334.
- Moon, C. M., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatrica*, *102*, 156–160.
- Moon, C. M., & Fifer, W. P. (2000). Evidence of transnatal auditory learning. *Journal of Perinatology*, *20*, S37–44.
- Moon, C. M., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*, 495–500.
- Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: a timeline of development. *International Journal of Audiology*, *46*, 460–478.
- Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*, *23*, 118–136.
- Mortimore, T., & Crozier, W. R. (2006). Dyslexia and difficulties with study skills in higher education. *Studies in Higher Education*, *31*, 235–251.
- Mueller, J. L., Friederici, A. D., & Mannel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences U S A*, *109*, 15953–15958.
- Nelson, D. G. K., Hirshpasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language-learning. *Journal of Child Language*, *16*, 55–68.
- Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research*, *16*, 492–495.
- Novitski, N., Huotilainen, M., Tervaniemi, M., Näätänen, R., & Fellman, V. (2007). Neonatal frequency discrimination in 250–4000-Hz range: electrophysiological evidence. *Clinical Neurophysiology*, *118*, 412–419.
- 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. *Clinical Neurophysiology*, *123*, 424–458.
- Näätänen, R., Kujala, T., Kreegipuu, K., Carlson, S., Escera, C., Baldeweg, T., & Ponton, C. (2011). The mismatch negativity: an index of cognitive decline in neuropsychiatric and neurological diseases and in ageing. *Brain*. 134:3435–53
- 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*, 4–22.

- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, *115*, 140–144.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences*, *24*, 283–288.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.
- Näätänen, R., & Alho, K. (1997). Mismatch negativity - The measure for central sound representation accuracy. *Audiology and Neuro-Otology*, *2*, 341–353.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huutilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434.
- 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. *Psychophysiology*, *24*, 375–425.
- Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, in press.
- Pakarinen, S., Huutilainen, M., & Näätänen, R. (2010). The mismatch negativity (MMN) with no standard stimulus. *Clinical Neurophysiology*, *121*, 1043–1050.
- Pakarinen, S., Lovio, R., Huutilainen, 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. *Biological Psychology*, *82*, 219–226.
- Pakarinen, S., Teinonen, T., Shestakova, A., Kwon, M. S., Kujala, T., Hämäläinen, H., . . . Huutilainen, M. (2013). Fast parametric evaluation of central speech-sound processing with mismatch negativity (MMN). *International Journal of Psychophysiology*, *87*, 103–110.
- Pardo, J. & Remez, R. (2006). The perception of speech. In Traxler, M. & Gernsbacher, M. (Eds.), *Handbook of Psycholinguistics* (pp. 201–248). The Netherlands: Academic Press.
- Peters, A., Abrams, R. M., Gerhardt, K., & Griffiths, S. (1993). Transmission of airborne sounds from 50–20000 Hz into the abdomen of sheep. *Journal of Low Frequency Noise Vibration*, *12*, 16–24.
- Pihko, E., Leppänen, P. H. T., 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*, 901–905.
- Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, *109*, 2190–2201.
- Preston, J., & Edwards, M. L. (2010). Phonological awareness and types of sound errors in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research*, *53*, 44–60.

- Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity. as a tool for studying higher cognitive processes. *Progress in Neurobiology*, *79*, 49–71.
- 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*, 607–616.
- Putkinen, V., Niinikuru, R., Lipsanen, J., Tervaniemi, M., & Huotilainen, M. (2012). Fast measurement of auditory event-related potential profiles in 2–3-year-olds. *Developmental Neuropsychology*, *37*, 51–75.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech-perception without traditional speech cues. *Science*, *212*, 947–950.
- 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. *Developmental Science*, *8*, 162–172.
- Sambeth, A., Pakarinen, S., Ruohio, K., Fellman, V., van Zuijen, T. L., & Huotilainen, M. (2009). Change detection in newborns using a multiple deviant paradigm: A study using magnetoencephalography. *Clinical Neurophysiology*, *120*, 530–538.
- Sambeth, A., Huotilainen, M., Kushnerenko, E., Fellman, V. & Pihko, E. (2006). Newborns discriminate novel from harmonic sounds: A study using magnetoencephalography. *Clinical Neurophysiology*, *117*, 496-503.
- Schulte-Körne, G., & Bruder, J. (2010). Clinical neurophysiology of visual and auditory processing in dyslexia: A review. *Clinical Neurophysiology*, *121*, 1794–1809.
- Shafer, V. L., Yu, Y.H., & Garrido-Nak, K.. (2012). Neural mismatch indices of vowel discrimination in monolingually and bilingually exposed infants: does attention matter? *Neuroscience Letters*, *526*, 10-14.
- Shafer, V. L., Yu, Y.H., & Datta, H. (2011). The development of english vowel perception in monolingual and bilingual infants: neurophysiological correlates. *Journal of Phonetics*, *39*, 527-545.
- Shafer, V. L., Yu, Y. H., & Datta, H. (2010). Maturation of speech discrimination in 4-to 7-yr-old children as indexed by event-related potential mismatch responses. *Ear and Hearing*, *31*, 735–745.
- Shafer, V. L., Morr, M. L., Datta, H., Kurtzberg, D., & Schwartz, R. G. (2005). Neurophysiological indexes of speech processing deficits in children with specific language impairment. *Journal of Cognitive Neuroscience*, *17*, 1168–1180.
- Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2000). Maturation of mismatch negativity in school-age children. *Ear and Hearing*, *21*, 242–251.
- Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V., & Huotilainen, M. (2004). Orderly cortical representation of vowel categories presented by multiple exemplars. *Cognitive Brain Research*, *21*, 342–350
- Shestakova, A., Huotilainen, M., Čeponienė, R., & Cheour, M. (2003). Event-related potentials associated with second language learning in children. *Clinical Neurophysiology*, *114*, 1507–1512.

- Shtyrov, Y., Kimpka, L., Pulvermüller, 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*, 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*, 657–663.
- Sorokin, A., Alku, P., & Kujala, T. (2010). Change and novelty detection in speech and non-speech sound streams. *Brain Research*, *1327*, 77–90.
- Spence, M. J., & Freeman, M. S. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior & Development*, *19*, 199–212.
- Steinberg, J., Truckenbrodt, H., & Jacobsen, T. (2010). Preattentive phonotactic processing as indexed by the mismatch negativity. *Journal of Cognitive Neuroscience*, *22*, 2174–2185.
- Stothard, S. E., Snowling, M. J., Bishop, D. V., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech Language and Hearing Research*, *41*, 407–418.
- Suomi, K., Toivanen, J., & Ylitalo, R. (2008). *Finnish sound structure : phonetics, phonology, phonotactics and prosody*. Oulu, Finland: University of Oulu.
- Sussman, E. S. (2007). A new view on the MMN and attention debate - The role of context in processing auditory events. *Journal of Psychophysiology*, *21*, 164–175.
- Takahashi, H., Rissling, A. J., Pascual-Marqui, R., Kirihara, K., Pela, M., Sprock, J., . . . Light, G. A. (2012). Neural substrates of normal and impaired preattentive sensory discrimination in large cohorts of nonpsychiatric subjects and schizophrenia patients as indexed by MMN and P3a change detection responses. *NeuroImage*, *66*, 594–603.
- Takegata, R., Heikkilä, R., & Näätänen, R. (2011). Sound energy and the magnitude of change: effects on mismatch negativity. *NeuroReport*, *22*, 171–174.
- Tamminen, H., Peltola, M. S., Toivonen, H., Kujala, T., & Näätänen, R. (2013). Phonological processing differences in bilinguals and monolinguals. *International Journal of Psychophysiology*, *87*, 8–12.
- Tervaniemi, M., Jacobsen, T., Rottger, S., Kujala, T., Widmann, A., Vainio, M., . . . Schröger, E. (2006). Selective tuning of cortical sound-feature processing by language experience. *European Journal of Neuroscience*, *23*, 2538–2541.
- Tervaniemi, M., Maury, S., & Näätänen, R. (1994). Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *NeuroReport*, *5*, 844–846.
- Tew, S., Fujioka, T., He, C., & Trainor, L. (2009). Neural representation of transposed melody in infants at 6 months of age. *Annals of the New York Academy of Sciences*, *1169*, 287–290.
- Todd, J., & Michie, P. T. (2000). Do perceived loudness cues contribute to duration mismatch negativity (MMN)? *NeuroReport*, *11*, 3771–3774.

- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*, 1245–1260.
- Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology, 51*, 5–15.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear, 22*, 79–90.
- Urban, A., Kremlacek, J., & Libiger, J. (2007). Mismatch negativity in patients with schizophrenia. *Acta Medica (Hradec Kralove), 50*, 23–28.
- Uwer, R., & von Suchodoletz, W. (2000). Stability of mismatch negativities in children. *Clinical Neurophysiology, 111*, 45–52.
- Valtonen, R. (2009). Kehityksen ja oppimisen ongelmien varhainen tunnistaminen Lene-arvioinnin avulla: kehityksen ongelmien päällekkäisyys ja jatkuvuus 4–6-vuotiailla sekä ongelmien yhteys kouluasuorittumiseen. Retrieved from <https://jyx.jyu.fi/dspace/bitstream/handle/123456789/19890/9789513935382.pdf?sequence=1>
- Valtonen, R., Mustonen, K., work group. (2003). Lene. Leikki-ikäisen lapsen neurologinen arvio. Jyväskylä: Niilo Mäki Instituutti.
- Vance, M., Rosen, S., & Coleman, M. (2009). Assessing speech perception in young children and relationships with language skills. *International Journal of Audiology, 48*, 708–717.
- Vuust, P., Brattico, E., Glerean, E., Seppänen, M., Pakarinen, S., Tervaniemi, M. & Näätänen R. (2011). New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. *Cortex, 47*, 1091–1098.
- Warren-Leubecker, A., & Bohannon, J. N. (1984). Intonation patterns in child-directed speech – mother-father differences. *Child Development, 55*, 1379–1385.
- Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: Neurophysiological evidence. *Cognitive Brain Research, 25*, 180–187.
- Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cognitive Brain Research, 18*, 149–161.
- Wechsler, D. (2004). *The Wechsler intelligence scale for children—fourth edition*. London, England: Pearson Assessment.
- Vestergaard, M. D., Haden, G. P., Shtyrov, Y., Patterson, R. D., Pulvermüller, F., Denham, S. L., . . . Winkler, I. (2009). Auditory size-deviant detection in adults and newborn infants. *Biological Psychology, 82*, 169–175.
- Willberg, N. (2002). Erityiset oppimisvaikeudet ja niistä seuraava syrjäytyminen – kahden vaihtoehdoisen erityisopetusmallin kustannus-hyötyanalyttinen tarkastelu. (Master's Thesis), Retrieved from <https://jyx.jyu.fi/dspace/bitstream/handle/123456789/17966/nilswill.pdf?sequence=1>

- Winkler, I., Korzyukov, O., Gumenyuk, V., Cowan, N., Linkenkaer-Hansen, K., Ilmoniemi, R. J., . . . Näätänen, R. (2002). Temporary and longer term retention of acoustic information. *Psychophysiology*, *39*, 530–534.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., . . . Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*, 638–642.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science*, *7*, 270–276.
- Xu, J. H., Yu, L. P., Cai, R., Zhang, J. P., & Sun, X. D. (2009). Early auditory enrichment with music enhances auditory discrimination learning and alters NR2B protein expression in rat auditory cortex. *Behavioural Brain Research*, *196*, 49–54.
- Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, *8*, 1971–1974.
- Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Research*, *1072*, 175–185.
- Zhang, L. I., Bao, S. W., & Merzenich, M. M. (2001). Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience*, *4*, 1123–1130.
- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage*, *26*, 703–720.