The contribution of spectral and temporal information to vowel length perception in German in children and adults

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr.rer.nat.)

vorgelegt beim Fachbereich 05
der Johann Wolfgang Goethe – Universität
in Frankfurt am Main

von
Bogdana Ulytska
aus Kiew

Frankfurt am Main, 2015
(D 30)

vom Fachbereich 05 der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Contents

Zusammenfassung

Sprache ist ein komplexes akustisches Signal, das in einzelne Merkmale zerlegt werden kann. Ungeklärt ist jedoch bislang wie genau diese vom Gehirn integriert werden. Denkbar ist beispielsweise, dass die Merkmale eines Sprachreizes zunächst aus dem Sprachsignal extrahiert, einzeln vom Gehirn untersucht und später wieder zu einer sprachlichen Repräsentation zusammengefügt werden. Vorhergehende Forschung legt zudem nahe, dass einzelne akustische „cues" im Sprachsignal in Abhängigkeit von der Muttersprache des Sprechers und dem Phonem, welches differenziert werden soll unterschiedlich informativ sein können und folglich unterschiedlich gewichtet werden (Bohn, 1995; Escudero, Benders, & Lipski, 2009; Kluender, Lotto, Holt,& Bloedel, 1998; Holt & Lotto, 2006; Toscano & McMurray, 2010).

Ereigniskorrelierte Potentiale (EKPs), wie beispielsweise die als Mismatch Negativity (MMN) bekannte EEG-Komponente, eignen sich besonders gut zur Untersuchung der Sensitivität des Gehirns auf bestimmte Ton- und Sprachreize. Die MMN kann mithilfe des sog. „Oddball Paradigmas" erfasst werden, bei welchem gleichartige Reize, sog. „Standards" kontinuierlich präsentiert werden. Zusätzlich werden hierbei selten (in 10-20% aller Fälle) einzelne in einem oder mehreren Merkmalen abweichende Reize, sog. „Deviants" eingestreut. Auf die Präsentation der „Deviants" reagiert das Gehirn mit einer Abweichung im EKP zwischen 100 und 300 ms nach dem Stimulusonset. Um die MMN, die sich als Negativierung an frontozentralen Elektroden äußert, zu bestimmen, werden die EKPs für „Standards" und „Deviants" subtrahiert (z.B., Kujala et al., 2007, Näätänen et al., 2007; Näätänen & Winkler, 1999; Schröger, 1998). Die MMN ist vor allem bei der Untersuchung von Kindern gut geeignet, da es sich bei dieser um eine aufmerksamkeitsunabhängige Komponente handelt, mit Hilfe derer präattentive Prozesse des Gehirns untersucht werden können (Alho et al., 1994).

Zusätzlich liegt bei Tönen Evidenz für MMN-Additivität vor, d.h. es entsteht eine stärkere MMN für Stimuli bestehend aus mehreren akustischen Merkmalen im Vergleich zu Stimuli bestehend aus einzelnen akustischen Merkmalen (z.B. Levänen et al., 1993; Schröger, 1995; Paavilainen et al., 2001). Hierbei wurde bereits gezeigt, dass Töne, die aus spektralen und temporalen Merkmalen bestehen eine stärkere MMN erzeugen als Töne, die entweder aus spektralen oder temporalen Merkmalen zusammengesetzt sind (Levänen et al., 1993). MMN Studien zeigen zudem, dass Phonemwahrnehmung charakterisiert ist durch die Gewichtung einzelner akustischer „cues" (z.B., Moberly et al., 2014; Lipski et al., 2012; Jacobsen et al., 2004). Es fehlen an dieser Stelle jedoch bislang systematische Untersuchungen, die veranschaulichen könnten, wie genau einzelne Merkmale innerhalb eines bestimmten Phonems integriert werden.

Die vorliegende Arbeit zielte darauf ab diese Lücke zu schließen. Hier sollte die Sensitivität des Sprachverarbeitungssystem hinsichtlich spektraler und temporaler Merkmale bei Kindern und Erwachsenen untersucht werden, d.h. es sollte geprüft werden wie genau temporale und spektrale Merkmale in natürlichen Sprachlauten vom Gehirn integriert werden und ob dieser Mechanismus entwicklungsabhängig ist. Bisherige empirische Studien legen hierbei nahe, dass es im Alter von etwa 10 Jahren im auditorischen System durch Hirnreifungsprozesse zu entscheidenden Veränderungen kommt (z.B., Bruder et al., 2011; Wetzel, Widmann & Schroeger, 2011). Folglich wurden im Rahmen der vorliegenden Untersuchung Kinder aus diesem Altersbereich akquiriert. Die Probanden wurden insgesamt zweimal untersucht, im Alter von 9-10 Jahren und ein Jahr später. Die Daten der Kinder wurden anschließend mit Daten von Erwachsenen verglichen. Eine weitere Fragestellung, die im Rahmen der vorliegenden Arbeit untersucht wurde, beschäftigt sich zudem mit der Ätiologie der Lese-Rechtschreibstörung (LRS). Zwar besteht Einigkeit darüber, dass phonologische Verarbeitungsdefizite ein Kernsymptom von LRS darstellen, dennoch ist es weiterhin

umstritten, ob diese auch die primäre Ursache der Störung sind oder ob sie doch nur als Folge eines nicht-sprachlichen auditiven Verarbeitungsdefizits auftreten, bei welchem eigentlich die Verarbeitung spektraler und temporaler Merkmale, d.h. einzelner akustischer „cues" innerhalb des Sprachsignals, defizitär ist (z.B., Ahissar et al., 2000; Corriveau et al., 2010).

Für die Untersuchung wurde ein Paradigma genutzt, dass den direkten Vergleich von phonologischer, temporaler und spektraler Verarbeitung erlaubt. Hierbei wurden deutsche Vokale unterschiedlicher Länge, die Vokalpaare /a/-/a:/ und /i/-/i:/, verwendet. Vokallängenkontraste sind an dieser Stelle besonders gut geeignet, da sie im Deutschen sowohl durch temporale als auch durch spektrale Merkmale charakterisiert sind. Neben natürlichen Vokalen wurden mithilfe des Phonetik-Programms „Praat" (Boersma, 2001) künstlich modifizierte Vokale generiert. Im Mismatch Negativity-Paradigma wurden dann durch die entsprechende Kombination der Stimuli die drei experimentellen Bedingungen phonologisch (spektrotemporal) vs. temporal vs. spektral hergestellt. In der phonologischen Bedingung wurden immer zwei natürliche Vokale unterschiedlicher Länge der gleichen Vokalkategorie als „Standard" und „Deviant" verwendet. Folglich war der Unterschied zwischen den Stimuli sowohl temporal als auch spektral. In der temporalen Bedingung hingegen wurde entweder ein natürlicher Langvokal mit einem verkürzten Langvokal kombiniert oder ein natürlicher Kurzvokal mit einem verlängerten Kurzvokal, so dass der Unterschied zwischen „Standard" und „Deviant" ausschließlich temporal war. In der spektralen Bedingung wiederum wurde entsprechend entweder ein natürlicher Langvokal mit einem verlängerten Kurzvokal kombiniert oder ein natürlicher Kurzvokal mit einem verkürzten Langvokal, so dass der Unterschied zwischen „Standard" und „Deviant" ausschließlich spektral war.

In Experiment 1 wurden zunächst erwachsene gesunde Probanden (N = 20) mit dem oben beschriebenen Paradigma untersucht. In allen Bedingungen, d.h. phonologisch,

temporal und spektral wurden reliable MMNs über frontozentralen Elektrodenpositionen gefunden, was zunächst bedeutet, dass das Vorhandensein einzelner akustischer „cues" für das Gehirn ausreichend ist um Vokale unterschiedlicher Länge zu differenzieren. Dennoch wurde in der phonologischen Bedingung eine deutlich stärkere MMN als in der temporalen oder spektralen Bedingung gefunden, wobei die MMN der phonologischen Bedingung nicht ganz der Summation der MMNs der temporalen und der spektralen Bedingung entsprach. Einzeln ausgewertet nach Vokalkategorien, zeigte sich dieses (unter-)additive Ergebnis zudem nur noch für Vokaltyp /i/. Entsprechend fanden sich hier auch vergleichbare MMNs in der spektralen und der temporalen Bedingung. Für Vokaltyp /a/ hingegen wurden eine starke MMN in der temporalen und nur eine schwache MMN in der spektralen Bedingung gefunden. Die MMN der temporalen Bedingung unterschied sich dabei nicht von der MMN der phonologischen Bedingung, d.h. es konnten keine additiven Effekte für Vokaltyp /a/ ausgemacht werden. Dieses Ergebnis zeigt zwar, dass unser auditives Verarbeitungssystem durchaus in der Lage ist MMNs für einzelne perzeptuelle Merkmale auch für Sprachreize zu summieren, demonstriert aber auch, dass diese Summation wiederum beeinflusst wird durch die Salienz der jeweiligen akustischen „cues" hinsichtlich des speziellen Sprachkontrastes.

In Experiment 2 wurde das Paradigma in abgekürzter Form auf Kinder im Alter von 9-10 Jahren angewendet. Die Hälfte der Kinder hatte LRS (N=15) und die andere Hälfte weder Schwierigkeiten im Lesen noch im Schreiben. Zudem wurden (um Vergleichbarkeit zu gewährleisten) erneut erwachsene gesunde Probanden (N=15) mit dem verkürzten Paradigma gemessen. Wie bereits in Experiment 2 wurden für Erwachsene additive MMN Effekte gefunden. Für Kinder konnte dieser Befund jedoch nicht repliziert werden. Zudem wurden in beiden Kindergruppen zwei MMN Komponenten (statt einer Komponente) in der phonologischen Bedingung gefunden. Hierbei entsprach die frühe zentrale MMN Komponente in der phonologischen

Bedingung zeitlich und topographisch der MMN in der spektralen Bedingung und die späte frontale zeitlich und topographisch der MMN in der temporalen Bedingung. Dieser Befund ist ein Hinweis darauf, dass Kinder im Alter von 9-10 Jahren im Gegensatz zu Erwachsenen einzelne akustische „cues" im Sprachsignal sequenziell verarbeiten. Zwischen Kindern mit und ohne LRS wurden neurophysiologisch überraschenderweise keinerlei Unterschiede gefunden.

Um zum einen mögliche Ursachen für dieses Ergebnis zu klären und zum anderen die Entwicklung der Sprachverarbeitung zu untersuchen, wurden in Experiment 3 ein Teil der Kinder mit (N = 13; Alter im Mittel 11.160) und ohne LRS (N = 12; Alter im Mittel 11.262) aus Experiment 2 erneut untersucht. Zusätzlich zum passiven Mismatch Paradigma wurde bei den Probanden die aktive Diskrimination von Vokalen unterschiedlicher Länge untersucht. Hierbei wurden dieselben Stimuli wie im passiven Design verwendet. Es wurden keine Unterschiede zwischen den Gruppen gefunden. Dieses Ergebnis ist folglich eine zusätzliche Validierung der Befunde aus Experiment 2. Interessanterweise konnte das MMN-Muster, das sich in Experiment 2 bei Kindern im Alter von 9-10 Jahren manifestierte, d.h. die zwei MMN Komponenten in der phonologischen und jeweils eine Komponente in der spektralen und der temporalen Bedingung, bei Kindern im Alter von 10-11 Jahren nicht mehr gefunden werden. Stattdessen zeigte sich eine starke Tendenz zur MMN-Additivität. Dieser Befund bedeutet wiederum, dass das MMN-Muster sich innerhalb eines Jahres deutlich an das der Erwachsenen angeglichen hat. Das Ergebnis bestätigt die Hypothese, dass das Alter von ca. 10 Jahren kritisch ist für die Entwicklung der auditiven Verarbeitung, veranschaulicht jedoch auch, dass diese Entwicklung mit 11 Jahren noch nicht abgeschlossen ist.

Zusammenfassend zeigt die vorliegende Arbeit, dass das ausgewachsene Gehirn einzelne akustische „cues" im Sprachsignal anders verarbeitet als das Heranreifende. Im

Alter von 10-11 Jahren nähern sich die auditiven Sprachverarbeitungsprozesse dann deutlich denen der Erwachsenen an.

List of Figures

List of Tables

Abbreviations

ANOVA	analysis of variance
d'	discrimination index
df(factor)	degrees of freedom of factor
df(error)	degrees of freedom of error
dB	decibel
EEG	electroencephalogram
EOG	electrooculography
ERP	event related potential
F1	first formant
F2	second formant
fMRI	functional magnet resonance imaging
Hz	hertz
ISI	inter-stimulus interval
MEG	magnetoencephalography
MMF	mismatch field
MMN	mismatch negativity
μV	microvolt
ms	milliseconds
RAN	rapid automized naming
PET	positron emission tomography
ROI	region of interest
RT	reaction time
s	seconds
SOA	stimulus onset asynchrony
SPL	sound pressure level

Preface

Speech is composed of multiple complex acoustic cues. So far, it is not clarified how exactly these cues are integrated within the speech signal. However, understanding the underlying speech perception mechanisms is especially crucial, as in our world speech is the most important and complex communication tool. Moreover, speech is indispensable for academic achievement. Correct speech perception is relevant for learning to read and to write properly and hence to acquire knowledge in school. Unfortunately, irrespective of a great amount of effort put into effective literacy acquisition in school, some children have deficits in learning to read (and to write). This condition is known as developmental dyslexia. Previous literature indicates that deficiencies in correct speech, i.e., phoneme, perception may be etiologically relevant for the emergence of dyslexia. However, it cannot be excluded that even more general low-level deficits in auditory processing underlie the deficit in speech perception in dyslexia (for more details, see Chapter 4).

All in all, the aim of the present work is first, the investigation of underlying principles of speech perception and second, the clarification of the nature of the auditory processing deficit underlying dyslexia. Furthermore, as precise speech perception is assumed to progress throughout childhood, this work explores developmental aspects of speech perception as well.

Chapter 1: Introduction

Mechanisms of speech perception

Plenty of theoretical and empirical work has been conducted to elucidate the mechanisms of speech processing. In the end, especially two theoretical approaches gained attention in this field: domain-specific and cue-specific models.

Domain-specific models postulate that speech is processed by a specialized left-hemisphere lateralized neural system. Referring to this view, speech sounds can be differentiated from non-speech sounds due to the storage of phoneme representations in the left auditory cortex (e.g., Alho et al., 1998a, Koyama et al., 2000, Shestakova et al., 2002). The existence of such memory traces for phonemes, in turn, makes our auditory system more effective in processing speech sounds of our native language (Kuhl, 2004). Following this, speech and non-speech sounds are analyzed differently. More precisely, whereas in speech the linguistic character triggers a unique neural response, e.g., a hemispheric specialization, non-speech sounds are processed on the basis of single acoustic features (for reviews, see Liberman & Mattingly, 1985 and Liberman & Whalen, 2000). Strictly speaking this model would predict that the same sound is processed differently depending on whether it is perceived as being speech or non-speech. Interestingly, this hypothesis has been investigated in studies of Dehaene-Lambertz et al. (2005) and Möötonen et al. (2006). In both studies sinewave analogues of speech have been used. People exposed to these stimuli first state to perceive them as non-speech sounds. Nonetheless, once they are informed that the sounds are artificially modified speech sounds their perception mode switches. Interestingly, Dehaene-Lambertz et al. (2005) and Möötonen et al. (2006) have found that, as the perception mode switches, the activation in the left temporal lobe (more precisely, in the left posterior superior temporal

sulcus) increases. Note that this effect was not observable in subjects who did not learn to perceive the sinewave stimuli as speech sounds.

Cue-specific theories, however, make the assumption that identical neural systems are responsible for speech and non-speech auditory processing and hence same mechanisms underlie the processing of any type of auditory input, irrespective of its linguistic character (for a review, see Diehl et al., 2004). Support for this approach comes for instance from evidence showing that speech-related and general auditory neural systems overlap (Price, Thierry, and Griffiths, 2005). Crucially, opponents of domain-specific models do not deny that left-hemisphere lateralized neural systems are involved in speech processing, but rather explain this by the fact that a speech stimulus is composed of single acoustic features, which again activate specialized neural systems, e.g., left auditory cortical regions, when they are processed. For instance, some protectionists of the cue-specific theory postulate that the processing of temporal features, i.e., rapidly changing information in sounds up to 200 ms, which is crucial for correct phoneme perception, is predominantly localized in the left hemisphere, whereas the processing of spectral cues organized in longer time windows is stationed in the right hemisphere (e.g., Zatorre, 2002; Ackermann & Riecker, 2004; Riecker, 2008; Basso et al., 1993).

Most important, domain-specific and cue-specific models are not entirely inconsistent with one another. Some researchers even argue that it is important to reconcile both theories, as empirical support has been found for certain predictions of both models (e.g., Zatorre & Gandour, 2008). They propose a hierarchically organized model of speech processing. Afferent bottom-up, i.e., cue-specific, pathways may be responsible for the processing of single acoustic cues, whereas top-down, i.e., domain-specific, mechanisms may adjust these processes (e.g., Zatorre & Gandour, 2008; Obleser et al., 2007). According to this, McGettigan & Scott (2012) agree that the left hemisphere

processes predominantly temporal information and the right hemisphere frequency changes. However, the authors further assume that while the left hemisphere forms domain-specific memory traces, e.g., memory traces for phonemes, the right hemisphere processes acoustic stimuli domain-independently. Correspondingly, Shtyrov et al. (2000) have demonstrated that consonants characterized by rapid transitions, but not non-speech stimuli characterized by rapid transitions, are associated with increased neural activity in the left hemisphere. Another set of empirical evidence has shown that single acoustic cues are weighted differently when speech-sounds, i.e., phonemes, are processed. This is not the case when non-speech sounds are processed (e.g., Repp, 1982). More precisely, more weight is given to acoustic cues especially informative for a phoneme contrast (Holt & Lotto, 2006; Maye, Werker, & Gerken, 2002; Toscano & McMurray, 2010). The informativeness of certain acoustic cues, in turn, depends on the language system the phonemes are embedded in (Bohn, 1995; Crowther & Mann, 1992; Escudero, Benders, & Lipski, 2009; Kluender, Lotto, Holt, & Bloedel, 1998). Past research has even shown that when a second language is learned, non-native listeners rather use acoustic cues relevant in their own native language to differentiate a phoneme contrast of a foreign language, which leads to perception deficiencies. This explains for instance the deficits Japanese listeners exhibit with the American English categories /l/ and /r/ (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004), as they are used to rely on the second formant in consonants and not on the third one, which is however the relevant cue for native Americans (Iverson et al., 2003; Miyawaki et al., 1975).

All in all, the findings cited above indicate that cue-specific and domain-specific mechanisms both are involved when speech stimuli are processed, i.e., that speech processing, in contrast to non-speech processing, requires the involvement of additional neural responses. However, to understand how exactly fundamental sensory driven processing mechanisms (bottom-up) interact with more abstract categorically driven

processing mechanisms (top-down) when speech is processed, far more research is needed. Furthermore, to the best of knowledge, there are no previous studies investigating these speech-processing mechanisms in the course of development. Crucially, the present work aims to do so by taking advantage of the properties of the German vowel system and by applying the mismatch negativity (MMN) – paradigm. Furthermore, the endeavor of the present work is to have a closer look on possible deviations in speech processing. Thus, as it is well established that developmental dyslexia (among others) is associated with deficits in speech processing, a subgroup suffering from this disorder is included in two experiments of the present work, investigating the development of speech perception (e.g., Adlard & Hazan, 1998; Ziegler et al., 2009; for more details, see Chapter 4).

German vowel length system

Vowel length is an important aspect of speech that can have strong influence on perceived word meaning in many languages. For example, the German words 'Miete' [/mi:te/, rent] – characterized by a long vowel 'i' – and 'Mitte' [/mɪtə/, middle] – containing a short 'i' – differ phonologically only minimally, i.e., in the length of the vowels. Nevertheless, the meanings of these words differ substantially. Similar differences in vowel length exist in other languages, as demonstrated by the contrast of the English words 'pull' and 'pool'. It is thus of great importance to understand how vowel length is processed and discriminated by our language processing system.

Vowel length discrimination is also critical for learning to write, as vowel length has to be mapped onto orthographical differences in a rule-based manner (e.g., Klicpera & Gasteiger-Klicpera, 1998). For instance, in German orthography, vowel length is not indicated by the letter of the vowel itself, but by the letters following the vowel letter. Thus, long vowels are often marked by adding a silent 'h' (as in Stahl, /ʃta:l/, [steel]) or

by doubling the vowel letter (as in Tee, /te:/, [tea]). Short vowels, in contrast, are often followed by consonants in German (e.g., Stall, /ʃtal/, [barn]).

The German vowel system is overall composed of fourteen vowel monophthongs which can be divided into seven pairs of short (lax) and long (tense) vowels:/i:/-/i/, /y:/-/ʏ/, /u:/-/ʊ/, /e:/-/ɛ/, /ø:/-/oe/, /o:/-/ɔ/, and /a:/-/a/ (Kohler, 1977; Moulton, 1962; Wiese, 2000). All vowels can be described as a function of height and frontness of tongue during their production (Speyer, 2007) (see Figure 1).
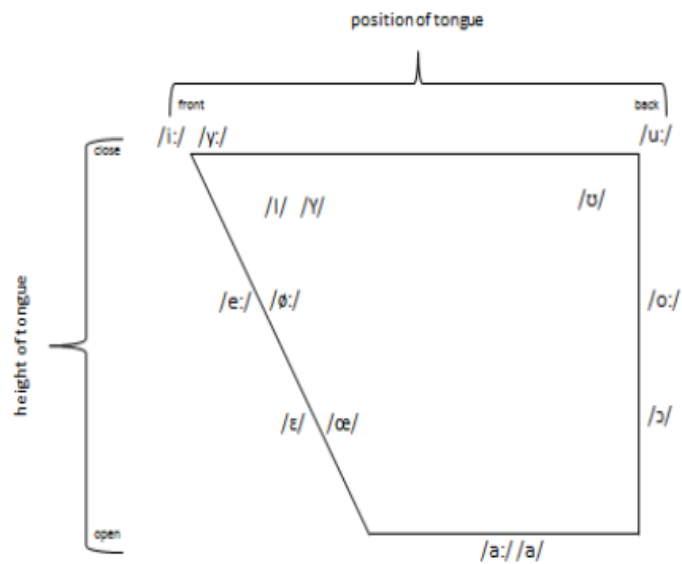
Figure 1: Position of the fourteen German monophthongs dependent on position of tongue and height of tongue (adapted from Christmann, 2014).

The frequency pattern of a vowel, with time on the horizontal axis and frequency on the vertical axis, can be depicted in a spectrogram (for vowel /a:/, see Figure 2). With higher intensity of a frequency at a given time point the shading skips from white over grey to black. Frequencies with the highest intensity are called formants (marked by black dots in Figure 2) (Carroll, 2004).
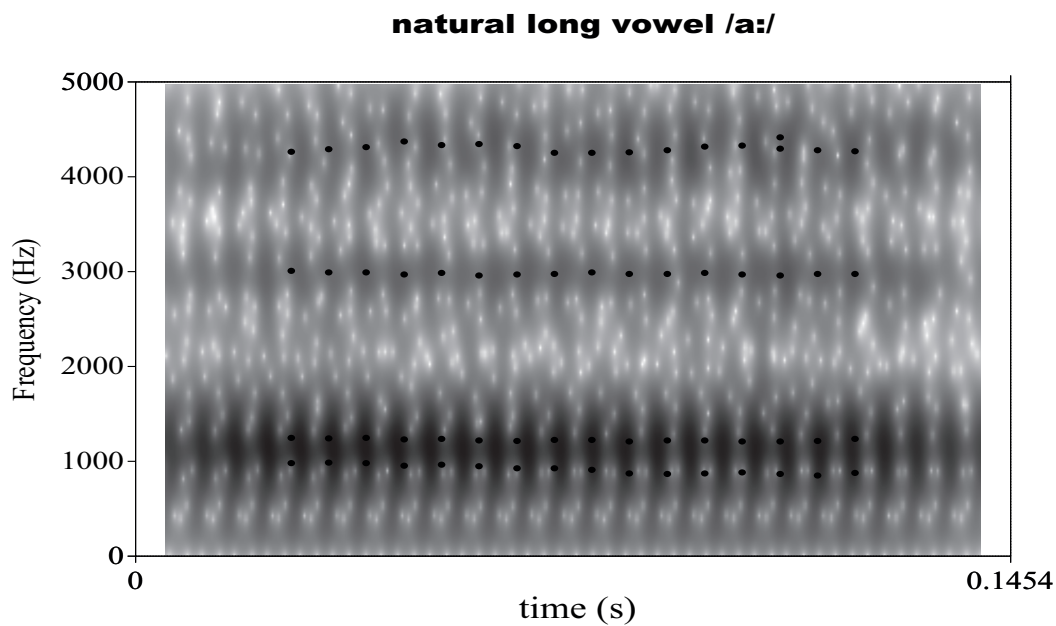
Figure 2: Spectrogram of vowel /a:/. The greyscale image depicts the strength of a certain frequency at a given time point. Formants are indicated by black dots.

In several languages, such as English and German, vowel length is represented in the speech signal by two distinct features, i.e., temporal and spectral differences in the acoustic signal. Figure 4 A shows long and short versions of two German vowels. These vowels not only differ in their duration, but also substantially in their spectral composition. The reason for spectral differences between vowels of different length is a different constellation of the speech-organs (Hertrich and Ackermann, 1997), which influences F1 and F2 formant frequencies, which again determines the timbre of the vowel.

Moreover, German vowels differ in vowel height. Previous empirical work has shown that the influence of temporal and spectral cues on vowel identification varies with vowel height. More specifically, to identify high vowels such as the German /i/, listeners rely more on spectral information, while temporal information is more important for

identifying low vowels such as the German /a/ (Sendlmeier, 1981; Strange & Bohn, 1998; Weiss, 1974; see Figure 3).
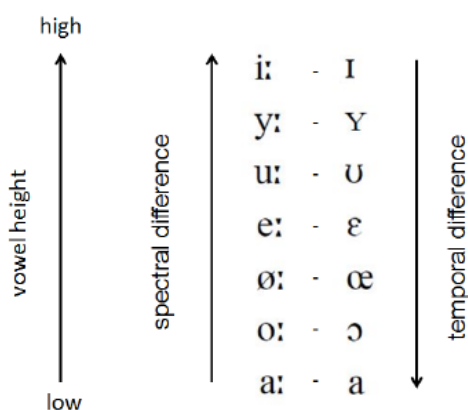
Figure 3: Relationship between vowel height and relevance of spectral and temporal cues for vowel length identification (adapted from Christmann, 2014).

To date there is only very little empirical evidence concerning the contribution of temporal and spectral information to vowel length discrimination in German. Bohn and Polka (2001), for instance, have shown that after the (partial) removal of temporal cues, German adults and even infants (mean age: 8.2 months) were still able to discriminate vowels by relying on spectral cues. Note however that the stimuli of Bohn and Polka (2001) always differed in both their spectral and temporal content, which was not the case in a study of our own research group (Groth et al., 2011; Steinbrink et al., 2012), outlined in the next section. Here, it was found that vowel length could also be discriminated based on temporal cues only, however at reduced performance levels, in a German adult sample. Taken together, evidence suggests that both acoustic features contribute to vowel length discrimination.

Vowel length discrimination paradigm (developed by Groth et al., 2011)

The vowel length discrimination paradigm was developed by Groth and colleagues (2011). They used the entirety of seven German vowel pairs differing in vowel length for their paradigm. Additionally, artificially modified stimuli were generated by shortening or lengthening the steady-state phase of the original stimulus so that, in the end, modified vowels were as long as their original counterparts. Spectral features of the stimuli remained unaltered. Thus, manipulated long vowels with the length of their short counterparts and manipulated short vowels with the length of their long counterparts were used. All vowels were embedded in CVC pseudo-word syllables: /fVp/ and /nVp/ (V = vowel), which again was in line with the German phonotactic rule system. One female professional speaker spoke the syllables without stress.

Table 1: Average durations (ms) of long and short vowels (differing in vowel height) used in the paradigm of Groth et al. (2011).

Vowel height	Vowel pair	Vowel duration (ms)		Difference (ms)
		Long	Short	Long − Short
High	/i: - ɪ/	91	51	40
	/y: - ʏ/	98	53	45
	/u: - ʊ/	102	57	45
Mid	/e: - ɛ/	110	66	44
	/ø: - œ/	121	70	51
	/o: - ɔ/	128	75	53
Low	/a: - a/	142	75	67

To ensure that the syllables correspond to the provided CVC category boundaries, ten healthy adult participants took part in an identification task. The syllables were presented auditorily to the participants, who again had to write the syllables down such as they were

perceived. The length of the vowels had to be marked by using the corresponding German orthographic conventions (i.e., "fahp" for /fa:p/ and "fapp" for /fap/). The criterion for syllable exclusion was set to 20% of judgements not fitting into the provided category.

The stimuli were presented to the participants in a forced-choice discrimination task, a so-called same-different task. Here, in each trial two syllables were delivered and the participants had to judge whether they are the same or whether they differ. The experiment consisted of two conditions: a phonological, i.e., spectrotemporal, and a temporal condition. In the phonological condition original vowels, embedded in syllables, were contrasted to their original counterparts. The difference between vowels was spectral and temporal in nature. In the temporal condition, however, natural vowels were contrasted to artificially modified ones. More precisely, an original long vowel was combined with a shortened one or an original short vowel was contrasted to a lengthened one. By doing so, the difference between the stimuli presented was exclusively temporal in nature, so that temporal auditory processing could be investigated in isolation.

In the current work, to shorten the experiment only a subset of two vowel categories was used, i.e., the low vowel /a/ and the high vowel /i/, each both in a long and a short version. In addition to natural vowels, artificially shortened long vowels and artificially prolonged short vowels were included. However, in contrast to Groth et al. (2011) isolated vowels, i.e., vowels that are not embedded in syllables or pseudo-words, were used in the present study. Moreover, to investigate the processing of spectral features in vowels, a spectral condition was appended to the design. The different stimuli were combined into pairs of long and short vowels such as to yield (i) a *phonological condition* in which vowels differed both spectrally and temporally, (ii) a *temporal condition* in which spectral content was identical, while vowel duration differed between vowels, and

(iii) a *spectral condition* in which the vowels were equally long, but differed in their spectral composition.

Interestingly, this paradigm has already been applied in a study of our research group, conducted with a sample of 9-year-old children (Steinbrink, Klatte, & Lachmann, 2014). Here, three vowel categories differing in vowel height were used: /iː/-/i/, /oː/-/ɔ/, and /aː/-/a/ (high vowel pair /iː/-/i/ and low vowel pairs /oː/-/ɔ/, and /aː/-/a/). In line with Groth et al. (2011), normally reading children had no problems to discriminate natural vowels of different length, as required in the phonological condition. However, the performance decreased in the temporal condition and this effect was associated with vowel height, as the decrease was highest for the vowel category /i/. In the spectral condition, in turn, a performance drop was observed exclusively for the vowel category /a/, but not for vowel categories /i/ and /o/. This was explicable by the negligibly small spectral difference between /aː/ and /a/, leading to the drop in discrimination performance. All in all, Steinbrink et al. (2014) have replicated and successfully extended (by adding a spectral condition) the results of Groth and colleagues (2011). This fact served as confirmation that isolated vowels were appropriate to be applied as stimuli in the present work as well.

Chapter 2: Mismatch Negativity in research of auditory processing

Mismatch negativity (MMN; Näätänen et al., 1978; Näätänen, 1992, 1990; Picton, 2000) is a cognitive event related potential (ERP), which is elicited within an oddball paradigm when a stream of repetitive stimuli, i.e., standards, is interrupted by the presentation of few rare stimuli, i.e., deviants, characterized by any deviation from the standard stimuli (Näätänen, 1992; 1991). With greater physical difference between standard and deviant, MMN becomes larger (Tiitinen et al., 1994). MMN is calculated by subtracting the ERPs to standards from ERPs to deviants and can be detected in healthy adults as a negative peak 100-250 ms after the onset of the deviant at frontal and central electrodes (Lieder et al., 2013).

MMN is assumed to reflect sensory memory processes. Each standard sound forms a memory trace in the auditory cortex, i.e., a neuronal memory representation, and when the deviant violates this representation, the mismatch is indicated by MMN (Näätänen, 1995). The length of a sensory memory trace can be probed by varying the interstimulus interval (ISI). Mäntysalo and Näätänen (1987) recorded the MMN response with an ISI of 1 or 2s, but not with an ISI of 4s. Hence, with an overly prolonged ISI, the memory trace of a stimulus might get replaced by the representation of the following stimulus. Consequently, no MMN is elicited by the presentation of the deviant. However, a shortened ISI can inhibit MMN elicitation as well, which is indicated by memory-deleting effects of masking stimuli. Here, the masking stimulus is presented after the stimulus that has to be discriminated from the other stimuli. With a short ISI between deviant and masking stimulus, no MMN is elicited anymore. It has been hypothesized that these backward-masking effects might reflect the ability of the auditory system to

integrate successive sounds to meaningful unities, which is crucial for correct speech perception and comprehension (Bregman, 1990). By means of further research, a temporal window of integration (TWI) between 150-200 ms was determined (Näätänen, 1990). Hence, an ISI of at least 150-200 ms between deviant and masking stimulus (or standard) is required for MMN elicitation. In a study of Tervaniemi and colleagues (1994), paired sounds were presented as standards, which again were contrasted with deviants, characterized by omissions of the second sound. Interestingly, MMN was only elicited when sounds within a pair were isolated from each other up to 140 ms, whereas no MMN could be generated with longer intervals.

The major source of MMN is localized in the auditory cortex. However, additional frontal sources, which have been associated with attention switch after deviant processing, have been reported as well (Alho, 1995; Giard et al., 1995; Rinne et al., 2000).

Moreover, MMN can be recorded in different modalities, e.g., auditory, visual (e.g., Tales et al., 1999; for a review, see Pazo-Alvarez et al., 2003) and even tactile (Kekoni et al., 1997; Shinozaki et al., 1998; Akatsuka et al., 2005), and can not only be recorded by electroencephalography (EEG), but as well by magnetoencephalography (MEG), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). However, the present work concentrates on the auditory system and on MMN recording with electrophysiological means.

MMN to different stimulus types

MMN can be elicited by simple stimulus types, e.g., pitch, duration or stimulus intensity (for a review, see Näätänen, 1992), as well as by more complex and abstract stimulus variations, e.g., tone order reversals (for a review, see Näätänen et al., 1993). That is to say, MMN reflects abstract representations of complex auditory rules and not only the

physical properties of an auditory stimulus. Paavilainen and colleagues (2001), for instance, have presented to their participants standard stimuli, always varying in both frequency and intensity. These standards followed the rule "the higher the frequency, the louder the intensity". Crucially, deviants violating this rule induced a mismatch response.

In addition, mismatch responses can be elicited by natural or artificial sounds (e.g., Shestakova et al., 2002; Dehaene-Lambertz and Pena, 2001), as well as by both speech and non-speech stimuli (e.g., Näätänen et al., 1978, Näätänen & Michie, 1979, Vihla and Eulitz, 2003; Dehaene-Lambertz and Dehaene, 1994). Nevertheless, studies comparing MMNs for speech and non-speech sounds have come to inconsistent results. Some have found larger MMNs for non-speech sounds as compared to speech sounds (e.g., Wunderlich et al., 2001), whereas others have found no differences at all (Nikjeh et al., 2009) or larger mismatch responses for speech sounds as compared to non-speech sounds (e.g., Jaramillo et al., 2001). However, not all studies matched their speech and non-speech stimuli in respect of complexity, although previous evidence indicates that the complexity of a stimulus, i.e., the number of various frequencies a sound is composed of, has an impact on MMN size (e.g., Tervaniemi et al., 2000; Zion-Golumbic et al., 2007; Takegata et al., 2008). Alho et al. (1996) have even demonstrated that MMN sources differ for stimuli composed of identical features, varying in complexity.

The impact of learning on MMN

MMN to both speech and non-speech stimuli underlies effects of learning-associated neural plasticity (Heim et al., 2000; Kujala et al., 2001; Kraus et al., 1995; Tremblay et al., 1998; Näätänen et al., 1993). Näätänen and colleagues (1993), for instance, have presented to their participants complex sound patterns, which were not discriminable and consequently, prior to training, induced no MMN at all. In participants, who learned to discriminate the sound patterns presented in the training period, an MMN response

appeared, whereas in participants who did not succeed while discrimination training, no MMN was found. Comparable effects have been found for acoustically differing (within-category) versions of the syllable /da/, which were not discriminable for the participants before training (Kraus et al., 1995). Moreover, training studies with special populations have been implemented in the past. Kujala et al. (2001), for instance, have investigated the impact of training on MMN in a dyslexic sample. Here, an audiovisual training was delivered to first-grade children, which enhanced both reading skills and MMN amplitudes. Another study of Ilvonen and colleagues (2003) investigated changes in MMN and speech-comprehension in stroke patients over 6 months after stroke occurence. With advance in speech comprehension an enhancement in MMN amplitude has been observed. Furthermore, Atienza et al. (2004) have shown that MMN might even reflect memory consolidation after a training period. Here, the MMN amplitude was significantly larger 48 h as compared to 24 h after training offset.

Language-specific memory traces and MMN

As outlined above, MMN is also elicited when speech sounds are processed (e.g., Sussman et al., 2004; Vihla and Eulitz, 2003, see also Dehaene-Lambertz and Dehaene, 2004). Furthermore, previous evidence indicates that language-specific memory traces may even boost the MMN response. Näätänen et al. (1997), for instance, presented to Estonian and Finnish subjects several vowels belonging to a vowel category in both languages and one vowel being part of the Estonian language, but not of the Finnish language system. For Estonian participants, vowels grouped in order of growing physical difference elicited increasingly larger MMN amplitudes (exactly in same order as the physical difference rised). Same pattern was found for vowels occurring in both languages in Finnish participants. However, a substantially diminished MMN amplitude was registered in Finnish participants for the Estonian vowel, which was not part of the

Finnish vowel system. In addition, in Finnish participants, different topographic MMN distributions have been found for native vowels and the non-native vowel, more precisely native vowels led to greater activations in the left temporal cortex, whereas no laterality effect was found for the non-native vowel. Comparable studies with similar results have been accomplished for Finnish–Hungarian (Winkler et al., 1999b), English–Japanese (Phillips et al., 1995), French–Japanese (Dehaene-Lambertz et al., 2000), English–Hindi (Shafer et al., 2004; see also Rivera-Gaxiola et al., 2000) and French–Hindi (Dehaene-Lambertz, 1997) language contrasts. In addition, Tervaniemi and colleagues (2006) have shown that Finns exhibit larger MMN amplitudes to durational changes (in non-speech stimuli) than Germans, as in Finnish durational aspects are more crucial for speech comprehension than in German. Interestingly, MMN amplitude differences between Germans and Finns corresponded to the behavioral discrimination performance in both groups.

Beyond that, memory traces for syllables (Shtyrov et a., 2000; Alho et al., 1998b) and even for whole words of the native language (Pulvermüller et al., 2001), as indexed by MMN, have been found. Moreover, Pulvermüller and colleagues (2004) found different topographical MMN distributions for different Finnish words in Finns, which might reflect the formation of separate neuronal representations for native syllables or words. In addition, MMN is also elicited by grammatical violations (e.g., Shtyrov et al., 2003) or boosted by higher phonotactic probability, i.e., the probability of a phoneme combination in a certain language (e.g., Bonte et al., 2005). To conclude, MMN is not solely elicited by sensory memory representations, but also affected by long-term memory representations.

Correspondence between discrimination accuracy and MMN

As correlations of MMN and behavioral measures have been found, MMN is suggested to reflect sound discrimination accuracy (see, for example, Aaltonen et al., 1994; Winkler et al., 1999; Amenedo and Escera, 2000; Atienza and Cantero, 2001; Kujala et al., 2001; Novitski et al., 2004). Aaltonen and colleagues (1994), for instance, have determined a correlation between the discrimination accuracy for subtle (within-category) changes in a Finnish vowel and the corresponding MMN amplitude.

However, the relationship is not definite, as some studies failed to find similar results (e.g., Bazana & Stelmack, 2002). Correspondingly, some previous studies comparing clinical and normal populations, have found group differences in MMN amplitude, but not group differences in active discrimination (Bradlow et al., 1999; Jaramillo et al., 2001; Kozou et al., 2005), whereas other studies have found group differences in behavioral discrimination performance without finding differing MMN amplitudes (e.g., Gaeta et al., 2001; Kujala et al., 2006a). In the study of Kujala and colleagues (2006a), for instance, dyslexic and control participants have been compared in regard to their temporal auditory processing abilities. As expected, the dyslexic sample did worse in comparison to the control sample in actively discriminating deviant sounds, whereas the MMN was similar in both groups.

Results gained from active behavioral discrimination tasks and passive MMN paradigms might diverge for several reasons. First, performance on behavioral measures might reflect motivational or attentional factors. Second, behavioral tasks indicate later stages of sound discrimination (than the MMN), which might be impaired without affecting the MMN. Alternatively, disturbances preceding MMN elicitation in the auditory cortex, e.g., on the brain stem level, can lead to an altered MMN without affecting behavioral performance. In addition, Gaeta and colleagues (2001) have proposed, that behavioral discrimination probably affects other memory processes than

the MMN does. Sensory memory traces might fade away with a prolonged ISI and correspondingly lead to a lack of MMN, whereas rehearsal of stimulus features during accomplishment of a behavioral task might keep the representation of a stimulus in working memory for a longer period of time, that is to say, behavioral tasks might affect the working memory system, while MMN might reflect the sensory memory system.

Concluding remarks and MMN additivity

MMN is attractive for studying auditory processing for several reasons. First, MMN is an attention independent measure (Alho et al., 1994). Thus, the sensitivity of the speech processing system to different features of the speech signal can be examined without explicitly directing attention to them. Accordingly, MMN is a popular measurement tool of auditory processing in clinical populations. MMN studies have been conducted for schizophrenia patients (Michie et al., 2000; Umbricht & Krljes, 2005), patients suffering from dyslexia (see Bishop, 2007 for a review) or specific language impairments (for a review, see Bishop, 2007) and even coma patients (Naccache, Puybasset, Gaillard, Serve, & Willer, 2005; Wijnen, van Boxtel, Eilander, & Gelder, 2007). Second, the occurrence of MMN has already been shown for various deviant types, e.g., duration or frequency (for a review, see Näätänen, 2001).

Additionally, it has been shown, e.g., for stimulus combinations such as frequency and stimulus-onset-asynchrony (SOA) (Levänen et al., 1993), frequency and location (Schröger, 1995), location and the conjunction between frequency and intensity (Takegata et al., 1999), that deviations in multiple features elicit stronger MMNs than deviations in single features and – of relevance for the present work – there is also evidence of MMN additivity for multiple-feature deviants composed of spectral and temporal cues (Lävenen et al., 1993; but see Jaramillo et al., 2001). Usually, to determine additive MMN effects, the modeled additive MMN, calculated as the sum of MMNs to

single-feature deviants has been compared to the true MMN to the multiple-feature deviant. In most of the cases, full or at least partial additivity has been found, which again leads to the assumption that separate neuronal populations are involved in the processing of single acoustic features.

In the study of Levänen and colleagues (1993), for instance, the MMF, the magnetoencephalographic (MEG) MMN correspondent, has been recorded. Here, a partial MMF additivity of frequency and duration as single-feature deviants compared to frequency + duration as double-feature deviant and of interstimulus interval (ISI) and frequency as single-feature deviants compared to ISI + frequency as double-feature deviant has been found in tones. Lävenen et al. (1993) suggested (in line with previous evidence) that different standard stimulus cues leave at once separate traces in the brain, so a multiple-feature stimulus representation is established. This multiple-feature stimulus representation, in turn, is represented by different neuronal populations in the auditory cortex. Hence, when a deviant, violating the multi-feature standard representation in terms of several acoustic cues, is presented, separate and independent neuronal activations summed together form the MMN. This assumption got support from the finding of different source locations for MMFs to ISI and frequency deviants. Intriguingly, such an effect was not found for frequency and duration deviants, i.e., the sources of duration and frequency MMFs did not differ, which however might be due to the fact that in the study of Levänen and colleagues (1993) tones of different durations differed in their spectral features as well. Hence, the MMF source of duration deviants might have been influenced by the MMF source of frequency deviants.

Interestingly, Jaramillo and colleagues (2001) observed no MMN additivity for complex tone stimuli and their deviants in duration and frequency. Note however that in that study durational differences emerged 200 ms later than spectral ones. In line with this, previous evidence shows additive MMN effects for deviants close in latency (up to

30 ms, Levänen et al., 1993), but not for deviants separated in latency of at least 75 ms (Czigler and Winkler, 1996). Note that Winkler and colleagues (1998) suggest a larger time window (of at least 200 ms) that is needed to prevent MMN additivity.

Paavilainen and colleagues (2001) carried forward the investigation of MMN additivity in non-speech stimuli by using deviants differing from standards in one, two and three features (frequency, intensity and SOA). In their study the authors probed whether the MMN to triple-feature deviants reflects the independent processing of single acoustic stimulus cues, just as the MMN to double-feature deviants (see, e.g., Levänen et al., 1993; Schröger, 1995; Takegata et al., 1999). Surprisingly, they found an additive double-feature deviant MMN, but an underadditive triple-feature deviant MMN. Paavilainen et al. (2001) suggested that the auditory system can only work dimension-specific with two deviations within one deviant, whereas with more than two deviations, fully or partially overlapping, inhibiting and interacting brain processes may prevent MMN additivity. However, an alternative explanation for the lack of additivity in triple-feature deviants might be a ceiling effect, i.e., MMN might have a maximum possible amplitude already reached for double-feature deviants.

Beyond that, Wolf and Schröger (2001) investigated whether MMN additivity vary for temporal and frontal MMN generators. The deviants used in their experiment differed in one, two or three dimensions. More precisely, deviations in duration and/or frequency and/or intensity were applied. Here, MMN additivity was found to duration + frequency and duration + intensity deviants for frontocentral and subtemporal recording sites. MMNs to frequency + intensity and duration + frequency + intensity deviants, however, showed additivity only for subtemporal recording sites.

Furthermore, Jaramillo and colleagues (2001) probed additive MMN effects in speech stimuli. The repetitive vowel /e/ was presented to Finnish participants as standard and stimuli differing in frequency (vowel /e/ with a higher fundamental frequency[1]) or duration (a shorter vowel /e/) or a vowel change (/o/ instead of /e/) were presented as deviants. No differences in MMN amplitude have been found between deviant types. Note however that despite the fact that phonemes, e.g., vowels, differ in spectral and temporal features from one another, the difference between the multiple-feature vowels /o/ and /e/ was not identical and thus not comparable to the differences between the multiple-feature vowel /e/ and the single-feature vowels taken together (vowel /e/ differing in frequency + vowel /e/ differing in duration from the original vowel /e/). Hence, it still remains unclear whether MMN additivity occurs in speech stimuli, just as in non-speech stimuli.

——

[1] Fundamental frequency is the lowest frequency of a vowel.

Aims of Experiment 1

The current study aims to investigate whether spectral and temporal features independently contribute to vowel length processing by means of the event-related potential MMN. In this case, according to previous research on non-speech auditory processing (see Chapter 2), an additive or at least subadditive summation of MMNs to both features is suggested when vowels of different length are processed. However, an increased MMN amplitude to multi-feature deviants, i.e., in the phonological condition, in comparison to single-feature deviants, i.e., to the temporal and the spectral condition, is also expected, supposing that natural phonemes, e.g., vowels, are processed in a unique speech-specific manner and only the multi-feature contrast entails a natural speech contrast, which is assumed to activate long-term representations enhancing MMN (see above). Nevertheless, if single features independently contribute to vowel length processing, the MMN amplitude in the phonological condition should reflect the sum of MMNs in the temporal and the spectral condition (in both vowel types), whereas if the MMN amplitude is boosted by the activation of long-term representations, the size of the MMN amplitude in the phonological condition should not be deducible from the size of MMNs in the spectral and the temporal condition (in both vowel types).

Moreover, as mentioned above, previous work has shown that the influence of temporal and spectral cues on vowel identification depends on vowel height. More specifically, to identify high vowels, such as the German /i/, listeners rely more on spectral information, while temporal information is more important for identifying low vowels, such as the German /a/ (Sendlmeier, 1981; Strange & Bohn, 1998; Weiss, 1974). The present study aims to prove this effect on a neurophysiological level. More precisely, referring to previous behavioral work, a larger MMN amplitude in the spectral as compared to the temporal condition is assumed for vowel type /i/, whereas the opposite

pattern, that is to say a larger MMN amplitude in the temporal as compared to the spectral condition is assumed for vowel type /a/. Note that such a sophisticated pattern would indicate pre-attentive cue-weighting processes in vowel, i.e., phoneme perception.

Experiment 1[2]

Summary

In the present study, mismatch negativity (MMN; n = 20) was used to examine the contribution of spectral vs. temporal perceptual features to vowel length discrimination in healthy adults. Natural (i.e., spectrotemporal) vowel length differences were compared with (artificially modified) stimulus pairs varying only in temporal or spectral characteristics.

Vowel length differences in all conditions (i.e., spectrotemporal, temporal, and spectral) produced reliable MMNs over frontocentral electrode sites, indicating that each acoustic cue in isolation is sufficient for the perception of vowel length. Moreover, MMN in the phonological condition was of greater amplitude than MMNs in the temporal and the spectral condition alone, suggesting summation of multiple speech features during speech perception. This summation, however, was not fully additive. In a vowel-specific analysis, MMN additivity was only observed in vowel pair /i/-/i:/, for which spectral information is more relevant for discrimination. In contrast, in vowel pair /a/-/a:/, for which temporal information is the salient cue, comparable MMNs were elicited in the temporal and the phonological condition, but a substantially weaker MMN in the spectral condition. All in all, the present study has demonstrated that the auditory system is definitely capable of summing perceptual cues when speech sounds are processed.

____

However, while being summed these perceptual cues are weighted differently depending on their saliency for the perception of a certain phoneme.

Methods

Participants

20 healthy young adults (age 20 to 30 years) participated in the study. All participants were students of psychology, had normal sense of hearing, average or above average reading and spelling abilities and were native speakers of German. Written and informed consent was obtained from all participants. Two participants were excluded due to missing data and two, as too many artifacts were determined in the data. Thus, the final sample size was 16 (mean age 23.75; 7 females).

Stimuli

In the current experiment a stimulus set was applied, which had already been used in previous behavioral and MMN studies of our group (Christmann, Lachmann & Steinbrink, submitted; Christmann, Berti, Steinbrink & Lachmann, under revision). The stimuli consisted of four German natural vowels, i.e., /i/, /i:/, /a/ and /a:/ (see Table 2), spoken by a trained female speaker. In addition, artificially prolonged and shortened versions of these vowels were used. The PSOLA algorithm of the phonetics software package *Praat* was applied to change the length of the vowels without modifying their spectral properties. The short vowel stimulus of each of the two vowel categories was lengthened to the length of the long one (lengthened short vowels, i.e., /a/ lengthened to 145ms and /i/ lengthened to 93ms) and the long vowel stimulus of each of the two vowel

categories was shortened to the length of the short one (shortened long vowels, i.e., /a:/ shortened to 75ms and /i:/ shortened to 51ms)[3].

Table 2: Durations (in ms) of long and short vowels used in the current study.

Vowel height	Vowel pair	Vowel duration Long vowel	Vowel duration Short vowel
High	/i/ - /i:/	93 ms	51 ms
Low	/a/ - /a:/	145 ms	75 ms

Conditions

The current experiment included three conditions, a phonological (i.e., spectrotemporal), temporal and a spectral condition. In the phonological condition, two natural vowels differing in vowel length were used as standard and deviant, i.e., /a/ vs. /a:/ (as in 'Bann', [/ban/,ban] and 'Bahn', [/ba:n/,train]) and /i/ vs. /i:/ (as in 'Mitte', [/mɪtə/,middle] and 'Miete', [/mi:te/,rent]). These vowels differed both with respect to their temporal and spectral features. In the temporal condition, however, natural vowels were combined with artificially modified ones, so that the difference between the stimuli was solely temporal. More precisely, a natural long vowel was paired with an artificially shortened long vowel or a natural short vowel with the same vowel in an artificially prolonged version. In the spectral condition same rationale was applied. One natural and one artificially modified stimulus were combined, so that the difference between the two stimuli was exclusively spectral. That is to say, a natural long vowel was paired with an artificially prolonged short natural vowel or a natural short vowel with an artificially shortened long vowel (for

_____

[3] For details concerning stimulus generation and manipulation, see Christmann et al. (under review).

illustration, see Figure 4). Note that only vowels that belong to the same vowel type were combined (e.g., /a/ was combined with /a:/, shortened /a:/ or lengthened /a/, but never with a vowel of the category /i/-/i:/).
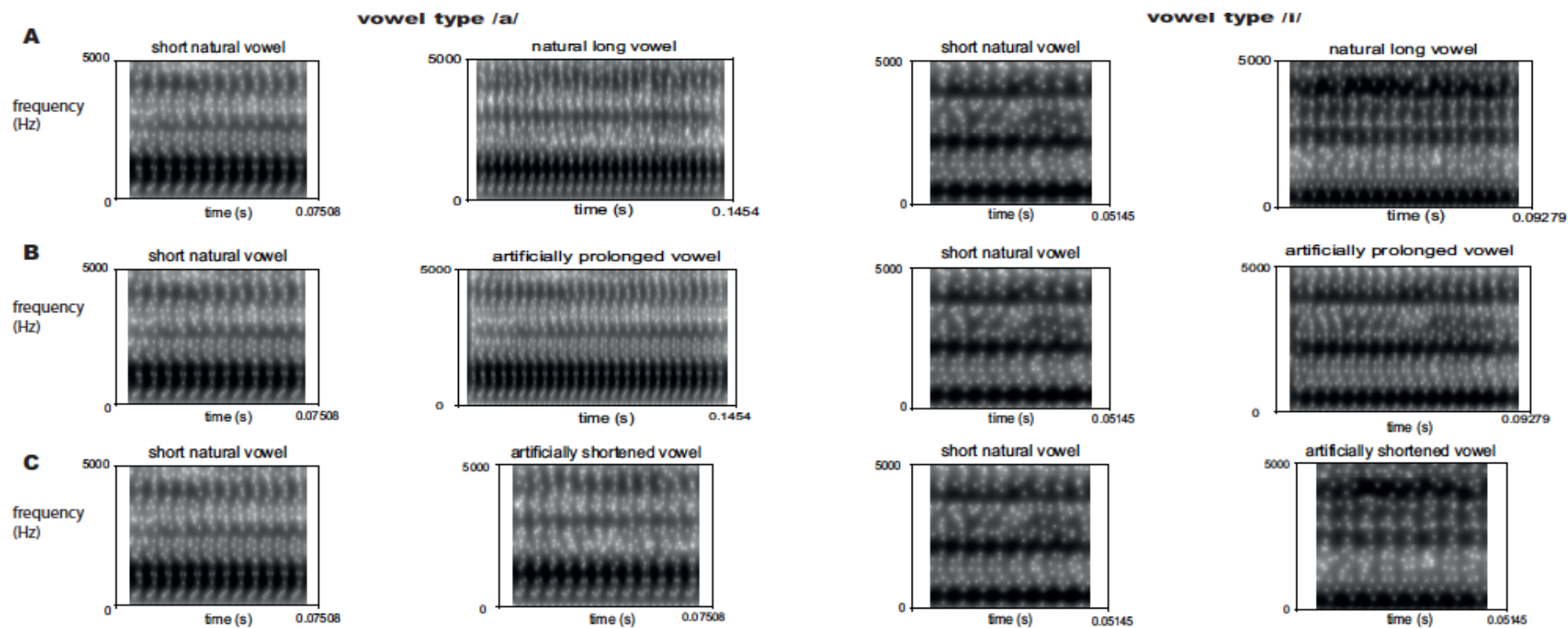
Figure 4: Illustration of stimuli (vowel type /i/ and /a/) used in the present study in all conditions. A. Phonological condition: two natural vowels of different length. Note that they differ in both their spectral and temporal features. B. Temporal condition: a natural short vowel combined with an artificially prolonged short vowel. Note that the temporal condition could also consist of a natural long vowel combined with an artificially shortened long vowel. C. Spectral condition: a natural short vowel combined with an artificially shortened long vowel. Note that the spectral condition could also consist of a natural long vowel combined with an artificially prolonged short vowel.

Experimental procedures

The stimuli were presented in a passive auditory oddball paradigm, involving the presentation of a large number of standard stimuli (82.22%) interspersed with a small number of deviant stimuli (17.78%). The paradigm was passive, so overt attention of the participants was not needed. The experiment was conducted in an acoustically and electrically shielded room. During stimulus presentation, participants were watching a self-selected silent movie. Participants were explicitly instructed not to pay attention to the auditory stimuli, which were delivered binaurally via headphones (Etymotic ER-2; Etymotic Research, Inc, Elk Grove Village, Illinois) at a fixed intensity of 67 dB SPL, controlled via Presentation software (Neurobehavioral Systems Inc., San Francisco, California).

All possible combinations of natural and artificial stimuli were used (for an overview of stimuli in different blocks, see Table 3), with each stimulus serving in different blocks as standard or deviant (for details concerning the MMN paradigm, see below). Thus, the experiment included a total of 20 different combinations of stimuli, 4 combinations for the phonological and 8 combinations each for the spectral and temporal conditions. The blocks were organized into five experimental runs, whereas one run lasted about 12 minutes. Each of these blocks was composed of 370 standard trials and 80 deviant trials. The sequence of stimuli within each block was pseudorandomized [controlling for the minimum (i.e., 3) and maximum (i.e., 5) interval between two successive deviants], and the order of blocks was randomized across participants. A constant interstimulus interval (ISI; offset to onset of successive stimuli) of 649 ms was used, whereas the last 100 ms served as baseline for the following stimulus. 10 additional standard stimuli were presented at the beginning of each block to establish an initial memory trace for standard stimuli.

Table 3: Overview of different blocks in the current study.

Condition	Block	Vowel	Combination	Function of the stimuli
Temporal	1	i	natural /i/ vs. shortened /i:/	natural /i:/ as deviant and shortened /i:/ as standard
Temporal	2	i	natural /i/ vs. shortened /i:/	natural /i:/ as standard and shortened /i:/ as deviant
Temporal	3	i	natural /i/ vs. prolonged /i/	natural /i/ as deviant and prolonged /i/ as standard
Temporal	4	i	natural /i/ vs. prolonged /i/	natural /i/ as standard and prolonged /i/ as deviant
Temporal	5	a	natural /a:/ vs. shortened /a:/	natural /a:/ as deviant and shortened /a:/ as standard
Temporal	6	a	natural /a:/ vs. shortened /a:/	natural /a:/ as standard and shortened /a:/ as deviant
Temporal	7	a	natural /a/ vs. prolonged /a/	natural /a/ as deviant and prolonged /a/ as standard
Temporal	8	a	natural /a/ vs. prolonged /a/	natural /a/ as standard and prolonged /a/ as deviant
Spectral	9	i	natural /i:/ vs. prolonged /i/	natural /i:/ as deviant and prolonged /i/ as standard
Spectral	10	i	natural /i:/ vs. prolonged /i/	natural /i:/ as standard and prolonged /i/ as deviant
Spectral	11	i	natural /i/ vs. shortened /i:/	natural /i/ as deviant and shortened /i:/ as standard
Spectral	12	i	natural /i/ vs. shortened /i:/	natural /i/ as standard and shortened /i:/ as deviant
Spectral	13	a	natural /a:/ vs. prolonged /a/	natural /a:/ as deviant and prolonged /a/ as standard
Spectral	14	a	natural /a:/ vs. prolonged /a/	natural /a:/ as standard and prolonged /a/ as deviant
Spectral	15	a	natural /a/ vs. shortened /a:/	natural /a/ as deviant and shortened /a:/ as standard
Spectral	16	a	natural /i/ vs. shortened /i:/	natural /a/ as standard and shortened /a:/ as deviant
Phonological	17	a	natural /a/ vs. natural /a:/	natural /a/ as deviant and natural /a:/ as standard
Phonological	18	a	natural /a/ vs. natural /a:/	natural /a/ as standard and natural /a:/ as deviant
Phonological	19	i	natural /i/ vs. natural /i:/	natural /i/ as deviant and natural /i:/ as standard
Phonological	20	i	natural /i/ vs. natural /i:/	natural /i/ as standard and natural /i:/ as deviant

EEG recording and analysis

Electroencephalogram (EEG) was continuously recorded using the 128 channel Geodesic Sensor Net System (EGI; Electrical Geodesics, Inc., Eugene, Oregon) and the EGI NET Amps 300 amplifier. The sampling rate was fixed at 1,000 Hz (for electrode positions, see Figure 5) and electrode impedances were kept below 50 kΩ. The reference electrode was initially placed at the vertex (corresponding to Cz in the international 10/20 system) and later re-referenced to an average reference. Virtual EOG channels were calculated to detect unexpected eye movements and eye blinks. For the horizontal EOG (virtual channel E132) the potential difference between channels E125 and E128 was computed (i.e., E132 = E125 – E128). Vertical EOGs were calculated as follows: E130 = E127 - (E21 + E25)/2; E131 = E126 – (E8 – E14)/2).
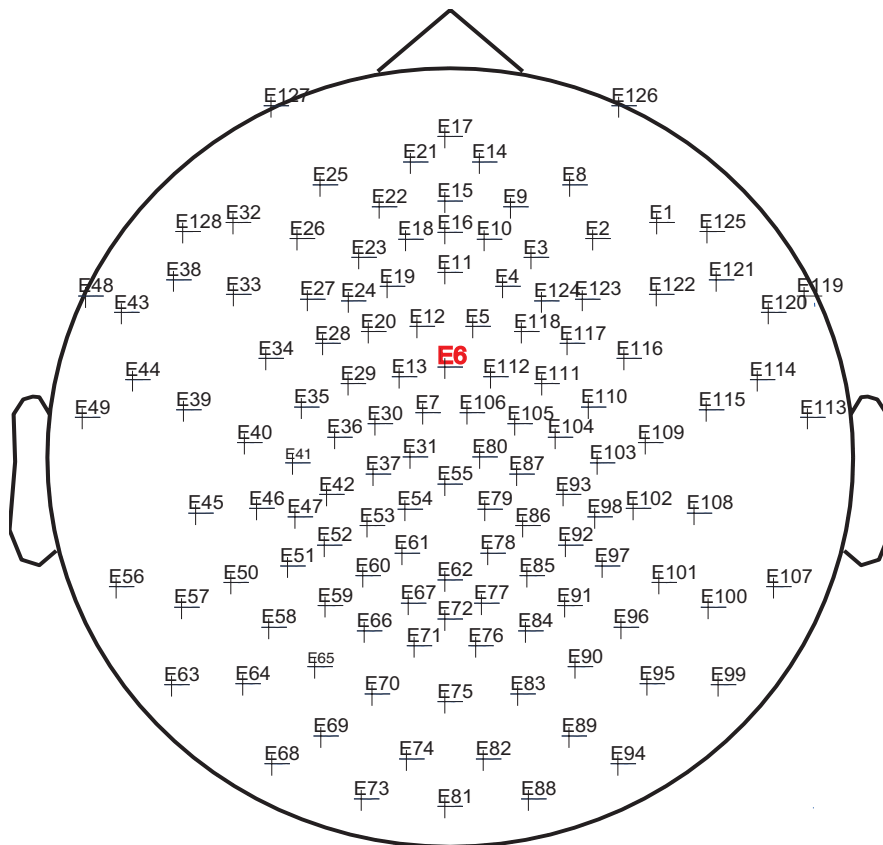
Figure 5: Electrode positions of the 128 channel Hydrocel Geodesic Net. Electrode position E6 (which corresponds to FCz in the 10-20-system) is depicted in red.

Data analysis was conducted with the open-source software Fieldtrip (http://fieldtrip.fcdonders.nl), a toolbox implemented in Matlab (The Mathworks Inc., Natick, Massachusetts). In the course of data preprocessing raw data were cut into trials that lasted from 100 ms before stimulus onset to 549 ms after stimulus offset, and bandpass-filtered using a lowpass filter of 30 Hz and a highpass filter of 1 Hz. Ocular artifacts, movement artifacts and amplifier artifacts were rejected both automatically by means of thresholding the z-transformed value of the preprocessed data (the cut-off z-value for EOG artifacts was $z = 5$ and for muscular artifacts $z = 12$; cf., Oostenveld et al., 2011) and manually by a trial-to-trial screening of all electrodes via visual inspection. ERPs were calculated by averaging trials from 0 to 549 ms after stimulus onset, using a

pre-stimulus baseline of 100 ms. Mismatch negativity (MMN) was calculated by subtracting average ERPs to standards from those elicited by deviants. ERPs from different blocks were averaged to calculate condition-specific MMNs. Importantly, the ERPs of different blocks were averaged across conditions irrespective of the function of each stimulus as standard or deviant within these conditions, as no a-priori hypotheses were established dependent on whether or not a stimulus served as standard or deviant. As an example, to calculate the MMN of the phonological condition for vowel type /a/, two blocks in which /a/ and /a:/ each served once as standard and once as deviant were taken together.

Two strategies were adopted to analyze the data. First, the mismatch activity over all 128 EEG channels, i.e., the whole scalp, was investigated by calculating paired cluster-based permutation t-tests (Maris & Oostenveld, 2007) for MMN mean amplitudes. The cluster-based permutation procedure was especially suitable for that purpose, as it minimizes the likelihood of false alarms in case of multiple comparisons. In detail, data were randomized between conditions 2,000 times and test statistics were calculated for each permutation. Significant samples (i.e., combinations of channels and time points) were clustered in connected sets based on their temporal characteristics and cluster-level statistics were calculated by taking the sum of the t-values within a cluster (clusteralpha = 0.05). The largest cluster-level statistic was taken to establish the null distribution. Then p-values were calculated as the proportion of randomizations with a more extreme test statistic than the observed one (by using the Monte Carlo Method; threshold $p = 0.05$; Maris & Oostenveld, 2007). Based on visual inspection of the data, mean amplitudes in the time window from 140 to 220 ms were taken for statistical comparisons across all conditions and vowel types, as well as between conditions and vowel types. This relatively broad time window was required because of MMN latency differences between conditions. More precisely, as temporal differences were detectable later in time than

spectral and spectrotemporal differences, MMN was elicited in the temporal condition later than in the phonological and the spectral condition. However, to explore whether differences between standards and deviants (in each condition and vowel type) were significant, a smaller time window of 40 ms around MMN peaks of interest was chosen. Note that outer electrodes (involving electrodes that were positioned at non-scalp sites) were excluded from cluster-based analyses and data visualization because of noisiness.

Second, as the time window for mean amplitudes of 80ms used in the first step of the analysis was large and no cluster-based permutation analysis of variance (ANOVA) can be computed with Fieldtrip, the results were validated with further statistical analyses of MMN peak amplitudes at electrode E6, which corresponds to the electrode FCz in the 10-20-system. The electrode E6 was selected for statistical analyses for two reasons. First, previous literature indicates that MMN is maximal at frontocentral electrode positions (e.g., Kujala et al., 2007, Näätänen, 2007; Näätänen & Winkler, 1999) and, second, a clear MMN was detected at E6 by means of visual inspection (see Figure 5). Thus, a repeated-measures ANOVA and subsequent t-tests were computed with IBM SPSS statistics 22 (International Business Machines Corporation, Armonk, New York) with the aim of testing the specific hypotheses of the present study.

Results

A significant MMN was observed over frontocentral scalp sides when averaging across all conditions and vowel types (t = -8.98; p < 0.001; time window of analysis was 140-220 ms) (see Figure 6). Subsequently, by calculating a repeated-measures ANOVA (on peak amplitudes of Fz) with the within-subject factors condition and vowel type significant main effects of condition (phonological vs. temporal vs. spectral: $F_{(2, 30)}$ = 19.34; p < 0.001) and vowel type (a vs. i: $F_{(1, 15)}$ = 23.2; p < 0.001), as well as a significant interaction between these two factors $F_{(2, 30)}$ = 12.17; p < 0.001), were found.

In the following, these effects will be resolved at a more fine-grained level. First, mismatch negativity results will be investigated separately for each condition and MMNs will be compared between conditions. Note that the comparison of MMN sizes between conditions is relevant for the investigation of additive MMN effects in speech sounds, i.e., in vowels of different length. Second, as on the one hand previous evidence indicates that the processing of /a/ and /i/, i.e., low vs. high vowels, respectively, relies to a variable extent on temporal vs. spectral information (for more details, see Chapter 1) and on the other hand a significant condition by vowel interaction was found in the above-reported ANOVA, condition effects will be tested separately for each vowel.
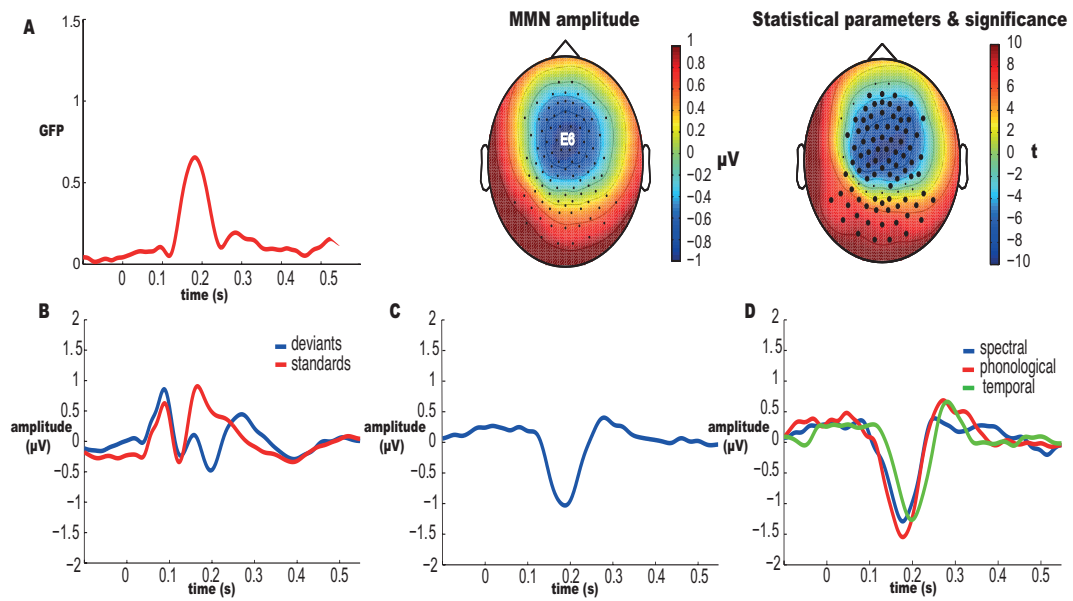
Figure 6: MMN over all conditions. A. Global field power (GFP[4]), correspondent amplitude distribution and distribution of t-values across the scalp for the statistical comparison between standards and deviants (MMN scalp topographies are shown for the time window of 170-210 ms after stimulus onset). B. ERP averages for standards and deviants. C. Corresponding MMN wave at electrode FCz. D. Condition-dependent MMNs at electrode FCz. Outer electrodes were removed for illustration reasons. The bold dots indicate significant electrodes (p < 0.05).

----

[4] Global Field Power (GFP) is a measure that characterizes global EEG activity (Lehmann & Skandries, 1980; Esser et al., 2006). Usually, GFP is calculated over subjects.

Condition-specific MMN effects
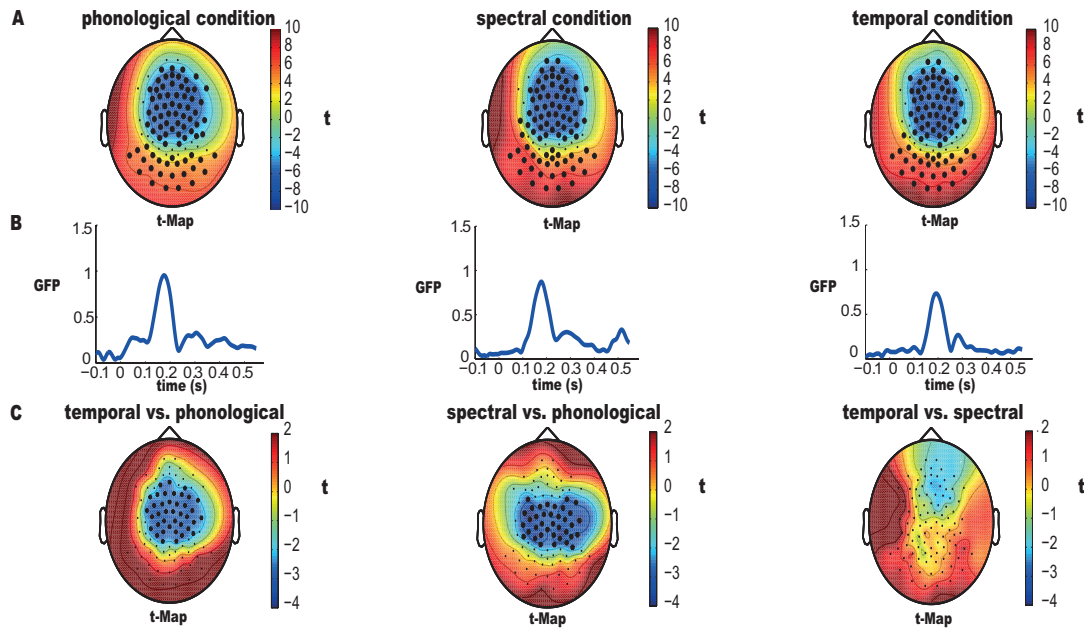
Figure 7: A. Statistical comparisons between standards and deviants for the phonological, the spectral (time windows of 160 – 200 ms post stimulus onset) and the temporal condition (time window of 180 – 220 ms post stimulus onset) (calculated over both vowel categories). Distribution of t-values across the scalp. B. Corresponding global field power (GFP). C. Pairwise statistical comparisons between conditions. Distribution of t-values across the scalp. Outer electrodes were removed for illustration reasons. The bold dots indicate significant electrodes (p < 0.05).

To resolve the main effect of condition and as a consequence, to investigate whether MMN additivity, i.e., a larger MMN in the phonological condition in comparison to the temporal and spectral conditions, is possible in speech, cluster-based permutation tests were calculated to examine MMN effects separately in each condition. As displayed in Figure 7 A, significant MMNs were found at frontocentral scalp regions for the phonological condition (t = -7.71; p < 0.001)), the spectral (t = -7.53; p < 0.001) and the temporal condition (t = -9.05; p < 0.001) (Figure 7 A). Moreover, as expected, a significantly larger mean MMN amplitude in the phonological condition in comparison to

the mean MMN amplitudes in the temporal (t = -2.6; p < 0.001) and the spectral condition (t = -3.13; p < 0.001) was found (see Figure 7 C). No significant differences were determined between the temporal and the spectral condition (t = -1.29). Furthermore, these findings were replicated with statistics calculated on peak amplitudes at electrode FCz (phonological vs. temporal t(15) = -2.702; p = 0.016); phonological vs. spectral t(15) = -3.275; p < 0.01; spectral vs. temporal t(15) = 0.61; p = 0.551).

Condition effects differ between vowels

As the next step, mean MMN amplitudes were examined separately for each vowel type in all conditions. Robust MMNs were found for the phonological (t= -8.55; p < 0.001), the temporal (t = -7.63; p < 0.001) and the spectral condition (t = -7.07; p < 0.001) for vowel type /i/-/i:/ and for the phonological (t= - 4.58; p < 0.001), the temporal (t = -5.69; p < 0.001) and the spectral condition (t = -4.63; p < 0.001) for vowel type /a/-/a:/ (see Figure 8). However, as expected due to the very small spectral difference between /a/ and /a:/, the spectral MMN for vowel type /a/-/a:/ was substantially smaller in amplitude than the MMNs in all other conditions.

It was predicted that temporal information is more relevant than spectral information for the correct discrimination of vowels of vowel type /a/, while spectral information is more crucial than temporal information for the correct discrimination of vowels of vowel type /i/. In line with this hypothesis, a condition by vowel interaction was found in the 3 x 2 repeated-measures ANOVA with the two within-subject factors condition and vowel type calculated on peak amplitudes of Fz (see above).
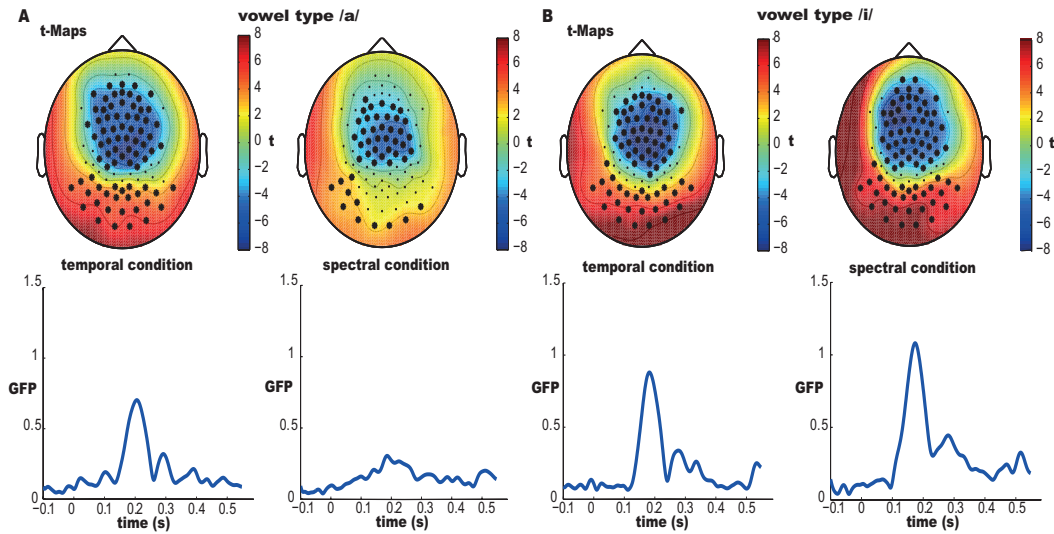
Figure 8: MMNs for A. vowel type /a/ and B. vowel type /i/ in the temporal and the spectral condition. Distribution of t-values across the scalp and corresponding GFP. All display conventions as described for Figure 7.

For the purpose of a more detailed consideration of vowel-specific MMN effects, an additional 2 x 2 repeated measures ANOVA with within-subject factors condition (spectral vs. temporal) and vowel (/a/ vs. /i/) was calculated on peak amplitudes. Significant main effects for the factors condition $F(1,15) = 5.33$; $p = 0.036$) and vowel $F(1,15) = 23.00$; $p < 0.001$), as well as a significant condition by vowel interaction ($F(1,15) = 15.99$; $p < 0.001$) were found. As expected, cluster-based permutation tests revealed a larger mean MMN amplitude in the temporal condition in comparison to the spectral condition for vowel type /a/ ($t = -4.07$, $p < 0.001$; see Figure 9A). For vowel /i/, in contrast, no difference was found between MMNs of the spectral and the temporal condition ($t = -0.07$; see Figure 9B). Same results were yielded with additional t-tests on peak amplitudes at electrode FCz (vowel type /a/-/a:/ $t(15) = -4.18$, $p < 0.001$; vowel type /i/-/i:/ $t(15) = 1.73$; $p = 0.1$).

Furthermore, cluster-based permutation tests revealed that additive MMN effects were present only for vowel type /i/-/i:/, as only for this vowel category, the phonological

condition elicited a larger MMN than both the temporal (t= -4.521; p < 0.001) and the spectral condition (t= -5.093; p < 0.001). This was different from vowel type /a-a:/, where the MMN amplitude did not differ between the phonological and the temporal condition (t = -0.847), and spectral information – even though eliciting a weak MMN when presented in isolation – did not seem to increase the phonological MMN beyond the temporal MMN. These results were replicated by additional t-tests on MMN peak amplitudes at FCz (vowel /i/: phonological vs. temporal t(15) = -4.854; p < 0.001; phonological vs. spectral t(15) = -4.230; p = 0.001; vowel /a/: phonological vs. temporal t(15) = 0.203; p = 0.842; phonological vs. spectral t(15) = -4.895 p < 0.001).
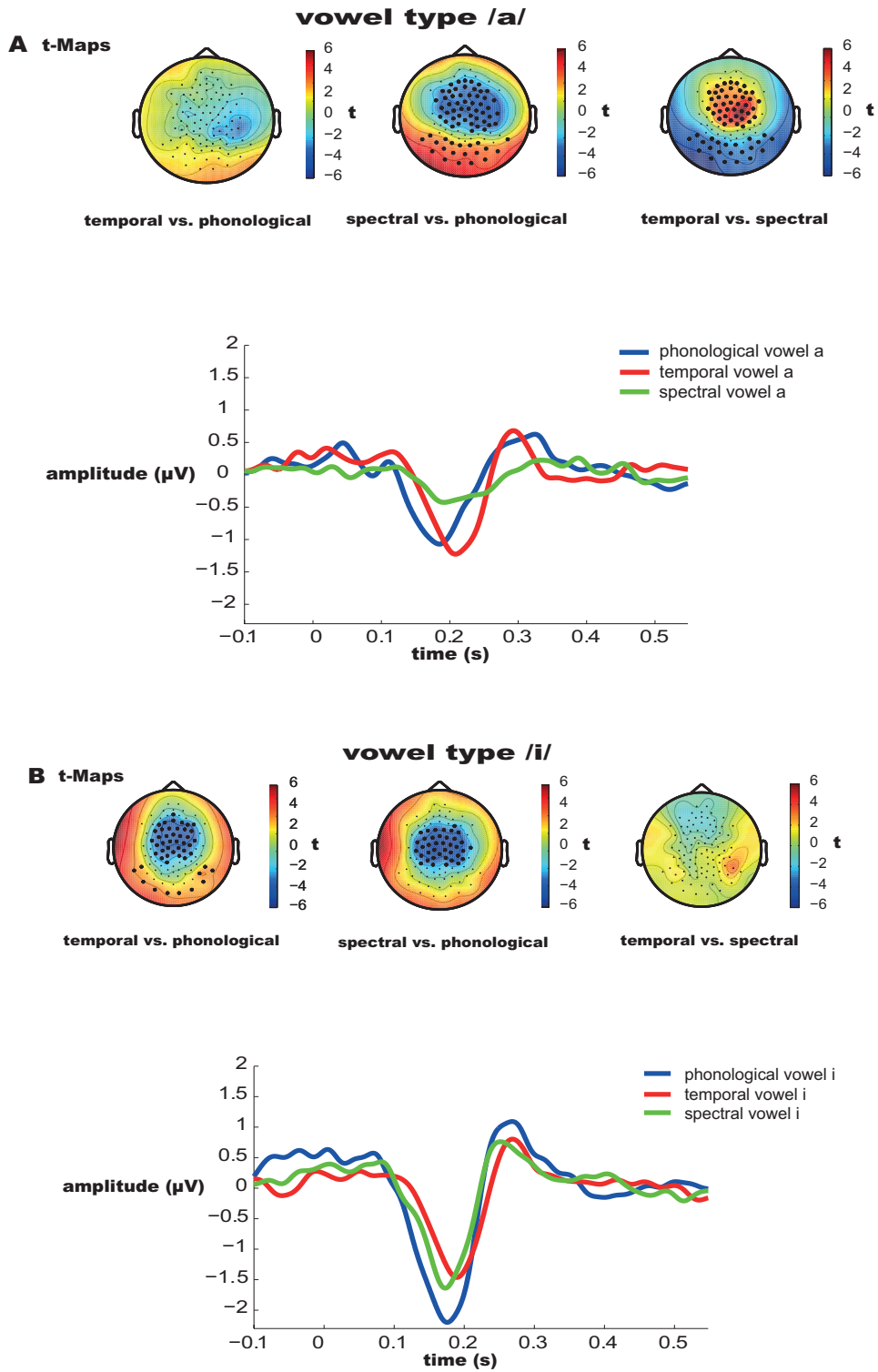
Figure 9: Condition effects varying between vowel types. Distribution of t-values across the scalp for pairwise statistical comparisons between conditions for vowel types A. /a/ and B. /i/. Outer electrodes were removed for illustration reasons. The bold dots indicate significant electrodes (p < 0.05).

Discussion

The present study investigated the contribution of spectral vs. temporal features of the speech signal to vowel length discrimination in German. For that purpose, natural vowel pairs, differing in spectral and temporal information and pairs of natural and artificially modified vowels, differing either in temporal or in spectral information were presented as standards and deviants in a passive oddball paradigm. MMN was assessed on one side as an indicator of the brain's ability to differentiate long vs. short vowels based on spectral or temporal information alone and on the other side as an indicator of whether or not the combined processing of both features within a speech contrast is characterized by a summation of the neural responses to these features. Furthermore, as mentioned above, two vowel pairs, differing with respect to the importance of temporal and spectral cues for vowel length discrimination (/a/-a:/: temporal information is salient; /i/-/i:/: spectral information is salient) were applied as stimuli in the present study to examine at a neurophysiological level how spectral and temporal information contribute to vowel length discrimination, dependent on saliency.

Processing of spectral and temporal information within speech

Combined analysis across vowel types showed that single features are sufficient for the adult healthy brain to differentiate vowels of different length, which was indicated by reliable MMNs in the spectral and the temporal condition. This result demonstrates that the human brain is highly robust against impaired speech signals. However, the phonological, i.e., spectrotemporal condition, characterized by spectrotemporal differences between long and short vowels, elicited a larger MMN than the other conditions, indicating that the presence of both features leads to even more efficient vowel length recognition. As outlined above, this MMN additivity effect for speech

sounds is particularly interesting, as previous studies have shown comparable effects only for non-speech stimuli (i.e., tones; Levänen et al, 1993; Schröger, 1995; Gomes et al. 1997; Winkler et al., 1998; Takegata et al., 1999), including tones composed of spectral and temporal cues (Levänen et al., 1993). However, to date, the MMN additivity results for non-speech multiple-feature deviants composed of spectral and temporal cues are inconsistent. Levänen and colleagues (1993; Experiment 2), for instance, measured the MMF (MEG correspondent of MMN) to deviants in frequency (i.e., 1 kHz vs. 1.1 kHz) and duration (i.e., 100 ms vs. 50 ms) and reported that the MMF amplitude is significantly increased by a factor of around 1.5 in double-feature deviants as compared to single-feature deviants. Jaramillo et al. (2001), however, found no MMN additivity for complex tone stimuli and deviants in duration (200 ms vs. 400 ms) and frequency (i.e., fundamental frequency of 105 vs. 117 Hz, and harmonics). Interestingly, the temporal interval separating the two deviant features seems to be critical for the appearance of additivity in multi-feature deviants: Winkler and colleagues (1998), for instance, suggest that additivity occurs when both deviants appear within a time window of 200 ms or less, whereas in case of a separation of more than 200 ms separate MMNs are elicited. In the current study, temporal (i.e., duration) deviants occurred later than spectral deviants, more precise, between 42 and 70 ms later. Referring to Winkler et al. (1998), this temporal difference definitely falls within the temporal window associated with additive summation of MMN responses.

Note that the summation in the present study was only slightly weaker than observed for non-speech auditory stimuli in the study of Levänen and colleagues (1993), i.e., the MMN amplitude of the phonological condition was on average around 1.4 times the amplitude of the single feature conditions. Summing up, the general observation of a greater MMN in the phonological condition in comparison to the purely spectral or the purely temporal condition, independent of the specific type of vowel, can be taken as

evidence for the ability of our auditory system to integrate single acoustic cues during phoneme processing – without however summing the neural responses to both cues in a strictly additive way.

Vowel-specific processing of spectral and temporal information

As outlined above, previous behavioral studies have shown that people rely more on spectral than on temporal information to discriminate high vowels of different length. The opposite is true for the discrimination of low vowels of different length (e.g., Groth et al., 2011, Sendlmeier, 1981; Strange & Bohn, 1998; Weiss, 1974). The present study aimed to investigate on a neurophysiological level whether or not spectral and temporal information contribute differently to vowel length processing in different vowel types and whether this differential contribution is associated with vowel height, as previously shown on a behavioral level.

When analyzing the two vowel types separately, indeed a greater temporal than spectral MMN was obtained for the low vowel /a/, but no MMN differences were found between the temporal and the spectral condition for the high vowel /i/. Most specifically, while a robust MMN was elicited in the temporal condition for vowel type /a/, the spectral MMN for this vowel type was negligible. For vowel type /i/, robust MMNs were elicited in both the spectral and the temporal condition. This may not be surprising, as in natural speech the spectral difference between the short and the long vowel /a/ is in fact rather small in comparison to the temporal difference between the short and the long vowel /i/. Consequently, when keeping the temporal difference for vowels of vowel type /a/ constant, as it was done in the spectral condition, the acoustical difference between the stimuli is negligible small and therefore hard to detect, resulting in an extremely weak spectral MMN. This is not the case, when keeping the spectral difference for vowels of vowel type /i/ constant, as it was done in the temporal condition, because we still have a

considerable temporal difference between the stimuli here, resulting in large MMNs in the spectral and the temporal condition. In sum, the results indicate that in high vowels in comparison to low vowels, spectral information plays a substantially greater role in vowel length discrimination.

This is in line with the behavioral results of Groth et al. (2011). Here, the behavioral impairment in the overt categorization task (Groth et al., 2011) indicated that temporal information alone might often not be sufficient for correct discrimination of high vowels. Note, moreover, that while in the current study no statistical difference between the spectral and the temporal condition was detected for vowel type /i/, descriptively the same trend was found as in previous studies. In line with past behavioral research (Weiss, 1974; Sendlmeier, 1981; Strange & Bohn, 1998; Groth et al., 2011; Steinbrink et al., 2014) Christmann and colleagues (2014), for instance, have shown in a preceding MMN study that spectral cues carry indeed a greater amount of information concerning vowel length discrimination for the vowel category /i/. In that study the same stimulus set was applied as in the present study. The differing results of the present study and the study of Christmann et al. (2014) may be due to methodological differences, e.g., because Christmann et al. (2014) used a multi-feature paradigm, a paradigm, which incorporates multiple deviants within one block, while the current study applied a classical oddball paradigm. Possibly, the multi-feature paradigm is more appropriate to capture subtle processing differences (for similar conclusions, see, e.g., Kujala et al., 2006b).

Furthermore, a larger MMN in the phonological condition in comparison to the spectral and temporal conditions was found only for vowel type /i/. For vowel type /a/, however, no difference between the phonological and the temporal condition could be detected. These results indicate that the summation of brain activation to various cues involved in phoneme perception might only be applied under certain conditions. The observation of an increased MMN amplitude for natural, i.e., spectrotemporal, vowel

length deviants in the vowel contrast /i/-/i:/ indicates that spectral and temporal features are presumably used in a comparable manner for vowel length discrimination, i.e., the summation of spectral and temporal information occurs when both acoustic cues are available. On the other hand, the fact that this summation is not strictly additive may indicate either that information transmitted by these two features is partly redundant or, alternatively that the integration of temporal and spectral cues weighs one of the two features (i.e., the more salient acoustic cue) more strongly. For the vowel category /a/-/a:/, in turn, the less salient spectral differences are processed to a considerably lesser extent than the temporal differences. Thus, under circumstances (here: for low vowel /a/-/a:/) in which both acoustic cues are available, the brain seems to rely to a larger extent on the salient cue (here: temporal information).

This result is in line with previous evidence demonstrating preattentive cue-weighting processes in phoneme perception (e.g., Moberly et al., 2014; Lipski et al., 2012; Jacobsen et al., 2004). Past MMN research indicates that the relative contribution of spectral and temporal cues may vary depending on the phoneme category and the considered language system (e.g., Moberly et al., 2014; Lipski et al., 2012). Moberly and colleagues (2014), for instance, have shown that native English listeners rely to a greater extent on spectral in comparison to temporal cues in the processing of the /ba/-/wa/-contrast. Moreover, Lipski and colleagues (2012) detected that Dutch listeners as compared to Spanish learners of Dutch exhibit greater MMN amplitudes to spectrally cued contrasts in comparison to temporally cued contrasts. Note however that none of these studies investigated the relative impact of certain acoustic cues within a phoneme contrast (composed of these acoustic cues), since MMNs have been compared between artificially modified speech stimuli, where single acoustic cues have been isolated. More specifically, when the MMN was larger in a certain artificially modified speech contrast, with the difference between stimuli being temporal in nature, in comparison to the very

same speech contrast, with the difference being spectral in nature, it has been concluded, that temporal cues are more crucial for this certain speech contrast and vice versa (e.g., Moberly et al., 2014; Lipski et al., 2012). Thus, the novel contribution of the present work relates to how different aspects of the speech signal are integrated during phoneme perception.

In sum, the vowel-specific results of the present study suggest that the mature speech perception system integrates different acoustic cues of the speech signal, while giving more weight to more salient cues, i.e., in the case at hand, spectral and temporal features contribute differentially to vowel length discrimination dependent on their saliency, which again varies with the vowel type under consideration. Interestingly, temporal information seems to be especially important, supporting vowel length processing even in vowels that rely more strongly on spectral information. By way of illustration, when temporal information was the salient cue (vowel pair /a/-/a:/), the temporal MMN was larger than the spectral MMN and equivalent to the phonological MMN, indicating a stronger reliance on temporal cues. However, when spectral information was the salient cue (vowel pair /i/-/i:/), the spectral MMN was equivalent to the temporal MMN and smaller than the phonological MMN, indicating that both temporal and spectral cues play an important role for vowel length processing.

Possible mechanisms of MMN summation in phoneme perception

MMN additivity in non-speech stimuli is assumed to result from the ability of the auditory cortex to represent different features of an auditory stimulus by different neuronal populations. The appearance of a deviant stimulus that violates more than one of these memory traces, leads to corresponding separate and independent neural activations, which taken together form the MMN. In support of this reasoning, evidence shows that different features of deviant non-speech stimuli elicit activations in different neuronal

populations, that is to say, in different areas of the auditory cortex (Giard et al., 1995; Levänen et al., 1996).

However, larger MMNs in multi-feature stimuli in comparison to single-feature stimuli in speech might emerge for reasons other than in tones. First, it is assumed that only native speech stimuli, but not artificially modified ones, activate long-term memory representations. Correspondingly, in the present study, an enlarged MMN was found in the phonological condition, but not in the temporal or in the spectral condition. In line with this account, numerous studies have shown that linguistic experience with and accordingly the involvement of long-term representations in the processing of speech segments like phonemes (e.g., Dehaene-Laembertz, 1997; Näätänen et al., 1997; Phillips et al., 2000) or syllables (Alho et al., 1998b; Shtyrov et al., 1998) results in an enlarged MMN. Correspondingly, native language stimuli elicit larger MMNs than stimuli of a foreign language (e.g., Nenonen et al., 2003; Winkler et al., 1999; Näätänen et al., 1997).

Alternatively, the phonological condition might have been the only condition including categorical phoneme differences. It is well established that speech perception involves categorical phonetic representations. According to the phonetic category boundary effect, the discrimination of different phonemes is easier in comparison to the discrimination of speech sounds within the same phoneme category, even when the acoustic difference is identical (Liberman et al., 1957; Baddeley et al., 1976; Philips et al., 1995; Winkler et al., 1999; Ylinen et al, 2006 etc.). For example, Winkler and colleagues (1999) have shown that MMN is elicited by both sensory and phonetic representations of speech stimuli. In that study phonemic within-category and across-category contrasts were taken to form natural and synthesized vowels. A larger MMN was found for across-category contrasts in comparison to within-category contrasts. With respect to the present study, the additional assumption can be plausibly made that natural vowel stimuli in the phonological condition belong to different phonemic categories,

whereas they would not be perceived as categorically different in the temporal and the spectral condition, as these contrasts do not occur in natural speech. Hence, the phonetic category boundary effect would indeed also predict a larger MMN amplitude for the phonological condition compared to the spectral and the temporal condition in the present study.

Note that even though all explanations for MMN additivity outlined in this section are plausible, they cannot fully account for the vowel-specific effects found in the present study. More precisely, if the activation of long-term memory representations or categorical perception of speech sounds alone is responsible for MMN additivity determined in the present study, MMN additivity should be present in both vowel types, which was not the case in the present study. Furthermore, the multiple-feature MMN was not fully deducible from the corresponding single-feature MMNs, in both vowel types (vowel type /a/: the MMN in the spectral condition did no increase the MMN in the phonological condition beyond the MMN in the temporal condition; vowel type /i/: the MMN in the phonological condition was smaller than the sum of MMNs in the spectral and the temporal condition). Hence, MMN additivity in speech seems indeed to be affected by cue-weighting mechanisms.

Conclusion

Summing up, the main finding of the study is the observation of the (not strictly) additive MMN effect for speech stimuli during – pre-attentive – vowel length discrimination in healthy adults. Furthermore, differential effects were found for different vowel types, indicating that spectral and temporal information is differentially important for length discrimination in different vowels, depending on saliency of different features for different vowels.

Note, moreover, that the findings of the present study demonstrate that single features (here: duration and spectral composition) of the speech signal are processed separately as different neuronal representations, i.e., in a cue-specific manner. However, unlike in non-speech (i.e., pure tone) sounds, they are weighted differently before their neural responses are summed up to the additive MMN. That is to say, cue-specific and domain specific mechanisms are both applied when vowels, i.e., phonemes, are processed (for more theoretical background on models of speech processing, see Chapter 1).

Chapter 3: Auditory processing mechanisms in children and adults

ERPs in children and adults

Event-related potentials (ERPs) change in the course of development. Hence, the complexity of a waveform may alter, old components may disappear or new components may arise (Cheour et al., 2001). Usually, latencies shorten and amplitudes increase when children are growing up (Thomas and Crow, 1994). So far, the neurobiological causes for these maturational changes have not been clarified. However, there are some interesting assumptions. For instance, maturational changes may be the result of an increase in neural conduction velocity due to myelination processes (Eggermonth, 1988, 1992) or they may be referable to an enhanced synaptic density in the auditory cortex (Eggermonth, 1988; Huttenlocher et al., 1982). Moreover, it should be always kept in mind that ERPs in children can be also significantly influenced by the stimuli used, the sample sizes, SOAs or simply the type of electrodes applied to record the data (Bishop et al., 2011; Maurer et al., 2003) and hence, should be interpreted carefully.

MMN in children and adults

Investigations in children and infants concerning the development of the event-related potential MMN are rare. However, it is well known that the maturation of auditory processing extends to the second decade of life (Ponton, Eggermont, Kwon & Don, 2000; Pang & Taylor, 2000; Shahin, Roberts & Trainor, 2004). Speech perception in noise, for instance, continue to develop up to teenage years due to the maturation of axons in layer II and upper layer III of the auditory cortex (Eggermont & Ponton, 2003).

Although most studies show that MMN is developmentally rather stable (Csépe, 1995; Kraus et al., 1992, 1993; for a review, see Cheour et al., 1999), MMN latency is slightly longer in children and infants compared to adults. The evidence base concerning the MMN amplitude is rather inconsistent. Here, the majority of studies have demonstrated that the MMN amplitude is larger in children than in adults or of comparable size in both groups (Csépe, 1995; Kraus et al., 1992, 1993; Sharma e al., 1993), which is surprising since, in general, ERPs tend to increase with age. Nevertheless, there is also evidence showing larger MMN amplitudes in adults compared to children (e.g., Bishop et al., 2011; Oades, Dittmann-Balcar, Zerbin, 1997). These contradictory findings may be explained by a variety of facts as methodological dissimilarities between studies, e.g., different stimulus types (synthesized vs. natural or vowels vs. consonants), or floor effects concealing differences in MMN amplitude between children and adults, as stimuli of low complexity were often used (Putkinen et al., 2013). Note however that larger MMN amplitudes in adults compared to children would be in line with previous behavioral studies showing that performance on auditory tasks improves from childhood to adulthood. Nevertheless, it is difficult to estimate in what extent the performance improvements in behavioral tasks reflect either changes in the auditory system or enhanced coping with task demands and better concentration skills (Banai & Ahissar, 2006; Sutcliffe & Bishop, 2005; Werner & Marean, 1996).

Additionally, the MMN scalp distribution differentiates children from adults. In adults, MMN is a predominantly frontocentral negativity (Alho, 1995), whereas in children and infants MMN cannot only be obtained over the frontocentral scalp area, but also over parietal regions (Cheour et al., 1996, 1998a; Cheour-Luhtanen et al., 1995, 1996; Maurer et al., 2003). This broader MMN scalp distribution in children may reflect the involvement of more or larger parts of the brain in auditory processing or may result from differences in skull thickness and accordingly in conductivity of the skull between

children and adults. Moreover, less specificity and more redundancy in connections between different parts of the brain is a plausible explanation as well (Neville, 1995). Furthermore, it should be kept in mind that frontal areas mature later than, e.g., temporal areas, and hence the frontal MMN component may develop later than, e.g., the temporal one, which again may have an impact on differences in MMN topography between children and adults (Alho, 1995; Giard et al., 1990). However, so far, there is no clarifying evidence on that account.

Taken together, it is still an open question to which extent the MMN measured in children resembles the one recorded in adults, e.g., in the underlying processes or topography (Picton & Taylor, 2007). However, as MMN has some undisputed advantages as an investigation tool of auditory processing in both children and adults, it is of special interest to learn to interpret MMN reactions appropriately as a function of age. First, behavioral measures require attention and cooperation, which is difficult to assure especially in studies with children. MMN is, however, an attention independent measure (Alho et al., 1994). Furthermore, as MMN is developmentally quite stable, i.e., can be even determined in infants (Cheour et al., 1997a; Kurtzberg et al., 1995; Leppänen, Eklund, and Lyytinen, 1997), it is a useful measure of auditory processing capacity in children. Moreover, both MMN studies with adults and those with children have shown that MMN is suitable for the investigation of neuronal plastic changes in the auditory system, which are associated with learning processes (Kraus et al., 1995b; Näätänen et al., 1997; Tremblay et al., 1998; Winkler et al., 1999; Dehaene-Lambertz & Baillet, 1998; Cheour et al., 1998).

Chapter 4: Developmental dyslexia

Developmental dyslexia is defined as a specific impairment in learning to read and/or to spell, which does not derive from an intelligence impairment, sensory deficits, a general brain dysfunction or inadequate schooling (American Psychiatric Association, APA, 2000). From a behavioral perspective, dyslexia is characterized by poor phonological processing skills (Wagner & Torgesen, 1987; McBride-Chang, 1996). First, dyslexics have deficits in phonological awareness, i.e., in the ability to perceive and manipulate correctly phonetic aspects of speech (Bradley & Bryant, 1983; Bruck, 1992; Swan & Goswami, 1997a). By way of example, this deficit can be captured in tasks where participants have to choose words starting with the same sound from a subset of words given (e.g., "hat", "bat", "hot", or "sun") or in tasks where words have to be segmented (e.g., "sun" into the sounds "s", "u" and "n") or single sounds have to be removed from words (e.g., "l" from "play") (Gabrieli, 2009). Second, impairments in phonological working memory are common in dyslexia (Jeffries & Everatt, 2004; Nelson & Warrington, 1980; Roodenrys & Stokes, 2001; Steinbrink & Klatte, 2008), which can be registered in tasks probing the immediate serial recall of unrelated items in the same order they were presented or in a reversed order (McBride-Chang, 1996). Third, dyslexics show a deficit in lexical access to phonological representations of the long-term memory, which has been often investigated with rapid automatized naming (RAN) (Denckla & Rudel, 1976; Fawcett & Nicolson, 1994; Swan & Goswami, 1997b). Here, participants have to name a subset of pictures, colors or letters as fast as possible. Dyslexics perform worse than controls on this task, which again is hypothesized to reflect a less efficient phonological recoding of well-known visually represented items already stored in verbal long-term memory (McBride-Chang, 1996).

In addition to impairments in phonological processing, previous studies have found deficits in several fields of speech perception in dyslexia (Adlard & Hazan, 1998; Manis et al., 1997; Ziegler et al., 2009). For instance, studies have shown that dyslexics perform worse when the discrimination of consonants (Schulte-Körne et al., 1999; Manis et al., 1998), consonant-vowel (CV-) syllables (Cornelissen et al., 1996) or vowels (e.g., Landerl, 2003) is required. However, the results are rather inconsistent indicating that speech perception deficits might be present only in a subgroup of dyslexics. Adlard and Hazan (1998), for instance, presented monosyllabic words, differing either in one consonant (e.g., met vs. net), in a consonant cluster (e.g., bow vs. blow) or in consonants surrounded by two vowels (e.g., aga vs. aba) to poor spellers and controls. A poor discrimination rate was found only for 30 percent of dyslexics taking part in that study. Accordingly, Landerl (2003) has determined deficits in vowel length discrimination in only between 41 and 55 percent of poor spellers.

To date, dyslexia is the most widespread learning disability and has a prevalence of approximately 5% in German population (Haffner et al., 1998; Schulte-Körne, 2002). Dyslexia is a neurobehavioral disorder with a genetic contribution persisting into adulthood (Bruck 1992; Felton et al 1990; Scarborough 1990; Francis et al 1996; Shaywitz et al 1995). However, as dyslexic children grow up the symptoms may change, more precisely dyslexics may learn to read words accurately, i.e., to assign graphemes to phonemes correctly, still having issues with reading at a sufficiently fast rate, i.e., with achieving automaticity in reading (Shaywitz & Shaywitz, 2005).

Developmental dyslexia is not language-specific (Schulte-Körne & Bruder, 2010), which is indicated by similar prevalences determined across languages. Furthermore, impairments in phonological processing seem to account for dyslexia in all languages. Though alphabetic languages (e.g., German, Spanish, Italian or English) are characterized by correspondence of letters to speech sounds, there is high variance in regularity. By

way of example, in German letters can be mapped to speech sounds more regularly, i.e., grapheme-phoneme decoding is easier, than in English. In languages with more regularity, (dyslexic) children above all have issues with reading fluency, whereas in languages with less regularity, e.g., non-transparent languages, both reading fluency and accuracy are difficult to master (Gabrieli, 2009). Accordingly, it was found that German dyslexic children are predominantly impaired in reading speed (Landerl, Wimmer, & Frith, 1997; Wimmer, 1993; Wimmer, Landerl, & Frith, 1999; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003), but not in reading accuracy (Goswami, 1999). Hence, varying diagnostic criteria are used across languages, with the result that diagnoses of dyslexia are difficult to compare. In English for instance dyslexia is diagnosed on the basis of a discrepancy between IQ or age and reading accuracy, which cannot be the right criterion for transparent language systems, where reading accuracy is rarely impaired (Vellutino et al., 2004).

Aetiology

To date, there is just little consensus about the underlying causes of dyslexia. However, a range of correlates and correspondent theories has been described in the past. The most popular theories leading to a great amount of research are for instance: the phonological processing deficit theory (Snowling, 1981; Snowling, 2000; Stanovich, 1988) and the auditory processing theory (e.g., Tallal, 1980), which both, as they are crucial for the present work will be described in more detail in the following sections. However, there are many more. The cerebellar theory (Denckla, 1985), for instance, grounded on findings of anatomic and metabolic cerebellar differences between dyslexics and controls, suggests that the dyslexic's cerebellum is mildly dysfunctional, which again leads to cognitive impairments. The visual theory (Lovegrove et al., 1980; Livingstone et al., 1991; Stein & Walsh, 1997), in turn, proposes that a visual deficit, i.e., abnormal

perception of visual motion or unstable binocular fixations, may cause problems with the processing of single letters and words. The magnocellullar theory (Demb, Boynton, Best, & Heeger, 1998; Galaburda & Livingstone, 1993; Stein, 2001) unifies visual, auditory, tactile, cerebellar and phonological findings by suggesting that magnocellular deviations are accountable for dyslexic symptoms. And last but not least, the double deficit hypothesis assumes three types of dyslexia, one with deficits in phonological awareness and grapheme-phoneme assignment, one with impairments in rapid automatized naming having an impact on reading fluency and a combination of both, whereas the presence of both deficits leads to an aggravated dyslexia manifestation.

Furthermore, a genetic contribution to developmental dyslexia is undisputed (Pennington and Gilger, 1996). One previous study has shown, for instance, that 23 to 65 percent of children with at least one parent with dyslexia have dyslexia as well (Scarborough, 1990).

Phonological processing deficit theory

The phonological theory proclaims a central and causal role of phonology in dyslexia and consequently conceives dyslexia as a disorder within the language system (Ramus, 2003). Undeniably, dyslexia is associated with deficient representation, storage and/or retrieval of speech sounds from long-term memory, which again affects phoneme perception (Ramus, 2003). Deficits in phoneme perception, as for instance demonstrated in categorical perception tasks (Adlard and Hazan, 1998; Godfrey et al., 1981, Manis et al., 1997; Mody et al., 1997), in turn, may account for the underspecification of phonological representations (Adlard and Hazan, 1998; Boada and Pennington, 2006; Elbro and Jensen, 2005; Manis et al., 1997; Mody et al., 1997), and eventually for the suboptimal access to these representations in long-term memory (Ramus & Szenkovits, 2008), which again impairs the capacity to segment and manipulate phonemes and to assign graphemes

to phonemes and vice versa. In conclusion, the ability to discriminate phonemes could have an impact on phonological processing, which again affects reading and spelling ability.

From the anatomical perspective, dysfunctions in left-hemispheric perisylvanian brain areas are suggested to be relevant in the context of deficits in phonological processing, as the perisylvanian brain regions are associated with the storage of phonological representations and the linkage between phonological and orthographical representations (Ramus, 2003).

The phonological processing theory is predominantly supported by findings showing deficits in dyslexic samples in phonological awareness, verbal short-term memory and in the recall of phonological representations from long-term memory (for more details, see above). Furthermore, previous evidence shows that training studies addressing phonological processing, i.e., letter-sound mapping, improve reading and spelling ability (e.g., Adams, 1990; Wise et al., 1999; Scanlon & Vellutino, 1996; Scanlon et al., 2000; Torgesen et al., 1999). In addition, functional brain imaging studies underpin the causal role of perisylvanian dysfunctions for phonological processing deficits in dyslexia (e.g., Paulesu et al., 1996; Shaywitz et al., 1998; Temple et al., 2001; Shaywitz et al., 2002).

The general auditory processing theory

Opponents of the phonological theory of dyslexia do not doubt the existence of phonological deficits and its impact on reading difficulties, but rather postulate underlying factors leading to them. As speech consists of single acoustic features, some researchers suggest that their correct and precise representation within the speech signal is a crucial precondition for the specification of phonological representations (Ahissar et al., 2000; Corriveau et al., 2010).

The rapid auditory processing theory, for instance, suggests that more basic (not speech-specific) deficits in the perception of short or rapidly varying sounds underly phonological deficits in dyslexia (Tallal, 1980; Tallal et al., 1993). Referring to this theory, the deficit in correctly representing short sounds and fast transitions has an impact on speech processing, especially when short and rapidly varying acoustic sounds are elements of speech contrasts (Ramus, 2003). Accordingly, Tallal (1980) has shown in a two-tone identification task that dyslexic children perform worse than controls when the ISI is < 0.25 s, while no group differences are present with longer ISIs. In addition, a correlation between two-tone identification and the performance in non-word reading has been found in that study, again confirming that dyslexia is associated with temporal auditory processing impairments.

However, the overall evidence on low-level temporal auditory processing deficits in dyslexia is mixed. Some behavioral studies with non-speech stimuli indeed show deficits in temporal auditory processing in dyslexia (for a review see Farmer & Klein, 2003), whereas others show none (Schulte-Körne et al., 1998a; Bretherton & Holmes, 2003). Same contradicting results have been collected when underlying temporal auditory deficits have been investigated with speech stimuli (Reed, 1989; Rey et al., 2002; Vandermosten et al., 2010; but see, Ziegler et al., 2009; Nittrouer, 1999). Furthermore, studies investigating both auditory temporal and phonological processing deficits within one sample have shown that phonological impairments can occur in the absence of deficiencies in temporal auditory processing as well (e.g., Ramus et al., 2003; White et al., 2006).

Note, moreover, that it is not clarified whether the assumed low-level auditory processing deficit in dyslexia is limited to the processing of temporal information or spills over to the processing of spectral information. Correspondingly, evidence has been found not only for temporal, but also for spectral auditory processing deficits in dyslexics (e.g.,

Ahissar et al., 2000; Walker et al., 2006). Other studies, in turn, found neither temporal nor spectral auditory processing deficits in dyslexics (e.g., Hill et al., 1999).

However, one major methodological issue with these studies is the fact that stimulus and task complexity often differed between experimental conditions testing phonological processing and those testing underlying auditory processing abilities as, e.g., sinusoidal tones have been used as non-speech stimuli and temporal order judgements and gap detection as tasks to investigate temporal processing deficits and speech stimuli and non-word repetition or rapid automatized naming (RAN) as tasks to assess phonological processing deficits (Boets et al., 2007; Bretherton and Holmes, 2003; Nittrouer, 1999; Ramus et al., 2003; White et al., 2006). According to this, Banai (2006) has shown that dyslexics have no impairments in identifying mild frequency changes in simple tones and small phonemic alterations in complex speech sounds when they are presented in a simple same-different task. However, when the task is more complicated, e.g., requiring the identification of the direction of a frequency change, dyslexics perform worse. The author suggested that dyslexics might have a deficit, which refers to the complexity of the task given and not to the stimulus type.

Furthermore, in numerous studies auditory deficits have been found only in a subgroup of dyslexics, varying from very few individuals to 50 percent of the dyslexic sample under investigation (Tallal, 1980; Reed, 1989; Manis et al., 1997; Mody et al., 1997; Adlard & Hazan, 1998; Lorenzi et al., 2000; Marshall et al., 2001; Rosen and Manganari, 2001). Summing up, the role of auditory processing in the etiology of dyslexia has not been clarified yet.

Phonological or general auditory processing deficits in dyslexia, as captured with MMN

Neurophysiological studies were conducted to shed light upon the controversially discussed debate on the validity of the auditory processing deficit theory of dyslexia.

Interestingly, even the absence of impairments on a behavioral level does not indicate the absence of deviations on a neurophysiological level (Stoodley et al., 2006). Note that event-related potentials like MMN reflect basic sensory processes, while behavioral responses involve cognitive, i.e., attentional and decisional processes, as well. Hence, the results of neurophysiological and behavioral measures may differ.

Molfese (2000), for instance, presented synthetic speech stimuli to newborns (36 h after birth) and classified them based on three ERPs (N1, P2 and N2) as being at risk of dyslexia or not. At the age of eight years, it could be verified that 81 percent (N = 48) indeed had been classified correctly.

MMN is especially appropriate to examine the exact nature of auditory deficits in children with dyslexia, as it is an attention-independent measure not affected by motivation, which is especially crucial when studies are implemented with children (for more theoretical background, see Chapter 2). However, MMN evidence on auditory processing in dyslexia is mixed as well. Both, altered MMN responses in dyslexics solely for speech stimuli, but not for non-speech stimuli (e.g., Schulte-Körne et al., 1998; 2001; Meng, 2005) and solely for non-speech stimuli, but not for speech stimuli (e.g., Corbera et al., 2006) have been determined. Furthermore, there are studies showing deficits in both speech (phonological) and non-speech auditory processing in dyslexia (e.g., Csepe, 2000). Schulte-Körne and colleagues, for instance, found for both dyslexic children (1998) and dyslexic adults (2001) evidence in favor of a specific speech perception deficit. Here, sinusoidal tones, differing in frequency served as non-speech stimuli and the syllables /da/, /ba/ and /da/, /ga/ as speech stimuli. Corbera (2006), however, presented to participants with and without dyslexia natural phonemes (vowels), phonemes with an altered second formant and tones differing in frequency. No group differences have been found for speech stimuli, whereas diminished MMN amplitudes have been determined for spectral differences in non-speech stimuli in dyslexics.

Moreover, evidence on the auditory processing of single acoustic cues in dyslexia is ambiguous as well. Baldeweg and colleagues (1999), for instance, have not found any differences between dyslexic adults and age-matched controls for tones of different length, whereas smaller MMN amplitudes have been found in dyslexics for variations in tone height. Huttunen et al. (2008), in contrast, have found differences in children with dyslexia in comparison to controls (age 8 to 14 years) in the processing of tones of different length. Moreover, Corbera and colleagues (2006) have not determined MMN alterations for tones of different height, but lowered MMN amplitudes for tones of different duration in dyslexic children (age 11 to 13 years). Accordingly, Maurer and colleagues (2003) have not found group differences for preschool-children at risk of dyslexia and controls for frequency differences in tones.

Additionally, the hypothesis was investigated, that not temporal or spectral auditory processing per se, but the auditory processing of complex spectro-temporal alterations is impaired in dyslexia. This deficit is supposed to be not limited to phonetic information, but rather extended to stimulus-unspecific auditory information (e.g., Schulte-Körne et al., 1999; Kujala et al., 2000; Kujala, 2003). Kujala et al. (2000), for example, presented to adults with and without dyslexia four or two successive tones, whereas the onsets of single tones were varied. In the four-tone condition, the standard pattern consisted of intervals between tones of 200, 150 and 50 ms. In the deviant condition, the second and the third tone interval were exchanged. Consequently, standards and deviants differed on two dimensions. First, the onset of the third tone was earlier in the deviant as in the standard condition and second, the third tone was absent in comparison to its expected position (see Figure 10). In controls, two MMN reactions were found, whereas dyslexics showed only one MMN. The authors suggested that dyslexics, in contrast to controls, treated two events as one and hence, only one MMN was elicited. In addition, the MMN response, which was present in both groups, was found predominantly in the right

hemisphere in controls, but bilaterally localized in dyslexics. The two-tone condition served as control condition. Here, as expected, no group differences were found.
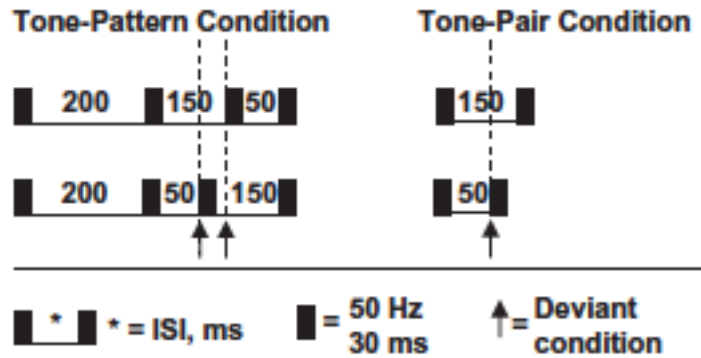
Figure 10: Stimuli used by Kujala et al. (2000). Arrows represent the onset of the deviant.

Furthermore, Lachmann et al. (2005) have found different patterns of auditory processing deficits in different subgroups of dyslexia. That study examined mismatch responses in 8 to 11 year old dyslexic children to syllable and tonal changes. Here, according to previous behavioral evidence outlined above (e.g., Tallal, 1980; Lorenzi et al., 2000; Marshall et al., 2001; Rosen and Manganari, 2001), MMNs to both tonal and syllable changes were decreased solely in a subgroup of dyslexics suffering from a deficit in frequent word reading, but not in non-word reading. This finding is particularly interesting, as children with deficits in non-word reading have greater deficits in phonological processing, especially in the assignment of graphemes to phonemes, compared to children with deficits in frequent word reading.

Developmental dyslexia and vowel length perception

As mentioned above, previous data indicates that children with dyslexia have difficulties in discriminating similar phonemes, e.g., "b" and "p" (Manis et al., 1997). However, just

few studies have investigated the association between dyslexia and the perception of phoneme duration. Richardson and colleagues (2003), for instance, have probed phoneme length perception in Finnish 6-month infants with and without risk of dyslexia. In Finnish, both consonants and vowels differ in length and are represented by double letters (e.g., tuli-tuuli-tulli). The participants were presented with the non-word stimulus /ata/. In half of the trials the middle consonant was lengthened. Intriguingly, in comparison to controls, children with risk of dyslexia showed a significantly weaker left-hemispheric activation when the modification in consonant duration occurred. In line with this, children with risk of dyslexia (and their parents as well) needed a longer voice onset time to perceive the consonant prolongation. Furthermore, Lehtonen and Bryant (2002) considered the impact of phoneme duration perception on spelling literacy, again in a Finnish children sample (1-3 and 7-9 years of age). The participants were given four words per trial, whereas one word differed from the others in sound duration (e.g., liima-kiire-piilo-tikku). Here, vowels and consonants varied in their length. Interestingly, the discrimination performance in this task correlated significantly with spelling ability.

As mentioned above, a distinction between vowels of different length is made in German as well (see Chapter 1). Moreover, correct spelling of vowel length is quite difficult, since several orthographic representations determine short and long vowels. In general, short vowels are marked by two following different or doubled consonants (e.g., Hand [/hant/, hand], Sommer [/zɔmɐ/,summer]). Only few short words, where short vowels are not marked, make an exception to this rule (e.g., in, im). Long vowels, however, are either not marked at all (e.g., Hof, Ofen) or marked by adding a silent h (e.g., Bahn [/baːn/, train], Mohn [/moːn/, poppy]). Alternatively, a long vowel can be represented by the diagraph "ie". In addition, for very few words the long vowel is indicated by a doubled vowel grapheme (e.g., Saal [/zaːl/, hall], Fee [/feːə/, fairy]). Thus, vowel length spelling is far from being consistent. That is to say, the child must first

determine the length of the vowel heard (within a word) and second, it has to choose between several orthographic representations. Some children not even manage it by eighth grade (Klicpera & Gasteiger-Klicpera, 1993). Hence, it is possible, that a deficit in correct perception of vowel length contributes to the difficulties that children with dyslexia have with spelling.

To date, only three studies indicate a relationship between dyslexia and vowel length perception in German. Landerl (2003), for instance, has shown in a behavioral study with a sample of 10-year-old children with and without spelling difficulties that vowel length perception deficits are indeed linked to spelling deficits. In contrast, Groth et al. (2011) found no increments in (natural) vowel length discrimination in adults with dyslexia. However, when one natural vowel was contrasted with an artificially prolonged or shortened one, so that the two contrasted vowels differed solely in their temporal features, while spectral features were kept constant, the discrimination performance of dyslexics decreased. This result indicates that dyslexics have no difficulties in vowel length discrimination per se, but rather in auditory temporal processing. Steinbrink and colleagues (2014) replicated the study of Groth and colleagues (2011) with 8 to 10 year old children with dyslexia (for more details, see Chapter 1). In that study an additional condition has been included where the difference between the stimuli was exclusively spectral in nature. Here, dyslexics showed lower performance in all conditions, which indicates deficits in phonological, i.e., vowel length, processing, as well as in the processing of spectral and temporal aspects of vowel length. Interestingly, this result confirms Landerl (2003) who has shown that dyslexics are impaired in discriminating German vowels differing in length. Note however that Groth et al. (2011), in contrast to Landerl (2003) and Steinbrink et al. (2014), have investigated adults (and not children) with dyslexia, which may account for the differing results between these studies.

Aims of Experiment 2

In previous studies with adults, stronger MMN reactions have been found for multiple-feature non-speech stimuli compared to single-feature non-speech stimuli (e.g., Schröger, 1995; Takegata et al., 1999; for more theoretical background on MMN additivity, see Chapter 2). Moreover, in Experiment 1, a study with adult participants, additive MMN effects were found for German vowels differing in length, i.e., for speech-sounds, as well. In German, vowels of different length differ in both their spectral and temporal features. Hence, by combining natural short or long vowels with their artificially shortened or lengthened counterparts, single features were isolated and vowels differing on both dimensions (spectral + temporal) could be contrasted with vowels differing on either the spectral or the temporal dimension. Interestingly however, there are no prior studies, neither with speech nor with non-speech stimuli, investigating additive MMN effects in children.

Moreover, previous evidence indicates that the maturation of the auditory system extends up to adult age (Ponton, Eggermont, Kwon & Don, 2000; Pang & Taylor, 2000; Shahin, Roberts & Trainor, 2004). Hence, it is plausible that the MMN, as an indicator of auditory processing, differs considerably between children and adults (e.g., in strength, latency or topography).

Thus, on the one hand, the current study addresses the research question, whether MMN responses in children of 9-10 years of age substantially differ from those in adults and on the other hand, whether MMN additivity for vowels, i.e., speech sounds, (irrespective of vowel type[5]) can be found in children as well as in adults.

Furthermore, the etiology of dyslexia is currently a highly debated topic (for more details, see above). Hence, MMN responses of dyslexic and control children and adults are contrasted with one another, on one side to figure out whether dyslexics suffer from a solely phonological (i.e., spectrotemporal), or rather from a more general auditory processing deficit (spectral and/or temporal) and on the other side to determine whether this deficit can or can not be regarded as a developmental lag, which could be reflected by a larger MMN amplitude and/ or a smaller MMN latency and/or an altered scalp topography in control children in comparison to children with dyslexia and, in turn, in adults in comparison to control children.

_____

[5] Note that in contrast to Experiment 1, no vowel-specific hypotheses were formulated, as the current study was conceptualized with the smallest possible number of trials (see Experiment 2) in single conditions to make the length of the experiment suitable for a children sample. Hence, no fine-grained analyses for single vowel types could be accomplished.

Experiment 2

Summary

Past research indicates that event related potentials (ERPs) in children might greatly differ from those in adults, making it uncertain whether underlying processes or components are analogous and therefore comparable. In addition, empirical evidence suggests a relationship between vowel length perception and dyslexia (e.g., Landerl, 2003). Thus, Experiment 2 investigated by means of mismatch negativity (MMN) how vowel length perception differs between dyslexic and control children of 9-10 years and adults. More precisely, on one hand, the present study aimed to contrast the processing of multiple-feature speech stimuli with the processing of single-feature speech stimuli in children with and/or without dyslexia and adults. On the other hand, it investigated whether the auditory processing deficit in dyslexia has a temporal, spectral or a spectrotemporal (i.e., phonological) character (see Chapter 4).

First, the results of Experiment 2 could be replicated, i.e., a larger MMN for vowels differing in both spectral and temporal features compared to vowels differing either in spectral or in temporal features was found in adults. However, unlike in adults, no MMN additivity could be detected in children (with and without dyslexia). Instead, two separate MMN components were determined in the phonological, i.e., spectrotemporal, condition corresponding temporally and topographically to the (one-peaked) MMNs in the temporal and the spectral condition. This result indicates that, while adults integrate spectral and temporal aspects of the speech signal in an additive way, children of 9-10 years of age sequentially process both features. Moreover, no indices were found for an auditory processing deficit in dyslexia, neither for a phonological nor for a temporal and/or a spectral processing deficit (for possible explanations, see below).

Methods

Participants

15 healthy young adults (mean age 23.5, 8 females), 15 typically developing children (mean age 9.33, 6 females) and 15 children with developmental dyslexia (mean age 9.28, 7 females) took part in the study. All participants had normal sense of hearing and were native speakers of German. Written and informed consent was obtained from all adult participants and from all parents of the children participating in the study. Different standardized metrics were used to classify children as dyslexics or controls. To get assigned into the dyslexics group an IQ $\geq$ 85 was required (assessed with CFT (Culture Fair Intelligence Test; German version, Weiß, 1997). In addition, reading literacy, i.e., reading and nonword-reading ability, was determined with SLRT-II (Salzburger Lese-Rechtschreibtest; Moll & Landerl, 2010) and spelling literacy with WRT-3+ (Weingartener Grundwortschatz Rechtschreibtest für dritte und vierte Klassen; Birkel, 2007). Phonological processing was addressed via RAN (Rapid Automatized Naming; Denckla et al., 1974), AGTB 5-12 (Working Memory Test Battery; subtests digit span forward and digit span backward; Hasselhorn et al., 2011) and a phonological awareness test battery including measures of vowel, consonant and initial sound discrimination (Klatte et al., in press). A percentage range of $\leq$ 16 was required in one of the subtests of SLRT-II to get assigned into the dyslexics group. However, an IQ $\geq$ 85 and a percentage range of > 16 in all tests and subtests on reading and spelling literacy were necessary to get assigned into the group of control children.

Table 4: Descriptive characteristics and corresponding statistical comparisons between groups investigated in Experiment 2. Note that for absolute numbers of errors no test statistics could be computed, as in both groups not enough errors were made.

Contents	Dyslexics		Controls		Statistical comparison		
	Mean	SD	Mean	SD	t-value	df	p-value
WRT3+ (raw value)	5.8	3.95	12.8	2.704	-5.664	28	< .001
SLRT_II word (raw value)	40	9.266	70.6	15.642	-6.519	28	< .001
SLRT_II non-word (raw value)	26.067	5.338	42.8	8.76	-6.317	28	< .001
RAN objects (absolute number of errors)	0.133	0.352	0.2	0.414			
RAN objects (time in s)	54.16	15.381	44.913	7.233	2.107	19.902	<.05
RAN colours (absolute number of errors)	0.267	0.458	0.2	0.561			
RAN colours (time in s)	56.62	22.266	44.06	11.786	1.931	21.274	.067
AGTB (digit span forward)	3.925	1.166	4.8	0.72	-2.472	23.334	< .05
AGTB (digit span backward)	3.208	0.474	3.541	0.691	-1.540	28	.135
Vowel discrimination (absolute number of errors)	1.333	0.976	1.333	1.589			
Initial sound discrimination (absolute number of errors)	8.8	3.052	5.067	3.411	3.159	28	< .01
Consonant discrimination (absolute number of errors)	6.4	2.586	4.400	3.602	1.747	28	.092
IQ	109	11.171	110.91	14.611	-.344	24	.734

Stimuli

Comparable to Experiment 1, two German natural vowels, i.e., 'i' and 'a' each in two different lengths, were used as stimuli. In addition, a phonetics software package (Praat; Boersma & Weenink, 2005) was applied to construct two types of artificial stimuli, i.e., (1) long vowels that were artificially shortened to the length of their short natural counterparts (i.e., /a:/ shortened to 75ms and /i:/ shortened to 51ms), and (2) short vowels that were lengthened to the length of their long counterparts (i.e., /a/ lengthened to 145ms

and /i/ lengthened to 93ms). Experimental conditions (see next section) resulted from different combinations of these natural and artificial vowels as standards and deviants in the MMN paradigm.

Conditions

As in Experiment 1, three conditions have been implemented: a phonological (i.e. spectrotemporal), temporal and a spectral condition. In the phonological condition, natural vowels of different length were presented, whereas in the temporal and the spectral condition, one natural and one artificially modified stimulus were combined in a way, so that the difference between the stimuli was solely temporal or spectral in nature (for more details, see Figure 4).

Experimental procedures

Equivalent to Experiment 1, the stimuli were presented in a passive auditory oddball paradigm, involving the presentation of a large number of standard stimuli (82.22%) interspersed with a small number of deviant stimuli (17.78%). The experiment was conducted in an acoustically and electrically shielded room. During stimulus presentation, participants were watching a self-selected silent movie. The movie was not subtitled. Participants were explicitly instructed not to pay attention to the stimuli, which were delivered binaurally via headphones (Etymotic ER-2; Etymotic Research, Inc, Elk Grove Village, Illinois) at a fixed intensity of 67 dB SPL, controlled via Presentation software (Neurobehavioral Systems Inc., San Francisco, California). All possible combinations of natural and artificial stimuli were used in the present study, with each stimulus serving as standard and deviant in different blocks (for more details, see Table 3). However, to make the study appropriate for children only a subset of stimuli of Experiment 1 was presented

to each participant in Experiment 2. Hence, two different randomizations were applied, each including 12 different combinations of stimuli (4 combinations per condition). Note that differences between combinations divided into different randomizations were not expected. The experiment was organized into three experimental runs, each lasting approximately 12 minutes. When necessary, small breaks of up to five minutes could be made between separate runs. Additionally, after two runs the children subgroup of the study was given a large break of 30 minutes. Each run consisted of four stimulus blocks and each of these blocks was composed of 280 standard trials and 70 deviant trials. The sequence of stimuli within each block was pseudorandomized [controlling for the minimum (i.e., 3) and maximum (i.e., 5) interval between two successive deviants]. The ISI was kept constant at 350 ms, whereas the last 100 ms served as baseline for the following stimulus. To establish an initial memory trace, 10 additional standard stimuli were presented at the beginning of each block. Moreover, for comparison reasons the same experimental procedure was applied for children and adults.

EEG recording and analysis

The same procedure as in Experiment 1 was applied. Electroencephalogram (EEG) was continuously recorded using the 128 channel Geodesic Sensor Net System (EGI; Electrical Geodesics, Inc., Eugene, Oregon), the amplifier EGI NET Amps 300 and a sampling rate of 1,000 Hz (for electrode positions, see Figure 5). Impedances were kept below 50 kΩ and the reference electrode was placed at the vertex (corresponding to Cz in the international 10/20 system). Virtual EOG channels (both vertical and horizontal) were calculated from channels E8, E14, E21, E25, E125, E126, E127 and E128. The data were preprocessed with FASTER (Nolan et al., 2010), a fully automatized toolbox implemented in EEGlab (Swatz Center of Computational Neuroscience, San Diego, CA). Here, the data were initially down sampled to 250 Hz, re-referenced to an average

reference and filtered with a lowpass filter of 30 Hz and a highpass filter of 1 Hz. Later on, ocular, movement and amplifier artifacts were rejected automatically by means of thresholding the z-transformed value of the preprocessed data (the cut-off z-value was z = 2) based on ICA (independent component analysis). To calculate event-related brain potentials (ERPs) the data were cut into trials that lasted from 100 ms before to 300 ms after stimulus onset.

Data analysis was conducted with the open-source software Fieldtrip (http://fieldtrip.fcdonders.nl), a toolbox implemented in Matlab (The Mathworks Inc., Natick, Massachusettes), ERP-lab (UC-Davis Center for Mind & Brain, Davis, California), a toolbox implemented in EEGlab, which again is implemented in Matlab and IBM SPSS statistics 22 (International Business Machines Corporation, Armonk, New York). ERPs were calculated by averaging trials from 0 to 300 ms after stimulus onset, using a pre-stimulus baseline of 100 ms. MMN was computed by subtracting average ERPs to standard stimuli from those to deviant stimuli. Condition-specific MMNs were derived by combining ERPs from different blocks. Importantly, the ERPs of different blocks were averaged across conditions irrespective of the function of each stimulus as standard or deviant within these conditions, as no a-priori hypotheses were established dependent on whether or not a stimulus served as standard or deviant.

In children with and without dyslexia, an early central MMN peak was found in the spectral and a late frontal MMN peak in the temporal condition. Interestingly, both peaks could be observed in the phonological condition as well. Additionally, visual inspection revealed a frontal positive mismatch reaction accompanied by a posterior negativity prior to the typical MMN in the phonological condition. This pattern differed from the one observed in adults. Here, visual inspection revealed one MMN peak in each condition (see Figure 12 and Figure 13).

Cluster-based permutation tests including all EEG channels were used (in children and adults) to probe the significance of the difference between standards and deviants in all conditions, i.e., to figure out whether reliable MMNs were present in the data (time windows used for the analyses: 140-200ms (early) and 200-260 ms (late)). Note that outer electrodes (involving electrodes that were positioned at non-scalp sites) were excluded from cluster-based analyses and data visualization because of noisiness.

To further investigate the association between the two MMN peaks (early central + late frontal) in the phonological condition and the single MMNs in the temporal (late frontal) and the spectral condition (early central) in children, a repeated-measures ANOVA with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological cs. temporal vs. spectral), time (early: 140-200ms vs. late: 200-260 ms) and group (dyslexics vs. controls) based on mean amplitudes was calculated. Additional Wilcoxon tests were computed to resolve the effects. Wilcoxon tests were used instead of t-tests because MMN amplitudes were not normally distributed in several conditions, in both children groups.[6] In addition, to further elucidate the association between MMN measures and reading and/or spelling ability and/or measures of phonological processing, Spearman's rho correlations were calculated between these measures, for all conditions separately.

Second, to compare MMN responses between children and adults and to avoid different sample sizes (children N = 30 vs. adults N = 15) an additional separate ANOVA for adults with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological cs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms) was calculated. Wilcoxon tests were attached (for the same reason as in children). The

——

[6] Note that Kolmogorov-Smirnov-tests were used to test for normal distribution.

central-posterior ROI encompassed the electrodes E55, E54, E53, E61, E86, E79 and E78 and the frontal ROI the electrodes E15, E16, E11, E18, E10, E19 and E4 (for an illustration of electrodes included in both ROIs, see Figure 11).
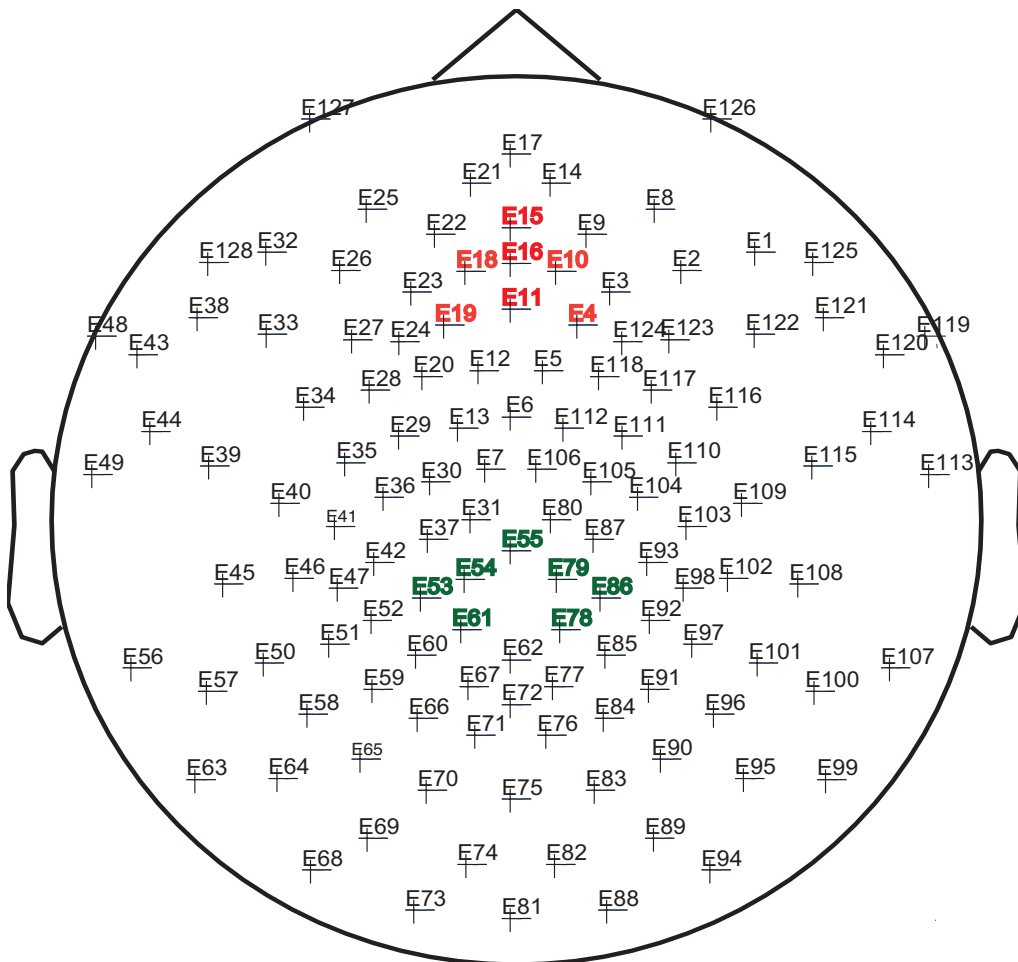
Figure 11: Electrodes included in ROIs of Experiment 2. Red electrodes illustrate the frontal ROI and green electrodes the central-posterior ROI.

A MMN topographies in all children (amplitude)

Phonological condition

Temporal condition

Spectral condition

B MMN topographies in adults (amplitude)

Phonological condition
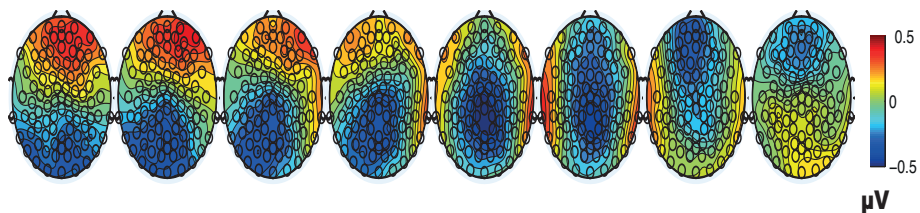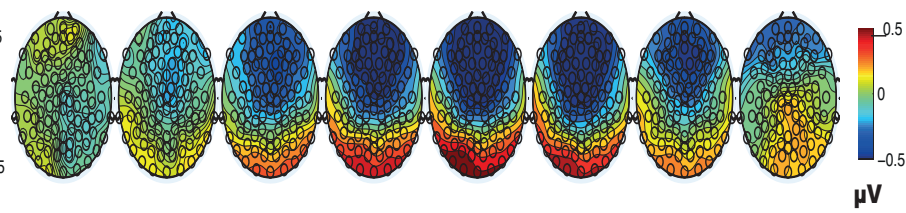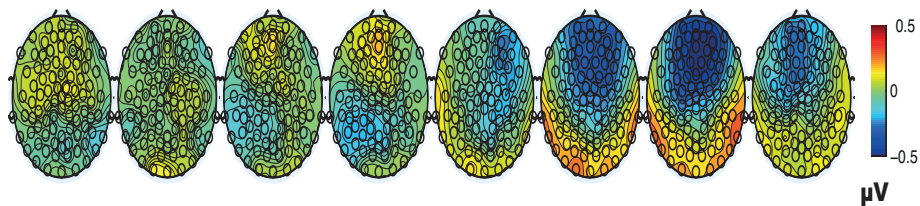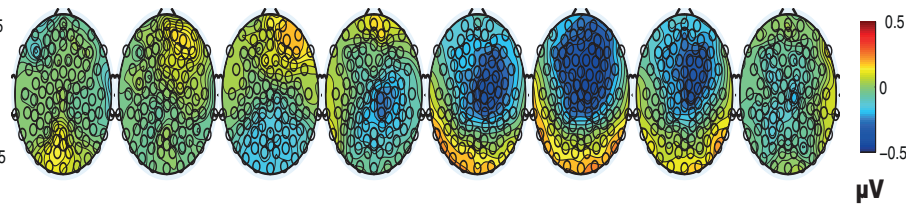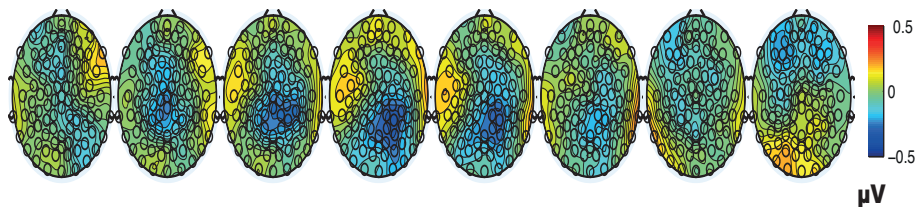
Temporal condition

Spectral condition
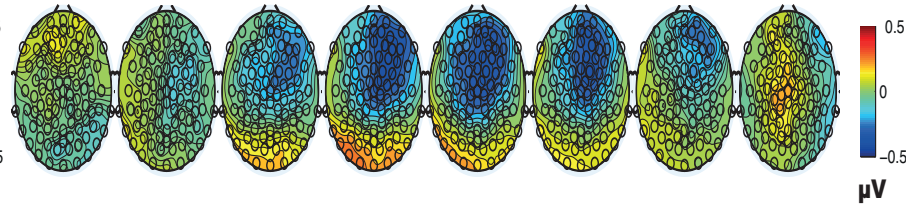
100 ms 120 ms 140 ms 160 ms 180 ms 200 ms 220 ms 240 ms

Figure 12: Time course of MMN topographies in all conditions. A. All children (N = 30). B. Adults (N = 15).

A Mismatch response over all electrodes (all children)

Phonological condition

Temporal condition

Spectral condition

B Mismatch response over all electrodes (adults)

Phonological condition

Temporal condition

Spectral condition
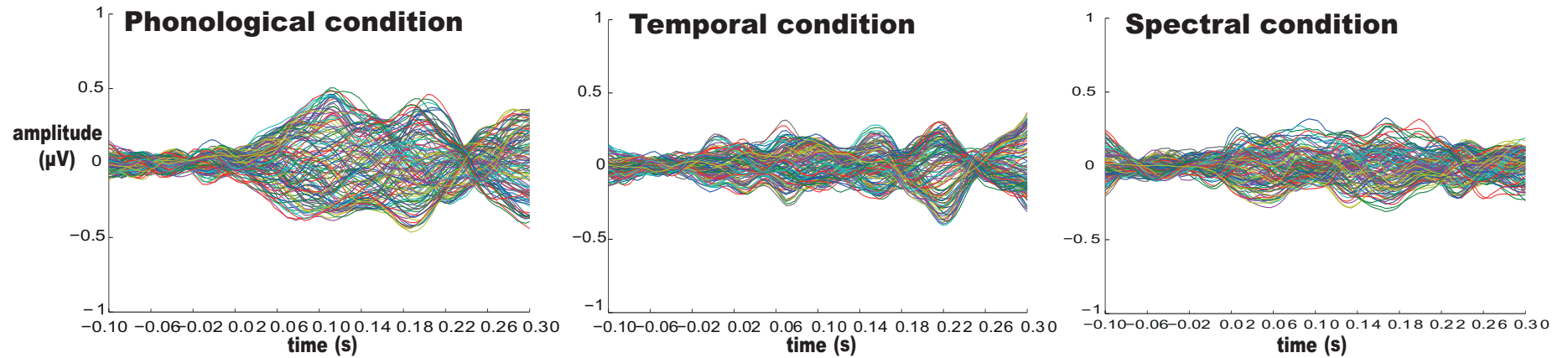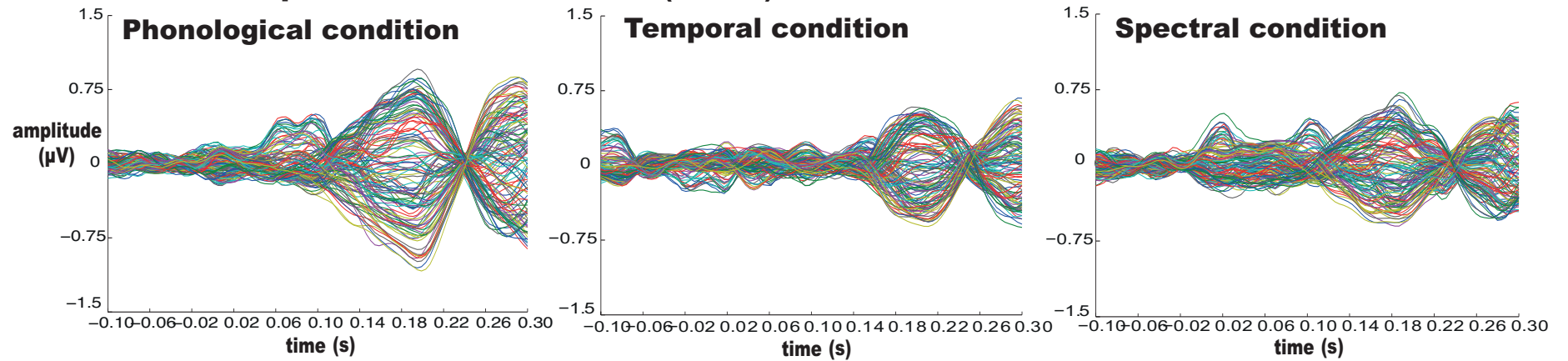
Figure 13: Mismatch responses over all electrodes in all conditions displayed in a butterfly plot for A. all children (N = 30) and B. adults (N = 15).

Results

First, cluster-based permutation tests were calculated to compare deviants and standards in all conditions and groups (for more details, see above). Equivalent to Experiment 1, for adults, reliable (one-peaked) MMNs were found in all conditions (time window of analysis 160-220 ms) (see Figure 14C).

For control and dyslexic children, however, an early central MMN peak (time window of analysis 140-200 ms) and a late frontal MMN peak (time window of analysis 200-260 ms) were detected in the phonological condition. Moreover, for the dyslexic group, but not for the control group, a late frontal MMN peak (time window of analysis 200 -260 ms; t-values are depicted in Figure 14A and B) was determined in the temporal condition. For both children groups considered separately, the difference between standards and deviants did not reach significance in the spectral condition (time window of analysis 140 -200 ms) (see Figure 14A and B). However, calculated over all children irrespective of group membership, a reliable early central spectral MMN (time window of analysis 140-200 ms) and a reliable late frontal temporal MMN (time window of analysis 200 – 260 ms) were found (see

Figure 15). That is to say, when all children were analyzed group-independently, both the early central and the late frontal MMN peak observed in the phonological condition were found in the spectral and the temporal condition as well.

In addition, in children with and without dyslexia, the statistical analysis revealed an early positive mismatch response with a posterior negativity (time window of analysis 80-140 ms) in the phonological condition (see Figure 14A and B).[7]
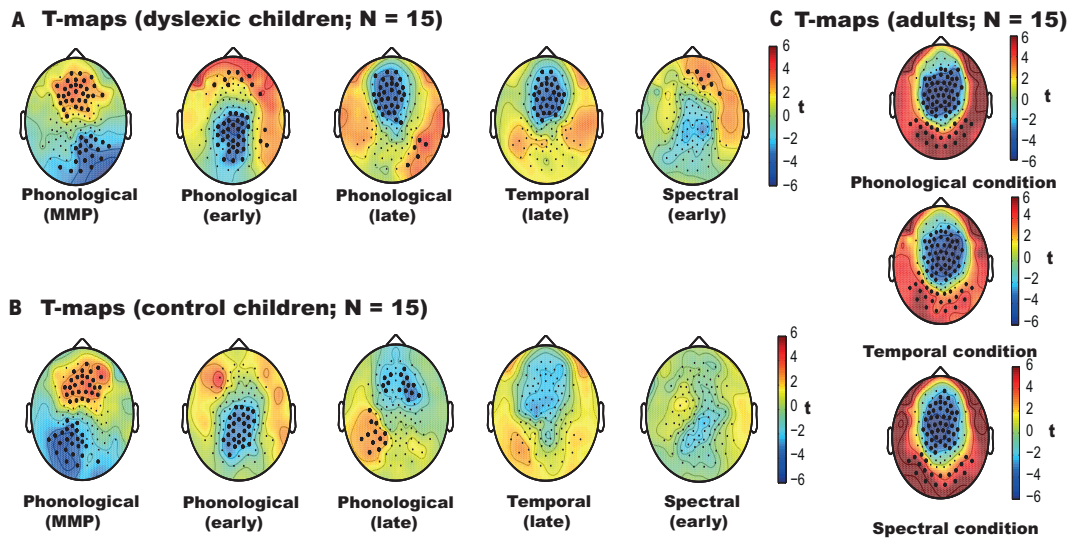
Figure 14: Distribution of t-values for statistical comparisons between standards and deviants in all conditions (early time window: 140-200 ms and late time window: 200-260 ms) for A. dyslexic children, B. control children and C. adults (MMP = Mismatch Positivity). The bold dots indicate significant electrodes (p < 0.05). Outer electrodes were removed for illustration reasons.

_____

[7] In contrast to Experiment 1, MMN over all conditions is not calculated in this section, as topographical differences in single condition-specific MMNs were found in Experiment 2.

T-maps

All children (N = 30)

Temporal condition   **Spectral condition**
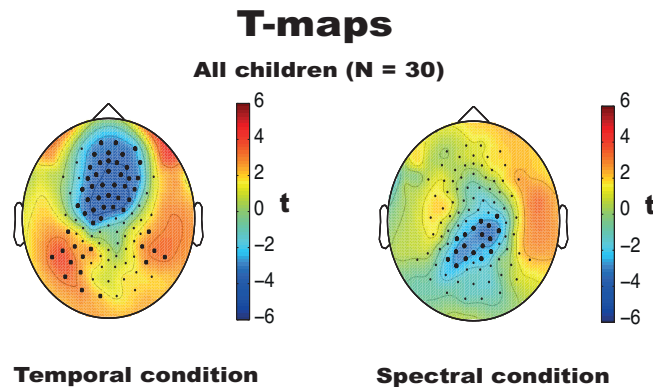
Figure 15: Distribution of t-values for statistical comparisons between standards and deviants in the spectral and the temporal condition, calculated over all children (N = 30) (the early time window (140-200 ms) was applied for the spectral condition and the late time window (200-260 ms) for the temporal condition). The bold dots indicate significant electrodes (p < 0.05). Outer electrodes were removed for illustration reasons.

Condition-specific MMN effects in children

First, a repeated-measures ANOVA with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological cs. temporal vs. spectral), time (early: 140-200ms vs. late: 200-260 ms) and group (dyslexics vs. controls) was calculated. A significant ROI*time*condition interaction was found ($F_{(2,27)} = 3,536$; $p = 0.043$; see Figure 16A). Interestingly, the ANOVA revealed no significant group effects (for more details, see Table 6). Correspondingly, no significant correlations between behavioral measures of reading and/or spelling literacy and/or phonological processing and MMN were found, confirming the lack of MMN differences between children with and without dyslexia. Due to the absence of group effects the significant ROI*time*interaction was interpreted group-independently.

To resolve the threefold interaction, subsequent Wilcoxon tests were computed for each condition separately. First, ROIs were compared for the early and the late time window. In the spectral condition an MMN was found in the early ($z = 2.191$; $p = 0.028$),

but not in the late (z = - 0.524; p = 0.6) time window, while in the temporal condition no early (z = 0.915; p = 0.36), but a late MMN (z = -2.684; p < 0.01) was determined. In the phonological condition descriptively the same pattern was found, whereas significance was only reached in the early time window (z = 3.178; p < 0.01). Furthermore, the early and late time windows were compared for the frontal and the central ROI. A central (z = - 2.458; p = 0.014), but not a frontal MMN (z = - 0.792; p = 0.428) was found in the spectral condition and a frontal (z = -3.096; p < 0.01), but not a central MMN (z = -1.244; p = 0.213) in the temporal condition. In the phonological condition, two MMN components were found, a central (z = 4.103; p < 0.00) and a frontal one (z = -3.198; p < 0.01). Hence, the results indicate two MMN components in the phonological condition, an early central and a late frontal one, and one early central MMN in the spectral and one late frontal MMN in the temporal condition.

Table 5: Results of the repeated-measures ANOVA for children with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological vs. temporal vs. spectral), time (early: 140-200ms vs. late: 200-260 ms) and group (dyslexics vs. controls).

Effect	F-value	df(factor)	df(error)	Significance (p-value)
Time	0.277	1	28	.603
Time*group	0.288	1	28	.596
ROI	0.945	1	28	.339
ROI*group	0.314	1	28	.580
Condition	2.273	2	27	.122
Condition*group	0.479	2	27	.625
Time*ROI	26.725	1	28	< .001
Time*ROI*group	3.097	1	28	.089
Time*condition	3.046	2	27	.064
Time*condition*group	0.418	2	27	.663
ROI*condition	3.502	2	27	.044
ROI*condition*group	0.578	2	27	.568
Time*ROI*condition	3.536	2	27	.043
Time*ROI*condtion*group	0.524	2	27	.598
group	0.278	1	28	.602

Furthermore, as outlined above, a larger MMN amplitude was expected in the phonological condition in comparison to the spectral and the temporal condition. As no hypotheses had been made referring to the temporal or the topographic distribution of MMNs in single conditions, the condition main effect (and not subsequent condition interactions) of the ANOVA was focus of interest to investigate this issue. However, the condition main effect did not reach significance for children ($F_{(2,27)} = 2.273$; $p = 0.122$),

indicating that the MMN in the phonological condition did not differ from MMNs in the spectral and the temporal condition.

Condition-specific MMN effects in adults

For comparison reasons, same analyses were calculated for children and adults. First, a repeated-measures ANOVA with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological vs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms) was computed for adults. Here, the threefold interaction ROI*time*condition did not reach significance ($F_{(2,13)} = 0.889$; $p = 0.435$; see Figure 16B). However, since on the one hand, the ROI*condition and the time*condition interaction became significant and on the other hand, it was aimed to parallelize the analysis steps for children and adults, it was legitimated to calculate identical Wilcoxon tests for adults, as it had been done in the prior section for children. The early and late time windows were compared for the frontal (phonological condition: $z = 2.726$; $p < 0.01$; temporal condition: $z =- 0.170$; $p = 0.865$; spectral condition: $z = 2.215$; $p = 0.027$) and the central ROI (phonological condition: $z = -1.988$; $p = 0.047$; temporal condition: $z =-1.817$; $p = 0.069$; spectral condition: $z =- 2.897$; $p < 0.01$) and the frontal and central ROIs were compared for the early (phonological condition: $z = - 2.272$; $p = 0.023$; temporal condition: $z =0.795$; $p = 0.427$; spectral condition: $z =-1.761$; $p = 0.078$) and the late time window (phonological condition: $z =-2.542$; $p = 0.011$; temporal condition: $z =- 0.057$; $p = 0.955$; spectral condition: $z =- 1.306$ $p = 0.191$). The results indicate one early frontal MMN in the phonological condition, one frontocentral MMN in the temporal condition positioned among the early and the late time window and one early frontocentral MMN in the spectral condition.

Furthermore, the analysis revealed a significant condition main effect ($F_{(2,13)} = 6.999$; $p < 0.01$). To resolve this condition main effect, three supplementary ANOVAs

---

with factors ROI (frontal vs. central-posterior), time (early vs. late) and condition (phonological vs. temporal and phonological vs. spectral and temporal vs. spectral) were calculated to compare the MMNs between the phonological and the spectral (condition main effect: $F(1,14) = 13.594$; $p < 0.01$), the phonological and the temporal ($F(1,14) = 5.933$; $p = 0.029$) and the temporal and the spectral conditions (condition main effect: $F(1,14) = 0.005$, $p = 0.944$). Note that ANOVAs instead of t-tests were calculated to resolve the main effect condition by keeping other factors, i.e., ROI and time, constant. A larger MMN in the phonological condition in comparison to the temporal and spectral conditions was found in adults. That is to say, the present study could replicate the findings of Experiment 1.

Table 6: Results of the repeated-measures ANOVA for adults with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological vs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms).

Effect	F-value	df(factor)	df(error)	p-value
Time	20.316	1	14	< .001
ROI	4.360	1	14	.056
Condition	5.933	1	14	.029
Time*ROI	0.003	1	14	.955
Time*condition	22.830	1	14	< .001
ROI*condition	8.735	1	14	.01
Time*ROI*condition	1.689	1	14	.215
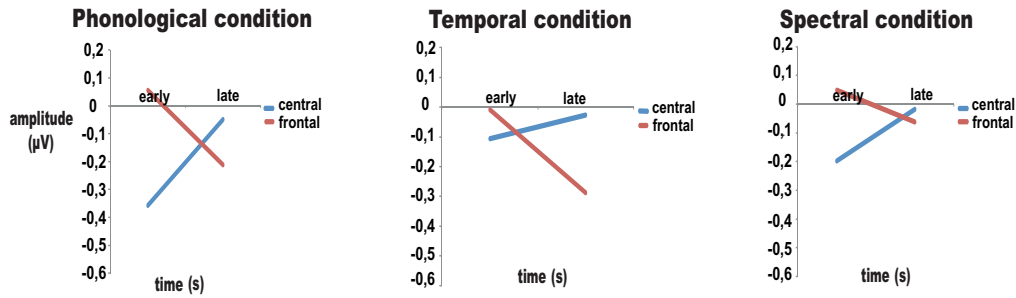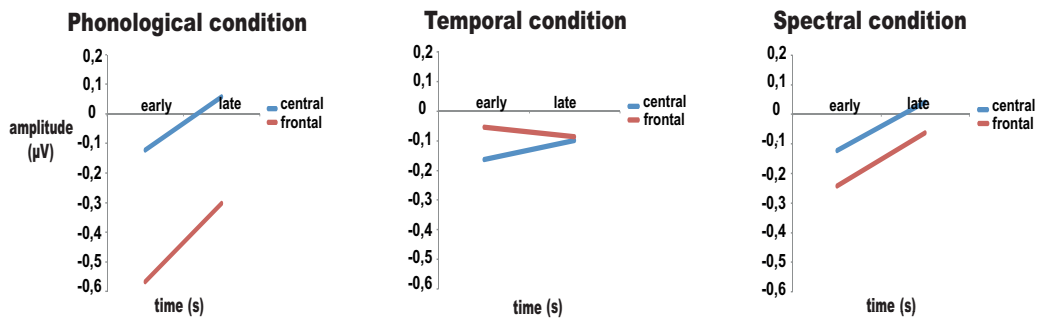
Figure 16: Illustration of the ANOVA interaction ROI*time*condition, which reached significance only in the children's sample. A. All children (N = 30) and B. adults (N = 15)[8].

Group differences in MMN amplitude and MMN latency

Since topographical differences (i.e., frontal and frontocentral MMNs in adults and frontal and central-posterior MMN responses in children) and differences in the number of relevant components were found in children and adults (i.e., two MMN components in the phonological condition in children, but only one in adults), indicating incomparable processes underlying the mismatch responses in these age-groups, condition-specific group comparisons could not be computed for MMN amplitudes and MMN latencies.

_____

[8] Note that the calculation of error bars (e.g., standard errors) was not legitimized in Figure 16, as the illustrated data were not normally distributed.

Discussion

The present experiment was implemented to compare MMN responses of children (with and without dyslexia) and adults. As expected, crucial differences were found. The main finding is that (equivalent to Experiment 1) a larger MMN in the phonological, i.e., spectrotemporal, condition in comparison to the spectral and temporal conditions was found in adults, whereas no MMN additivity could be determined in children. Here, two MMN components instead of one (determined in adults) were found in the phonological, i.e., spectrotemporal, condition and one MMN, respectively, in the temporal and the spectral condition. Since these two MMN components in the phonological condition corresponded topographically and temporally to the MMNs in the temporal and the spectral condition, the conclusion is drawn that they reflect the sequential processing of spectral and temporal information. That is to say, while adults process single features of the speech signal in an additive way, children process them sequentially. In the following sections, the reader will be supplied with more details.

Phonological, i.e., spectrotemporal, temporal and spectral MMNs in children and adults

In adults, reliable MMNs were found in all conditions, i.e., in the phonological (spectrotemporal), the temporal and the spectral condition, indicating that the presence of single acustic cues is sufficient for correct vowel length perception. Additionally, adult's MMN was significantly larger in the phonological condition than in the spectral and temporal conditions, confirming the additive MMN effect found in Experiment 1. In addition to Experiment 1, however, the current study aimed to investigate condition-specific MMNs in children and adults not only in terms of amplitude, but also in terms of their temporal and topographical distribution. Hence, a slightly different topographical

MMN distribution in the phonological in comparison to the temporal and the spectral condition, i.e., frontal vs. frontocentral, was detected in adults. It is conceivable that these condition-specific differences in MMN topography indicate differences between speech-specific and general auditory processing.

In children, in turn, reliable MMNs were found in all conditions only when taking both children groups (dyslexics and controls) together. However, when considering children groups separately reliable MMN reactions were determined exclusively in the phonological condition (and in the dyslexic's sample in the temporal condition). It is possible that greater statistical power is needed to find reliable MMNs for spectral and temporal differences in both children groups calculated separately. That is to say, MMNs to single-feature deviants seem to be less robust in children of 9-10 years in contrast to adults. Alternatively, the weak temporal and spectral MMNs in children (but not in adults) might be due to the artificial character of the stimuli used in the current study. Artificially manipulated vowels are potentially easier to process for a mature auditory system, as maturation of the auditory system may be accompanied by greater flexibility in dealing with unusual stimulus types.

Complex multi-peaked phonological MMN instead of MMN additivity in children

First, an early frontal positive mismatch reaction accompanied by a posterior negativity was determined in the phonological condition prior to the typical negative mismatch response, in children. Several previous studies have shown positive deflections instead of negative ones in studies with infants or small children using speech and non-speech stimuli (Dehaene-Lambertz, 2000, Morr et al., 2002), whereas speech contrasts seem to foster the appearance of a positive mismatch response. This fits with the results of the

current study, where a positive mismatch response was detected exclusively in the phonological condition, where natural speech contrasts were used.

However, by using a very large deviance, Morr and colleagues (2002) reported a typical MMN in nearly all of the children. The authors hypothesized that the mismatch positivity could mask the MMN in infants and this masking effect could be varying as a function of deviation between standards and deviants and/or as a function of children's age. To some extent this account may have an impact on the results of the present study as well. The stimuli of the present study were isolated natural or artificially modified vowels. Hence, they may have been difficult for children to process, i.e., the deviation may have been too small, which again may have caused the additional positivity prior to the typical negative mismatch response in children.

In contrast to studies cited above, Maurer and colleagues (2003) investigated older children between 6-7 years and compared them to adults. Maurer et al. (2003) found a typical frontocentral MMN in adults, but a frontal positive mismatch response with posterior negativity in children. In that study, frequency (pure tones) and phoneme stimuli were used. Additionally, a short SOA (onset to onset interval) of 383 ms was applied, as it was demonstrated in a study with 7-9 year old children using a frequency paradigm that event related potentials (ERP) to standards are affected by the choice of SOA (Ceponiene et al., 1998). With shorter SOAs, children show an ERP sequence typical for children (Sharma et al., 1997; Albrecht et al., 2000; Pang and Taylor, 2000; Ponton et al., 2000), whereas with longer SOAs additional components appear, more typical for adults. Note that the SOAs of the current study were varying from 401 to 495 ms, i.e., were quite short as well, which may have had an impact on the emergence of large differences between MMNs in children and adults. Interestingly, all previous studies with children of comparable age used longer SOAs (> 450 ms) and larger deviances (Cheour et al., 1997; Holopainen et al., 1997; Gomes et al., 1999; Kraus et al.,

1999; Gomot et al., 2000; Shafer et al., 2000; Korpilahti et al., 2001; Cheour et al., 2002). Furthermore, an inspection of the MMN figures in these studies revealed additional positive MMN deflections that were not discussed (Holopainen et al., 1997; Gomes et al., 1999; Gomot et al., 2000; Shafer et al., 2000; Korpilahti et al., 2001; Cheour et al., 2002). Maurer et al. (2003) assume that a typical MMN replaces the positive mismatch reaction in the course of development. Interestingly, the results of the present study, where both a positive and a negative mismatch response were found, confirm this assumption by indicating that it is more likely that the positive mismatch response gets replaced or overlapped by the typical MMN with maturation than that the positive mismatch response undergoes a polarity reversal. Note that it is conceivable that the present study captures a characteristic age window where both components, a positive and a negative mismatch response, are still present.

However, to draw final conclusions e.g., to exclude that the positive component found in the present study in children not simply reflects a certain stimulus property, as for instance the speech character of the stimuli, additional experiments with younger children (than 9-10 years of age) are needed (showing a positive mismatch response instead of a typical MMN). In addition, it has to be kept in mind that in the study of Maurer and colleagues (2003) the positive mismatch response was found in the time window of 179 - 207 ms after stimulus onset, which clearly deviates from the time window in the current study where the mismatch positivity was found (80 - 140 ms). Furthermore, the children in the present study were far older (9-10 years) compared to those in the study of Maurer and colleagues (6-7 years). This is also true for other studies cited above reporting a positive mismatch response in children or infants (e.g., Dehaene-Lambertz, 2000, Morr et al., 2002). All of them found positive mismatch reactions in later time windows and used younger children as participants than it was done in the present study. In addition, to the best of knowledge, there are no previous studies

showing a negative and a positive mismatch response side by side. Hence, the results have to be interpreted carefully in favor of a positive mismatch response, as alternative interpretations such as the option that we have to deal with a moving mismatch negativity dipole (from posterior to central) may be more appropriate. Note however that this particular interpretation is questionable, as the butterfly plot in Figure 13A shows a clear additional component in the time window of 80-140 ms in the phonological condition.

In addition to the positive mismatch reaction in the phonological condition, typical negative mismatch responses were found in children in all conditions. Interestingly, no MMN additivity, i.e., a larger MMN amplitude in the phonological condition in comparison to the temporal and spectral conditions, could be detected. Instead, a double-peaked MMN was determined in the phonological condition in contrast to the temporal and the spectral condition, where only single MMN peaks were found. The MMN in the spectral condition corresponded temporally and topographically to the early central MMN peak in the phonological condition and the MMN in the temporal condition to the late frontal MMN peak in the phonological condition, suggesting that the two negative mismatch components in the phonological condition reflect the sequential processing of temporal and spectral features in natural vowels of different length.

All in all, the results of the current study suggest differing speech-processing mechanisms in children and adults. The results of Experiment 1 and 2 combined indicate that adults integrate spectral and temporal aspects of the speech signal in an additive way after weighting these features differently according to their saliency for a certain phoneme (for more details, see Experiment 1), whereas children of 9-10 years of age sequentially process them.

It can be hypothesized that the additive MMN effect in adults is due to the activation of long-term representations, which may be not sufficiently differentiated in children of 9-10 years. In line with this assumption, several studies have shown that linguistic

experience and accordingly the activation of long-term representations results in an enlarged MMN amplitude in adults (e.g., Dehaene-Laembertz, 1997; Näätänen et al., 1997; Alho et al., 1998b; Phillips et al., 2000). However, previous studies have shown larger MMNs for speech stimuli that are prototypical in one's native language, i.e., that are accompanied by the activation of long-term representations, compared to non-native speech stimuli, in children as well (Cheour et al., 1998; Dehaene-Lambertz & Baillet, 1998). Dehaene-Lambertz and Baillet (1998), for instance, presented syllables belonging to the same or to different phoneme categories in French to infants. The acoustical difference between standards and deviants was kept constant in both conditions. Interestingly, infants at the age of 3 months displayed larger MMN responses to syllables belonging to different phoneme categories compared to syllables belonging to same phoneme categories. Moreover, Cheour and colleagues (1998) found larger MMN responses for vowel prototypes compared to non-prototypes in Finnish infants at the age of 1 year. These findings are in contrast to the result of the current experiment, showing no increase in MMN amplitude for natural speech contrasts (represented in long-term memory) in comparison to artificially modified speech contrasts (not represented in long-term memory), in children. Hence, they indicate that the activation of long-term representations in the phonological condition can hardly be the reason for MMN additivity found in adults, but not in children, as in such a case MMN additivity should be detectable in both children and adults. Note however that these findings cannot exclude the option that long-term representations are less differentiated in children than in adults, as no direct comparisons between children and adults have been implemented. Furthermore, it is conceivable that long-term representations of syllables and Finnish vowels boost the MMN of children and adults, while they are not sufficiently differentiated in children for isolated German vowels. Hence, further evidence is needed to draw final conclusions.

Another possible explanation for differing MMN additivity effects in children and adults may be connected to the temporal interval separating the two deviant features in the phonological condition of the present study. Here, the temporal (i.e. durational) deviance was later identifiable in comparison to the spectral one, more precise, between 42 and 70 ms later. Past research indicates a temporal window of integration (TWI) of 150 - 200 ms in adults (Näätänen, 1900). Hence, stimuli presented within one TWI are integrated into one perceptual unity, i.e., MMN, by the auditory cortex (Näätänen & Winkler, 1999). Accordingly, Winkler et al. (1998) posit that additivity occurs when both deviants take place within 200 ms or less, whereas more than one MMN results when deviants are separated by more than 200 ms. Interestingly, a previous study indicates a divergent TWI for children in comparison to adults (Wang et al., 2005), which is congruent with the observation that two distinguishable MMN components were elicited in the phonological condition in children, but not in adults. In that study, TWIs of children belonging to two age groups (age 5-8 and 9-11 years) were compared with those of adults. Wang and colleagues (2005) have shown that the length of TWI is shorter for adults than for children (of both age groups), which is eventually the opposite of the result of the present study. However, it has to be kept in mind that Wang and colleagues (2005) used non-speech stimuli (frequency and intensity deviants) and speech and non-speech processing could be characterized by other developmental trajectories.

Auditory processing in children with and without dyslexia

Interestingly, no differences between dyslexic and typically developing children could be detected in none of the conditions of the present study. Hence, auditory processing in children with and without dyslexia did not differ from one another in the present study. Consequently, our data cannot be taken as evidence neither for the phonological theory of dyslexia nor for the auditory processing deficit theory of dyslexia.

Several reasons can account for this result. On one hand, isolated natural and artificially modified vowels of different length were presented to the participants in the present study. Both types of stimuli do not appear in this specific (i.e., isolated) form in natural speech. Hence, it is conceivable that both groups, children with and without dyslexia, had difficulties with these stimulus types. This fact could have led to a floor effect in both groups, with the consequence that no MMN (amplitude) group differences could have been detected. Note however that in a previous behavioral study with identical stimuli and conditions as in the current experiment, our research group has found differences between dyslexic and control children in all conditions, i.e., phonological, temporal and spectral (Steinbrink et al., 2014), indicating that our stimulus set is suitable for the investigation of auditory processing deficits in dyslexia. However, the inclusion criteria to the dyslexic group were much more strict in that study in comparison to the present study, since dyslexic children were recruited from German schools specialized on treatment of severe dyslexia. Moreover, active discrimination can differ from brain responses, indexed by MMN, as neurophysiological reactions reflect solely low-level auditory processes, whereas in behavioral reactions other processes, like attentional or decisional ones have an impact as well. In addition, evidence indicates that behavioral measurements may capture memory processes different from those captured by psychophysiological measures, such as MMN. Gaeta et al. (2001), for instance, have shown that behavioral detection is possible without MMN elicitation. The authors suggested that the sensory memory trace gathered by MMN might vanish in dependence of the duration of ISI used, whereas rehearsal during active discrimination might keep the representation of the stimulus features in working-memory storage for a longer time period.

Alternatively, many studies investigating auditory processing deficits in children with dyslexia found those solely in a subgroup of dyslexics (e.g., Tallal, 1980; Reed,

1989; Manis et al., 1997; Marshall et al, 2001; Rosen and Manganari, 2001). Hence, it cannot be excluded that the present sample is composed of a subgroup of children not affected by auditory processing deficits. Lachmann et al. (2005), for instance, have demonstrated auditory processing deficits, indexed by MMN aberration, only in a subgroup of dyslexics (age 8-11 years), suffering from a frequent word reading deficit, but not from a non-word reading deficit. The subgroup with deficits in non-word reading or both non-word reading and frequent word reading did not show any MMN abnormalities. Indeed, overall 12 dyslexic children out of 15 measured in the present sample had a deficit in non-word reading or both non-word and frequent word reading, which may account for the lack of group differences found in the present study[9].

Conclusion

The results demonstrate that children of 9-10 years of age process single features of the speech signal sequentially, more precise, feature-by-feature. The mature brain, however, processes single cues of the speech signal in an additive manner, that is to say, as a single complex neural representation, which for several reasons may reflect higher speech processing efficiency. For instance, additive processing of single acoustic features in speech may facilitate fast mapping of speech representations stored in long-term memory

____

[9] Note that in Experiment 3, a behavioral vowel length discrimination paradigm was implemented, in which same stimuli (vowels of different length) and conditions as in the passive MMN paradigm of Experiment 2 were used. Same children as in Experiment 2 took part in Experiment 3. However, no differences were found between control children and dyslexics in the active vowel discrimination task, indicating that the present group of children with dyslexia indeed does not suffer neither from attentive nor from preattentive deficits in auditory processing.

and representations of just processed speech stimuli stored in sensory memory and hence, may foster effective decoding of speech. Alternatively, additive processing of speech may be a prerequisite for more advanced cue-weighting mechanisms. Indeed, previous studies show that children and adults weight single features of the speech signal differently (e.g., Nittrouer et al., 2007; Mayo et al., 2005; Mayo et al., 2003). Nittrouer and colleagues, for instance, have demonstrated that children give more weight to formant transitions, i.e., spectral cues (e.g., Nittrouer, 1992; Nittrouer & Miller, 1997; Nittrouer & Studdent-Kennedy, 1987). This is fully in line with the results of the present study, as children processing single acoustic features within the speech signal sequentially, process spectral features (i.e., frequency patterns of sounds) in most cases prior to temporal features (i.e., transitions between sounds or the length of them), and hence give them greater weight. However, as a consequence of maturation of additive cue processing in speech, this pattern may change and cues more informative for a certain speech contrast may gain more weight.

All in all, the current study demonstrates that the maturation of the auditory system is not accomplished by middle childhood, which is in line with the assumption that auditory processing gradually improves from childhood to adulthood, like it is indicated by the ability to discriminate speech in noise that develops throughout teenage years (Talarico et al., 2007) or by changes in myelination and synaptic pruning in the secondary auditory cortex that take place up to adolescence (Devous et al., 2006). Hence, MMNs in children and adults are not equivalent and may reflect different processes. This again questions the common procedure of comparing mismatch amplitudes or latencies between children and adults.

Furthermore, as in the present experiment a higher density of electrodes was used than it was common in prior MMN studies, additional effects were found, i.e.,

fundamental differences between children and adults, which otherwise may have remained undetected.

Chapter 5: The impact of maturation on auditory processing in children at the age of approximately 10 years

As illustrated in Chapter 3, auditory function and consequently auditory ERPs in general and MMN in particular change in the course of development due to maturation. Interestingly, previous evidence indicates that these maturational effects are especially powerful at the age of approximately 10 years.

Wetzel and colleagues (2011), for instance, presented to their participants (normally developing children of 7-8 years and adults) task-irrelevant novel sounds varying in identifiability and duration. First, in children and adults, identifiable stimuli elicited larger ERPs than non-identifiable ones. However, exclusively in children this effect was extended for short stimuli after 300 ms. Furthermore, long stimuli in comparison to short stimuli led to more positive amplitudes in children than in adults at the time window of 400-600 ms. These results confirm that children of 7-8 years on one hand, process physically rich sounds in a different way than adults and on the other hand, are less efficient in inhibiting meaningful acoustic information than adults.

Bishop and colleagues (2011), in return, investigated children's auditory processing skills in the framework of a longitudinal study. Half of the children were initially 7 years old and 9 years old at the follow-up and the other half were initially 9 years old and 11 years old at the follow-up. Here, sinusoidal tones were presented to the participants in an oddball paradigm, while they were silently reading or playing electronic games. Bishop et al. (2011) determined maturational changes in auditory ERPs between 7 and 11 years. Interestingly, this developmental effect, here conceptualized as an increase in ERP amplitude, differed between temporal and fronto-central electrode sites, i.e., whereas a

maturational change was observed fronto-centrally, no developmental effects were found temporally, indicating that different cortical regions may mature independently from one another.

Furthermore, maturational changes at the age of approximately 10 years have been demonstrated in dyslexic children as well. As outlined above, at least a subgroup of dyslexics exhibits temporal auditory processing impairments (for more theoretical background, see Chapter 4). In addition, previous evidence demonstrates that certain temporal auditory processing deficits in dyslexia may vanish up to the age of 10 years (Hautus et al., 2003) and the development of other auditory processing abilities may proceed more slowly in dyslexics in comparison to normally reading children up to the age of 10 and end prematurely at the onset of puberty (Wright & Zecker, 2004).

As an example for the former case, so far, no differences between adult dyslexic and control participants have been found in auditory gap detection (McAnally & Stein, 1996; Protopapas et al., 2002; Schulte-Körne et al., 1998). Hautus and colleagues (2003), however, determined a deficit in auditory gap detection in dyslexic children of 6-9 years of age, but corresponding to previous evidence, not in three other age groups, including an adult group (10-11 years; 12-13 years and 23-25 years). Gap detection is a measure of temporal acuity (Irwin et al., 1985). Here, a short gap in a burst of Gaussian noise has to be detected. The shortest gap that can be detected provides the threshold of acuity, which can be compared between participants. This result is relevant, as although the early deficit in temporal auditory processing may disappear with age, this very same deficit may still have a great impact or be causal for speech-related, i.e., phonological, auditory processing deficits arising later on.

As an example for the latter case, Wright and Zecker (2004) found deviations in the developmental trajectory of auditory processing in children with dyslexia. Here, brief tones, which were presented against the background of a noise masker had to be detected.

Previously, deficits in this task had been found in a sample consisting of 8-year-old children with language impairments (Wright et al., 1997). In the study of Wright and Zecker (2004), 115 participants with either a diagnosis of dyslexia (N = 27), of a central auditory processing deficit (N = 15) or of a specific language deficit (N = 12), belonging to five age groups, took part in the testing. 61 children formed the control group. First, no performance differences between disorder subgroups have been found. Second, the pattern of results indicated a slower development of auditory functions in impaired children in comparison to control children. Moreover, approximately at the age of 10 years, the development of auditory functions halted in impaired children, whereas in normally reading children developmental performance ameliorations continued into adolescence. The authors suggested an interaction with puberty to be causal for this halt in development of auditory processing in children with learning problems.

Aims of Experiment 3

The present study is directly linked to Experiment 2, in which substantially different MMN reactions were found in children of 9-10 years vs. adults, whereas no differences were detected between children with and without dyslexia. To address maturational changes, Experiment 3 aims to replicate Experiment 2 one year later with a subset of the sample of Experiment 2 (N = 13 dyslexic and N = 12 control children).

As previous evidence (e.g., Bishop et al., 2011; Bruder et al., 2011) indicates that the age of approximately 10 years is relevant for the development of auditory processing (see above), it is expected to find MMN differences between children at the age of 9-10 years and children at the age of 10-11 years. To be exact, Experiment 2 demonstrated that children at the age of 9-10 years (in contrast to adults) process single features of the speech signal in a sequential manner. One year later, it is expected that this pattern changes, i.e., that single features of the speech stimuli are not processed in a strictly sequential manner anymore. More precisely, a greater temporal and/or topographical overlap between MMN components of the spectral and the temporal condition and correspondingly a greater temporal and/or topographical overlap between single MMN components of the phonological condition, maybe even MMN additivity comparable to the one in adults, is expected to be found.

Furthermore, in addition to the passive oddball paradigm an active vowel length discrimination task is added to Experiment 3. Identical stimuli will be presented to the participants in both paradigms. This is done to prove definitely, whether we have to deal with a dyslexic sample suffering from auditory processing deficits or whether we have not, which can be doubted for two reasons. First, no differences in auditory processing, as indexed by MMN, were found between dyslexics and controls in Experiment 2 and second, previous literature indicates that auditory processing deficits may be detectable

and consequently causally relevant for the emergence of dyslexia solely in a subgroup of dyslexics and it stands to reason (based on the results of Experiment 2) that the present sample does not belong to such a subgroup (for more theoretical background, see Chapter 4).

Experiment 3

Summary

The previous study (for more details, see Experiment 2) compared MMNs between children of 9-10 years of age (with and without dyslexia) and adults. Here, MMN additivity, i.e., a larger MMN amplitude for multiple-feature speech stimuli (spectral + temporal) in comparison to single-feature speech stimuli (spectral or temporal), was found in adults only. In children, no differences in MMN magnitude between conditions were found. Instead, it was shown that natural phonological, i.e., spectrotemporal, vowel length differences elicit two temporally and topographically distinct MMNs. This result indicates that, while children of 9-10 years sequentially process single acoustic features of the speech signal, adults integrate them additively. Furthermore, no evidence for differences between dyslexic and control children was found in Experiment 2.

Both studies with normally reading (Bishop et al., 2011; Bruder et al., 2011; Wetzel et al., 2011) and dyslexic children (Hautus et al., 2003; Wright & Zecker, 2004) document crucial alterations in auditory processing due to brain maturation at the age of around 10 years. Accordingly, it is possible that the pattern found in Experiment 2 for children at the age of 9-10 years shifts one year later. To probe this assumption, the same paradigm as in Experiment 2 was applied in Experiment 3, i.e., natural and artificially prolonged or shortened vowels were presented to the participants in a way, so that the difference between them was either spectrotemporal (phonological), temporal or spectral. Overall 13 dyslexic and 12 control children from the sample of Experiment 2 took part in the study. In addition, as no differences were found between dyslexic and control children in the passive oddball paradigm of Experiment 2, an active discrimination paradigm was implemented in Experiment 3. Here, identical stimuli as in the passive paradigm were

presented to the participants. However, corresponding to Experiment 2, dyslexic and control children did not differ in the active discrimination paradigm. Group specific analyses of MMN data were not possible, as due to drop out effects not enough participants remained in single groups. For this reason, in the following both children groups were analyzed together.

An altered MMN pattern was indeed determined in children at the age of 10-11 years in comparison to the very same children at the age of 9-10 years. On one hand, additive MMN effects were found and on the other hand, the presence of two MMN components in the phonological condition was still visually detectable, but not statistically reliable anymore in children at the age of 10-11 years. Crucially, these results are in line with previous evidence and confirm the hypothesis of significant maturational changes in auditory processing, as indexed by MMN, at the age around 10 years.

Methods

Participants

12 normally developing children (mean age 11.262, 5 females) and 13 children with developmental dyslexia (mean age 11.160, 6 females), already participating in Experiment 2, took part in the present study. Written and informed consent was obtained from all parents. Different standardized metrics were applied to assure that the classification of children as dyslexics or controls in Experiment 2 was still correct for Experiment 3, i.e., one year later. Reading literacy, i.e., frequent word and nonword reading ability, was determined with SLRT-II (Salzburger Lese-Rechtschreibtest; Moll & Landerl, 2010). To address spelling literacy, WRT3+ was replaced by P-ITPA (Potsdam-Illinois Test für Psycholinguistische Fähigkeiten; Esser et al., 2010), as no appropriate norms are available for the fifth grade for WRT3+. Phonological processing was addressed via RAN (Rapid Automatized Naming; Denckla et al., 1974), AGTB 5-12 (Working Memory Test Battery; subtests digit span forward and digit span backward; Hasselhorn et al., 2011) and a phonological awareness test battery including measures of vowel, consonant and initial sound discrimination (Klatte et al., in press). Same assignment criteria as in Experiment 2 were used in Experiment 3. To reiterate, a percentage range of ≤ 16 was required in one of the subtests of SLRT-II to be assigned to the dyslexic group and a percentage range of > 16 in all reading and spelling tests was assumed to be assigned to the control group (for more details, see Experiment 2).

Table 7: Descriptive characteristics of groups investigated in Experiment 3 and corresponding statistical comparisons. Note that for absolute numbers of errors no test statistics could be computed, as in both groups not enough errors were made.

Contents	Dyslexics		Controls		Statistical comparison		
	Mean	SD	Mean	SD	t-value	df	p-value
P-ITPA word (absolute number of errors)	9.923	3.774	4.667	4.997	2.983	23	<.01
P-ITPA word grapheme (absolute number of errors)	13.231	7.166	6	7.580	2.452	23	<.05
P-ITPA non-word (absolute number of errors)	4.846	1.819	4.25	2.301	0.722	23	.478
P-ITPA non-word grapheme (absolute number of errors)	8.154	5.970	6.75	4.495	0.660	23	.516
SLRT_II word (raw value)	51.692	13.634	78.083	28.385	-3.002	23	<.01
SLRT_II non-word (raw value)	31	7.083	48.833	11.424	-4.733	23	<.001
RAN objects (absolute number of errors)	0.461	0.776	0	0			
RAN objects (time in s)	52.308	21.963	40.417	5.596	1.819	23	.082
RAN colours (absolute number of errors)	0.538	0.967	0.083	0.289			
RAN colours (time in s)	49.538	11.657	39.042	9.292	2.476	23	<.05
AGTB digit span forward	4.433	0.93	5.259	0.804	-2.367	23	<.05
AGTB digit span backward	3.221	0.709	4.00	0.877	-2.453	23	<.05
Vowel discrimination (absolute number of errors)	1.538	1.266	0.583	0.668			
Initial sound discrimination (absolute number of errors)	5.846	2.703	4	2.594	1.739	23	.095
Consonant discrimination (absolute number of errors)	3.385	2.434	2.75	2.179	0.685	23	.5

Stimuli

The same stimulus set as in Experiment 1 and 2, encompassing two German natural vowels, i.e., 'i' and 'a' each in two different lengths, and their artificially lengthened and shortened counterparts, was used in Experiment 3 (for more details, see Experiment 1 and 2).

Conditions

Equivalent to Experiments 1 and 2, three conditions were established: a phonological (i.e. spectrotemporal), a temporal and a spectral one (for more details on stimulus structure in each condition, see Experiment 1; Figure 4).

Experimental procedures

First, the stimuli were presented in a passive auditory oddball paradigm, involving the presentation of a large number of standard stimuli (82.22%) interspersed with a small number of deviant stimuli (17.78%) (for more details, see Experiment 2).

In addition, the stimuli of the passive auditory oddball paradigm were embedded into an active auditory two-alternative forced-choice discrimination paradigm. Vowel pairs, differing either in their phonological, i.e., spectrotemporal, spectral or temporal features, were presented to the participants, which had to decide, whether the stimuli differed from one another or if they were identical. The experiment was composed of overall 96 trials. 48 trials (with 24 trials per vowel type) with same vowels and 48 trials (with 24 trials per vowel type) with vowels differing from one another were included. Five types of different trials were integrated into the paradigm. In the phonological condition, a natural long vowel was combined with a natural short vowel. In total, 8 phonological trials per vowel type were part of the experiment. In the temporal condition, an original long vowel was paired with an artificially shortened long vowel or a natural short vowel with an artificially lengthened short vowel. Each version was provided with 4 trials per vowel type. In the spectral condition, a natural long vowel was combined with an artificially lengthened short vowel in 4 trials per vowel type and a natural short vowel with an artificially shortened long vowel in 4 further trials per vowel type. In both paradigms, stimuli were presented at a fixed intensity of 67 dB SPL via Presentation software

(Neurobehavioral Systems Inc., San Francisco, California). An ISI of 250 ms between two vowels presented in succession was applied. To respond, the participants had to press one of two buttons. The assignment between response type and button was switched for the second half of participants. Response accuracy was measured. A practice phase (including overall 18 trials) preceded the experiment, to make sure that the participants had understood the task requirement.

Note that the mismatch response was captured solely during the passive oddball paradigm.

EEG recording and analysis

Comparable to Experiment 1 and 2, EEG was continuously recorded using the 128 channel Geodesic Sensor Net System (EGI; Electrical Geodesics, Inc., Eugene, Oregon). The data were preprocessed with FASTER (Nolan et al., 2010), a fully-automatized toolbox implemented in EEGlab (Swatz Center of Computational Neuroscience, San Diego, CA). Data analysis was conducted with the open-source softwares Fieldtrip (http://fieldtrip.fcdonders.nl), a toolbox implemented in Matlab (The Mathworks Inc., Natick, MA), ERP-lab (UC-Davis Center for Mind & Brain, Davis, CA), a toolbox implemented in EEGlab, that is again implemented in Matlab and IBM SPSS statistics 22 (International Business Machines Corporation, Armonk, New York) (for more detail, see Experiment 2)

Note that the aim of the current work was the appreciation of maturational changes in auditory processing occurring within one year in children (9-10 years vs. 10-11 years). Hence, the statistical analysis of Experiment 3 was paralleled to the one of Experiment 2. A separate statistical analysis was conducted with children of 10-11 years, since age groups could not be integrated into one analysis for several reasons. On one hand, sample sizes differed in Experiment 2 and 3 (N = 30 in Experiment 2 and N = 25 in Experiment

3) and on the other hand, more crucially, visual inspection revealed that the frontal MMN in children of 10-11 years (in contrast to children of 9-10 years; see Experiment 2) was localized over the right scalp side in all conditions (see Figure 18). Hence, the frontal ROI had to be relocated in Experiment 3. Consequently, a statistical analysis with both age groups (children of 9-10 years vs. 10-11 years) was not legitimized. Note that, as the statistical analysis lacks a direct comparison between age groups, the results have to be interpreted with caution.

To analyze the data, initially, cluster-based permutation tests were calculated in each condition over all EEG electrodes in two time windows (early: 140-200ms vs. late: 200-260 ms), to examine the reliability of mismatch responses in single conditions, which were determined by means of visual inspection. Note that outer electrodes (involving electrodes that were positioned at non-scalp sites) were excluded from cluster-based analyses and data visualization because of noisiness.

Second, a repeated-measures ANOVA based on mean amplitudes with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological cs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms) was computed. As mentioned above, the frontal ROI was relocated. Thus, equivalent to Experiment 2, the central-posterior ROI encompassed the electrodes E55, E54, E53, E61, E86, E79 and E78 and the frontal ROI encompassed the electrodes E118, E5, E112, E111, E117, E124 and E4 (see Figure 17). Note that the factor group was excluded from statistical analysis, since due to drop out effects not enough participants per group (12 normally developing children vs. 13 children with developmental dyslexia) were left to find reliable group-specific MMNs in single conditions, neither by means of visual inspection nor by application of statistical tests. Additional ANOVAs and Wilcoxon tests were computed to resolve the effects.

Wilcoxon tests were used instead of t-tests, since in several conditions MMN amplitudes were not normally distributed[10].

Moreover, Spearman's rho correlations were calculated for all conditions separately between reading and spelling ability, measures of phonological processing and mean MMN amplitudes, to validate the results of Experiment 2.
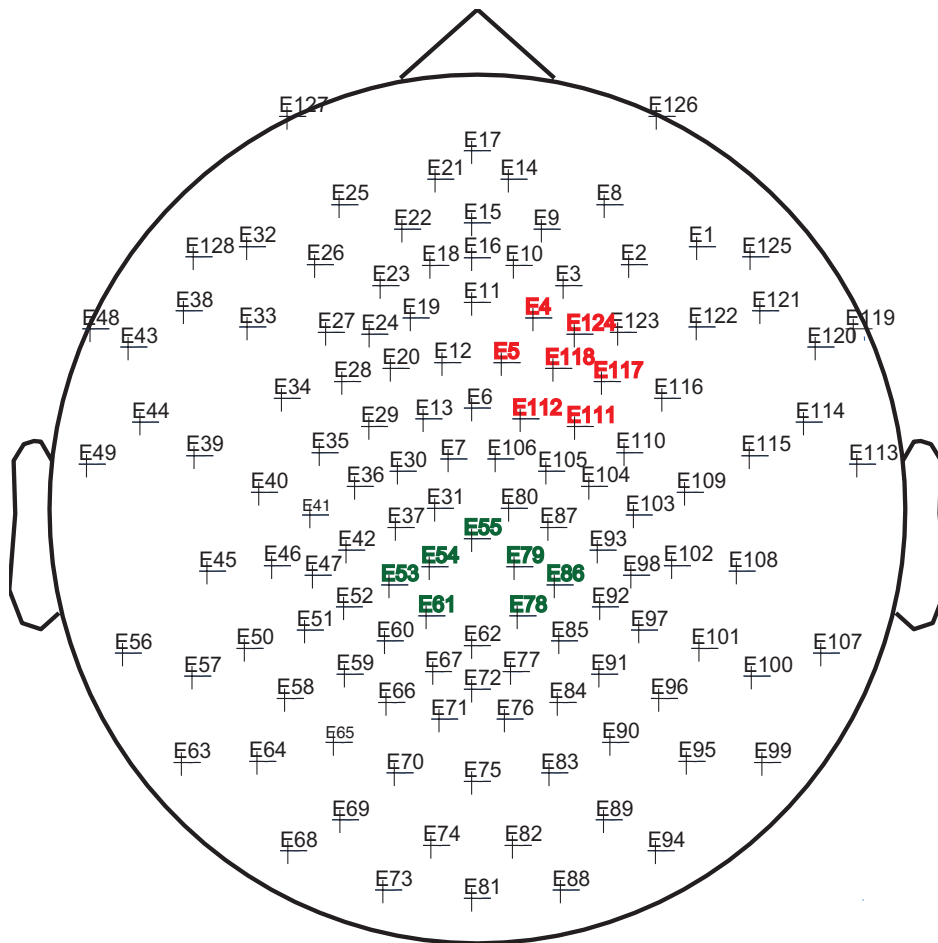
Figure 17: Electrodes included in ROIs of Experiment 3. Red electrodes illustrate the frontal ROI and green electrodes the central-posterior ROI.

---

[10] Note that Kolmogorov-Smirnov-tests were used to test for normal distribution.

Behavioral data analysis

To analyze differences between dyslexic and control children in active discrimination of stimuli belonging to the phonological, the temporal and the spectral condition, first, the discrimination index (d') was computed. Macmillan and Creelman (1991) developed d' exclusively for the analysis of same-different tasks. Here, the exact quantity of subject's hits and false alarms is included: a hit captures the correct detection of a difference and a false alarm the incorrect indication of a difference between stimuli presented. To compute d', false alarms are subtracted from hits, values are z-transformed and corrected for possible response biases, e.g., the bias that subjects more often judge the stimuli being the same (for more details, see Macmillan & Creelman, 1991). The discrimination index varies from 0, i.e., lowest performance accuracy to 5, i.e., highest performance accuracy.

Second, an ANOVA with factors condition (phonological vs. temporal vs. spectral) and group (dyslexic children vs. control children) was calculated. To resolve significant effects additional Wilcoxon tests were computed. Wilcoxon tests were used because the data were not overall normally distributed[11].

----

[11] Note that Kolmogorov-Smirnov-tests were used to test for normal distribution.

A   MMN topographies in all children (N =25; amplitude)

Phonological condition
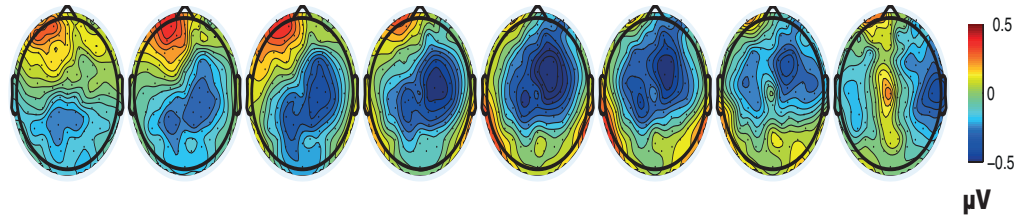
Temporal condition
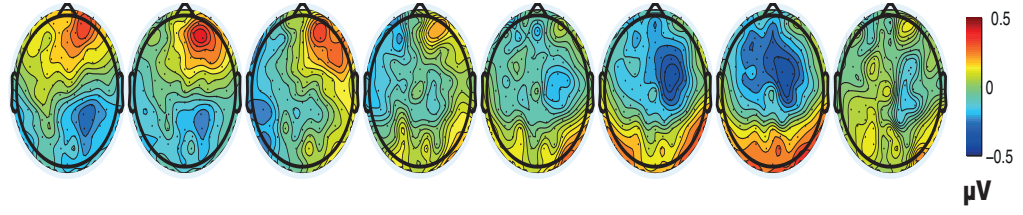
Spectral condition
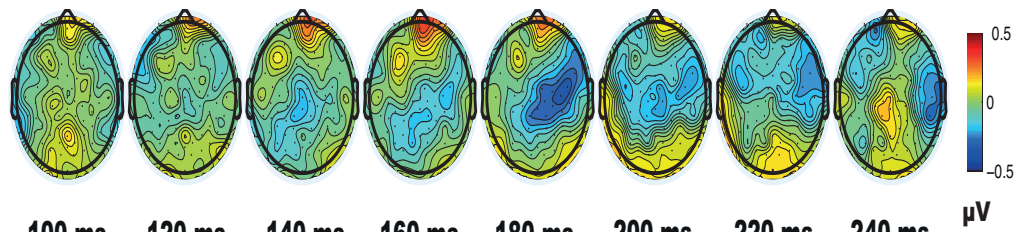
100 ms   120 ms   140 ms   160 ms   180 ms   200 ms   220 ms   240 ms

B   Mismatch response over all electrodes (all children)
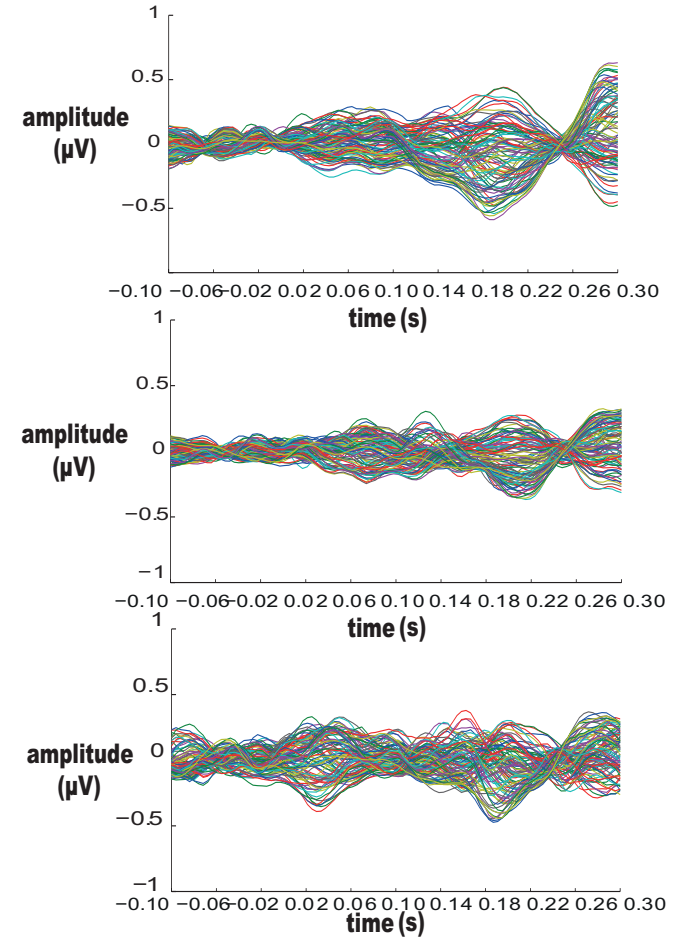
Figure 18: A. Time course of MMN topographies in all conditions and B. mismatch responses at all electrodes for the whole sample (N = 25) in all conditions displayed in a butterfly plot.

Results

EEG data

Equivalent to Experiment 2, initially, deviant ERPs were compared to standard ERPs in all conditions separately, by means of cluster-based permutation tests (calculated over all electrodes). Moreover, same time windows of analysis were applied (140-200 ms and 200-260 ms). Unfortunately, as already mentioned above, group specific statistical analyses were not legitimized, since due to drop out effects not enough participants were left per group (12 normally developing children vs. 13 children with developmental dyslexia) to find reliable MMNs in single conditions. Hence, the EEG data of children belonging to different groups were analyzed together.

Reliable MMNs could be found in both time windows as well in the phonological as in the temporal condition. Indeed, the significant MMN in the early time window in the temporal condition was unexpected. However, as mentioned above, to parallelize the results between the current study and Experiment 2 same time windows were used for statistical analyses in both experiments. Hence, as the MMN peak in the temporal condition was present around 200 ms after stimulus onset, the temporal MMN was detectable in both time windows applied in the analysis, the early (140 - 200 ms) and the late time window (200-260 ms). Note that, as the present study aimed to show that the principles of auditory processing have changed within 1 year and not to further investigate the nature of MMNs in children at the age of 10-11 years, the use of identical time windows in Experiment 2 and 3 was fully legitimized.
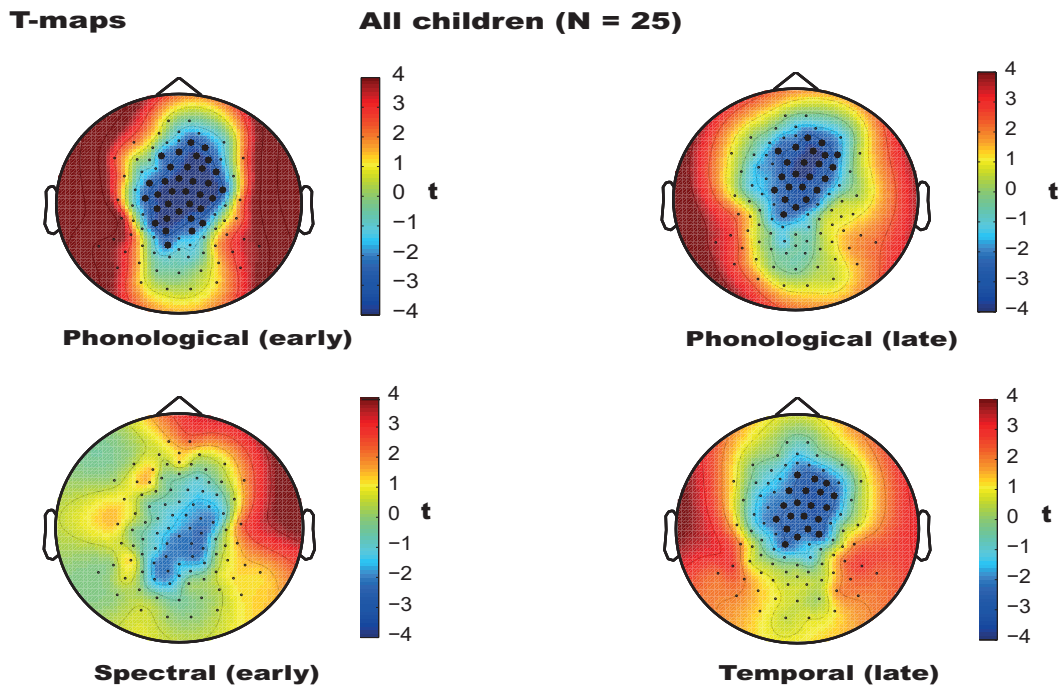
Figure 19: Distribution of t-values for comparisons between standards and deviants in all conditions (early time window: 140-200 ms and late time window: 200-260 ms). Note that no distinction was made between children with and without dyslexia, as for this purpose, i.e., for the group-specific analysis of MMN data, due to drop-out effects not enough participants were present in single groups. The bold dots indicate significant electrodes (p < 0.05). Outer electrodes were removed for illustration reasons.

In the spectral condition, a reliable MMN was found in the early time window only when merging central-posterior ROI electrodes (E55, E54, E53, E61, E86, E79 and E78) into one statistical analysis (t = 1.944; p = 0.025). However, when calculating cluster-based permutation tests on all EEG electrodes separately, no significant differences between

standards and deviants were found in the spectral condition, neither for the early nor for the late time window[12]. Corresponding t-maps are depicted in Figure 19.

Condition-specific MMN effects

To parallelize the current analysis with the one of Experiment 2, a repeated-measures ANOVA with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological cs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms) was calculated. Note however that, as opposed to the statistical analysis of Experiment 2, group was not included as factor (for more details, see above). A significant condition*time interaction ($F_{(2,23)} = 7,113$; $p < 0.01$) and significant condition ($F_{(2,23)} = 3,648$; $p = 0.042$) and time ($F_{(1,24)} = 7,183$; $p = 0.013$) effects were found. In addition, a marginally significant ROI*time interaction was determined ($F_{(1,24)} = 3,326$; $p = 0.081$). In contrast to Experiment 2, no ROI*time*condition interaction could be found ($F_{(2,23)} = 0,381$; $p = 0.687$) (for more details, see Table 8).

On one hand, to contrast the effects of the current study with those of Experiment 2 and on the other hand, to investigate the ROI*time and condition*time interactions determined in the ANOVA further Wilcoxon-Tests were calculated. For the phonological condition, equivalent to Experiment 2, ROIs were compared for the early ($z = -0.780$; $p = 0.435$) and the late time window ($z = -1,332$; $p = 0.183$) and time windows were compared for the frontal ($z = 2.005$; $p = 0.045$) and the central ROI ($z = 3.323$; $p < 0.01$).

____

[12] Note that this analysis step, i.e., the calculation of cluster-based permutation tests on all EEG electrodes separately, was also conducted with a smaller time window around the MMN peak in the spectral condition (160-200 ms and 170-190 ms). However, the difference between standards and deviants did not reach significance.

The results indicate a larger MMN for the early (time window 140-200 ms) compared to the late time window in the phonological condition, which is, however, present in both ROIs. Identical analyses were computed for the temporal and the spectral condition, i.e., ROIs were compared for the early (temporal condition: z = 0.377; p = 0.706; spectral condition: z = 0.605; p = 0.545) and the late time window (temporal condition: z = -0.794; p = 427; spectral condition: z = - 0.081; p = 0.936). In addition, time windows were compared for the frontal (temporal condition: z = -0.901; p = 0.367; spectral condition: z = 0.471; p = 0.638) and the central ROI (temporal condition: z = -0.363; p = 0.716; spectral condition: z = -3.068; p < 0.01). Summing up, the results demonstrate no MMN differences between time windows and ROIs for the temporal condition and a greater MMN in the central ROI in the early time window for the spectral condition.

Since additive MMN effects, i.e., a larger MMN in the phonological condition in comparison to the temporal and spectral conditions, were plausible to emerge in children of 10-11 years in comparison to children of 9-10 years, a further exploration of the main effect condition was of great interest. For further investigations, additional ANOVAs with factors condition (phonological vs. temporal and phonological vs. spectral and temporal vs. spectral), ROI (frontal vs. central) and time (140-200ms vs. 200-260ms) were computed to compare MMNs between single conditions, irrespective of their temporal or topographical distribution. Note that, ANOVAs instead of t-tests were calculated to compare the main effect condition by keeping other factors, i.e., ROI and time, constant. Indeed, a significant difference was found between the phonological and the spectral condition ($F(1,24) = 7,613$; p = 0.011) and a tendency to significance between the phonological and the temporal condition ($F(1,24) = 2,631$; p = 0.118). In addition, as expected, the temporal and the spectral condition did not differ from one another (F < 1) (for illustration, see Figure 20).

Furthermore, comparable to Experiment 2, no significant correlations (calculated for all conditions separately) were found between reading and spelling ability, measures of phonological processing and mean MMN amplitudes.

Table 8: Results of the repeated-measures ANOVA with factors region of interest (ROI: frontal vs. central-posterior), condition (phonological vs. temporal vs. spectral) and time (early: 140-200ms vs. late: 200-260 ms).

Effect	F-value	df(factor)	df(error)	Significance (p-value)
Time	7.183	1	24	.013
ROI	0.479	1	24	.496
Condition	3.648	2	23	.042
Time*ROI	3.326	1	24	.081
Time*condition	7.113	2	23	< .01
ROI*condition	0.581	2	23	.567
Time*ROI*condition	0.381	2	23	.687
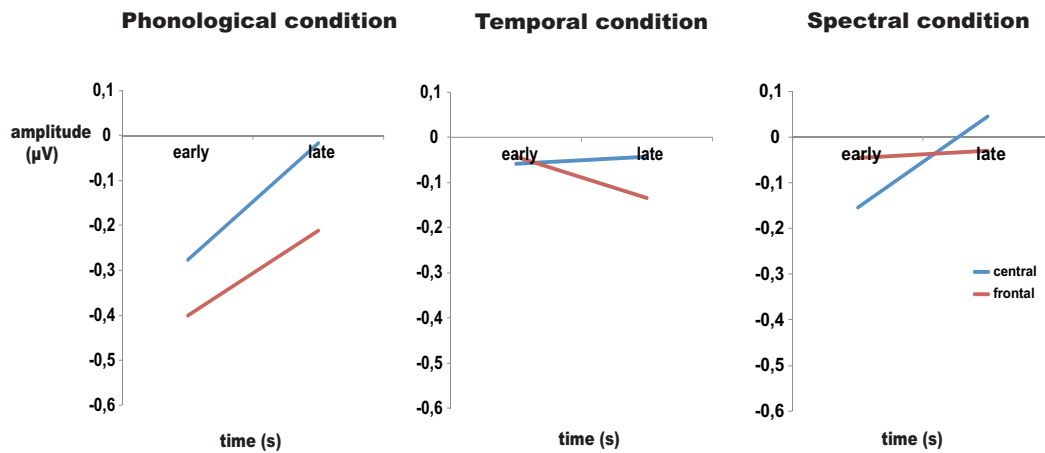
Figure 20: Illustration of the ANOVA interaction ROI*time*condition calculated for all children taken together (N = 25).[13]

Group specific d' effects

To compare the active discrimination accuracy, indexed by d', between children with and without dyslexia, a repeated-measures ANOVA with factors condition (phonological vs. temporal vs. spectral) and group (dyslexic children vs. control children) was calculated. Here, no significant group effects were found (group: $F_{(1,23)} > 1$; group*condition: $F_{(2,22)} < 1$). However, a significant condition effect was determined ($F_{(2,22)} = 10,308$; $p = 0.001$). Further Wilcoxon Tests revealed, as expected, significant differences between the phonological and the temporal ($z = 3.516$; $p = < 0.001$) and between the phonological and the spectral condition ($z = 3.025$; $p = < 0.01$). No differences were found between the temporal and the spectral condition ($z = -0,149$; $p = 0.881$) (for illustration, see Figure 21).

_____

[13] Note that the calculation of error bars (e.g., standard errors) was not legitimized in Figure 20, as the illustrated data were not normally distributed.

A Discrimination performance over all children (N=25)

B Discrimination performance of dyslexic (N = 13) and control children (N=12)
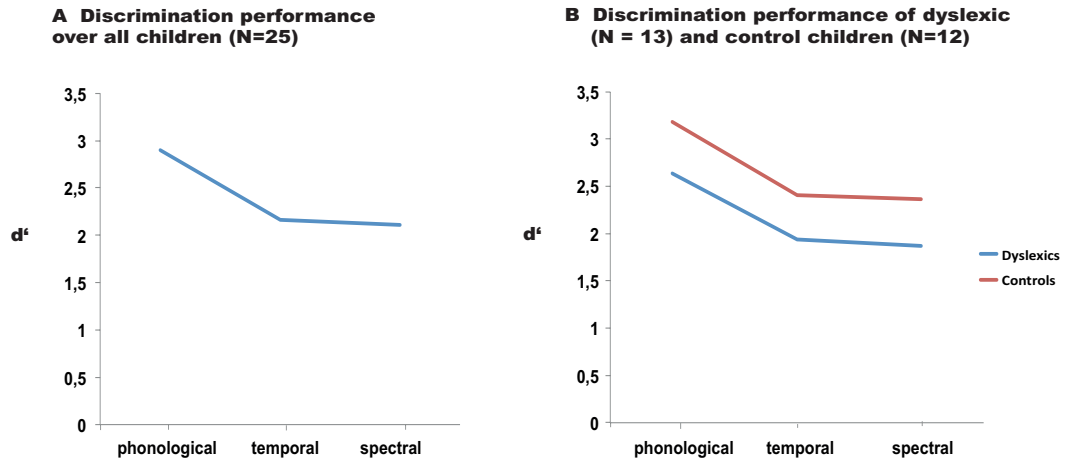
Figure 21: Discrimination index in all conditions, illustrated A. for all children taken together (N =25) and B. for dyslexic (N =13) and control children (N = 12) separately[14].

_____

[14] Note that the calculation of error bars (e.g., standard errors) was not legitimized in Figure 16, as the illustrated data were not normally distributed.

Discussion

The present study aimed to investigate additive MMN effects in speech in children of 10-11 years and to interrelate the results to those of Experiment 2, where children of 9-10 years and adults were investigated by means of the same paradigm, as in the present study. In contrast to children of 9-10 years, a tendency to MMN additivity was found in children of 10-11 years, indicating significant changes in auditory processing at approximately 10 years of age. In the following sections, these results will be discussed in more detail.

Maturation of auditory speech processing - a comparison between children of 9-10, 10-11 years and adults

It is widely accepted that auditory ERPs like MMN change throughout development. Over the course of maturation, alterations in latency, amplitude or topography may arise (for more details, see Chapter 5). Moreover, as outlined in Chapter 5, previous evidence indicates that the age of approximately 10 years is especially critical for the maturation of auditory processing. However, to the best of knowledge, there are no previous studies examining additive MMN effects from a developmental perspective. Hence, the current work aimed to investigate additive MMN effects with phonological, i.e., spectrotemporal, spectral and temporal deviants in children of 10-11 years of age and to interrelate the results to those of Experiment 2, where adults and children of 9-10 years of age were investigated. To do so, equivalent to Experiment 2, German vowels, differing either in both their spectral and temporal features or in spectral or temporal features only, were used as stimuli. In addition, identical statistical analyses were calculated in Experiment 2 and 3. Note however that due to divergent ROIs and sample sizes in different age groups, no direct statistical comparisons were calculated between age groups (for more details,

see the Methods section of Experiment 3). Hence, the results have to be interpreted with caution and further investigations are needed to make more reliable conclusions.

In Experiment 1 and 2, additive MMN effects were found in adults, i.e., adult's MMN was larger in the phonological condition in comparison to both the temporal and the spectral condition, showing that adults integrate spectral and temporal features of the speech signal in an additive way. In children of 9-10 years of age, by contrast, as examined in Experiment 2, no MMN additivity was determined. Instead, two MMN components were found in the phonological condition, corresponding (temporally and topographically) to MMNs in the temporal and the spectral condition and referring to this, reflecting the sequential processing of temporal and spectral information within the speech signal. Furthermore, an early frontal positive mismatch response accompanied by a posterior negativity was found prior to the two typical negative MMN responses in the phonological condition. This finding was surprising, as so far, positive mismatch deflections have not been detected in children exceeding the age of 6-7 years (for more theoretical background, see Experiment 2).

A shift in auditory speech processing was expected between 9-10 and 10-11 years, and perhaps even an approximation to the auditory speech processing mechanism of adults, that is to say, an emerging shift from sequential processing of single acoustic features in speech (characteristic for children) to summative processing of single acoustic features in speech (characteristic for adults). Correspondingly, an MMN pattern differing from both the one of adults and the one of younger children was found in children of 10-11 years (see Figure 22). First, frontocentral MMNs were detected in all conditions, with MMN in the temporal condition emerging later in time than MMNs in the phonological and the spectral condition. However, this latency difference between conditions was not surprising (see also Experiment 1 and 2), as it demonstrates the later onset of deviants in the temporal condition in comparison to deviants in the phonological and the spectral

condition. That is to say, in contrast to children of 9-10 years, the scalp topography in children of 10-11 years did not differ between conditions. Moreover, a strong tendency to MMN additivity was found in children of 10-11 years of age. More precisely, the MMN in the phonological condition (irrespective of its temporal or topographical distribution) was significantly larger than the MMN in the spectral condition and tended to be significantly larger than the MMN in the temporal condition, indicating summative integration of MMNs to single features of the speech signal. In addition, in contrast to Experiment 2, statistically no indices were found for two (temporally and topographically) distinct MMN components in the phonological condition with temporally and topographically corresponding MMNs in the temporal and the spectral condition, which would reflect sequential processing of temporal and spectral features within the speech signal. Nevertheless, visual inspection still revealed an MMN pattern, comparable to the one of children of 9-10 years. More precisely, an early predominantly centrally localized MMN component and a late predominantly frontally localized MMN component, both components (topographically and temporally) overlapping to a higher degree than in children of 9-10 years, were present in the phonological condition. Moreover, a late frontal MMN in the temporal condition and an early central MMN in the spectral condition, both components temporally less distributed than it was the case in children of 9-10 years in Experiment 2, could be visually inspected in children of 10-11 years of age (see Figure 22). However, as this MMN pattern was not supported statistically anymore, it is legitimized to conclude that essential differences between children of 9-10 and 10-11 years are present. In line with this, as expected due to the old age of the children, the positive mismatch reaction, still detected in children of 9-10 years of age, disappeared one year later.

To sum up, the current study supports the hypothesis, that the age of around 10 years is critical for substantial changes in auditory processing, which again corresponds to previous literature (for more theoretical background, see Chapter 5).

Several reasons may account for the pattern of results found in children of 10-11 years of age, i.e., the visually detectable, but not statistically provable double-peaked MMN in the phonological condition, which corresponds to the one-peaked MMNs in the spectral and the temporal condition. On one hand, a lack of statistical power may be responsible for the absence of statistical effects, since due to dropout effects only 25 participants (out of 30 participating in Experiment 2) took part in the current study. On the other hand, however, the result may reflect maturational changes, which would be in line with predictions done beforehand (see Chapter 5). Hence, further investigations are needed before drawing final conclusions.
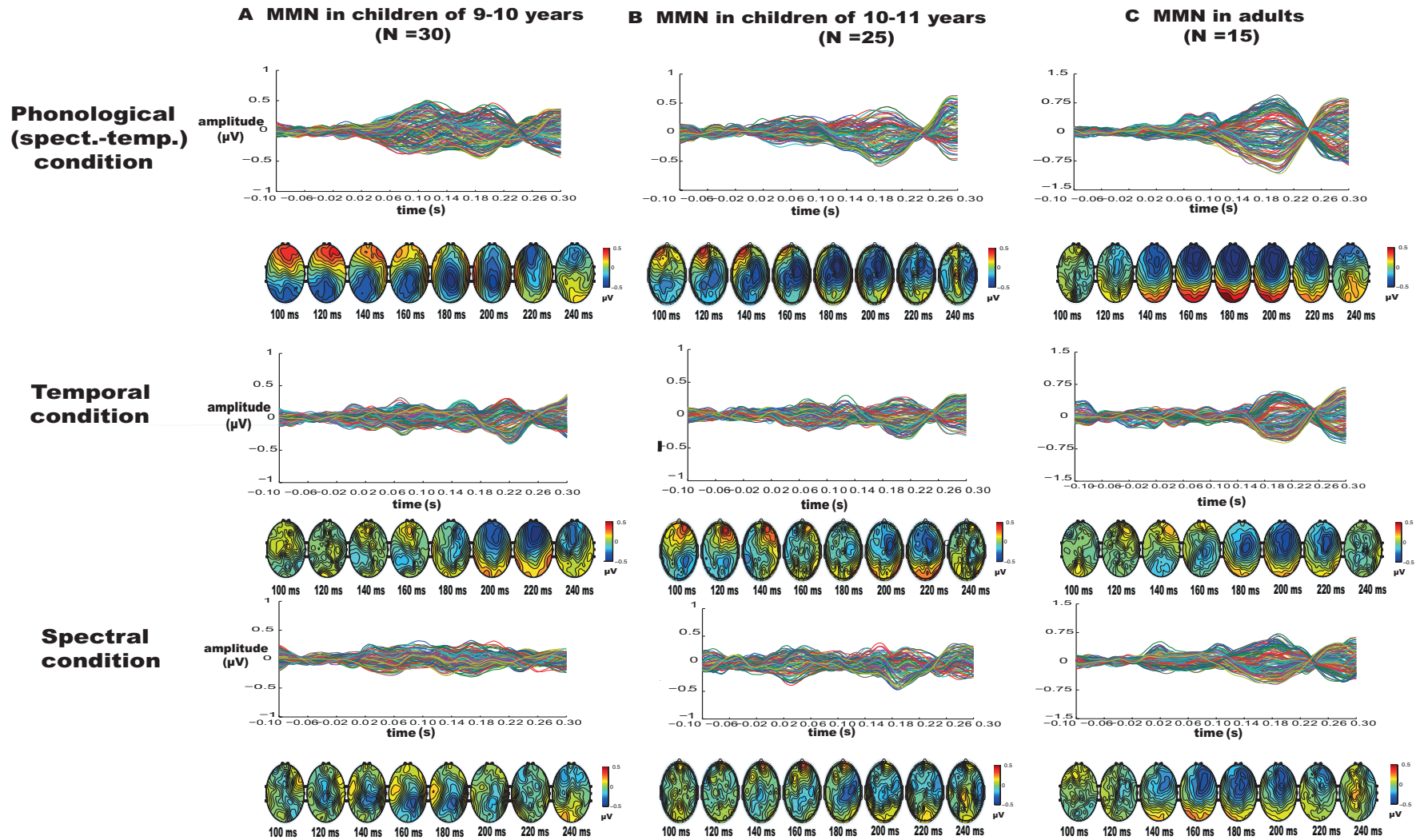
Figure 22: Overview over MMN results (Time course of MMN topographies and mismatch responses over all electrodes, displayed in butterfly plots in all conditions) in all age groups, i.e., in A. children of 9-10 and B. 10-11 years of age and C. adults.

Auditory processing in children with and without dyslexia II

In contrast to Experiment 2, the calculation of MMN differences between children with and without dyslexia was not permitted in Experiment 3, since due to drop out effects not enough statistical power was present in single groups and reliable MMNs could not be detected for single groups.

Nevertheless, for further validation of the results of Experiment 2, i.e., for further investigation of differences in auditory processing between dyslexic children and controls, an active vowel length discrimination paradigm was implemented in Experiment 3. Here, same stimuli and conditions as in the passive oddball paradigm were applied[15]. This transfer of the passive paradigm in an active one was especially important, as previous evidence shows that though neurophysiological measures and perception are highly correlated, they may also diverge from one another (for more details, see Chapter 2). However, the discrimination accuracy (in all conditions) did not differ between children with and without dyslexia. That is to say, a similar pattern was determined in both paradigms, the active discrimination paradigm of Experiment 3 and the passive oddball paradigm of Experiment 2.

This result can be considered as additional evidence that the dyslexic children of the present sample belong to a subgroup without (or at least without substantial) deficits in auditory processing. This interpretation is consistent with previous studies showing that

----

[15] Note that for the group-specific statistical analysis of MMN data more statistical power was required than for the group-specific statistical analysis of behavioral data. Consequently, the group-specific statistical analysis of data derived from the active vowel length discrimination paradigm was possible despite the fact that the group-specific statistical analysis of the data derived from the MMN-paradigm was not allowed.

dyslexia is not firmly linked to the presence of speech and/ or non-speech related auditory processing deficits (e.g., Lachmann et al., 2005; for more details, see Chapter 4).

Conclusion

The present work confirms the assumption represented in the literature that auditory processing, as indexed by MMN, takes a crucial turn at the age of approximately 10 years. However, the maturation of MMN is far from being concluded at this age. This finding is especially interesting, as it indicates the existence of a fixed age window characterized by increased neuronal plasticity. Consequently, in future it would be interesting to probe how training and special intervention programs at the age of 10 years affect the course of maturation.

However, although no differences between children with and without dyslexia were found for variables of interest, it should be kept in mind that a special subgroup of participants, not representative for the statistical population took part in the current work. Hence, in future, to derive final conclusions a replication with another participant pool (without any type of learning difficulties) is required.

Chapter 6: General discussion

Speech is a highly complex auditory input, which can be decomposed in several acoustic cues, whereas temporal and spectral cues are especially relevant for speech perception. As already mentioned in Ulytska et al. (submitted), speech is organized in temporal sequences and includes fast transitions between sounds differing in length. Hence, fast temporal integration of single speech units is essential for correct speech perception. Furthermore, the speech signal is characterized by a specific composition of frequencies, varying in the course of time. Thus, correct identification of frequencies included in the speech signal is important for correct phoneme perception. By way of example, the identification of the first two formants of a vowel determines its identification (Nawka & Wirth, 2008).

Previous research has shown that the processing of temporal features, i.e., information organzed in time windows up to 20-25 ms, is predominantly localized in the left hemisphere, whereas the processing of spectral cues, i.e., information organized in longer time windows up to 150-250 ms, is stationed in the right hemisphere (e.g., Zatorre, 2002; Ackermann & Riecker, 2004; Riecker, 2008; Basso et al., 1993). Moreover, several previous studies indicate that spectral and temporal cues may be weighted differently when phonemes are processed. This weighting process, in turn, is dependent on the phoneme type and the language system the certain phoneme is embedded in (e.g., Repp, 1982). However, to the best of knowledge the present work is the first contribution systematically investigating how exactly spectral and temporal cues are integrated by the brain when natural phonemes, consisting of both features, are processed (for more details, see Experiment 1). Interestingly, this research question is firmly associated with the controversially discussed dispute whether speech is processed either in a cue-specific or in a domain-specific manner or whether both mechanisms are involved, i.e., whether

single acoustic cues of the speech signal are processed independently from one another or rather in an integrated speech-specific manner or whether the separate processing of individual acoustic cues is simply adjusted by speech-specific long-term representations when speech is processed (e.g., Zatorre & Gandour, 2008; for more details, see Chapter 1).

In German, vowels of different length differ in both their duration and their spectral composition. Hence, to investigate the mechanisms of speech perception based on temporal and spectral information, (natural and artificially modified) vowels differing in length were presented as standards and deviants in a passive oddball paradigm. Two vowel categories, i.e., 'i' and 'a' differing in vowel height, were used. Note that previous evidence indicates that vowel height determines the relevance of spectral and temporal information for vowel identification (Sendlmeier, 1981; Strange & Bohn, 1998; Weiss, 1974; for more details, see Chapter 1). The stimuli were combined in a way, so that the difference between them was either phonological, i.e., spectrotemporal, temporal or spectral in nature. The MMN response was recorded as a neurophysiological correlate of sound discrimination accuracy.

Note that the experimental design of the current work was especially suitable for a systematic investigation of how single acoustic features of the speech signal are processed by the human brain. First, temporal and spectral features were isolated without altering the complexity of the stimuli. Hence, the stimulus complexity could be kept invariant in all conditions. Furthermore, temporal and spectral features were isolated without altering the speech character of the stimuli. Thus, by comparing the MMN to natural phoneme deviants, including spectral and temporal information, with the MMNs to artificially modified phoneme deviants, including either temporal or spectral information, MMN additivity could be investigated in speech. As outlined in Chapter 2, MMN additivity, i.e., a larger MMN for multiple-feature stimuli in comparison to single-

feature stimuli, has already been reported for non-speech stimuli in adult samples (e.g., Levänen et al., 1993; Schröger, 1995; Takegata et al., 1999). That is to say, it was shown that single acoustic cues within a tone are processed separately, i.e., in a cue-specific manner. The investigation of MMN additivity in speech, in turn, can substantially contribute to the debate of how speech is processed, either in a cue-specific (like it is the case in tones) or in a domain-specific manner and is hence of high relevance. More specifically, in case that no MMN additivity is found, i.e., that no association between the multiple-feature MMN with the corresponding single-feature MMNs is detectable, or the additive MMN effect is influenced by differences in phoneme type (e.g., vowel 'i' or 'a'), it can be concluded that cue-specific mechanisms cannot be alone involved in speech processing and speech and other types of auditory input are processed differently. Alternatively, in case that full or at least partial MMN additivity is found and multiple-feature MMNs (in both vowel types) are firmly associated with their corresponding single-feature MMNs, it can be inferred that same mechanisms are applied when speech or non-speech input is processed. However, in case that enlarged multiple-feature MMNs not associated with their corresponding single-feature MMNs are found (in both vowel types), it can be reasoned that speech and non-speech processing mechanisms differ substantially from one another and speech is processed in a unique speech-specific manner.

The phonological, i.e., spectrotemporal, spectral and temporal processing in children and adults

Initially, the present work has shown that the presence of single features, i.e., temporal or spectral, is sufficient for the human brain to identify vowels of different length (for more details, see Experiment 1, 2 and 3). This finding confirms that speech perception mechanisms are robust against variability of the auditory speech input, which again can

become especially important when the speech signal is impaired, e.g., by surrounding noise or due to a poor telephone connection. Moreover, speakers differing in age, gender or migration background produce speech with varying acoustic parameters, e.g., in terms of fundamental frequency, intensity or spectral and harmonic structure (see e.g., Iverson & Kuhl, 2000). To comprehend spoken speech anyhow, the speech perception system has to be enormously flexible.

Furthermore, additive MMN effects were found in a phoneme-unspecific data analysis, in adults. This finding confirms that, in principle, the mature brain is capable of summing MMNs to single features into one MMN, when speech sounds are processed. However, when analyzing the data phoneme-specifically, i.e., separately for vowel type /a/ and /i/, additive MMN effects were found for vowel type /i/, but not for vowel type /a/, indicating a complex integration process of single-feature MMNs to the cumulative multiple-feature MMN. More precise, as expected for (the low) vowel type /a/, the more salient temporal cue was weighted to a notably greater extent than the less salient spectral cue, preventing MMN additivity for vowel type /a/. This interpretation is supported by the clearly weaker MMN in the spectral than in the temporal condition for vowel type /a/ and by the fact that the spectrotemporal MMN was not differing in amplitude from the temporal MMN for vowel type /a/. For vowel type /i/, on the contrary, temporal and spectral cues were processed to a similar extent, which was reflected by equal MMNs in the temporal and the spectral condition, leading to MMN additivity for vowel type /i/. Hence, as for both vowel types, i.e., /i/ and /a/, the multiple-feature MMN was deducible from the corresponding single-feature MMNs, it could be concluded that particular speech features are processed separately by different neuronal populations before being integrated into one neuronal representation. However, as MMN additivity was found solely for vowel type /i/, but not for vowel type /a/ where the spectral MMN did not extend the size of the spectrotemporal MMN beyond the size of the temporal MMN, it

could be further concluded that speech perception entails an additional mechanism with single acoustic features being weighted differently depending on their relevance for a certain phoneme, which is not taking place when non-speech stimuli are processed. Note that this conclusion indicates that both cue- and domain-specific mechanisms are applied when speech is processed (e.g., Zatorre & Gandour, 2008).

In contrast to adults, no MMN additivity was found in children of 9-10 years. Instead, two MMN components were detected in the phonological condition topographically and temporally corresponding to the MMNs in the spectral and the temporal condition. Accordingly, it was concluded that the MMN components in the phonological condition reflect the sequential processing of spectral and temporal features in vowels of different length. This interpretation suggests that children of 9-10 years of age in contrast to adults do not integrate single speech features within one unifying neuronal representation, but rather process them sequentially. Interestingly, an altered MMN pattern was found in children at the age of 10-11 years. Here, no statistically reliable indices for sequential processing of single acoustic speech features could be determined, i.e., one-peaked MMNs were found in all conditions. Moreover, a strong tendency to additive processing of multiple-feature stimuli was found.

Taken together, the results indicate that children at the age of 9-10 years (and presumably before this age) process speech sounds in a cue-specific manner, whereas adults apply both, cue-specific as well as domain-specific speech-processing mechanisms: on the one hand by processing single acoustic cues within the speech stimulus in an integrated manner and on the other hand by weighting them differently depending on the extent of their informativeness for a certain speech contrast. Furthermore, between 9-10 and 10-11 years a crucial change in auditory processing occurs, i.e., speech processing shifts from a cue-specific to a domain- and cue-specific mode. This finding indicates that the child's auditory cortex is characterized by increased

neuronal plasticity at about 10 years of age. Hence, in future studies, it would be especially interesting to figure out how a direct training on speech perception would affect speech processing dependent on the child's age.

So far, the majority of trainings that aim to improve auditory perception are based on the identification or the discrimination of acoustically similar tones, syllables or words, whereas their acoustic content, e.g., the spectral, temporal or intensity information, is varied (for a review, see Sweetow et al., 2005). It is assumed, that current neuronal representations of certain acoustic cues are impaired, while training modifies these neuronal representations and hence improves the physiological ability to identify and to discriminate single acoustic cues within the auditory signal. If the training is successful, improved perception is assumed to generalize beyond the training content. That is to say, when a participant is trained to differentiate 'thee' and 'fee', he or she would be less likely to exchange 'th' and 'f' in other contexts, i.e., words (Tremblay, 2007). Ylinen et al. (2009), for instance, have investigated the impact of training on cue-weighting processes in adults. In general, Finns rely more on durational and less on spectral cues when vowels have to be identified, whereas the opposite is true for native English speakers. Interestingly, Ylinen and colleagues (2009) have shown that Finn's ability to rely on spectral cues for the identification of English vowels can be enhanced with training. A minimal-pair perceptual identification task has been used in the training period. The participants were supplied with 10 training sessions within 3 weeks, which lasted about 20-25 min. Different words, e.g., 'heat' or 'hit', were presented auditory to the participants who were instructed to indicate which word they had just heard. In case of a correct answer, a positive feedback (a happy animated bunny) followed and vice versa. The durational information was varied in different trials. By making the durational cues, Finns are used to rely on when vowels are processed, unreliable, the authors aimed to guide Finns to use spectral cues to a greater extent.

Comparable training procedures are conceivable in children at (and before) the age of 10 years as well. As a consequence, it might be that the shift in speech processing between 9-10 and 10-11 years can be preponed to an earlier age by means of a specific training on speech perception. Moreover, provided that the speech perception training would have an impact on reading and spelling literacy, formal education could be adjusted accordingly. On the one hand, it is for instance possible to include additional German lessons in which speech processing is explicitly trained, e.g., before and/or simultaneously with reading and spelling literacy acquisition. On the other hand, it might be that additional lessons in school, where spelling and reading are directly practiced, are especially effective at the age of approximately 10 years, as at this age the speech processing system is characterized by a certain plasticity. In that case, pupils around the age of 10 years have to be provided with such additional reading and spelling lessons.

All in all, the findings of the present work suggest that the maturation of auditory speech processing is far from being completed by middle childhood, which certainly questions the comparability of MMNs in children and adults. Given the assumption, that mature speech processing mechanisms are more efficient than immature, e.g., as speech processing is assumed to develop with linguistic experience, which is associated with age (e.g., Shtyrov et al., 1998; Dehaene-Laembertz, 1997; Näätänen et al., 1997), for future research it is without doubt of great interest to figure out whether children and adults differ only in speech-related categorical processing or in all types of auditory categorical processing, which would, for instance, have implications on the choice of the best age window to start musical education. Note that categorical perception is not unique to speech. Musical notes or intervals are perceived categorically as well (Zatorre, 2008). Hence, if the findings of our study can be generalized on categorical perception per se, it should be considered to start (or to intensify) the musical training at the age around 10 years. Moreover, (musical) training studies are needed with children younger that 10

years, investigating whether it is possible to relocate this time window at about 10 years of age, in which auditory functions are characterized by increased plasticity, to an earlier age and hence to facilitate musical education.

The phonological, i.e., spectrotemporal, spectral and temporal processing in children with and without dyslexia

As additional research question, the present work aimed to investigate possible deviations in speech perception. To date, it is well established that developmental dyslexia (among others) is associated with deficits in auditory perception (for more details, see Chapter 4). However, so far, no clarity is established about the quality of these auditory processing deficits in dyslexia, that is to say, whether these deficits are speech-specific or whether they refer to single acoustic cues within any type of auditory input, e.g., within speech. Hence, the present work aimed to compare phonological, i.e., spectrotemporal, temporal and spectral auditory processing in children with and without dyslexia.

The vowel length discrimination paradigm was especially suitable to investigate auditory processing deficits in dyslexia for several reasons. One of the greatest advantages was certainly the opportunity to investigate phonological (i.e., spectrotemporal), temporal and spectral auditory processing in dyslexics strictly separated from one another. Moreover, in most previous studies, different tasks have been used to investigate auditory processing in dyslexia, e.g., temporal order judgments (e.g., Tallal, 1980), gap detection (van Ingelghem et al., 2001), same different judgments (e.g., Hill et al., 1999; Groth et al., 2011; Vandermosten et al., 2010) or high-low discrimination (e.g., Banai & Ahissar, 2006). However, on one side, working memory load varies with task complexity and on the other side, the impact of motivational and attentional factors on task performance may vary in different tasks. Correspondingly, Banai and Ahissar (2006) have shown that dyslexics perform worse on tasks with larger demands on working

memory. The vowel length discrimination paradigm applied in the current work, however, was embedded in a passive oddball paradigm where participants were not instructed to work on an explicit task at all. Additionally, a simple same-different discrimination task not making high demands on working memory was appended to contrast attentive and preattentive speech processing abilities in dyslexic and control children. Also important is that stimuli of identical complexity were used in different conditions of the present study, which was not the case in most previous studies (e.g., Parviainen, Helenius, & Salmelin, 2005).

Nevertheless, dyslexic and control children did not differ from each other (in none of the conditions) in the present study, neither in MMN magnitude nor in the topographical or temporal MMN distribution. The active discrimination paradigm yielded same results (for a detailed discussion, see Experiment 2 and 3).

In general, the group assignment procedure in studies of dyslexia can be questioned. In a majority of studies, a certain cut-off value is chosen and children below this cut-off are assigned into the group of dyslexics and children above this cut-off into the group of controls (e.g., Bretherton & Holmes, 2003; Schulte-Körne, 1999; 1998). In the present study, for instance, all children with average or above average IQ with a percentage range of ≤ 16 in at least one of the reading subtests were assigned into the dyslexic's sample and children of comparable IQ and a percentage range > 16 in all reading and spelling subtests into the group of controls. However, it is highly debatable, whether a child with a percentage range of, for instance, 18 in a reading test differs substantially from a child with a percentage range of 16. That is to say, it is an open question whether reading or spelling deficits should be considered categorically or rather, as varying on a continuum. As a consequence of this debate, additional correlations were computed between reading and spelling abilities, phonological processing skills and MMN measures in the present study. Nevertheless, significant correlations between these measures were not found,

confirming the assumption that we have to deal with dyslexic children without auditory processing deficits. Note however that the dyslexic children in the present work were not severely affected by dyslexia and hardly any participant had a firm diagnosis of dyslexia, whereas in other studies, e.g., Steinbrink et al. (2014), dyslexic children were acquired from special schools with severe cases of dyslexia. This fact could be one reason for the lack of differences in auditory processing between children with and without dyslexia in the present study. Hence, an interesting research question for future studies would refer to the investigation of the degree in which severity of dyslexia is connected to the appearance of auditory processing deficits, as it is possible that auditory deficits in dyslexia are not so much causal for the disorder, but rather influence its gravity.

Regardless of the absence of auditory processing differences between dyslexic and control children in the present study, this work is nonetheless important for the research of dyslexia, as the finding that children process single features in speech sounds sequentially explains, at least partly, why deficiencies in auditory low-level and auditory speech processing so frequently coincide in children with dyslexia. The results of the present work indicate that, in both dyslexic and control children, single features of the speech stimulus are processed in a sequential manner, that is to say, cue-specific mechanisms are applied. When single features of non-speech stimuli are processed, cue-specific processing is presumably applied as well (for more theoretical background, see Chapter 1). Hence, it is likely that the processing of speech and non-speech stimuli does not differ from one another in children. As a consequence, auditory processing deficits in dyslexic children manifest theirself both, as speech and non-speech related. Note however that, to date, this conclusion is rather speculative, as the present work lacks a direct comparison between speech and non-speech processing. Hence, further work is needed for validation.

General Conclusion

Summing up, it can be concluded that the present work considerably contributes to the research on basic principles of speech perception, not least because it illuminates an essential mechanism of speech perception and its susceptibility to developmental influences. However, it has to be kept in mind that the present work has also some methodological limitations. First, a direct statistical group comparison, examining speech perception in children and adults is lacking. Second, to investigate the development of auditory processing, the data of children with and without dyslexia were analyzed together. Hence, a replication with a sample of children representative for the statistical population, i.e., without systematic spelling or reading deficits, is required, to make final conclusions about the development of auditory processing. In addition, to prove the generalizability of the present findings beyond the German vowel system, studies investigating other phoneme classes are needed.

References

Aaltonen, O., Eerola, O., Lang, A. H., Uusipaikka, E., & Tuomainen, J. (1994). Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected bymismatch negativity. *The Journal of the Acoustical Society of America*, *96*(3), 1489 1493.

Ackermann H., Riecker A. (2004). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language, 89(2),* 320-328.

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Adlard A., & Hazan, V. (1998). Speech perception in children with specific reading difficulties (dyslexia). *The Quarterly Journal of Experimental Psychology: Section A, 51*(1), 153-177.

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(12), 6832-6837.

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M.(2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13367-13372.

Akatsuka, K., Wasaka, T., Nakata, H., Inui, K., Hoshiyama, M., & Kakigi, R. (2005). Mismatch responses related to temporal discrimination of somatosensory stimulation. *Clinical Neurophysiology, 116* (8), 1930–1937.

Albrecht, R., Suchodoletz, W. V., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clinical Neurophysiology*, *111* (12), 2268-2276.

Alho, K., Woods, D. L., Algazi, A., Knight, R. T., & N, R. (1994). Lesions of frontal cortex diminish the auditory mismatch negativity. *Electroencephalography and Clinical Neurophysiology*, *91*(5), 353–362.

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

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., Knuutila, J., & Näätänen, R. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology*, *33*(4), 369-375.

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskeläinen, I. P., et al. (1998a). Processing of novel sounds and frequency changes in the human auditory cortex: Magnetoencephalographic recordings. *Psychophysiology, 35(2),* 211–224.

Alho, K., Connolly, J. F., Cheour, M., Lehtokoski, A., Huotilainen, M., Virtanen, J., Aulanko, R., & Ilmoniemi, R. J. (1998b). Hemispheric lateralization in preatten- tive processing of speech sounds. *Neuroscience Letters, 258*(1), 9–12.

Amenedo, E., & Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioral perception. *European Journal of Neuroscience*, *12*(7), 2570-2574.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4. vollständig überarbeitete Auflage). Washington, DC: Author.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R.,&Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /4/ and English /l/ and /R/. *Journal of Phonetics, 32*, 233–250.

Atienza, M., & Cantero, J. L. (2001). Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity (MMN). *Brain Research*, *901*(1), 151-160.

Atienza, M., Cantero, J. L., & Stickgold, R. (2004). Posttraining sleep enhances automaticity in perceptual discrimination. *Journal of Cognitive Neuroscience*, *16*(1), 53-64.

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*(4), 495-503.

Banai, K., & Ahissar, M. (2006). Auditory processing deficits in dyslexia: task or stimulus related? *Cerebral Cortex*, *16*(12), 1718-1728.

Basso, Amusia, in: F. Boller, J. Grafman (Eds.), Handbook of Neuropsychology, Vol. 8, Oxford University Press, Oxford, 1993, pp. 373–390.

Bazana, P. G., & Stelmack, R. M. (2002). Intelligence and information processing during an auditory discrimination task with backward masking: an event-related potential analysis. *Journal of Personality and Social Psychology*, *83*(4), 998-1008.

Birkel, P. (2007). *Weingartener Grundwortschatz Rechtschreibtest für dritte und vierte Klassen (WRT3+;* 2. neu normierte und vollständig überarbeitete Auflage). Göttingen: Hogrefe.

Bishop, D. V. M. (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*(4), 651-672.

Bishop, D. V. M., Anderson, M., Reid, C., & Fox, A. M. (2011). Auditory Development between 7 and 11 Years : An Event-Related Potential (ERP) Study. *PloS one*, *6*(5), e18993.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology*, *95*(3), 153-193.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International* *5*(9/10), 341-345.

Boets, B., Wouters, J., van Wieringen, A., & Ghesquière, P. (2007). Auditory processing, speech perception and phonological ability in pre‑school children at high-risk for dyslexia: A longitudinal study of the auditory temporal processing theory. *Neuropsychologia, 45*(8), 1608–1620.

Bohn, O. S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(03), 303-328.

Bohn, O.-S. (1995). Cross-language speech perception in adults: L1 transfer doesn't tell it all. In W. Strange (Hrsg.). *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (S. 275-300). Timonium, MD: York Press.

Bohn, O.-S., & Polka, L. (2001). Target spectral, dynamic spectral and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America, 110*(1), 504–515.

Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., & Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology, 116*(12), 2765-2774.

Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge: MIT Press.

Bradley, L., & Bryant, P.E. (1983). Categorizing sounds and learning to read – a causal connection. *Nature, 301*, 419-421.

Bradlow, A. R., Kraus, N., Nicol, T. G., McGee, T. J., Cunningham, J., Zecker, S. G., & Carrell, T. D. (1999). Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. *The Journal of the Acoustical Society of America, 106*(4), 2086-2096.

Bretherton, L., & Holmes, V.M. (2003). The relationship between auditory temporal processing, phonemic awareness, and reading disability. *Journal of Experimental Child Psychology, 84*(3), 218-243.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*(5), 874-886.

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*(8), 1107–18.

Carroll, D. W. (2004). *Psychology of language* (4. Aufl.). Australia, Belmont, CA: Thomson/Wadsworth.

Ceponiene, R., Cheour, M., & Näätänen, R. (1998). Interstimulus interval and auditory event related potentials in children: Evidence for multiple generators. *Electroencephalography and Clinical Neurophysiology*, 108(4), 345–354.

Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O., & Näätänen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research*, *82*(1), 53-58.

Cheour-Luhtanen, M., Alho, K., Sainio, K., Rinne, T., Reinikainen, K., Pohjavuori, M., Renlund, M., Aaltonen, O., Eerola, O., Näätänen, R. (1996). The onto- genetically earliest discriminative response of the human brain. *Psychophysiology, 33*(4), 478-481.

Cheour, M., Alho, K., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O., & Näätänen, R. (1997). The mismatch negativity to changes in speech sounds at the ageof three months. *Developmental Neuropsychology, 13*(2), 167-174.

Cheour, M., Haapanen, M. L., Hukki, J., Ceponiene, R., Kurjenluoma, S., Alho, K., Tervaniemi, M., Ranta, R., & Näätänen, R. (1997). The first neurophysiological

evidence for cognitive brain dysfunctions in children with CATCH. *Neuroreport,* *8*(7), 1785-1787.

Cheour, M., Alho, K., Čeponiené, R., Reinikainen, K., Sainio, K., Pohjavuori, M., Aaltonen, O., & Näätänen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology*, *29*(2), 217-226.

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

Cheour, M., Čeponiené, R., Hukki, J., Haapanen, M. L., Näätänen, R., & Alho, K. (1999). Brain dysfunction in neonates with cleft palate revealed by the mismatch negativity. *Clinical Neurophysiology*, *110*(2), 324-328.

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 and Neurotology*, *6*(1), 2-11.

Cheour, M., Čëponiené, R., Leppänen, P., Alho, K., Kujala, T., Renlund, M., Fellman, V. & Näätänen, R. (2002). The auditory sensory memory trace decays rapidly in newborns. *Scandinavian Journal of Psychology, 43*, 33–39.

Cheour, M., Shestakova, A., Alku, P., Ceponiene, R., & Näätänen, R. (2002). Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native speech sounds within two months. *Neuroscience Letters*, *325*(3), 187-190.

Christmann, C.A. (2014). The role of stimulus complexity in auditory research of speech and non speech on the behavioral and electrophysiological level. urn:nbn:de:hbz:386-kluedo-37497.

Christmann, C.A., Lachmann, T. & Steinbrink, C. (under review). Evidence for a general auditory processing deficit in developmental dyslexia from a discrimination

paradigm using speech vs. non-speech sounds matched in complexity. Manuscript submitted for publication.

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

Cornelissen, P. L., Hansen, P. C., Bradley, L., & Stein, J. F. (1996). Analysis of perceptual confusions between nine sets of consonant-vowel sounds in normal and dyslexic   adults. *Cognition, 59*(3), 275-306.

Corriveau, K. H., Goswami, U., & Thomson, J. M. (2010). Auditory processing and early literacy skills in a preschool and kindergarten population. *Journal of Learning Disabilities, 43*(4), 369-382.

Crowther, C. S., & Mann, V. (1992). Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America, 92*, 711–722.

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

Csépe, V., & Gyurkócza, E. E. (2000). Abnormal mismatch negativity to phonetic deviations in developmental dyslexia. *Developmental Neuropsychology*.

Czigler, I., & Winkler, I. (1996). Preattentive auditory change detection relies on unitary sensory memory representations. *NeuroReport, 7*(15-17), 2413-2418.

Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature, 370*(6487), 292-295.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8*(4), 919-924.

Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport, 9*(8), 1885-1888.

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

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, *12*(3), 449-460.

Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *NeuroReport, 12*(14), 3155-3158.

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger- Charolles, L., Jobert, A. & Dehaene, S. (2005). Neural correlates of switching from auditory to speech perception. *NeuroImage, 24(1),* 21–33.

Demb, J. B., Boynton, G. M., Best, M., & Heeger, D. J. (1998). Psychophysical evidence for a magnocellular pathway deficit in dyslexia. *Vision Research, 38*(11), 1555–1559.

Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex*, *10*(2), 186-202.

Denckla, M. B., & Rudel, R. G. (1976). Naming of object-drawings by dyslexic and other learning disabled children. *Brain and Language*, *3*(1), 1-15.

Denckla, M. B. (1985). Motor coordination in dyslexic children: theoretical and clinical implications. In F. H. Duffy & N. Geschwind (Hrsg.), *Dyslexia: a neuroscientific approach to clinical evaluation* (S. 184–195). Boston, MA: Little, Brown, & Co.

Devous, M. D., Altuna, D., Ngai, W. T., Chiu, S., Harris, T. S., & Tobey, E. A. (2006). Maturation of speech and language functional neuroanatomy in pediatric normal controls. *Journal of Speech, Language, and Hearing Research*, *49*(4), 856-866.

Diehl, R. L., Lotto, A. J. & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology, 55,* 149–179.

Eggermont, J. J. (1988). On the rate of maturation of sensory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *70*(4), 293-305.

Eggermont, J. J. (1991). Development of auditory evoked potentials. *Acta oto-laryngologica*, *112*(2), 197-200.

Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: correlations with changes in structure and speech perception. *Acta oto-laryngologica*, *123*(2), 249-252.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*(4), 375-384.

Escudero, P., Benders, T.,&Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics, 37*, 452–465.

Esser, S. K., Huber, R., Massimini, M., Peterson, M. J., Ferrarelli, F., & Tononi, G. (2006). A direct demonstration of cortical LTP in humans: A combined TMS/EEG study. *Brain Research Bulletin*, *69*(1), 86–94.

Esser, G., Wyschkon, A., Ballaschk, K. & Hänsch, S. (2010). *Potsdam-Illinois Test für Psycholinguistische Fähigkeiten - P-ITPA*. Göttingen: Hogrefe, Testzentrale.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, *2*(4), 460-493.

Fawcett, A. J., & Nicolson, R. I. (1994). Naming speed in children with dyslexia. *Journal of Learning Disabilities*, *27*(10), 641-646.

Felton, R. H., Naylor, C. E., & Wood, F. B. (1990). Neuropsychological profile of adult dyslexics. *Brain and Language, 39*(4), 485-497.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437-470.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, *88*(1), 3-17.

Gabrieli, J.D.E. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science, 325*, 280-283.

Gaeta, H., Friedman, D., Ritter, W., & Cheng, J. (2001). The effect of perceptual grouping on the mismatch negativity. *Psychophysiology*, *38*(2), 316-324.

Galaburda, A., & Livingstone, M. (1993). Evidence for a Magnocellular Defect in Developmental Dyslexia. *Annals of the New York Academy of Sciences*, *682*(1), 70 82.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain Generators Implicated in the Processing of Auditory Stimulus Deviance: A Topographic Event-Related Potential Study. *Psychophysiology*, *27*(6), 627-640.

Giard, M., Lavikahen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience, 7*(2), 133-143.

Godfrey, J. J., Syrdal-Lasky, K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology*, *32*(3), 401-424.

Gomes, H., Bernstein, R., Ritter, W., Vaughanm, H. G., & Miller, J. (1997). Storage of feature conjunction in transient auditor memory. *Psychophysiology*, *34*(6), 712-716.

Gomes, H., Sussman, E., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan Jr, H. G. (1999). Electrophysiological evidence of developmental changes in the duration of auditory sensory memory. *Developmental Psychology*, *35*(1), 294-302.

Gomot, M., Giard, M. H., Roux, S., Barthélémy, C., & Bruneau, N. (2000). Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *NeuroReport*, *11*(14), 3109-3112.

Goswami, U. (1999). The relationship between phonological awareness and orthographic representation in different orthographies. In M. Harris & G. Hatano (Hrsg.). *Learning to read and write* (S. 134–156). Cambridge: Psychology Press.

Groth, K., Lachmann, T., Riecker, A., Muthmann, I., & Steinbrink, C. (2011). Developmental dyslexics show deficits in the processing of temporal auditory information in German vowel length discrimination. *Reading and Writing*, *24*(3), 285–303.

Haffner, J., Zerahn-Hartung, C., Pfüller, U., Parzer, P., Strehlow, U., & Resch, F. (1998). Auswirkungen und Bedeutung spezifischer Rechtschreibprobleme bei jungen Erwachsenen–empirische Befunde in einer epidemiologischen Stichprobe. *Zeitschrift für Kinder-und Jugendpsychiatrie und Psychotherapie*, *26*(2), 124-135.

Hasselhorn, M., Schumann-Hengsteler,R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., Seitz-Stein, K., & Zoelch C. (2011). *Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12)*. Göttingen: Hogrefe-Verlag.

Hautus, M. J., Setchell, G. J., Waldie, K. E., & Kirk, I. J. (2003). Age-related improvements in auditory temporal resolution in reading impaired children. *Dyslexia*, *9*(1), 37-45.

Heim, S., Eulitz, C., Wienbruck, C., Elbert, T.(2000). Effects of syllabic training on literacy skills, phonological processing, and cortical organization in children with language impairment. In: Society for Psychophysiological Research (Hrsg.). *Abstracts of the 40th Annual Meeting, October, 18-22,2000*. San Diego, CA, USA.

Hertrich, I., Ackermann, H. (1997). Articulatory control of phonological vowel length contrasts: kinematic analysis of labial gestures. *Journal of the Acoustical Society of America, 102*(1), 523–536.

Hill, N. I., Bailey, P. J., Griffiths, Y. M., & Snowling, M. J. (1999). Frequency acuity and binaural masking release in dyslexic listeners. *The Journal of the Acoustical Society of America*, *106*(6), L53-L58.

Holopainen, I. E., Korpilahti, P., Juottonen, K., Lang, H., & Sillanpää, M. (1997). Attenuated auditory event-related potential (mismatch negativity) in children with developmental dysphasia. *Neuropediatrics*, *28*(05), 253-256.

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America, 119,* 3059–3071.

Huttenlocher, P. R., de Courten, C., Garey, L. J., & Van der Loos, H. (1982). Synaptogenesis in human visual cortex—evidence for synapse elimination during normal development. *Neuroscience Letters*, *33*(3), 247-252.

Huttunen-Scott, T., Kaartinen, J., Tolvanen, A., & Lyytinen, H. (2008). Mismatch negativity (MMN) elicited by duration deviations in children with reading disorder, attention deficit or both. *International Journal of Psychophysiology*, *69*(1), 69-77.

Ilvonen, T. M., Kujala, T., Tervaniemi, M., Salonen, O., Näätänen, R., & Pekkonen, E. (2001). The processing of sound duration after left hemisphere stroke: event-related potential and behavioral evidence. *Psychophysiology*, *38*(04), 622-628.

Irwin, R. J., Ball, A. K., Kay, N., Stillman, J. A., & Rosser, J. (1985). The development of auditory temporal acuity in children. *Child Development*, *56*(3), 614-620.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87,* B47– 57.

Jacobsen, T., & Schro, E. (2004). Pre-attentive perception of vowel phonemes from variable speech stimuli. *Psychophysiology, 41,* 654–659.

Jaramillo, M., Ilvonen, T., Kujala, T., Alku, P., Tervaniemi, M., & Alho, K. (2001). Are different kinds of acoustic features processed differently for speech and non-speech sounds?. *Cognitive Brain Research*, *12*(3), 459-466.

Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia*, *10*(3), 196-214.

Kekoni, J., Hämäläinen, H., Saarinen, M., Gröhn, J., Reinikainen, K., Lehtokoski, A., & Näätänen, R. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biological Psychology*, *46*(2), 125-142.

Tremblay, K. L. (2007). Training-Related Changes in the Brain : Evidence from Human Auditory-Evoked Potentials. *Seminars in Hearing, 28*(2), 120–132.

Klicpera, C., & Gasteiger-Klicpera, B. (1998). *Psychologie der Lese- und Schreibschwierigkeiten* (Psychology of reading and spelling disabilities). Weinheim: Beltz, PVU.

Klicpera, C., & Gasteiger–Klicpera, B. (1993). *Lesen und Schreiben: Entwicklung und Schwierigkeiten [Reading and spelling: Development and difficulties]*. Bern: Hans Huber.

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America, 104*, 3568–3582.

Kohler, K. J. (1977*). Einführung in die Phonetik des Deutschen* (1. Aufl.). Berlin: E. Schmidt.

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*(3), 332-339.

Koyama, S., Gunji, A., Yabe, H., Oiwa, S., Akahane-Yamada, R., Kakigi, R., et al. (2000). Hemispheric lateralization in an analysis of speech sounds: Left hemisphere dominance replicated in Japanese subjects. *Cognitive Brain Research, 10(1-2),* 119–124.

Kozou, H., Kujala, T., Shtyrov, Y., Toppila, E., Starck, J., Alku, P., & Näätänen, R. (2005). The effect of different noise types on the speech and non-speech elicited mismatch negativity. *Hearing Research*, *199*(1), 31-39.

Kraus, N., McGee, T., Sharma, A., Carrell, T., & Nicol, T. (1992). Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing*, *13*(3), 158-164.

Kraus, N., McGee, T., Micco, A., Sharma, A., Carrell, T., & Nicol, T. (1993). Mismatch negativity in school-age children to speech stimuli that are just perceptibly different. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(2), 123-130.

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience*, *7*(1), 25-32.

Kraus, N., McGee, T., Carrell, T. D., & Sharma, A. (1995). Neurophysiologic bases of speech discrimination. *Ear and Hearing*, *16*(1), 19-37.

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*(5), 1042-1060.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843.

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

Kujala, T., Karma, K., Ceponiene, R., Belitz, S., Turkilla, P., Tervaniemi, M., Näätänen, R. (2001). Plastic neural changes and reading improvement caused by audio-visual training in reading-impaired children. *Proceedings of the National Academy of Sciences of the United States of America, 98*(18), 10509–10514.

Kujala, T., Kallio, J., Tervaniemi, M., & Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology*, *112*(9), 1712-1719.

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

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

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

Kurtzberg, D., Vaughan Jr, H. G., Kreuzer, J. A., & Fliegler, K. Z. (1995). Developmental studies and clinical application of mismatch negativity: problems and prospects. *Ear and Hearing*, *16*(1), 105-117.

Lachmann, T., Berti, S., Kujala, T., & Schröger, E. (2005). Diagnostic subgroups of developmental dyslexia have different deficits in neural processing of tones and phonemes. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *56*(2), 105–20.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: a German-English comparison. *Cognition, 63*(3), 315–334.

Landerl, K. (2003). Categorization of vowel length in German poor spellers: An orthographically relevant phonological distinction. *Applied Psycholinguistics*, *24*(4), 523- 538.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*(6), 609–621.

Lehtonen, A., & Bryant, P. (2002, June). *Awareness of phoneme duration predicts literacy skills.* Poster presented at the conference of the Society of the Scientific Study of Reading, Chicago, IL.

Leppänen, P. H., Eklund, K. M., & Lyytinen, H. (1997). Event‑related brain potentials to change in rapidly presented acoustic stimuli in newborns. *Developmental Neuropsychology*, *13*(2), 175-204.

Levänen, S., Hari, R., McEvoy, L., & Sams, M. (1993). Responses of the human auditory cortex to changes in one versus two stimulus features. *Experimental Brain Research, 97*(1), 177–183.

Liberman, A.M., Harris, K.S., Hoffman, H.S., Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme categories. *Journal of Experimental Psychology 54*(5), 358–368.

Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21(1)*, 1–36.

Liberman, A. M. & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4(5),* 187–196.

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., & Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *PLoS computational biology*, *9*(2), e1002911.

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.

Livingstone, M.S., Rosen, G.D., Drislane, F.W., Galaburda, A.M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences, 88*(18*),* 7943-7.

Lorenzi, C., Dumont, A., & Fullgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *43*(6), 1367-1379.

Lovegrove, W.J., Bowling, A., Badcock, B., Blackwood, M. (1980). Specific reading-disability: differences in contrast sensitivity as a funstion of spacial frequency. *Science, 210*, 439-40.

Mayo, C., & Scobbie, J. M. (2003). The Influence of Phonemic Acoustic Cue Weighting Strategies in Children's Speech Perception. *Journal of Speech, Language, and Hearing Reasearch, 46,* 1184–1196.

Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *The Journal of the Acoustical Society of America, 118*(3).

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, New York: Cambridge University Press.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen, A. (1997). Are Speech Perception Deficits Associated with

Developmental Dyslexia? *Journal of Experimental Child Psychology, 66*(2), 211-235.

Mäntysalo, S., & Näätänen, R. (1987). The duration of a neuronal trace of an auditory stimulus as indicated by event-related potentials. *Biological Psychology, 24*(3), 183 195.

Marshall, C. M., Snowling, M. J., & Bailey, P. J. (2001). Rapid auditory processing and phonological ability in normal readers and readers with dyslexia. *Journal of Speech, Language, and Hearing Research, 44*(4), 925-940.

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*(5), 808-817.

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

Maye, J. C., Werker, J. F., & Gerken, L.A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82,* B101–B111.

Mcanally, K. I., & Stein, J. F. (1996). Auditory temporal coding in dyslexia. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 263*(1373), 961-965.

McBride-Chang, C. (1996). Models of speech perception and phonological processing in reading. *Child Development, 67*(4), 1836-1856.

McGettigan C., Scott, S.K. (2012). Cortical asymmetries in speech perception: What's wrong, what's right and what's left? *Trends of Cognitive Sciences, 16(5),* 269–276.

Meng, X., Sai, X., Wang, C., Wang, J., Sha, S., & Zhou, X. (2005). Auditory and speech processing and reading development in Chinese school children: Behavioral and ERP evidence. *Dyslexia, 11*(4), 292-310.

Michie, P. T., Budd, T. W., Todd, J., Rock, D., Wichmann, H., Box, J., & Jablensky, A. V. (2000). Duration and frequency mismatch negativity in schizophrenia. *Clinical Neurophysiology, 111*(6), 1054–1065.

Miyawaki, K., Strange, W., Verbrugge, R. R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*, 331–340.

Moberly, A. C., Bhat, J., Welling, D. B., & Shahin, A. J. (2014). Neurophysiology of spectrotemporal cue organization of spoken language in auditory memory. *Brain and Language*, *130*, 42–49.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech Perception Deficits in Poor Readers: Auditory Processing or Phonological Coding? *Journal of Experimental Child Psychology, 64*(2), 199–231.

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language*, *72*(3), 238-245.

Moulton, W. G. (1962). *The sounds of English and German.* Chicago: University of Chicago Press.

Moll, K., Landerl, K. (2010*). SLRT II - Salzburger Lese- und Rechtschreibtest* (2. Korrigierte Auflage mit erweiterten Normen). Bern: Verlag Hans Huber.

Möttönen, R., Calvert, G. A., Jääskelainen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J. & Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *NeuroImage, 30(2),* 563–569.

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*(2), 118-136.

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

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

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*(02), 201-288.

Näätänen, R. (1991). Mismatch negativity outside strong attentional focus: a commentary on Woldorff et al. (1991). *Psychophysiology*, *28*(4), 478-484.

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale: Erlbaum.

Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., & Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *NeuroReport*, *4*(5), 503-506.

Näätänen, R., & Alho, K. (1995). Mismatch negativity-a unique measure of sensory processing in audition. *International Journal of Neuroscience*, *80*(1-4), 317-337.

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

Näätänen R, Winkler I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*(6), 826–859.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence'in the auditory cortex. *Trends in Neurosciences*, *24*(5), 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), 1–21.

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*(12), 2544–2590.

Naccache, L., Puybasset, L., Gaillard, R., Serve, E., & Willer, J.‑C. (2005). Auditory mismatch negativity is a good predictor of awakening in comatose patients: a fast and reliable procedure. *Clinical Neurophysiology, 116*(4), 988–989.

Nagarajan, S., Mahncke, H., Salz, T., Tallal, P., Roberts, T., & Merzenich, M. M. (1999). Cortical auditory signal processing in poor readers. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(11), 6483-6488.

Nawka, T., & Wirth, G. (2008). *Stimmstörungen: Für Ärzte, Logopäden, Sprachheilpädagogen und Sprechwissenschaftler; mit 30 Tabellen* (5th ed.). Köln: Dt. Ärzte‑Verl.

Nelson, H. E., & Warrington, E. K. (1980). An investigation of memory functions in dyslexic children. *British Journal of Psychology*, *71*(4), 487-503.

Neville, H. (1995). Developmental specificity in neurocog- nitive development in humans. In M. S. Gazzaniga (Hrsg.): *The Cognitive Neurosciences* (S. 219-232). Mass: Bradford Book.

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). Preattentive Cortical-Evoked Responses to Pure Tones, Harmonic Tones, and Speech: Influence of Music Training. *Ear and Hearing, 30*(4), 432–446.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30,* 319–329.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*, 351–382.

Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America, 101*, 2253–2266.

Nittrouer, S. & Lowenstein, J.H. (2007). Children's weighting strategies for word-final stop voicing are not explained by auditory sensitivities. *Journal of Speech, Language, and Hearing Research, 50*(1), 58-73.

Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems?. *Journal of Speech, Language, and Hearing Research*, *42*(4), 925-942.

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162.

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, *20*(1), 26-36.

Oades, R. D., Dittmann-Balcar, A., & Zerbin, D. (1997). Development and topography of auditory event-related potentials, mismatch and processing negativity from 8 to 22 years of age. *Psychophysiology, 34*(6), 677 – 693.

Obleser, J., Wise, R. J. S., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(9), 2283–9.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience 2011*, 156869.

Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology, 38*(02), 359-365.

Paavilainen, P., Valppu, S., & Näätänen, R. (2001). The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity: 1+1 approximately 2 but 1+1+1<3. *Neuroscience Letters*, *301*(3), 179–82.

Pang, E., & Taylor, M. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clinical Neurophysiology, 111*(3), 388– 397.

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S., & Frith, C.D. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain*, *119*(1), 143-157.

Parviainen, T., Helenius, P., & Salmelin, R. (2005). Cortical differentiation of speech and nonspeech sounds at 100 ms: implications for dyslexia. *Cerebral Cortex*, *15*(7), 1054 1063.

Pazo- Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: a review. *Biological Psychology, 63*(3), 199–236.

Phillips C., Marantz A., McGinnis M., Pesetsky D., Wexler K., & Yellin, E. (1995). Brain mechanisms of speech perception: a preliminary report. In: T. Schütze , J. B. Ganger, K. Broihier (Hrsg.). *Papers on language processing and acquisition* (S. 153-191). Cambridge, MA: MIT Press.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., Mcginnis, M., & Roberts, T. (2000). Auditory Cortex Accesses Phonological Categories : An MEG Mismatch Study. *Journal of Cognitive Neuroscience*, *12*(6), 1038–1055.

Pennington, B. F., & Gilger J. W. (1996). How is dyslexia transmitted? In: C. H. Chase, G. D. Rosen, G. F. Sherman (Hrsg.). *Developmental dyslexia. Neural, cognitive, and genetic mechanisms* (S. 41-61). Baltimore, MD: York Press.

Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. E. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology, 5*(3-4), 111-139.

Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Developmental Neuropsychology, 31*(3), 249-278.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology, 111*(2), 220-236.

Price, C., Thierry, G., & Griffiths, T. (2005). Speech-specific auditory processing: Where is it? *Trends in Cognitive Sciences, 9(6),* 271–276.

Protopapas, A., Ahissar, M., & Merzenich, M. M. (2002). *Auditory processing deficits in adults with a history of reading deficits*. Scientific Learning Corp. URL: http://www.scilearn.com.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Sami, M., Ilmoniemi, R., J., & Näätänen, R. (2001). Memory traces for words as revealed bythe mismatch negativity. *NeuroImage, 14*(3), 607-616.

Pulvermüller, F., Shtyrov, Y., Kujala, T., & Näätänen, R. (2004). Word specific cortical activity as revealed by the mismatch negativity. *Psychophysiology, 41*(1), 106-112.

Putkinen, V., Tervaniemi, M., Saarikivi, K., Ojala, P., & Huotilainen, M. (2014). Enhanced development of auditory change detection in musically trained school-aged children: a longitudinal event-related potential study. *Developmental Science, 17*(2), 282-297.

Ramus, F. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841–865.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit?. *The Quarterly Journal of Experimental Psychology*, *61*(1), 129-141.

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of experimental child psychology*, *48*(2), 270-292.

Repp, B. H. (1982). Phonetic trading relations and context effects: New evidence for a phonetic mode of perception. *Psychological Bulletin, 92*, 81–110.

Rey, V., De Martino, S., Espesser, R., & Habib, M. (2002). Temporal processing and phonological impairment in dyslexia: Effect of phoneme lengthening on order judgment of two consonants. *Brain and Language*, *80*(3), 576-591.

Richardson, U., Leppänen, P. H., Leiwo, M., & Lyytinen, H. (2003). Speech perception of infants with high familial risk for dyslexia differ at the age of 6 months. *Developmental Neuropsychology*, *23*(3), 385-397.

Riecker, A., Brendel, B., Ziegler, W., Erb, M., & Ackermann, H. (2008). The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain and Language, 107(2),* 102–113.

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

Rivera-Gaxiola, M., Johnson, M. H., Csibra, G., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of category goodness. *Behavioral Brain Research*, *112*(1), 1-11.

Roodenrys, S., & Stokes, J. (2001). Serial recall and nonword repetition in reading disabled children. *Reading and Writing, 14*(5-6), 379-394.

Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia?. *Journal of Speech, Language, and Hearing Research*, *44*(4), 720-736.

Scanlon, D.M., & Vellutino, F.R. (1996) Prerequisite skills, early instruction, and success in first grade reading: Selected results from a longitudinal study. *Mental Retardation and Development Disabilities, 2,* 54–63.

Scanlon, D.M., Vellutino, F.R., Small, S.G., & Fanuele, D.P. (2000). *Severe reading difficulties: Can they be prevented? A comparison of prevention and intervention approaches.* Paper presented at the Annual Conference of The American Educational Research Association. New Orleans, April.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*(6), 1728-1743.

Schröger, E. (1995). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology*, *32*(1), 55–65.

Schröger, E. (1998). Application of ERP measures in psychological research. *Behavior Research Methods, Instruments and Computers, 30*, 131-145.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998a). Role of auditory temporal processing for reading and spelling disability. *Perceptual and Motor Skills, 86*(3), 1043-1047.

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

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1999). The role of phonological awareness, speech perception, and auditory temporal processing for dyslexia. *European Child & Adolescent Psychiatry, 8*(3), 28-34.

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

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (2001). Speech perception deficit in dyslexic adults as measured by mismatch negativity (MMN). *International Journal of Psychophysiology*, *40*(1), 77–87.

Schulte-Körne, G. (2002). *Legasthenie: Zum aktuellen Stand der Ursachenforschung, der diagnostischen Methoden und Förderkonzepte*. Bochum: Winkler.

Schulte-Körne, G., & Bruder, J. (2010). Clinical neurophysiology of visual and auditory processing in dyslexia: a review. *Clinical Neurophysiology*, *121*(11), 1794-1809.

Sendlmeier, W.F.(1981). Der Einfluss von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen (The influence of quality and quantity on perception of stressed German vowels). *Phonetica, 38*(5-6), 291-308.

Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2000). Maturation of mismatch negativity in school-age children. *Ear and Hearing*, *21*(3), 242-251.

Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, *18*(3), 242-254.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., Ilmoniemi, R. J., & Näätänen, R. (2002). Abstract phoneme representations in the lefttemporal cortex: magnetic mismatch negativity study. *NeuroReport*, *13*(14), 1813 1816.

Shahin, A., Roberts, L.E., & Trainor, L.J. (2004). Enhancement of auditory cortical development by musical experience in children. *NeuroReport, 15*(12), 1917–1921.

Sharma, A., Kraus, N., Mcgee, T., Carrell, T., & Nicol, T. (1993). Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-

related potential. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(1), 64-71.

Sharma, A., Kraus, N., J McGee, T., & Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *104*(6), 540-545.

Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J., & Shaywitz, S. E. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly*, *30*, 894-906.

Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine*, *338*(5), 307-312.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., Constable, R. T., Marchione, K. E., Fletcher, J. M., Lyon G. R., & Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*(2), 101-110.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry*, *57*(11), 1301-1309.

Shinozaki, N., Yabe, H., Sutoh, T., Hiruma, T., & Kaneko, S. (1998). Somatosensory automatic responses to deviant stimuli. *Cognitive Brain Research, 7*(2), 165–171.

Speyer, A. (2007). *Germanische Sprachen: Ein historischer Vergleich*. Göttingen: Vandenhoeck & Ruprecht.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., et al. (2002). Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study. *NeuroReport, 13*, 1813–1816.

Shtyrov, Y., Kujala, T., Lyytinen, H., Kujala, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Lateralization of speech processing in the brain as indicated by mismatch negativity and dichotic listening. *Brain and Cognition*, *43*(1-3), 392-398.

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

Shtyrov, Y., Pulvermüller, F., Näätänen, R., & Ilmoniemi, R. J. (2003). Grammar processing outside the focus of attention: an MEG study. *Journal of Cognitive Neuroscience*, *15*(8), 1195-1206.

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

Snowling, M. J. (2000). *Dyslexia*. Oxford: Blackwell.

Stanovich, K. E. (1988). Explaining the Differences Between the Dyslexic and the Garden Variety Poor Reader: The Phonological-Core Variable-Difference Model. *Journal of Learning Disabilities, 21*(10), 590–604.

Stein, J., Walsh, V. (1997). To see but not to read: the magnocellular theory of dyslexia. *Trends in Neurosciences, 20,* 147-52.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia, 7*(1), 12–36.

Steinbrink, C., & Klatte, M. (2008). Phonological working memory in German children with poor reading and spelling abilities. *Dyslexia, 14* (4), 271-290.

Steinbrink, C., Groth, K., Lachmann, T., & Riecker, A. (2012). Neural correlates of temporal auditory processing in developmental dyslexia during German vowel length discrimination: an fMRI study. *Brain and Language, 121*(1), 1–11.

Steinbrink, C., Klatte, M., Lachmann, T. (2014). Phonological, temporal and spectral processing in vowel length discrimination is impaired in German primary school children with developmental dyslexia. *Research in Developmental Disabilities, 35(11),* 3034–45.

Stoodley, C. J., Hill, P. R., Stein, J. F., & Bishop, D. V. (2006). Auditory event-related potentials differ in dyslexics even when auditory psychophysical performance is normal. *Brain Research*, *1121*(1), 190-199.

Strange, W., & Bohn, O.-S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *Journal of the Acoustical Society of America, 104*(1), 488–504.

Sussman, E., Kujala, T., Halmetoja, J., Lyytinen, H., Alku, P., & Näätänen, R. (2004). Automatic and controlled processing of acoustic and phonetic contrasts. *Hearing Research*, *190*(1), 128-140.

Sutcliffe, P., & Bishop, D. (2005). Psychophysical design influences frequency discrimination performance in young children. *Journal of Experimental Child Psychology*, *91*(3), 249-270.

Swan, D., & Goswami, U. (1997a). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology, 66*(1), 18-41.

Swan, D., & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language, 56*(3), 334-353.

Sweetow, R. & Palmer C.V. (2005). Efficacy of individual auditory training in adults: a systematic review of the evidence. *Journal of the American Academy of Audiology. 16*(7), 494– 504.

Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, *266*(2), 109–12.

Takegata, R.T.M., Alku, P., Ylinen, S., & Näätänen, R. (2008). Parameter-specific modulation of the mismatch negativity to duration decrement and increment: Evidence for asymmetric processes. *Clinical Neurophysiology, 119*(7), 1515–1523.

Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., … & Paolini, A. G. (2006). Effect of age and cognition on childhood speech in noise perception abilities. *Audiology and Neurotology*, *12*(1), 13-19.

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

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

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, *682*(1), 27-47.

Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *NeuroReport*, *12*(2), 299-307.

Tervaniemi, M., Saarinen, J., Paavilainen, P., Danilova, N., & Näätänen, R. (1994). Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biological Psychology*, *38*(2), 157-167.

Tervaniemi, M., Ilvonen, T., Sinkkonen, J., Kujala, A., Alho, K., Huotilainen, M., & Näätänen, R. (2000). Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence. *Neuroscience Letters, 279*(1), 29–32.

Tervaniemi, M., Jacobsen, T., Röttger, S., Kujala, T., Widmann, A., Vainio, M., Näätänen, R., & Schröger, E. (2006). Selective tuning of cortical sound feature processing by language experience. *European Journal of Neuroscience*, *23*(9), 2538-2541.

Thomas, D. G., Crow, C. D. (1994). Development of evoked electrical brain activity in infancy; in Dawson G, Fischer KW (eds): *Human Behavavior and the Developing Brain* (S. 207-231). New York: Guilford Press.

Tiitinen H., May, P., Reinikainen, K., Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature, 327*, 90-92.

Torgesen, J.K., Rose, E., Lindamood, P., Conway, T., & Garvan C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91,* 579-594.

Toscano, J. C., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science, 34,* 434–464.

Tremblay, K., Kraus, N., & McGee, T. (1998). The time-course of auditory perceptual learning: Which comes first, the chicken or the egg. *NeuroReport*, *9*, 3557-3560.

Ulytska, B., Steinbrink, C., Lachmann, T., Christmann, A.C., Linkersdörfer, J., Fiebach, C.J. (submitted). The contribution of spectral and temporal information to vowel length discrimination in German: a mismatch negativity study.

Umbricht, D., & Krljes, S. (2005). Mismatch negativity in schizophrenia: a meta-analysis. *Schizophrenia Research*, *76*(1), 1-23.

Vandermosten, M., Boets, B., Luts, L., Poelmans, H., Golestani, N, Wouters, J. & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences, 107*(23), 10389-10394.

Van Ingelghem, M., van Wieringen, A., Wouters, J., Vandenbussche, E., Onghena, P., & Ghesquière, P. (2001). Psychophysical evidence for a general temporal processing deficit in children with dyslexia. *NeuroReport, 12*(16), 3603–3607.

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

Vihla, M., & Eulitz, C. (2003). Topography of the auditory evoked potential in humans reflects differences between vowels embedded in pseudo-words. *Neuroscience Letters*, *338*(3), 189-192.

Wagner, R.K., & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192-212.

Weiss, R. (1974). Relationship of vowel length and quality in the perception of German vowels. *Linguistics,123*(12), 59–70.

Walker, M.M., Givens, G.D., Cranford, J.L., Holbert, D., & Walker, L. (2006). Auditory pattern recognition and brief tone discrimination of children with reading disorders. *Journal of Communication Disorders, 39*(6), 442-455.

Wang, W., Datta, H., & Sussman, E. (2005). The development of the length of the temporal window of integration for rapidly presented auditory information as indexed by MMN. *Clinical Neurophysiology*, *116*(7), 1695–1706.

Werner, L. A., & Marean, G. C. (1996). *Human auditory development*. Boulder, CO: Westview Press.

Wetzel, N., Widmann, A., & Schröger, E. (2011). Processing of novel identifiability and duration in children and adults. *Biological Psychology*, *86*(1), 39-49.

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., & Ramus, F. (2006). The role of sensorimotor impairments in dyslexia: a multiple case study of dyslexic children. *Developmental Science, 9*(3), 237-255.

Wiese, R. (2000). *The phonology of German*. Oxford: Oxford University Press.

Wijnen, V. J. M., van Boxtel, G. J. M., Eilander, H. J., & Gelder, B. de. (2007). Mismatch negativity predicts recovery from the vegetative state. *Clinical Neurophysiology, 118*(3), 597–605.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, *14*(01), 1-33.

Wimmer, H., Landerl, K., & Frith, U. (1999). Learning to read German: normal and impaired acquisition. In M. Harris & G. Hatano (Hrsg.). *Learning to read and write* (S. 34–50). Cambridge: Psychology Press.

Winkler, I., Czigler, I., Jaramillo, M., Paavilainen, P., & Näätänen, R. (1998). Temporal constraints of auditory event synthesis: evidence from ERPs. *NeuroReport*, *9*(3), 495-499.

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*(3), 357-369.

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*(5), 638-642.

Winkler, I., Lehtokoski, a, Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., et al. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*(3), 357–69.

Wise, W., Olson, R.K., Wise, B., & Ring, J. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72*, 271–304.

Wolff, C., & Schröger, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neuroscience Letters*, *311*(1), 37–40.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language impaired children. *Nature, 387*, 176–178.

Wright, B.A. & Zecker, S.G. (2004). Learning problems, delayed development, and puberty. *Proceedings of the National Academy of Sciences, 101*(26), 9942-9946.

Wunderlich, J. L., & Cone-Wesson, B. K. (2001). Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory evoked potential. *Journal of the Acoustical Society of Amercia, 109*(4), 1526-1537.

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*(1), 175–85.

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-yamada, R., & Näätänen, R. (2009). Training the Brain to Weight Speech Cues Differently : A Study of Finnish Second-language Users of English. *Journal of Cognitive Neuroscience, 22*(6), 1319–1332.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6(1),* 37–46.

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch : moving beyond the dichotomies Neural specializations for speech and pitch : moving beyond the dichotomies. *Philosophical Transactions of the Royal Society, Biological Sciences, 363(1493),* 1087–1104.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: language specific or universal? *Journal of experimental child psychology, 86*(3), 169–193.

Ziegler, J., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science, 12*(5), 732-745.

Zion-Golumbic, E., Deouell, L. Y., Whalen, D. H., & Bentin, S. (2007). Representation of harmonic frequencies in auditory memory: A mismatch negativity study. *Psychophysiology, 44*(5), 671–679.

Curriculum Vitae

Schulische Ausbildung

1997 – 2004	Goethe Gymnasium in Freiburg
2004 – 2006	Friedrich Gymnasium in Freiburg
	Abschluss: Abitur (Note 1,6)

Studium

seit 2012	Beginn der Ausbildung für Psychoanalyse und tiefenpsychologisch fundierte Psychotherapie am Institut für Psychoanalyse und analytische Psychotherapie in Würzburg
seit 2011	Promotion an der Goethe Universität Frankfurt in der Abteilung der Neurokognitiven Psychologie bei Prof. Dr. Christian Fiebach zum Thema „The contribution of spectral and temporal information to vowel length perception in German in children and adults"
2011	Abschluss des Studiums der Psychologie (Note 1,4)
2010	Diplomarbeit bei Prof. Dr. Andrea Kübler mit dem Thema „Auswirkung von Stimmung und Intuition auf die Lernleistung, im über sensomotorische Rhythmen gesteuerten, Brain Computer Interface"
2007 – 2011	Studium der Psychologie an der Universität Würzburg. Zusätzlich zum Fach Klinische Psychologie Vertiefung des Fachs: Verhaltensanalyse und Verhaltensregulation

Nebentätigkeiten

2011	Sechswöchiges Praktikum an der Klinik und Poliklinik für Psychiatrie, Psychosomatik und Psychotherapie Würzburg
2008	Tätigkeit am Lehrstuhl II als Studentische Hilfskraft bei Dr. Sascha Topolinski
2008	Sechswöchiges Forschungspraktikum am Institut für Psychologie, Lehrstuhl II bei Prof. Dr. Roland Deutsch