The role of stimulus complexity in auditory research of speech and non-speech on the behavioral and electrophysiological level

Vom Fachbereich Sozialwissenschaften
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Philosophie (Dr. phil.)
genehmigte

D i s s e r t a t i o n

vorgelegt von

Dipl.-Psych. Corinna Anna Christmann
aus Frankenthal

Tag der mündlichen Prüfung:	22.01.2014
Dekan:	Prof. Dr. Thomas Schmidt
Vorsitzender:	apl. Prof. Dr. Maria Klatte
Gutachter:	1. Prof. Dr. Thomas Lachmann
	2. apl. Prof. Dr. Stefan Berti

D 386
(2014)

Für meine Eltern

Contents

List of Figures

List of Tables

Abbreviations

ANOVA	analysis of variance
BF	bands of formants
cg	control group
cla	clarinet
d’	discrimination index
df(factor)	degrees of freedom of factor
df(error)	degrees of freedom of error
dB	decibel
dg	dyslexic group
EEG	electroencephalogram
ERP	event related potential
F0	pitch
F1	first formant
F2	second formant
fMRI	functional magnet resonance imaging
F_R	rotation frequency
Hz	hertz
ISI	inter-stimulus interval
ITI	inter-trial interval
LVC	low pass filtered vowel center stimulus
MMN	mismatch negativity
µV	microvolt
ms	milliseconds
RAN	rapid automized naming
PET	positron emission tomography
RT	reaction time
RVC	spectrally rotated vowel center stimuli
s	seconds
sax	saxophone
SD	standard deviation

sin	sinus
SOA	stimulus onset asynchrony
VC	vowel center stimuli
*	significant on the 5% level
**	significant on the 1% level

Preface

The work outlined in this dissertation was carried out in the Department of Cognitive and Developmental Psychology, TU Kaiserslautern, over the period from June 2011 to January 2014. During this period, I was a research assistant for a project funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), called „The influence of temporal and spectral stimulus features on the processing of German vowels and complex tones in developmental dyslexia: Behavioural and fMRI experiments (principal investigators: Claudia Steinbrink, Axel Riecker & Thomas Lachmann, grant number: STE 1699/2-1)". The research reported in Chapter 2 and 3 was conducted in the context of this DFG project.

Chapter 1:

The models of speech perception

Speech is one of the most complex acoustic signals in our daily environment (Saberi & Perrott, 1999). Therefore, it is hardly surprising that numerous studies have addressed the issue of how speech is processed by the human brain. There are two classes of models for speech perception (Zatorre & Gandour, 2008). The first class, the so-called *domain specific models*, assumes that the speech signal by-passes the normal acoustic pathway. As a result, speech is processed differently compared to non-speech, even at very early stages. One prominent example for this kind of theory is the motor theory of speech perception (Liberman & Mattingly, 1985) which states that the objects of speech perception are not the sounds per se but instead the intended phonetic gestures of the speaker. The second class consists of *cue specific models* which question the specificity of speech sounds (see Diehl, Lotto, & Holt, 2004, for a review). *Cue specific models* assume that speech is processed by the same neural network which is responsible for general auditory processing. These models explain differences in processing of speech and non-speech sounds in terms of low-level features (e.g. temporal and frequency resolution), as speech and non-speech sounds are mostly incomparable concerning their physical properties.

Ongoing research within the various fields of psychology and neuroscience continue to address these issues of distinction within speech. Examples of some of these issues include: categorical perception (Eimas, 1963; Miller et al., 1976; Whalen & Liberman, 1987), hemispheric asymmetries in the auditory cortices for speech and non-speech sounds (e.g. Sorokin, Alku, & Kujala, 2010 for EEG; Binder, Frost, Hammeke, Cox, Rao, & Prieto, 1997 for fMRI; Belin, Zilbovicius, Crozier, Thivard, Fontaine, Masure, & Samson, 1998 for PET), the processing of speech and non-speech sounds in Aphasia (e.g. Aaltonen, Tuomainen, Laine, & Niemi, 1993), specific language impairment (SLI, e.g. McArthur & Bishop, 2005) and developmental dyslexia (e.g., Lachmann, Berti, Kujala, & Schröger, 2005). Despite the large number of studies, there is no sufficient answer to the question of which class of models is more suitable in which context.

Within this thesis both classes of models will be taken into consideration. The *cue specific models* explain differences between the processing of speech and non-speech sounds as a function of their physical properties, particularly regarding temporal and spectral ones. If this assumption is true, speech and non-speech sounds would be processed in the same way whenever the physical properties of both stimulus classes are identical. This is why the first goal of this thesis is the creation of a stimulus set including speech and non-speech stimuli which are comparable concerning their physical properties. This approach presents specific challenges concerning the required stimuli, as the complex spectro-temporal pattern of an original speech sound has to be imitated. Therefore, the stimulus set must be in line with two requirements:

The first requirement involves the complexity of the speech and non-speech stimuli. In this thesis, complexity is defined as the number of different frequencies within the sound at a given time point. The complexity of speech and non-speech sounds has yet to be comparable or controlled for in most studies (but see Scott & Wise, 2004) and therefore, it remains unclear whether stimulus complexity might moderate the differences between the processing of speech and non-speech sounds. This is why, non-speech stimuli with either the same or lower complexity compared to speech are used in this thesis to control for stimulus complexity.

The second requirement for the stimulus set concerns the fact that the perception of speech is associated with the extraction of both temporal and spectral features. However, many of the studies have focused on the processing of only either the temporal or the spectral feature. To circumvent this limitation, a recently developed paradigm in the context of developmental dyslexia which enables the investigation of the processing of spectral and temporal features within one phoneme category (Groth, Lachmann, Riecker, Muthmann, & Steinbrink, 2011; Steinbrink, Groth, Lachmann, & Riecker, 2012; Steinbrink, Klatte, & Lachmann, in preparation) will be used for creating the speech sounds. In this thesis, this paradigm will be called the “German vowel length discrimination paradigm”. The creation and the properties of the stimulus set (speech and non-speech stimuli) will be described in further detail in Chapter 2. These methodological questions are very important in the context of the *cue specific mod*els of speech perception, as they provide a more suitable way to control for the influences of stimulus features and therefore, they enable to test the models.

The following chapter (Chapter 3) is application-orientated while maintaining the focus on the specificity of speech and the role of stimulus complexity during the comparison of speech and non-speech sounds. It is still unclear to date whether the auditory impairments which are regularly found in children and adults with the diagnosis of developmental dyslexia are restricted to speech sounds. Many studies failed to detect general auditory deficits in dyslexia. One reason for these findings could be that these studies used non-speech stimuli which were incomparable to the speech-like ones. Therefore, the same paradigm as introduced in Chapter 2 will be used to compare the discrimination performance for speech and non-speech stimuli in a group of dyslexic adults. Importantly, stimulus complexity will be controlled for in this paradigm. This approach circumvents the above mentioned short coming of prior research and adds to the understanding of the causes of developmental dyslexia.

According to the *domain specific models* of speech perception, speech should not only be processed in a special way in the dyslexic brain. Differences between the processing of speech and non-speech are also expected in healthy children and adults. The aim of the next chapter (Chapter 4) is to test both the *cue specific* and the *domain specific models* in an EEG study with a sample of healthy adults. The *domain specific models* assume that the differences in processing between speech and non-speech are observable even at early stages of auditory processing. One suitable way to investigate the early processes of the auditory system is via an EEG component called the mismatch negativity (MMN; Näätänen, Gaillard, & Mäntsyalo, 1978). According to the *domain specific models*, the differences between speech and non-speech sounds can already be observed in the early stages of processing. Contrary to this, no differences between the speech and non-speech sounds are predicted by the *cue specific models*, if both stimulus types are identical concerning their physical properties. Additionally, *cue specific models* would assume that non-speech sounds with lower complexity would not be processed in the same way as other stimulus types with higher complexity. This is the first time that the MMN elicited by speech and non-speech sounds are compared while controlling for stimulus complexity by using spectrally rotated vowels.

In the last experiment (Experiment 4), the role of harmony is investigated by comparing spectrally rotated tones with tones. Vowels and tones both show a harmonic structure while this is not the case for spectrally rotated tones and spectrally rotated vowels. According to

the *cue specific models*, harmony could moderate differences between the processing of vowels and spectrally rotated vowels. If this assumption is right, the harmonic structure of the stimuli should also result in differences in the processing of tones and spectrally rotated tones, whereas no differences are expected according to the *domain specific models*.

A general discussion is presented in Chapter 5, involving a short summary of the main results, their impact and some short comings of the experiments. In the end, an outlook on the requirements of future research will be given.

Chapter 2:

Creating an optimal non-speech analogue to German vowels

> "Vowels are the eyes through which words look at you.
> A word that has lost its vowels has become blind. Yes, blnd."
>
> Rainer Kohlmayer

Vowel length in German

There are seven pairs of monophthongs (from the Greek word monóphthonogos with monós meaning "single" and phthóngos meaning "sound"; Liddell & Scott, 1996) within the German language (Lühr, 1993): /iː/ - /ɪ/, /yː/ - /ʏ/, /uː/ - /ʊ/, /eː/ - /ɛ/, /øː/ - /œ/, /oː/ - /ɔ/, and /aː/ - /a/. Each pair consists of a lax (or short) and a tense (or long) version of the respective vowel (Kohler, 1977; Moulton, 1962; Wiese, 2000). They can be presented within a trapezium which represents the height of the tongue in the vertical dimension and the tongue's frontness in the horizontal dimension during the production of the vowel (Speyer, 2007). Figure 1 illustrates the vowel trapezium, including the fourteen German monophthongs.

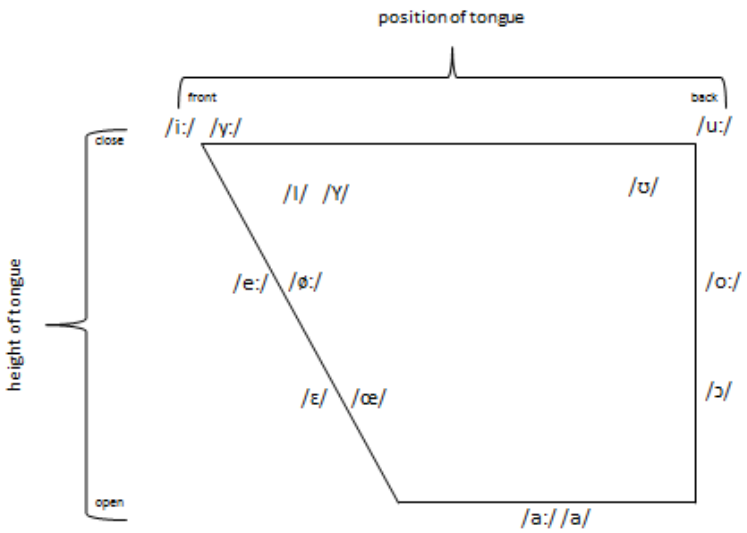

Figure 1: The vowel trapezium of the German language, representing the height of the tongue in the vertical dimension and the tongue's frontness in the horizontal dimension during the production of the vowel (adapted from Mangold, 2005).

In German, short and long vowels do not only differ with respect to their temporal duration, but also concerning their spectral quality. The spectral information of a sound can be illustrated within a spectrogram, showing the strength of each frequency at each time point. Figure 2 depicts the spectrogram of an /i:/ produced by a male speaker.

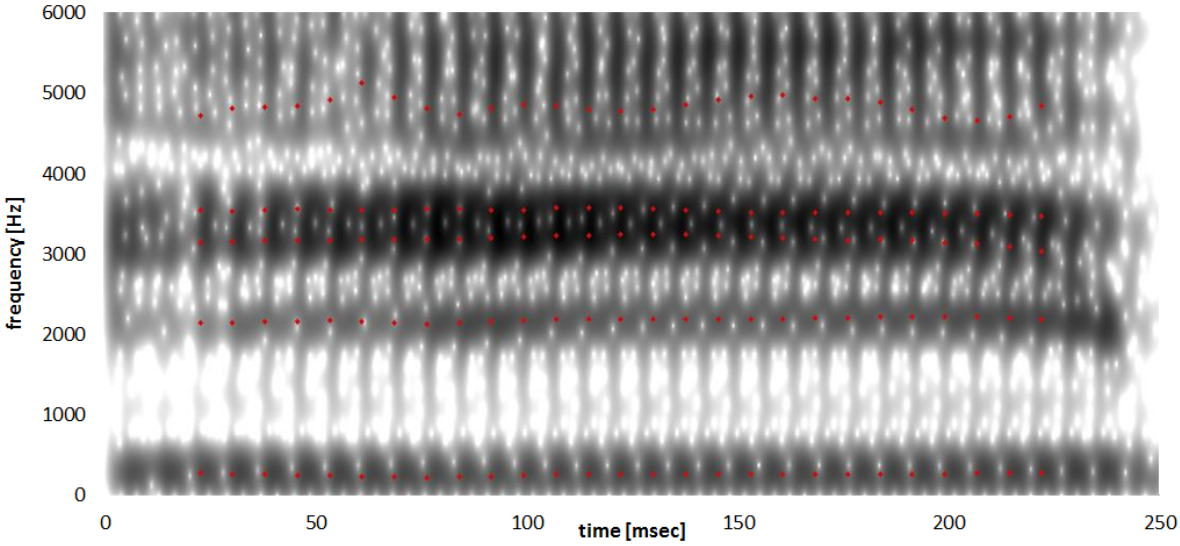

Figure 2: Spectrogram of the vowel /i:/. Time is displayed on the x-axis, frequency along the y-axis. The first five formants are marked by red dots.

There are several frequencies concurrent at each time point, each differing in intensity. The shade of grey represents the intensity of each frequency, getting darker with increasing intensity. The horizontal dark bands in the signal are called formants (Carroll, 2004). The formants are the frequencies with the highest intensity at a given time point. The first five formants are marked by red dots in Figure 2. They can be measured by carrying out a Fourier analysis (Bregman, 1995). As a result, the intensity of each frequency is provided at a given time point. The formants are the extrema of this function (Pfister & Kaufmann, 2008). The result of a Fourier analysis of the vowel /i:/ at 25ms is illustrated in Figure 3. The region of frequencies with a power of at the most three dB beneath the power of the formant is defined as the bandwidth of the formant (Fant, 1960). Formants are systematically changed by moving the articulatory organs. The first two formants are essential for the correct identification of the vowel (Nawka & Wirth, 2008).

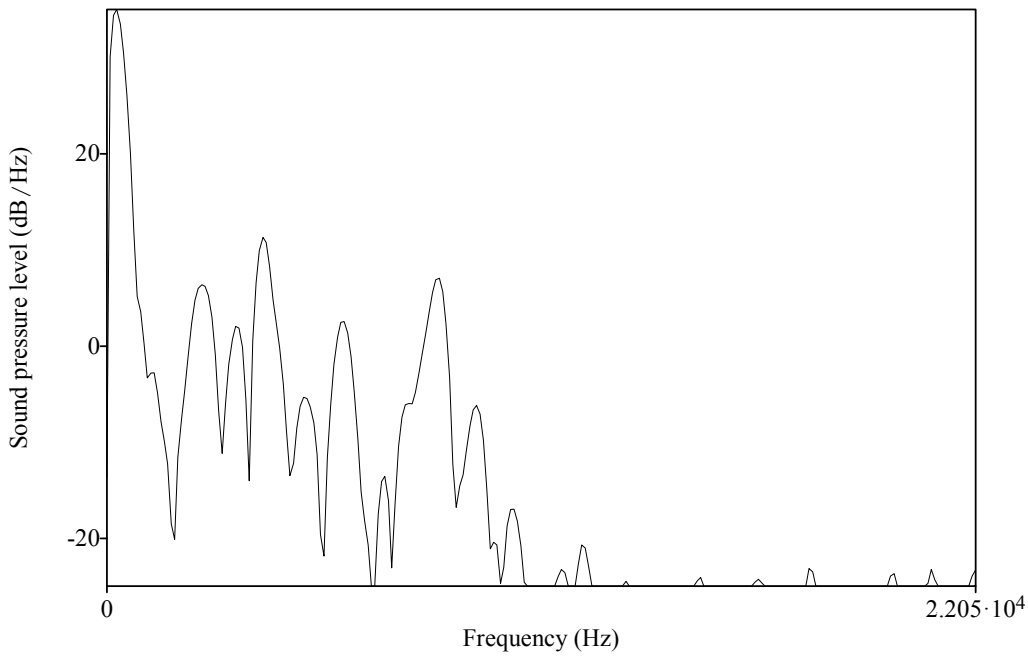

Figure 3: Result of a Fourier analysis of the vowel /i:/ at 25ms. Frequencies are arranged along the x-axis and the sound pressure level is illustrated on the y-axis. The formants are the local extrema of the function.

Sendlmeier and Seebode (2006) identified the first two formants of all German monophtongs, using naturally spoken words produced by 58 female speakers. Their results are presented in Figure 4. The ellipses represent the area of the observed values for the first (F1) and second formant (F2).

Several studies in the late 1970s and 1980s dealt with the question of which of the two properties (temporal/durational or spectral quality) might be more important for the identification of German vowels. In these studies, the vowels were embedded within words. The monophtongs were manipulated in their temporal duration by extending or reducing their steady state parts. There are some observations supporting the special role of temporal cues: when /e:/, /o:/, and /a:/ are shortened in duration, they are judged as /ɪ/, /ʊ/, and /a/ (Heike, 1970; Heike, 1971) and a stretched /ʊ/ is perceived as /u:/ (Sendlmeier, 1981). Becker (1998) summarized studies with synthetically produced vowels and technically manipulated vowels with reduced vowel duration. He came to the conclusion that vowel perception changes when temporal length is decreased: shortened /e:/, /i:/, and /o:/ are perceived as /ɪ/, /ʏ/, and /ʊ/. On the other hand, Sendlmeier (1981) reported that a shortened /y:/ is still judged as /y:/. As nearly the complete steady state part of the vowel

was removed, the author concluded that the main information must lie in the spectral pattern of this vowel and not in its temporal duration. It is also apparent that the reduced vowels were not perceived as the short version of the original vowel, but as the short version of another vowel type in most cases.

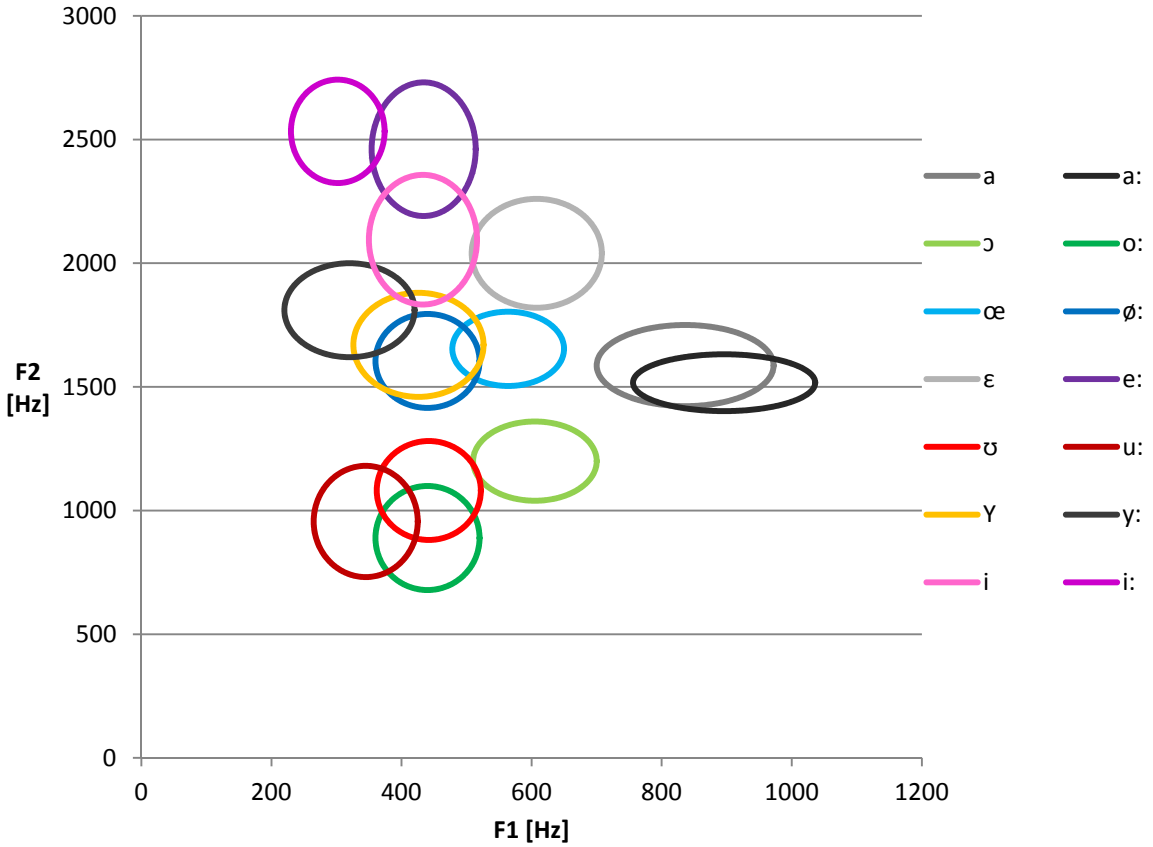

Figure 4: Positions of the first two formants for the German vowels spoken by German females (adapted from Sendlmeier & Seebode, 2006). The frequency of the first formant (F1) is displayed on the x-axis, the frequency of the second one (F2) on the y-axis.

Strange and Bohn (1998) chose another approach to identify information which is more noticeable for the correct identification of vowels. In their first condition, they used silent center stimuli, in which the steady state part of the vowels is silenced. In the second condition, vowel center (VC) stimuli which consisted only of the steady-state part of the vowel were presented. In both conditions, durational cues were still available and the participants performed reasonably well. Error rates ranged from 0 to 34% in the silent center condition and from 0 to 41% in the vowel center condition. When asked to identify silent center and vowel center stimuli with fixed duration, participants' performance dropped

dramatically, underpinning the importance of durational information for vowel perception. Nevertheless, this drop was not observed for all monophtongs. Error rates were especially high for /e:/, /ø:/, /o:/, and /a:/. Contrary to this, performance was not reduced for the high vowels /i:/, /y:/, /u:/ and most of the short ones.

Considering these heterogeneous results, it seems unlikely that it is possible to rely only on either temporal or spectral information to identify German vowels correctly. Bennet (1968) proposed that the impact of the temporal information is inversely proportional to the difference between the spectral properties of the vowel pair, meaning that participants tend to rely on temporal cues especially when there is hardly any difference in quality. The difference of quality is illustrated in Figure 4: Vowels represented by the ellipses that are close together are harder to discriminate than those with distant ellipses. Weiss (1974) was able to show that the relative importance of temporal versus spectral cues decreases with increasing vowel height. This relation is illustrated in Figure 5. The durational difference between /a:/ and /a/ (low vowel pair) is more noticeable than the spectral distinctions (Ungeheuer, 1969), whereas /i:/ and /ɪ/ (high vowel pair) differ mainly with regard to their spectral information.

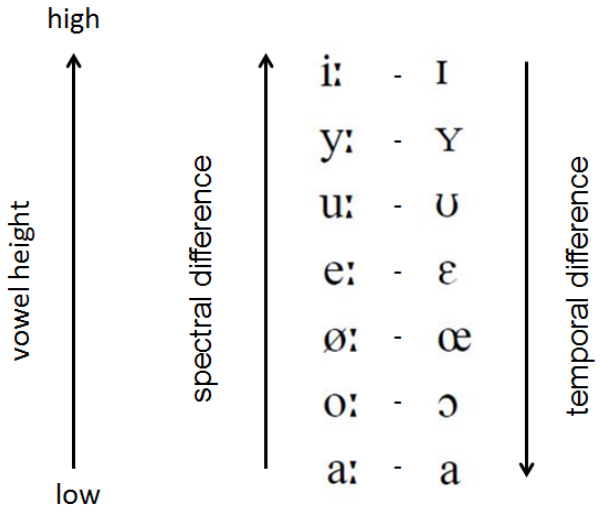

Figure 5: The influence of vowel height on the relative impact of spectral and temporal information during the identification of German vowels. The temporal difference decreases with increasing vowel height. Contrary to this, spectral information is more salient in high vowels compared to low ones.

The German vowel length discrimination paradigm

Based on the finding that auditory discrimination was found to be impaired in some dyslexic children and adults, Groth and colleagues (2011) compared a group of dyslexic adults to a matched control group with the so-called German vowel length discrimination paradigm. This term will be adopted in this thesis. The goal of the above mentioned study was to examine whether dyslexic adults are able to detect short durational differences. They had two conditions of a same-different discrimination task: the different pairs in the spectro-temporal condition comprised the seven lax-tense vowel pairs. These stimuli were naturally spoken and embedded within two pseudo words, /fVp/ and /nVp/ (V = vowel, e.g., /fap/). As a result, each stimulus pair did not only differ with respect to vowel duration, but also with regard to its spectral information. The temporal duration of the steady state part was measured for each vowel pair. The results of this analysis (see Table 1) are concurrent with the model presented by Weiss (1974), as the difference in temporal length decreases as vowel height increases. Neither the dyslexic adults nor the members of the control group had any difficulties to discriminate these spectro-temporal contrasts.

Table 1: Results of the durational analysis of the steady state part of the vowels used by Groth and colleagues (2011). The durational difference decreases with increasing vowel height.

Vowel pair	Duration of the vowel [ms]		Difference [ms]
	Long vowel	Short vowel	
/a:/ -/a/	142	75	67
/o:/ - /ɔ/	128	75	53
/ø:/ - /œ/	121	70	51
/e:/ - /ɛ/	110	66	44
/u:/- /ʊ/	102	57	45
/y:/- /ʏ/	98	53	45
/i:/ - /ɪ/	91	51	40

In their second condition, the so-called temporal condition, Groth and colleagues (2011) removed the spectral contrast. There were two ways to achieve this goal. The first way was to compare the tense vowel of each pair to a shortened version of itself, providing the same spectral information. The second one was to compare the lax vowel of each pair to a

lengthened version of itself. There was no systematical difference between two versions of the temporal contrast. Interestingly, performance dropped in both groups with increasing vowel height. This means that discrimination performance was better for lower vowel pairs, like /a:/ - /a/ compared to higher ones, like /i:/ - /ɪ/.

The same paradigm was used to test primary school students (mean age: 9 years) with and without diagnosed developmental dyslexia in a follow-up study (Steinbrink, Klatte, & Lachmann, in preparation). In this experiment, the authors chose only three vowel contrasts: /i:/ - /ɪ/, /o:/ - /ɔ/, and /a:/ - /a/. With regard to vowel height /i:/ - /ɪ/ represents the high extreme, whereas /o:/ - /ɔ/ and /a:/ - /a/ form the low extreme. The authors were able to replicate the results of the spectro-temporal and the temporal condition of their former study for the children of the control group. These children did not have any problems discriminating vowel length when both the temporal and spectral information of the contrast were available. Comparable to the results of the previous study, performance decreased when the spectral information was removed (temporal condition). This drop was also associated with vowel height, as the drop was highest for the vowel pair /i:/ - /ɪ/. In this experiment, a third condition was also included in which the durational contrast of each vowel pair was removed (spectral condition). This was realized by comparing an originally tense vowel to the lengthened lax vowel and by comparing the originally lax vowel to the shortened tense vowel, respectively. These stimuli were equally long, but differed with regard to their spectral information. Therefore, it is called the spectral condition. As there is hardly any spectral difference between /a:/ and /a/, the deletion of the temporal contrast is expected to lead to a reduced discrimination in performance, which is what the authors observed in both groups. Contrary to this, performance was not impaired for the /i:/ - /ɪ/ and /o:/ - /ɔ/ contrasts, as these vowels can also be distinguished by comparing their spectral quality.

Matching the complexity of non-speech stimuli to German vowels

The major goal of this chapter is the creation of non-speech stimuli which show the same physical properties as German vowels. However, speech is the most complex acoustic signal in our daily environment (Saberi & Perrott, 1999). Therefore, it is a big challenge to create a non-speech signal with comparable properties, especially with comparable complexity. Scott and Wise (2004) gave a detailed overview about which non-speech stimuli are normally used when speech processing is compared to the basic acoustic processing in experiments applying imaging techniques. The following list of non-speech stimuli is an extended summary of these stimuli, as material of behavioral experiments and EEG studies is also included. The aim of the present chapter is to find an appropriate non-speech analogue for vowels. This is why the properties of every stimulus type will be discussed briefly. The stimulus types are ordered according to the increase in stimulus complexity.

Single sinusoidal tones

Single sinusoidal tones, also called sine waves, are frequently used in auditory experiments that compare the processing of non-speech to speech sounds (e.g. behavioral: Fujisaki, Nakamura, & Imoto, 1975; Jones & Macken, 1993; EEG: Aaltonen, Tuomainen, Laine, & Niemi, 1993; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998a; Uwer, Albrecht, & Suchodoletz, 2002; fMRI: Binder et al., 1997). They are composed of a single frequency which remains constant over time (Eichler, 2011). Therefore, they show the least possible physical complexity of acoustic sounds. Every acoustical signal can be decomposed into a couple of sine waves at a given time point by applying the Fourier transformation (Carstensen, 2004).

One advantage of sine waves is that they can be created without much effort. One disadvantage is that, although differences of only 1Hz can be detected by the human ear (Fastl & Zwicker, 2007; Hellbrück & Ellermeier, 2004), pitch discrimination is harder for single sine waves compared to non-speech sounds containing several frequencies (Sidtis, 1980) or compared to vowels (Flanagan, 1958), which show a complete spectrum of frequencies. Moreover, in contrast to vowels, sine waves do not evoke the perception of timbre, as they do not contain any harmonic overtones (Bayerdörfer, 2002). Another difference between sine waves and vowels is that sine waves are reasonably artificial and rarely found in our

daily environment (Pollmann, 2008). This is why sine waves cannot be used as an appropriate non-speech analogue in the context of *cue specific models* of speech perception.

Multiple sine waves

Stimulus complexity can be increased by combining several sine waves. In a harmonic sound, higher frequencies are in a ratio of a whole number compared to the lowest frequency. These higher frequencies are called overtones or harmonics. The number and position of the overtones modulate the timbre of a sound (Hagendorf, Krummenacher, Müller, & Schubert, 2011). The position of the fundamental frequency determines the perceived pitch (Friesecke, 2007). Therefore, two harmonic sounds may differ concerning timbre, pitch, or both. As vowels also show a harmonic structure (Wirth, Ptok, & Schönweiler, 2000), sine waves which include several overtones were used in a couple of studies to compare vowels to a harmonic non-speech analogue (e.g., Dehaene-Lambertz, 2000; Jaramillo et al., 2001). In other studies, the frequencies of the sine waves are based on the formants of the vowel (Čeponiené et al., 2002). Nevertheless, the resulting stimuli show a lower complexity than the vowels, as they do not consist of a full spectrum of frequencies.

Musical sounds

Musical sounds have also been chosen as non-speech analogue (e.g., Benson et al., 2001; Molfese, Freeman, & Palermo, 1975). The production of a vowel is comparable to the production of a tone with a wind instrument (Pahn, 2000). The resulting harmonic structure and complexity of tones and vowels is relatively similar, but not exactly identical. One big advantage of a music stimulus is that it is not artificial. By choosing tones of different instruments, timbre can also be varied (Riggenbach, 2000). Nevertheless, most consonants do not show a harmonic structure (Behrends et al., 2010). Therefore, it is difficult to compare musical sounds to syllables or words. Tones will be used in the fourth experiment of the present thesis (see Chapter 4).

Noise

A sound is called noise whenever it consists of many different frequencies without any harmonics. The most prominent examples are the white, pink and brown noise (Weinzierl, 2008). White nose is characterized by a continuous spectrum in which all frequencies occur

with the same probability and power. In contrast to white noise, high frequencies are weaker compared to low frequencies in pink and brown noise. Therefore, they are consistently rated as to be more pleasant compared to white noise (Derry, 2006; Möser, 2007). As already mentioned, consonants do not show a harmonic structure (Behrends et al., 2010). This is why it is hardly surprising that noise bursts were used as non-speech analogue in numerous studies (e.g., Molfese et al., 1975; Zatorre, Evans, Meyer, & Gjedde, 1992). Due to the fact that noisy sounds can be manipulated in many ways, only some examples will be presented here. Amplitude-modulated noise (e.g. Zatorre, Evans, & Meyer, 1994) is generated by matching the amplitude envelope of the original speech stimulus to the amplitude of white noise (Budinger, Heil, König, & Scheich, 2005). Another approach is to scramble the spectrogram of the original sound within several time windows by randomly intermixing phase and amplitude components in the frequency domain (e.g., Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Belin, Zatorre, & Ahad, 2002; Budinger et al., 2005; Stoppelman, Harpaz, & Ben-Shachar, 2013). The resulting scrambled or signal correlated noise (Schröder, 1968) has the same energy and envelope as the original sound. Another option is to replace only a band of frequencies with noise. This type of stimulus is called noise-vocoded speech (e.g., Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). In summary, noise-like stimuli are a pretty good non-speech analogue for consonants, but not for vowels. Therefore, they were not used in the present thesis.

Reversed speech

Reversed speech (Meyer-Eppler, 1950) is a result of reversing the stimulus across the time domain (Scott & Wise, 2004). The average spectrum and amplitude modulation variation is equal compared to the original stimulus. Therefore, it has been used in a lot of studies as non-speech analogue (e.g., Binder et al., 2000; Hickok, Love, Swinney, Wong, & Buxton, 1997; Howard et al., 1992; Stoppelman et al., 2013). However, Scott and Wise (2004) describe some of the problems which are associated with using reversed speech. The most important disadvantage of reversed speech is that the steady state part of vowels and fricatives is not affected by the transformation at all. This means that these parts of the resulting stimulus evoke a speech-like impression.

Sine wave speech

A completely different approach was presented by Remez and colleagues (1981), as their sounds can be used as the speech and non-speech stimuli within the same experiment (e.g., Benson, Richardson, Whalen, & Lai, 2006; Serniclaes, Sprenger-Charolles, Carré, & Demonet, 2001; Tremblay, Nicholls, Alford, & Jones, 2000). They produced three-tone sinusoidal replica of natural speech sounds. These sinusoidal tones follow the frequencies of the formants, which change over time. The resulting stimulus is perceived as non-speech, until someone is told what the original signal sounds like. From that moment on it is impossible to perceive it as non-speech again. Therefore, it is crucial to start with the non-speech condition in the experiment, or to work with two separate groups, one for the non-speech and one for the speech condition. Although sine wave speech can be perceived as speech, it is less complex than natural speech. Further problems concerning the usage of sine-wave speech are described by Rosen and Iverson (2007). These are for example weird intonation and the missing harmonics of the vowel sounds.

Phonemes of foreign languages

The advantages and disadvantages of using sounds from foreign languages are discussed in great detail by Scott and Wise (2004). The most important disadvantage of sounds coming from foreign languages is that they are still speech. Therefore, they can be used to study unintelligibility, but they cannot be used to investigate how the brain deals with non-speech auditory stimuli. Nevertheless, phonemes of foreign languages play a crucial role to investigate language specific representations in the brain (Näätänen et al., 1997) (see Chapter 4 for details).

Animal sounds and human non-speech sounds

Some studies used animal sounds as non-speech stimuli (e.g., Marcus, Fernandes, & Johnson, 2007). However, this approach does not control for complexity. This potentially confounding factor was circumvented in a study by Neath and colleagues (1993). They used the same stimulus in two conditions: The first experimental group was told that the sound was the syllable /bæ/ spoken by a man, whereas the second group was told that they would hear the bleat of a sheep. Another possibility would be to use human non-speech sounds like laughs, cries, moans, or sighs (e.g., Belin et al., 2002); however, these are normally very

emotional and should therefore only be compared to speech stimuli that express the same emotion. They also do not control for complexity.

Spectrally rotated speech

One effective solution that circumvents most problems of the above mentioned non-speech stimuli was presented by Blesser (1972): the spectral rotation of speech. Starting with a study by Scott and colleagues (2000), spectral rotation is commonly used to compare speech to non-speech in imaging studies (e.g., Abrams et al., 2012; Awad, Warren, Scott, Turkheimer, & Wise, 2007; Lachs & Pisoni, 2004; Obleser, Wise, Dresner, & Scott, 2007; Peelle, Gross, & Davis, 2013; Sabri et al., 2008; Sauter & Eimer, 2010; Scott, Blank, Rosen, & Wise, 2000; Scott, Rosen, Beaman, Davis, & Wise, 2009; Scott, Rosen, Lang, & Wise, 2006; Sjerps, Mitterer, & McQueen, 2011; Spitsyna, Warren, Scott, Turkheimer, & Wise, 2006). The spectral rotation of a speech sound can be conducted with Matlab (version R2011a; Mathworks) using a script provided by Scott and colleagues (2000). The process consists of several steps:

a) Low pass filter

The highest frequency of the speech signal is dependent on the rotation frequency (F_R). Therefore, a low pass filter is used to modify the original speech signal. The cut-off frequency of the low pass filter (F_L) can be calculated with the following formula: $F_L = 0.95 \cdot 2\ F_R$.

The most important frequencies of the speech signal are supposed to lie between 500 and 4000Hz (Wilmanns & Schmitt, 2002). This is why 4000Hz was used as the cut-off frequency for the low pass filter in most studies dealing with spectrally rotated speech (e.g., Davids et al., 2011; Evans et al., 2013; Narain et al., 2003; Okada et al., 2010; Scott et al., 2000; Scott et al., 2009; Scott et al., 2006; Sörqvist, Nöstl, & Halin, 2012; Vandermosten et al., 2011; Vandermosten et al., 2010). One disadvantage of this procedure is that the original speech sound has to be low pass filtered. The intelligibility of the signal is not reduced in this way (Scott & Wise, 2004), but its naturalness could be impaired (Moore & Tan, 2003).

b) Equalizer

As a result of the rotation, high frequencies of the stimulus will become low and low frequencies will become high. As the human auditory system is more sensitive to high compared to low frequencies within the speech signal (Baumann, 2010), the low frequencies of the original speech signal must be reduced in their intensity as they would be too

intensive after the rotation. The solution to this problem is to use a high-pass filter (Byrne et al., 1994) which has been included in the Matlab script by Scott and colleagues (2000).

c) Mirroring at F_R

The next step is to mirror all frequencies at each time point at F_R. The mathematical formula for this procedure is: $\sin(2\pi 2F_R)$.

d) Adjusting the root mean square level

The intensity of the spectrally rotated speech signal is controlled for by matching its root mean square level to that of the original speech signal.

As both stimuli show the same spectro-temporal pattern, their complexity is completely matched. This property is the reason why spectrally rotated speech is supposed to be more suitable as non-speech analogue compared to the other non-speech types presented above (Scott & Wise, 2004). The course of the pitch in the original speech signal is also taken into consideration in the spectrally rotated speech stimulus (Blesser, 1972). This means that intonation is preserved after the spectral inversion, which is for example important for distinguishing between statements and questions. Two speech stimuli with the same pitch will also have an equal pitch after spectral rotation.

Blesser (1972) was able to show that one can learn to understand spectrally rotated speech after intensive training, and so spectrally rotated stimuli should only be used whenever participants do not have any prior experience with this type of stimulus. In his study, participants were asked to discriminate and identify spectrally rotated phonemes. The spectral rotation did not affect the perception of fricatives. This finding can be explained by the fact that fricatives consist of almost all frequencies and all of them nearly show the same intensity. Therefore, the spectral composition is the same before and after the spectral inversion. The identification of spectrally rotated nasals was hardly impaired as well. Spectrally rotated plosives were identified as plosives, but often confused with another phoneme, e.g., the spectrally rotated /p/ was not only perceived as /p/ but also as /t/ and /k/, and vice versa. Figures 6 and 7 show the spectrograms of the syllable /fap/ and of its spectrally rotated counterpart. The phonemes /f/ and /p/ look quite similar before and after the spectral rotation.

The pattern of results of the phoneme identification was also dependent on the vowel in the middle of the word: A spectrally rotated /p/ followed by a back vowel (e.g., /u:/) was more

often identified correctly compared to a spectrally rotated /p/ which was followed by a front vowel (e.g., /i:/). The opposite pattern of results was found for the spectrally rotated /k/.

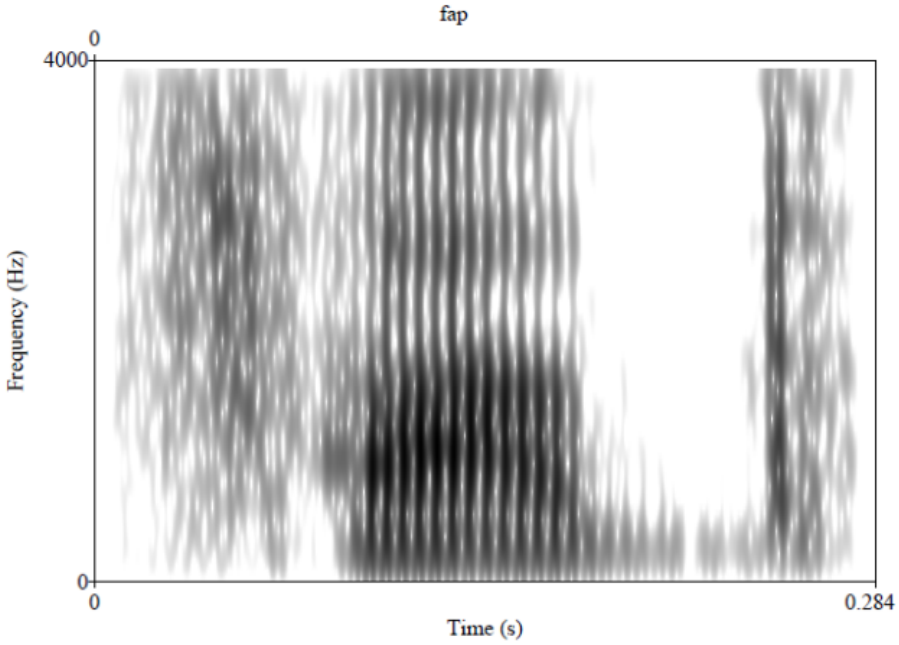

Figure 6: Spectrogram of the syllable /fap/. Time [s] is displayed along the x-axis, frequency [Hz] along the y-axis. Frequencies with higher intensity are illustrated darker.

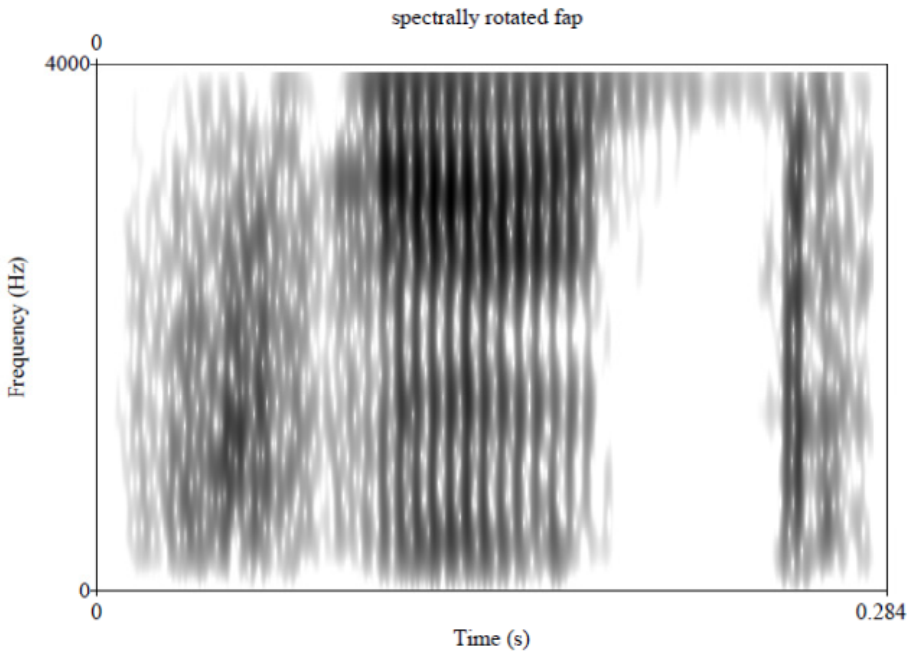

Figure 7: Spectrogram of the spectrally rotated syllable /fap/. Time [s] is displayed along the x-axis, frequency [Hz] along the y-axis. Frequencies with higher intensity are illustrated darker. The phonemes /f/ and /p/ look similar for the syllable and the spectrally rotated syllable.

The discrimination performance for spectrally rotated vowels was extremely accurate. Even before the training, participants achieved 90% correct responses. Identification was much worse. /u:/ was perceived as /i:/ and vice versa for instance. Nevertheless, it was a forced identification task and all vowels were embedded into a word. As previously mentioned, some spectrally rotated consonants were not impaired by the inversion. Therefore, it might be possible that the spectrally rotated vowels are only perceived as speech when being embedded within a word. This assumption is supported by the fact that the identification of the spectrally rotated vowels was highly dependent of the surrounding context (see Blesser, 1972 for details).

In summary, due to the fact that spectrally rotated consonants can be perceived as speech-like sounds, they will not be used in the present thesis. Spectrally rotated vowels which are presented in isolation will be used as non-speech stimuli with the same complexity as German vowels.

However, there is one property in which the spectrally rotated sound is not equal to the original speech stimulus; the harmonic structure of a vowel will not be preserved, as the integral ratio of the frequencies will be disrupted as a result of the transformation. This will be clarified by means of an example with a sinusoidal tone of 700Hz with two harmonic partials of 1400 and 2100Hz. If one choses the standard rotation frequency of 2000Hz the resulting stimulus will consist of the following frequencies:

1) $(2F_R)$ - 700Hz = 3300Hz
2) $(2F_R)$ - 1400Hz = 2600Hz
3) $(2F_R)$ - 2100Hz = 1900Hz

The three tones do not form a harmonic stimulus, as 1900, 2600, and 3300Hz cannot be expressed by the integral ratio of the same fundamental frequency. To test the influence of harmony, an additional experiment (Experiment 4) will be presented in Chapter 4.

Experiment 1

The general goal of this chapter is to provide a complete paradigm which enables the testing of the *domain specific* and the *cue specific models* of speech perception. This will be achieved by considering the following aims:

The speech stimuli used in this thesis will be created following the German vowel length discrimination paradigm, as temporal, spectral and spectro-temporal aspects of speech perception can be investigated within the same stimulus set and within the same phoneme category (Groth et al., 2011; Steinbrink et al., 2012; Steinbrink et al., in preparation). The vowels were originally embedded in CVC syllables in this paradigm. However, consonants have been shown to be hardly impaired by the spectral rotation (Blesser, 1972). Therefore, the aim of this experiment is to modify the paradigm used by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation) to vowel center stimuli. Only two vowel pairs will be used in this thesis: /a/ - /a:/ and /ɪ/ - /i:/. These pairs form the upper and lower extremes concerning vowel height. As a result, the relative impact of spectral and temporal information for the vowel discrimination, which is dependent on vowel height, will be preserved.

The second aim is to expand the paradigm with two non-speech conditions. The stimuli of the first one non-speech class should be comparable to the complexity of the vowel center stimuli. The second non-speech condition is expected to represent non-speech stimuli with lower complexity while maintaining the most important frequencies of the vowels.

As spectrally rotated speech can only be matched to low pass filtered speech, there is a third aim. It is important to find an answer to the following question: Does it make any difference to use low pass filtered vowels instead of the full spectrum with respect to discrimination performance and perceived naturalness?

All stimuli will be presented within a same-different task (see Groth et al., 2011 and Steinbrink et al., in preparation) in order to estimate the difficulty of the temporal, spectral and spectro-temporal contrasts for each stimulus type. The aim is to rule out bottom and ceiling effects in discrimination performance. To prove whether the speech and non-speech stimulus types are really perceived as speech and non-speech, each participant will be questioned about the stimuli after the experiment.

Participants

Twenty-five young adults (14 female) took part in the experiment. The mean age was 21.72 years with a standard deviation of 2.30 years. The age range was 17 to 25 years. All of them were students of the University of Kaiserslautern, except one person who was a trainee. All of them were paid after having completed the experiment. None of them reported impaired hearing. All of them were German native speakers.

Material

Five different stimulus types were used, which will be explained in detail. Two of them were speech-like as they were based on German vowels (vowel center stimuli and low pass filtered vowel center stimuli). The other three stimulus types were non-speech-like with different levels of complexity (spectrally rotated vowel center stimuli and two types of bands of formants).

The name of each stimulus depends on the stimulus type (V = Vowels, L = Low pass filtered vowels, R = spectrally Rotated vowels, B = Bands of formants based on the vowels, BL = Bands of formants based on the Low pass filtered vowels) and the vowel type ("a" for the vowel pair /a/ - /a:/ and "i" for the vowel pair /ɪ/ - /i:/). The last letter describes whether the stimulus is based on the original vowel ("o") or whether the vowel was modified ("m"). The numbers at the end of the name are identical to the duration of the stimulus in milliseconds. For example, vao75 means that it is the vowel center stimulus, based on the vowel pair /a/ - /a:/. The duration of this stimulus is 75ms.

Vowel center stimuli: full spectrum and low pass filtered vowels

These stimuli were based on four naturally spoken vowels: /a/, /a:/, /i:/, and /ɪ/. The vowels were spoken in isolation by a female German native speaker. To obtain only the static spectral information of each vowel, all but the steady state portion was removed. Pitch was kept constant within one vowel pair. The durations of the long and short vowels were chosen following those reported by Groth and colleagues (2011) (see Tables 1 and 2). It is not recommended to cut a vowel within one pitch period, as this would result in an artificial audio impression. As a result, the duration of the vowels did not perfectly match to those of Groth and colleagues (2011), but the deviation did not exceed 3ms. The intensity was kept constant by setting the "scale intensity" in Praat to 75dB (; Boersma, Weenink, 2013).

The PSOLA (Pitch Synchronous Overlap and Add) algorithm of Praat was used to change the length of the vowels without distorting their spectral properties. The short vowel center stimulus was lengthened to the duration of the long one and vice versa. As a result, there were four stimuli for each of the two vowels: the original tense-lax pair (vao75 and vao145 for the vowel pair /a/ – /a:/ and vio51 and vio93 for the vowel pair /ɪ/ – /I:/) and the two modified stimuli (vam75 and vam145 for the vowel pair /a/ – /a:/ and vim51 and vim93 for the vowel pair /ɪ/ – /i:/) (see Figures 10 and 15). This procedure is identical to the one used by Groth and colleagues (2011) with two exceptions: only the two extreme vowel pairs concerning vowel height were used (/a/ – /a:/ and /ɪ/ – /i:/) and there is no change of spectral information within the stimuli, as they are restricted to the vowel center. The first and last five milliseconds of each stimulus were faded with Audition (version CS5.5; Adobe). The duration, pitch (F0), and the first and second formant (F1 and F2) of each vowel center stimulus are illustrated in Table 2. The pitch (F0) and formants (F1 and F2) were established with Praat.

Table 2: Results of the analysis of the vowel center stimuli based on the vowels /a/ (vao75 and vam145), /a:/ (vao145 and vam75), /ɪ/ (vio51 and vim93), and /i:/ (vio93 and vim51). The temporal length in milliseconds, the pitch (F0), and the first two formants (F1 and F2) in hertz (Hz) are provided.

Name	Length [ms]	F0 [Hz]	F1 [Hz]	F2 [Hz]	Modification
Vao75	75	186	792	1302	original short
Vao145	145	186	922	1272	original long
Vam75	75	186	918	1253	shortened
Vam145	145	186	785	1298	lengthened
Vio51	51	194	406	2117	original short
Vio93	93	194	338	2439	original long
Vim51	51	194	325	2416	shortened
Vim93	93	194	415	2128	lengthened

The second speech-like type of stimulus was produced by low pass filtering all vowel center stimuli at 4000Hz. This was carried out in Matlab (version R2011A; Mathworks) using the script provided by Scott and colleagues (2000). The properties of all eight low pass filtered

vowel center stimuli are given in Table 3. The spectrograms are illustrated in Figure 11 for the vowel pair /a/ – /a:/ and in Figure 16 for the vowel pair /ɪ/ – /i:/.

Table 3: Results of the analysis of the low pass filtered vowel center stimuli based on the vowels /a/ (lao75 and lam145), /a:/ (lao145 and lam75), /ɪ/ (lio51 and lim93), and /i:/ (lio93 and lim51). The temporal length in milliseconds, the pitch (F0), and the first two formants (F1 and F2) in Hz are provided.

Name	Length [ms]	F0 [Hz]	F1 [Hz]	F2 [Hz]	Modification
Lao75	75	186	775	1257	original short
Lao145	145	186	770	1192	original long
Lam75	75	186	757	1178	shortened
Lam145	145	186	758	1267	lengthened
Lio51	51	194	401	2130	original short
Lio93	93	194	298	2419	original long
Lim51	51	194	323	2608	shortened
Lim93	93	194	411	2128	lengthened

Spectrally rotated vowels

For each of the eight vowel center stimuli one spectrally rotated counterpart was produced. The whole procedure was carried out in Matlab (version R2011A; Mathworks) using the script provided by Scott and colleagues (2000). The spectrograms are illustrated in Figure 12 for the vowel pair /a:/ - /a/ and in Figure 17 for the vowel pair /ɪ/ - /i:/.

Bands of formants on the basis of vowels and low pass filtered vowels

The last type of stimulus should also be perceived as non-speech, while maintaining the most important information of the speech signal. It is composed only of the first two formants of the vowel including all bandwidth frequencies. To make it more comparable to the formant bands of the vowel, the power of the frequencies in the middle of the bands are highest and decrease towards the two borders. The relative power of the two formants was also considered. All information that is necessary to produce the bands of formants is provided in Table 4.

The two bands were produced separately in Matlab (version R2011A; Mathworks) with a continuous Fourier synthesis on the basis of a Gaussian function with the middle frequency

corresponding to the formant of the vowel and the half width corresponding to the band width of the formant. This function is transformed numerically to the time domain by means of the Fast Fourier Transformation (FFT). As a result one obtains a stimulus with a limited band of frequencies. The middle frequency shows the highest power and the power of the remaining frequencies decrease with increasing distance to the center. The resulting band is very short in duration. In light of this, phase noise is added to the frequency domain in order to lengthen the stimulus to the desired temporal duration.

Table 4: Summary of the most important information for creating the bands of formants based on the vowel center stimuli. The length of the stimulus is comparable to those of the vowel center stimuli. The middle of the two bands is formed by the first two formants, F1 and F2. The relative intensity of the two bands is adapted to the formants' intensity of the vowel center stimuli. The width of the bands corresponds to the bandwidth of the formants.

Name	length [ms]	F1 [Hz]	F2 [Hz]	Difference of intensity between F1 and F2 [dB]	B1 [Hz]	B2 [Hz]
Bao75	75	792	1302	4.06	166	161
Bao145	145	922	1272	1.28	284	225
Bam75	75	922	1272	1.28	284	225
Bam145	145	792	1302	4.06	166	161
Bio51	51	406	2117	16.92	89	124
Bio93	93	338	2439	27.31	262	197
Bim51	51	338	2439	27.31	262	197
Bim93	93	406	2117	16.92	89	124

In the second step, the two bands were mixed together in Audition (version CS5.5; Adobe). The difference in intensity between the two formants was also considered, which is why the first band shows a higher power than the second one.

The spectrograms of the bands of formants based on the vowel center stimuli are illustrated in Figure 13 for the vowel pair /a/ - /a:/ and in Figure 18 for the vowel pair /ɪ/ - /i:/.

The bands of formants based on the low pass filtered vowel center stimuli were created in the same way, based on the values provided in Table 5.

Table 5: Summary of the most important information for creating the bands of formants based on the low pass filtered vowel center stimuli. The length of the stimulus is comparable to those of the low pass filtered vowel center stimuli. The middle of the two bands is formed by the first two formants F1 and F2. The relative intensity of the two bands is adapted to the formants' intensity of the low pass filtered vowel center stimuli. The width of the bands corresponds to the bandwidth of the formants.

Name	Length [ms]	F1 [Hz]	F2 [Hz]	Difference of intensity between F1 and F2 [dB]	B1 [Hz]	B2 [Hz]
Blao75	75	775	1257	3.04	182	242
Blao145	145	770	1192	-2.83	407	195
Blam75	75	770	1192	-2.83	407	195
Blam145	145	775	1257	3.04	182	242
Blio51	51	401	2130	16.43	78	80
Blio93	93	298	2419	28.02	274	99
Blim51	51	298	2419	28.02	274	99
Blim93	93	401	2130	16.43	78	80

The spectrograms of the bands of formants based on the low pass filtered vowel center stimuli are illustrated in Figure 14 for the vowel pair /a/ - /a:/ and in Figure 19 for the vowel pair /ɪ/ - /i:/.

Sinusoidal tones

The stimuli of the demo trials were supposed to be easily discriminable. Therefore, only two sinusoidal tones corresponding to the first two formants of the original vowel pair /a/ – /a:/ and with the same temporal duration were chosen. The properties of the four stimuli are summarized in Table 6.

Table 6: Properties of sinusoidal tones used in the demo trials. The length was matched to the vowel center stimuli of the vowel pair /a/ - /a:/. The tones were composed of two sinusoidal tones corresponding to the first two formants (F1 and F2) of the vowel center stimuli.

Name	Length [ms]	F1 [Hz]	F2 [Hz]
tao75	75	792	1302
tao145	145	922	1272
tam75	75	922	1272
tam145	145	792	1302

Task

All stimuli were presented within a same-different task. Two stimuli were presented sequentially, separated by an inter-stimulus interval (ISI) of 600ms. Participants were asked to decide whether the two stimuli were equal or different. They were instructed to respond as fast and correctly as possible by pressing the correct button out of two: "=" for "same" responses, "≠" for "different" answers. In order to rule out any effects of handedness on reaction time, key assignments were counterbalanced. There was a short practice block with 8 trials to familiarize participants with the task. During these trials, acoustic feedback was given following incorrect responses. During the experimental block no feedback was given. There was no time limit for the participants' responses. The inter-trial interval (ITI) lasted 2000ms in each block. The sequence for a practice trial and for an experimental trial is illustrated in Figures 8 and 9.

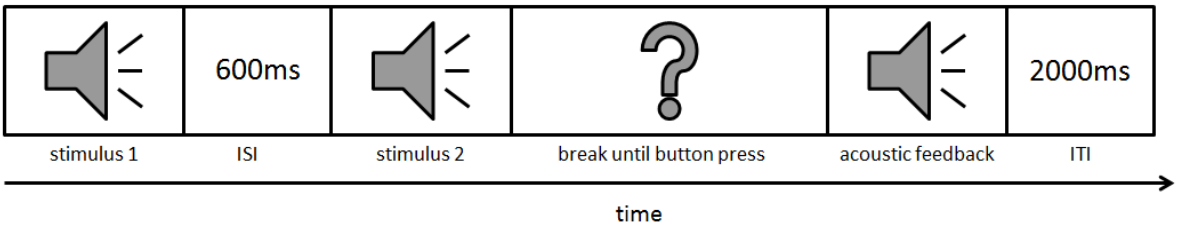

Figure 8: Sequence for a practice trial. Two stimuli were presented sequentially, separated by an inter-stimulus interval (ISI) of 600ms. Participants responded as fast and correctly as possible by pressing the correct button out of two: "=" for "same" responses, "≠" for "different" answers. Acoustic feedback was given following incorrect responses. The inter-trial interval (ITI) lasted 2000ms.

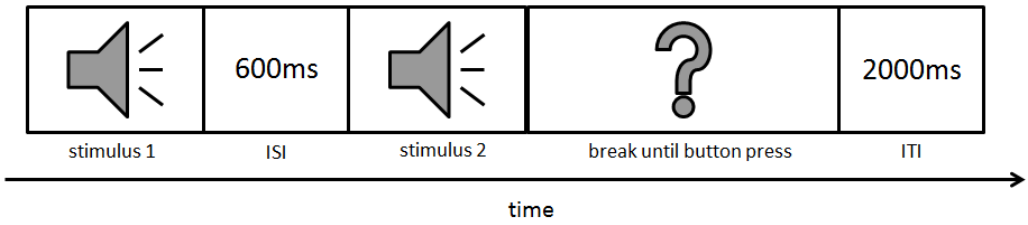

Figure 9: Sequence for an experimental trial of the same-different task. Two stimuli were presented sequentially, separated by an inter-stimulus interval (ISI) of 600ms. Participants responded as fast and correctly as possible by pressing the correct button out of two: "=" for "same" responses, "≠" for "different" answers. The inter-trial interval (ITI) lasted 2000ms. No feedback was provided.

Apparatus

All stimuli were presented with an external soundcard (UGM96, ESI Audiotechnik GmbH, Leonberg, Germany) binaurally via two closed headphones (Beyerdynamic DT 770) with an intensity of 86 dB(SPL), equivalent to 80 dB(A). The intensity was measured with an artificial head (HSM III.0, HEAD acoustics, Aachen, Germany). One headphone was provided for the participant, the other one for the experimenter. The operating system on the laptop was Windows XP. Presentation (version 14.5, Neurobehavioral Systems, Albany, California) was used to control the experimental protocol. All sessions took place in an acoustically shielded room.

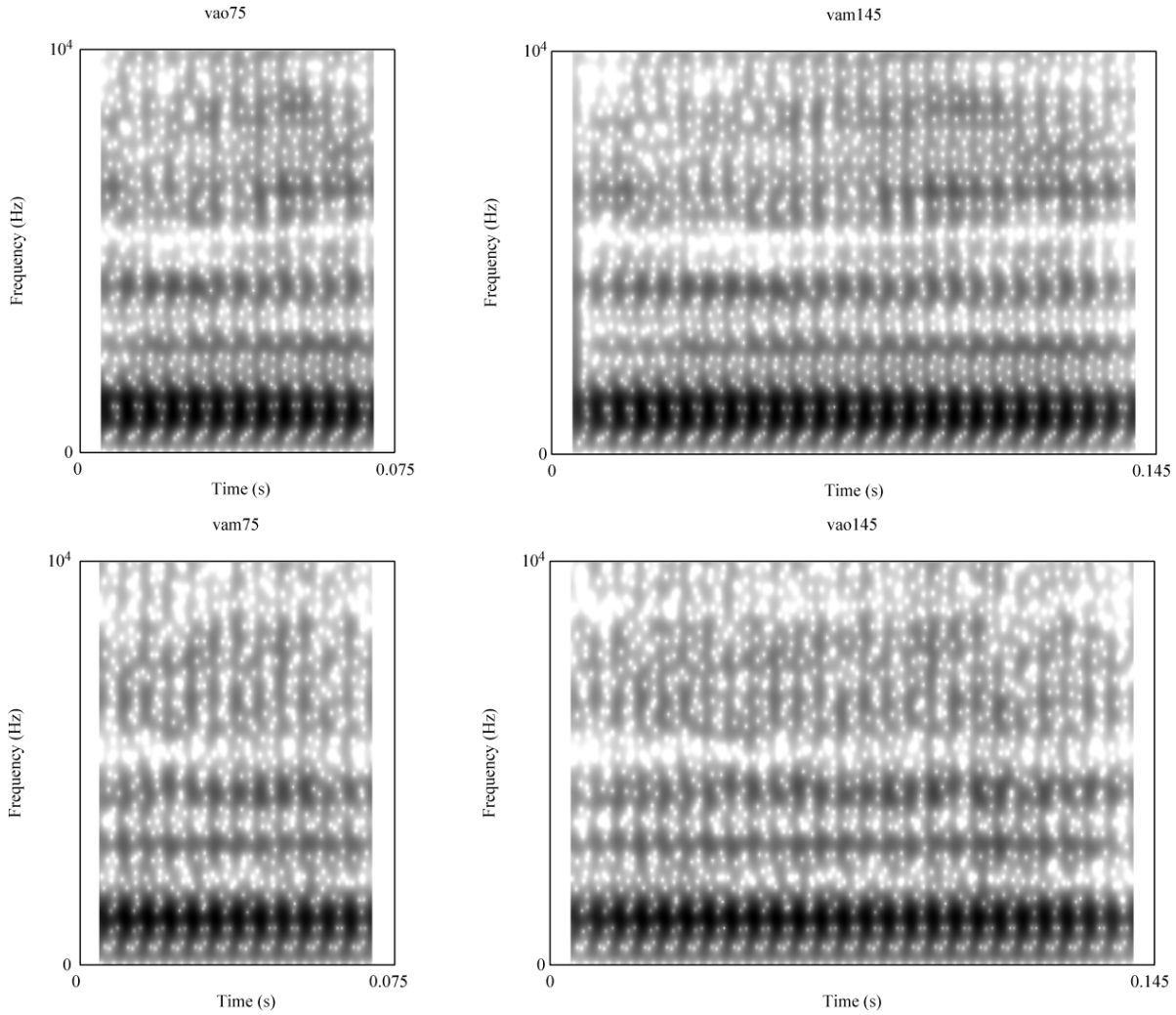

Figure 10: Spectrograms of the four vowel center stimuli based on /a/ - /a:/. Vao75 and vao145 are based on the original lax-tense pair. They differ with respect to both temporal and spectral information. Vam75 is the shortened version of vao145 and vam145 is the lengthened version of vao75.

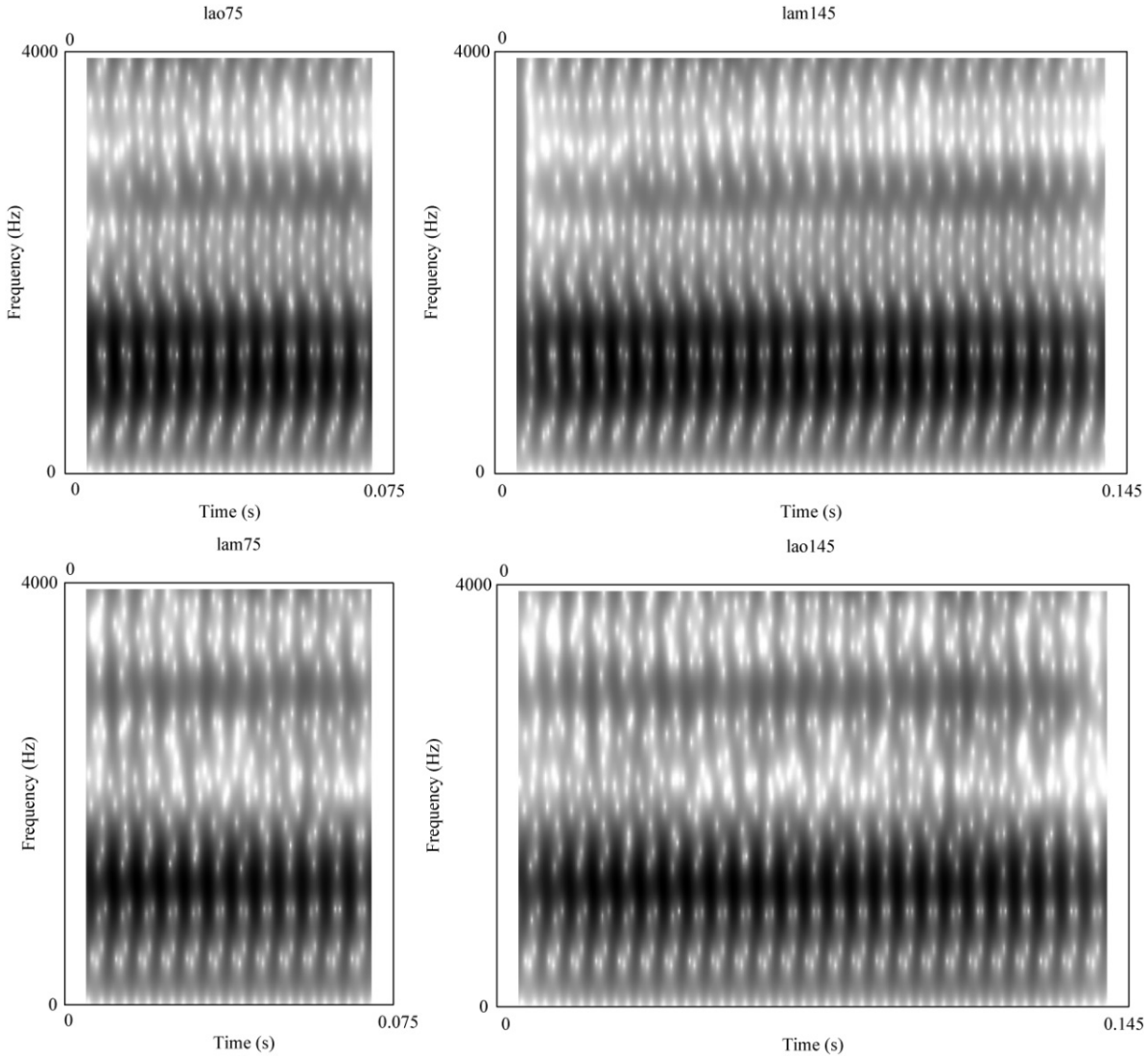

Figure 11: Spectrograms of the four low pass filtered vowel center stimuli based on /a/ - /a:/. Lao75 and lao145 are based on the original lax-tense pair. They differ with respect to both temporal and spectral information. Lam75 is the shortened version of lao145 and lam145 is the lengthened version of lao75.

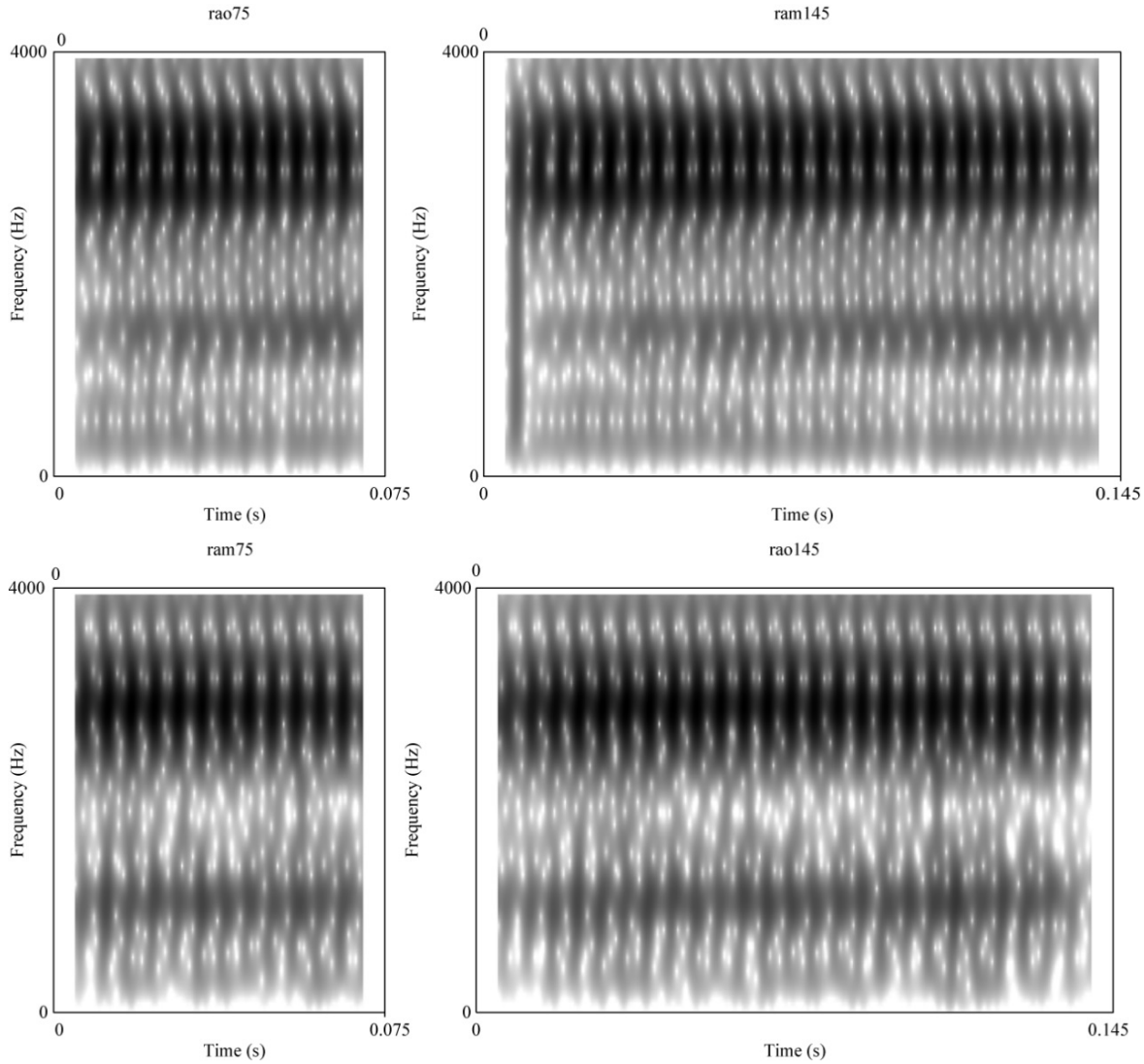

Figure 12: Spectrograms of the four spectrally rotated vowel center stimuli based on /a/ - /a:/. Rao75 and rao145 are based on the original lax-tense pair. They differ with respect to both temporal and spectral information. Ram75 is the shortened version of rao145 and ram145 is the lengthened version of rao75.

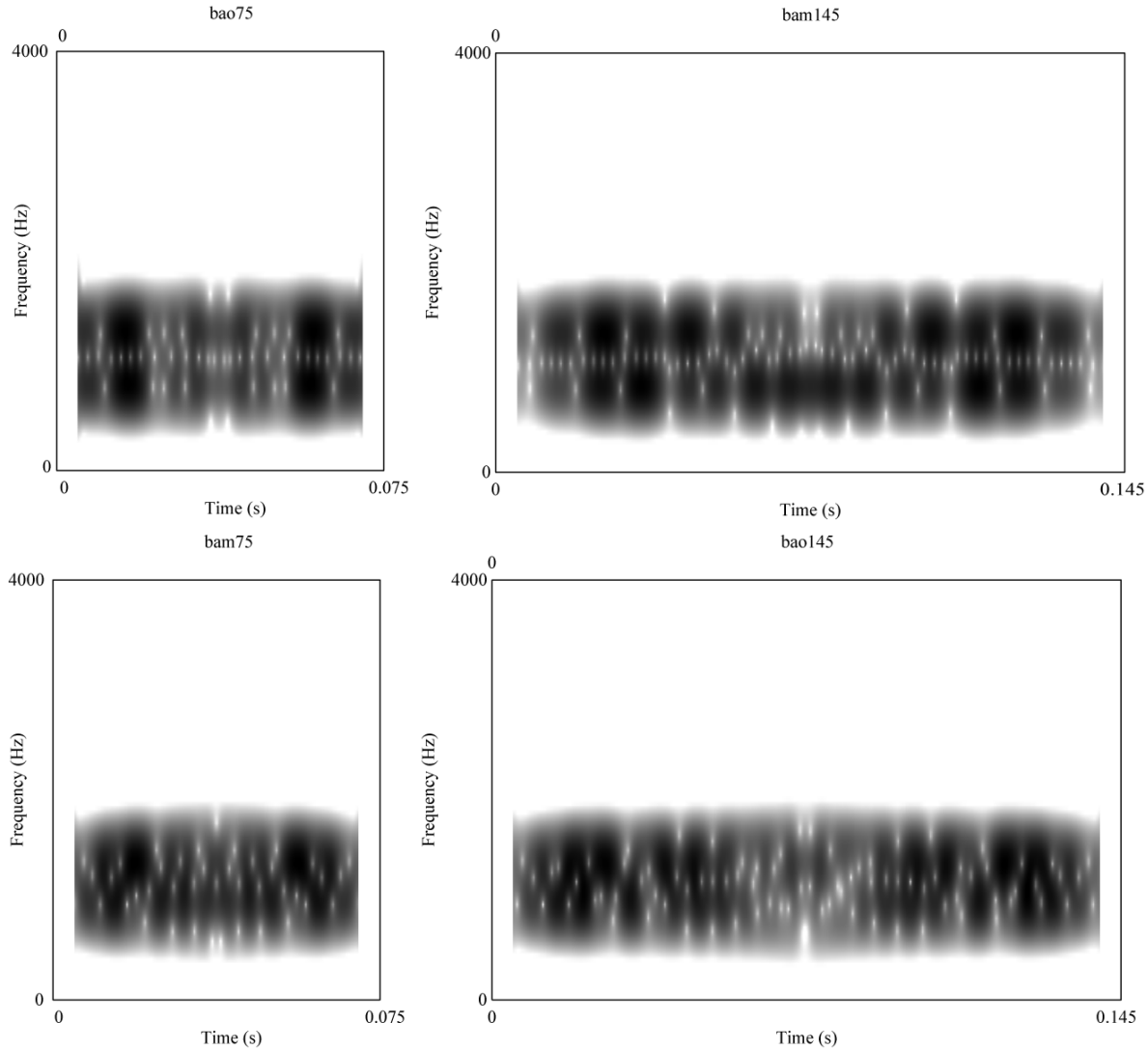

Figure 13: Spectrograms of the four bands of formants based on the vowel center stimuli vao75, vao145, vam75 and vam145.

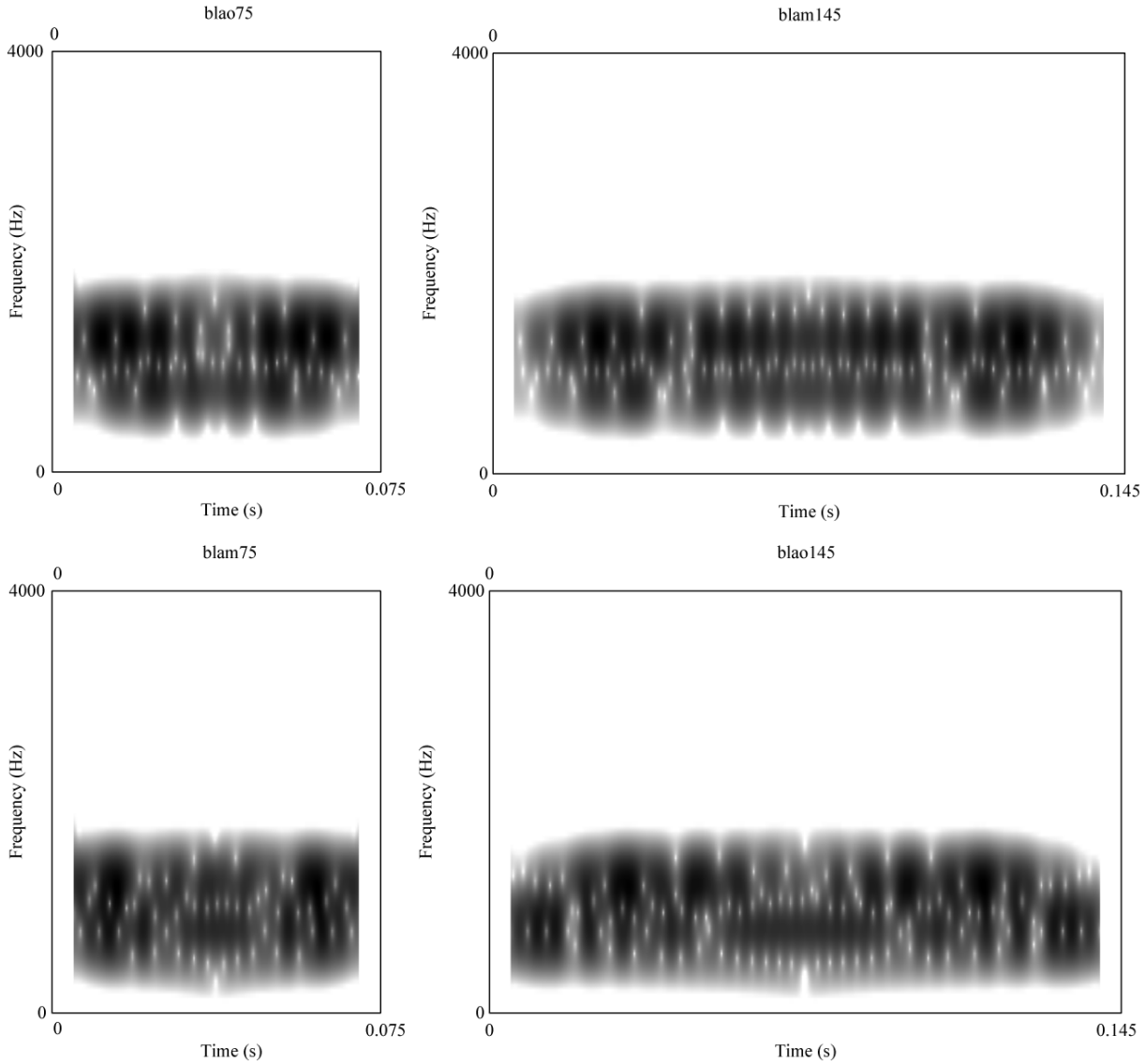

Figure 14: Spectrograms of the four bands of formants based in the low pass filtered vowel center stimuli lao75, lao145, lam75 and lam145.

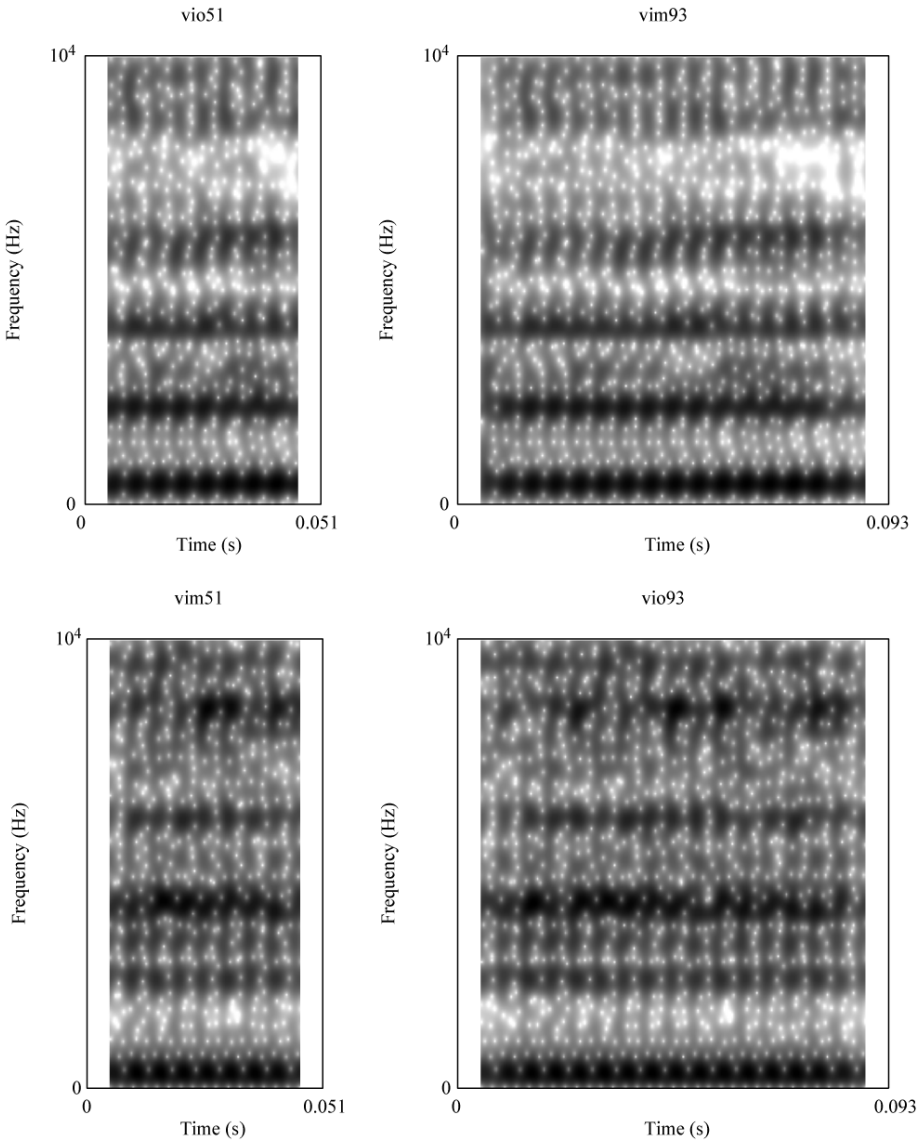

Figure 15: Spectrograms of the four vowel center stimuli based on /ɪ/ - /i:/. Vio51 and vio93 are based on the original lax-tense pair. They differ with respect to both temporal and spectral information. Vim51 is the shortened version of vio93 and vim93 is the lengthened version of vio51.

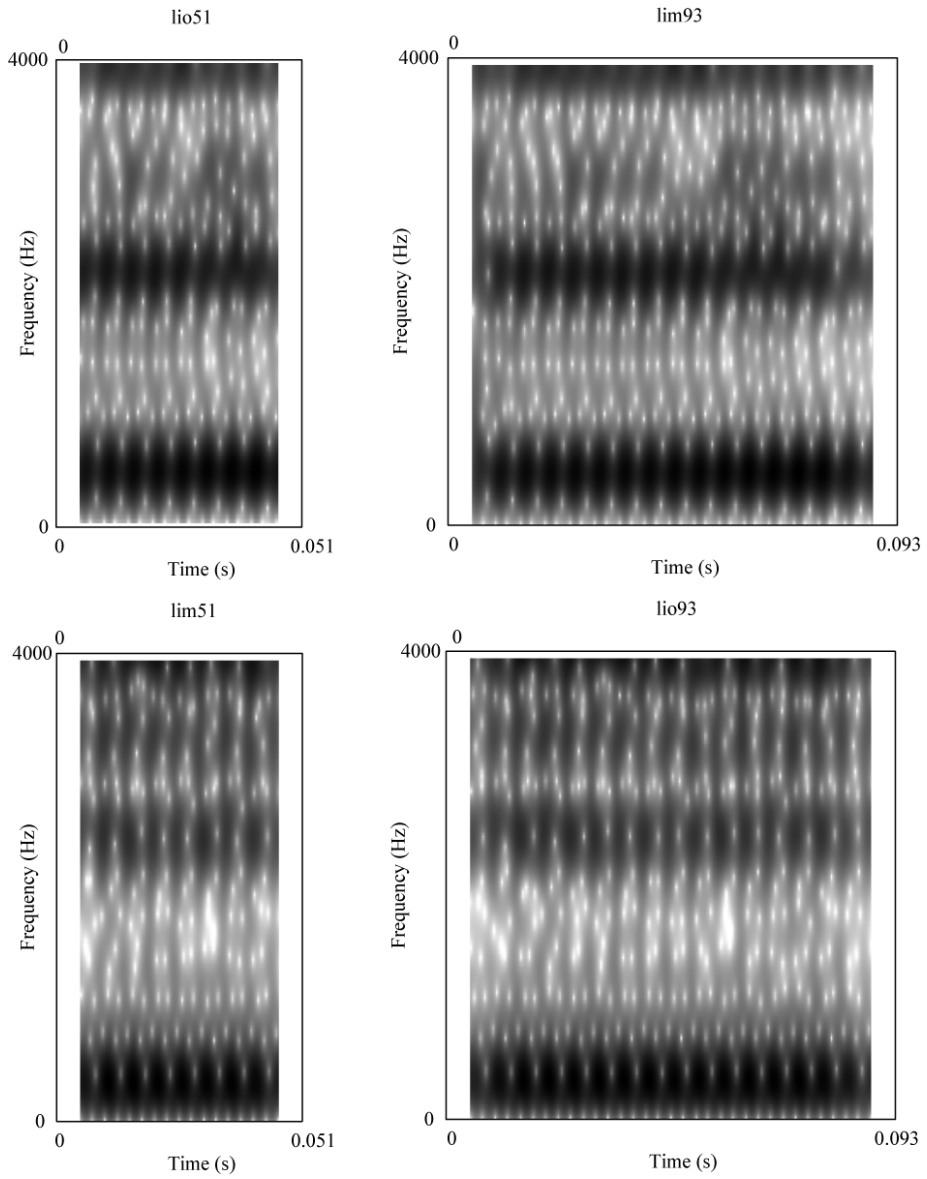

Figure 16: Spectrograms of the four low pass filtered vowel center stimuli based on /ɪ/ - /i:/. Lio51 and lio93 are based on the original lax-tense pair. They differ with respect to both temporal and spectral information. Lim51 is the shortened version of lio93 and lim93 is the lengthened version of lio51.

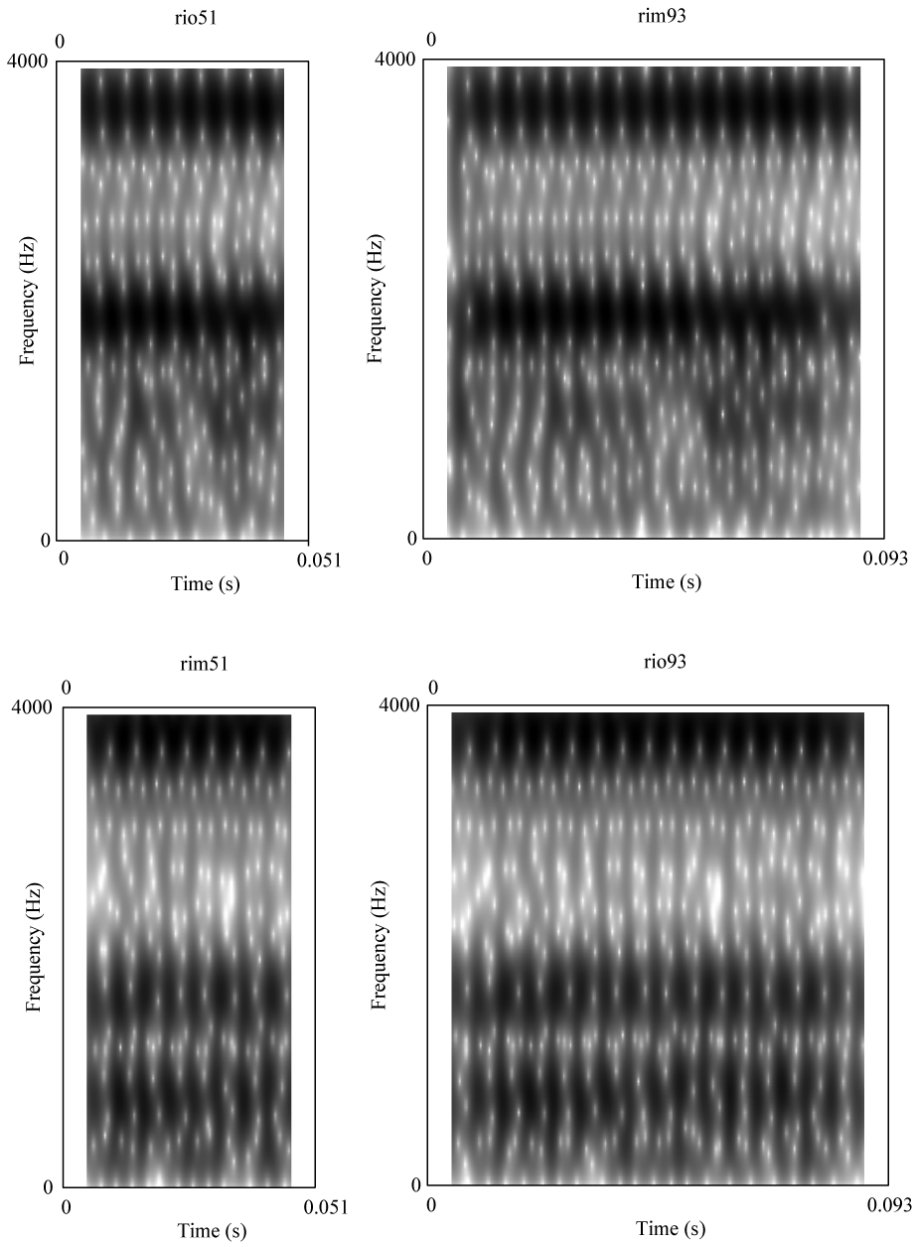

Figure 17: Spectrograms of the four spectrally rotated vowel center stimuli based on /ɪ/ - /i:/. Rio51 and lio93 are based on the original low pass filtered lax-tense pair. They differ with respect to both temporal and spectral information. Rim51 is the shortened version of rio93 and rim93 is the lengthened version of rio51.

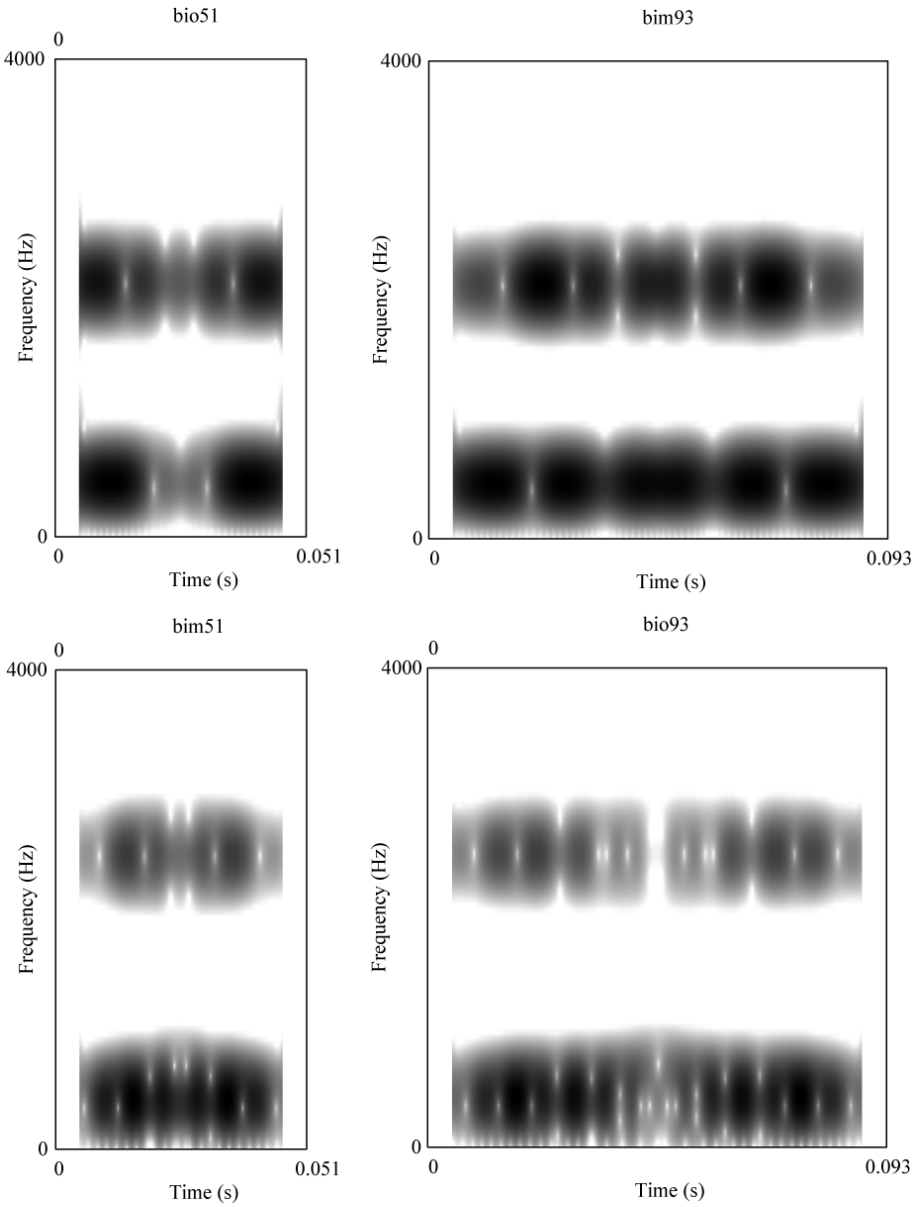

Figure 18: Spectrograms of the four bands of formants based on the vowel center stimuli vio51, vio93, vim51 and vim93.

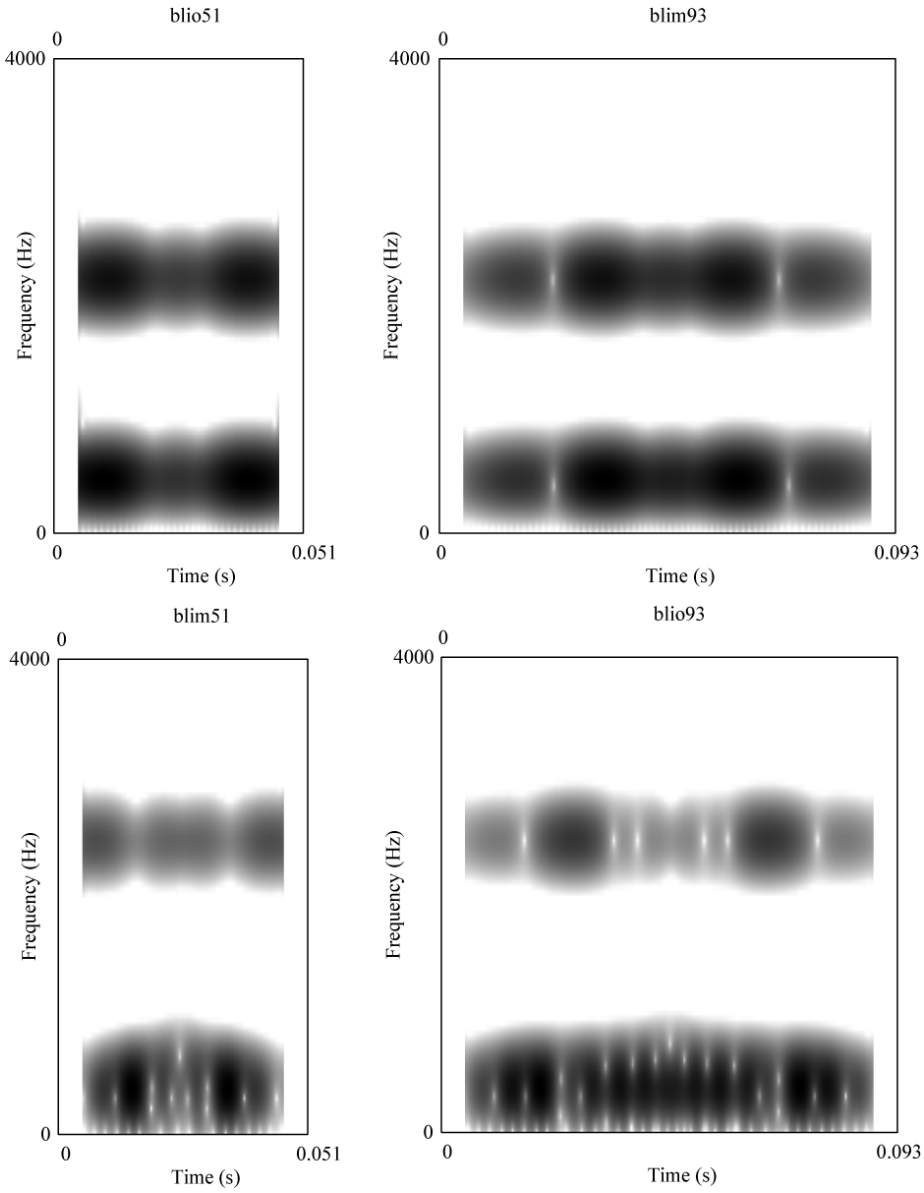

Figure 19: Spectrograms of the four bands of formants based on the low pass filtered vowel center stimuli lio51, lio93, lim51 and lim93.

Table 7: Experimental design of all trials with stimuli based on /a/ - /a:/ in Experiment 1.

/a/ - /a:/	**different condition (24x)**			**same condition (24x)**
	Temporal (8x)	Spectral (8x)	Both (8x)	
vowel center (VC)	Vao75 vs. Vam145 (4x) Vao145 vs. Vam75 (4x)	Vao75 vs. Vam75 (4x) Vao145 vs. Vam145 (4x)	Vao75 vs. Vao145 (8x)	Vao75 vs. Vao75 (6x) Vao145 vs. Vao145 (6x) Vam75 vs. Vam75 (6x) Vam145 vs. Vam145 (6x)
low pass filtered vowel center (LVC)	Lao75 vs. Lam145 (4x) Lao145 vs. Lam75 (4x)	Lao75 vs. Lam75 (4x) Lao145 vs. Lam145 (4x)	Lao75 vs. Lao145 (8x)	Lao75 vs. Lao75 (6x) Lao145 vs. Lao145 (6x) Lam75 vs. Lam75 (6x) Lam145 vs. Lam145 (6x)
spectrally rotated vowel center (RVC)	Rao75 vs. Ram145 (4x) Rao145 vs. Ram75 (4x)	Rao75 vs. Ram75 (4x) Rao145 vs. Ram145 (4x)	Rao75 vs. Rao145 (8x)	Rao75 vs. Rao75 (6x) Rao145 vs. Rao145 (6x) Ram75 vs. Ram75 (6x) Ram145 vs. Ram145 (6x)
bands of formants based on the vowel center (BFCV)	Bao75 vs. Bam145 (4x) Bao145 vs. Bam75 (4x)	Bao75 vs. Bam75 (4x) Bao145 vs. Bam145 (4x)	Bao75 vs. Bao145 (8x)	Bao75 vs. Bao75 (6x) Bao145 vs. Bao145 (6x) Bam75 vs. Bam75 (6x) Bam145 vs. Bam145 (6x)
bands of formants based on the low pass filtered vowel center (BFLVC)	Blao75 vs. Blam145 (4x) Blao145 vs. Blam75 (4x)	Blao75 vs. Blam75 (4x) Blao145 vs. Blam145 (4x)	Blao75 vs. Blao145 (8x)	Blao75 vs. Blao75 (6x) Blao145 vs. Blao145 (6x) Blam75 vs. Blam75 (6x) Blam145 vs. Blam145 (6x)

Table 8: Experimental design of all trials with stimuli based on /ɪ/ - /i:/ in Experiment 1.

/ɪ/ - /i:/	Different condition (24x)			Same condition (24x)
	Temporal (8x)	Spectral (8x)	Both (8x)	
vowel center (VC)	Vio51 vs. Vim93 (4x) Vio93 vs. Vim51 (4x)	Vio51 vs. Vim51 (4x) Vio93 vs. Vim93 (4x)	Vio51 vs. Vio93 (8x)	Vio51 vs. Vio51 (6x) Vio93 vs. Vio93 (6x) Vim51 vs. Vim51 (6x) Vim93 vs. Vim93 (6x)
low pass filtered vowel center (LVC)	Lio51 vs. Lim93 (4x) Lio93 vs. Lim51 (4x)	Lio51 vs. Lim51 (4x) Lio93 vs. Lim93 (4x)	Lio51 vs. Lio93 (8x)	Lio51 vs. Lio51 (6x) Lio93 vs. Lio93 (6x) Lim51 vs. Lim51 (6x) Lim93 vs. Lim93 (6x)
spectrally rotated vowel center (RVC)	Rio51 vs. Rim93 (4x) Rio93 vs. Rim51 (4x)	Rio51 vs. Rim51 (4x) Rio93 vs. Rim93 (4x)	Rio51 vs. Rio93 (8x)	Rio51 vs. Rio51 (6x) Rio93 vs. Rio93 (6x) Rim51 vs. Rim51 (6x) Rim93 vs. Rim93 (6x)
bands of formants based on the vowel center (BFCV)	Bio51 vs. Bim93 (4x) Bio93 vs. Bim51 (4x)	Bio51 vs. Bim51 (4x) Bio93 vs. Bim93 (4x)	Bio51 vs. Bio93 (8x)	Bio51 vs. Bio51 (6x) Bio93 vs. Bio93 (6x) Bim51 vs. Bim51 (6x) Bim93 vs. Bim93 (6x)
bands of formants based on the low pass filtered vowel center (BFLVC)	Blio51 vs. Blim93 (4x) Blio93 vs. Blim51 (4x)	Blio51 vs. Blim51 (4x) Blio93 vs. Blim93 (4x)	Blio51 vs. Blio93 (8x)	Blio51 vs. Blio51 (6x) Blio93 vs. Blio93 (6x) Blim51 vs. Blim51 (6x) Blim93 vs. Blim93 (6x)

Design

The complete design is illustrated in Tables 7 and 8. All in all, one block consisted of 96 trials. There were 5 different blocks with one for each *stimulus type*: the vowel center stimuli with full spectrum, the low pass filtered vowel center stimuli, the spectrally rotated vowel center stimuli, the bands of formants based on the vowel center stimuli with the full spectrum and the bands of formants based on the low pass filter vowel center stimuli. The order of blocks was mixed between participants. There was one block for each type of stimulus. Within each block there were *two vowel types*: /a/ – /a:/ and /ɪ/ – /i:/.

During one half of the trials one stimulus was presented twice (same condition), whereas two different stimuli could be distinguished during the second half of the trials (different condition). There were *three types of auditory difference*: temporal, spectral and spectro-temporal. The order of the trials in each block was pseudo randomized in accordance with the following rules: there were maximally three trials in sequence which required the same response and in addition, vowel identity changed at least after every third trial.

Dependent variables

Two dependent variables were used for the data analysis: the discrimination index d' and mean reaction time of correct responses. D' was calculated as reported by Macmillan and Creelman (1991) for same-different designs. D' does not consider hits only, but also the number of false alarms. A hit is observed when a person realizes that there is a difference between two distinctive stimuli. A false alarm means that a person classifies two equal stimuli as different. The discrimination index increases with the number of hits and decreases with the number of false alarms. The relative frequencies of both, the hits and the false alarms, are transformed into z values based on the normal distribution. As relative frequencies of 0 and 1 cannot be transformed into z values, a value of 0 was replaced by .01 and 1 was replaced by .99 (Macmillan & Creelman, 1991, page 10). D' is the difference of the two z values (d' = z(hits) – z(false alarms)) in a simple yes-no experiment, when participants' responses are not biased. Unfortunately, responding behavior is biased in most same-different tasks, as participants tend to choose the 'same' response more often. To circumvent this problem, Macmillan and Creelman (1991, page 145) provide two correction formulas which include this bias and additionally expand the model to same-different tasks:

(1) $p(c) = \Phi\{[z(hit) - z(false\ alarm)]/2\}$

(2) $d' = 2z[0.5 \cdot \{1 + [2p(c) - 1]^{1/2}\}]$

P(c) is the estimated proportion of correct responses, which would be expected from an unbiased observer. This information is sufficient to calculate d' with the second formula.

The second dependent variable was the mean reaction time to correct responses. Reaction times which were longer than three seconds were excluded from the analysis (less than 5% of the trials).

Hypotheses

(1) It was proposed that the intelligibility of low pass filtered speech is not reduced (e.g., Scott & Wise, 2004). In accordance with this assumption, the performance for the low pass filtered vowel center stimuli should not be reduced compared to the vowel center stimuli.

(2) For the vowel center stimuli, discrimination scores should be dependent on the type of vowel and on the auditory contrast, as these stimuli are based on the German vowel system:

   a) For the vowel pair /a/ – /a:/, performance should be better in the temporal compared to the spectral condition.

   b) For the vowel pair /ɪ/ – /i:/, performance should be better in the spectral compared to the temporal condition.

   c) For the temporal contrast, performance should be better for the vowel pair /a/ – /a:/ compared to the vowel pair /ɪ/ – /i:/.

   d) For the spectral contrast, performance should be better for the vowel pair /ɪ/ – /i:/ compared to the vowel pair /a/ – /a:/.

   e) As two different auditory cues are provided within the spectro-temporal contrast, performance should be better in this condition compared to the performance in conditions where only a temporal or spectral cue is available.

(3) The same pattern of results should be observed for the spectrally rotated vowel center stimuli, as they are equally complex:

   a) For the spectrally rotated vowel pair /a/ – /a:/, performance should be better in the temporal compared to the spectral condition.

   b) For the spectrally rotated vowel pair /ɪ/ – /i:/, performance should be better in the spectral compared to the temporal condition.

c) For the temporal contrast, performance should be better for the spectrally rotated vowel pair /a/ - /a:/ compared to the spectrally rotated vowel pair /ɪ/ – /i:/.

d) For the spectral contrast, performance should be better for the spectrally rotated vowel pair /ɪ/ – /i:/ compared to the spectrally rotated vowel pair /a/ – /a:/.

e) Because two different auditory cues are provided in the spectro-temporal contrast, performance should be better in this condition compared to the performance in conditions where only a temporal or spectral cue is available for the spectrally rotated stimuli.

(4) The two types of the bands of formants were created with the same procedure and with similar values (compare Table 4 and 5), so there should be no systematic difference between the performance for the bands of formants on the basis of the vowel center stimuli and the low pass filtered vowel center stimuli.

(5) The bands of formants are based on the vowel center stimuli and the low pass filtered vowel center stimuli. Nevertheless, they are less complex and there is one crucial difference in the spectral contrast: The vowel center stimuli and low pass filtered vowel center stimuli have the same pitch and differ only with respect to timbre. In contrast, the two stimuli of a spectral contrast in the bands of formants differ with respect to pitch. Because the human ear is able to distinguish very small differences between the pitch of two sounds (Hellbrück & Ellermeier, 2004), performance should be enhanced when more information about pitch differences is available. This additional information is only provided in the spectral condition and not in the temporal one. As follows:

a) The spectral condition of the bands of formants should be easier to discriminate compared to the spectral condition of the vowel center stimuli, low pass filtered vowel center stimuli or spectrally rotated vowel center stimuli.

b) The temporal contrast for the vowel pair /ɪ/ – /i:/ should be harder to discriminate compared to the vowel pair /a/ – /a:/.

c) Because two different auditory cues are provided in the spectro-temporal contrast, performance should be better in this condition compared to performance in conditions where only a temporal or spectral cue is available for the bands of formants.

Results

A 5*2*3 analysis of variances (ANOVA) with repeated measurements was conducted, including: *Stimulus Type* (5: vowel center stimuli vs. low pass filtered vowel center stimuli vs. spectrally rotated vowel center stimuli vs. bands of formants based on the vowel center stimuli vs. bands of formants based on the low pass filtered vowel center stimuli), *Vowel Type* (2: /a/ vs. /ɪ/), and *Auditory Difference* (3: temporal vs. spectral vs. spectro-temporal). An overview of the data is given in Table 9 and Figure 20. Every time the assumption of sphericity was rejected as revealed by the Mauchly's test, the degrees of freedom (df) were corrected according to Greenhouse-Geisser.

A significant main effect of *Stimulus Type* was found ($F(4,96) = 4.60$, $p < .01$). Bonferroni-corrected t-tests revealed the following pattern of results: There was no difference between the vowel center stimuli and the low pass filtered vowel center stimuli ($t(24) = -0.52$, $p = .61$) (see Hypothesis 1). There was no difference between the two versions of the bands of formants ($t(24) = 1.19$, $p = .25$) (see Hypothesis 4) and both did not differ from the spectrally rotated vowel center stimuli ($t(24) = -0.12$, $p = .90$). Both vowel stimuli were significantly less accurately discriminated compared to the three non-speech conditions ($t(24) = -3.97$, $p < .01$, $d = 0.81$). This pattern of results is illustrated in Figure 21.

Table 9: Results of the analysis of variances based on the discrimination index d' in Experiment 1.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Stimulus Type	4.70	4	96	< .01	.16
Vowel Type	2.13	1	24	.16	.08
Auditory difference	34.62	2	48	< .01	.59
Stimulus * Vowel Type	1.47	4	96	.22	.06
Stimulus Type * Auditory Difference	7.33	8	192	< .01	.23
Vowel Type* Auditory Difference	109.94	2	48	< .01	.82
Stimulus * Vowel Type * Auditory Difference	17.48	8	192	< .01	.42

There was no significant main effect of type of vowel ($F(1,24) = 2.13$, $p > .16$). The main effect of auditory difference reached significance ($F(2,48) = 34.62$, $p < .01$). The temporal condition was more difficult than the spectral condition ($t(1) = -3.46$, $p < .01$, $d = 0.69$). Performance was significantly better when both temporal and spectral information were available, compared to spectral information alone ($t(1) = -6.60$, $p < .01$, $d = 1.32$) or temporal information alone ($t(1) = -7.54$, $p < .01$, $d = 2.56$) (see Figure 22 and Hypotheses 2e, 3e and 5c).

There was a significant interaction between the type of stimulus and the auditory difference ($F(8,192) = 7.33$, $p < .01$). A drop of performance in the spectral and temporal condition was observed only for the two vowel center stimulus types. The contrasts of the two vowel center stimulus types compared to the three non-speech stimulus types revealed significant differences for the temporal ($t(24)$ -3.67, $p < .01$, $d = 0.73$) and the spectral condition ($t(24) = -4.17$, $p < .01$, $d = 0.83$) but not for the "both" condition ($t(24) = 0.80$, $p = .43$).

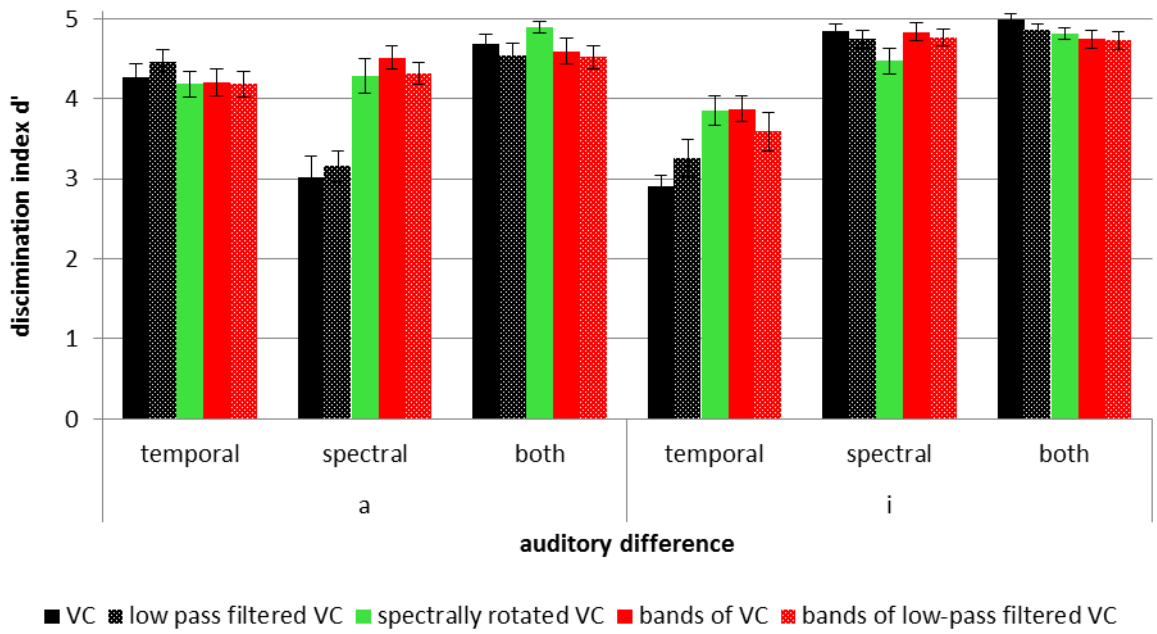

Figure 20: Means and standard errors of the discrimination index d' for each auditory difference (temporal, spectral and spectro-temporal), vowel type (a vs. i), and stimulus type in Experiment 1: vowel center stimuli (VC) = black, low pass filtered vowel center stimuli = black hatched, spectrally rotated vowel center stimuli = green, bands of formants based on the vowel center stimuli = red, bands of formants based on the low pass filtered vowel center stimuli = red hatched.

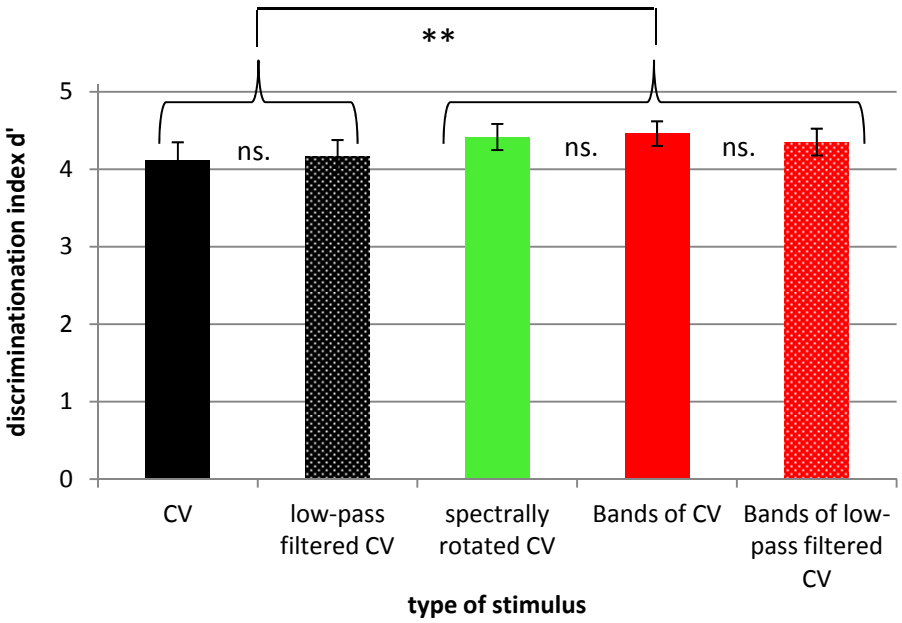

Figure 21: Comparison of the two vowel center (VC) stimulus types (vowel center stimuli = black, low pass filtered vowel center stimuli = black hatched) and the three non-speech conditions (spectrally rotated vowel center stimuli = green, bands of formants based on the vowel center stimuli = red, bands of formants based on the low pass filtered vowel center stimuli = red hatched) in Experiment 1 for the discrimination index d'. The bars represent the means including standard errors.

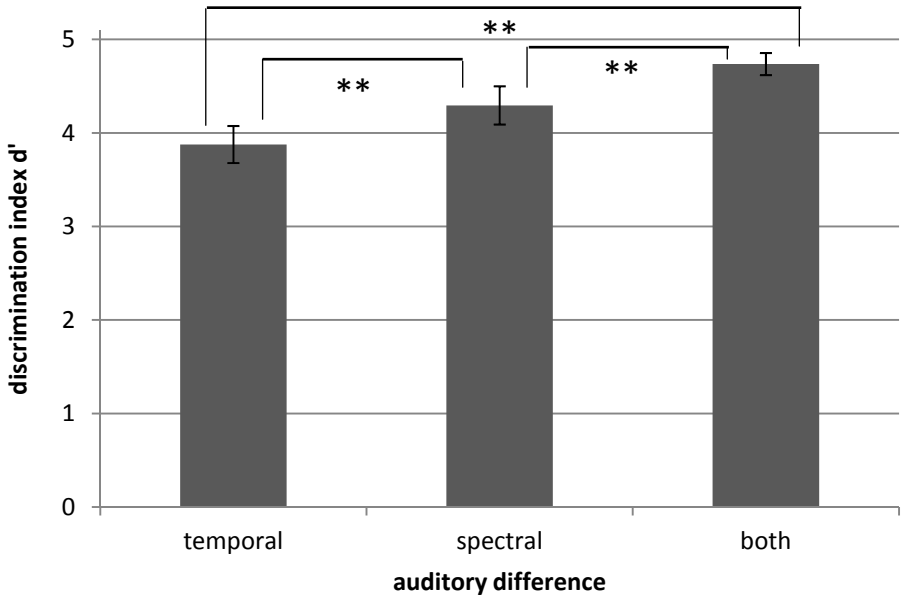

Figure 22: Comparison of the three auditory contrasts (temporal, spectral and spectro-temporal) in Experiment 1 for the discrimination index d'. The bars represent the means including standard errors.

The interaction between stimulus and vowel types did not reach significance ($F(4,96) = 0.47$, $p = .22$). A significant interaction between type of vowel and auditory difference was found ($F(2,48) = 109.94$, $p < .01$). Performance in the temporal condition dropped especially for the vowel pair /ɪ/ - /i:/ compared to the vowel pair /a/ - /a:/ ($t(24) = -7.29$, $p < .01$, $d = 1.46$). This finding was not only observed for the two vowel center stimulus types ($t(24) = -10.41$, $p < .01$, $d = 1.42$) (see Hypothesis 2c), but also for the three non-speech stimulus types ($t(24) = 7.05$, $p < .01$, $d = 0.75$) (see Hypotheses 3c and 5b). In addition, the discrimination index was significantly lower in the spectral condition for the vowel pair /a/ - /a:/ than for the vowel pair /ɪ/ - /i:/ ($t(24) = 9.64$, $p < .01$, $d = 1.93$). This observation seems to be a consequence of the averaging over all stimulus types. Performance was not reduced for the non-speech stimuli in the spectral condition of the vowel pair /a/ - /a:/ compared to the other two auditory differences ($t(24) = -0.49$, $p = .63$). The significant triple interaction between type of stimulus, type of vowel and auditory difference ($F(8,192) = 17.48$, $p < .01$) can be explained by the fact that performance especially dropped for the two vowel center stimuli, but only for the temporal condition of the vowel pair /ɪ/ - /i:/ and the spectral condition of the vowel pair /a/ - /a:/ (see Hypotheses 2a-d).

A second ANOVA was conducted with the mean reaction time as the dependent variable. The results are shown in Table 10 and Figure 23.

A significant main effect of stimulus type was found ($F(4,96) = 5.47$, $p < .01$). The difference between the vowel center stimuli and the low pass filtered vowel center stimuli did not reach significance ($t(24) = -1.67$, $p = .11$). The two versions of the bands of formants did not differ significantly either ($t(24) = 0.76$, $p = .45$). Both vowel center stimulus types were discriminated more slowly compared to the spectrally rotated vowels ($t(24) = 4.03$, $p < .01$, $d = 0.81$) and the two versions of the bands of formants ($t(24) = 4.24$, $p < .01$, $d = 0.84$). No difference was found between the spectrally rotated vowels and the bands of formants ($t(24) = -0.20$, $p = .84$).

The main effect of vowel type did not reach significance ($F(1,24) = 0.24$, $p = .62$). However, a significant main effect of auditory difference was found ($F(2,48) = 45.13$, $p < .01$). Response times in the temporal condition were longer compared to the spectral ($t(24) = 4.92$, $p < .01$, $d = 0.99$) or spectro-temporal condition ($t(24) = 8.05$, $p < .01$, $d = 1.61$). Faster responses were found in the spectro-temporal condition compared to the spectral condition ($t(24) = 6.10$, $p < .01$, $d = 1.22$).

Table 10: Results of the analysis of variance based on reaction times in Experiment 1.

Factor (RT)	*F*	df(factor)	df(error)	*p*	partial eta²
Stimulus Type	5.47	4	96	< .01	.19
Vowel Type	0.26	1	24	.62	.01
Auditory difference	45.13	2	48	< .01	.65
Stimulus * Vowel Type	3.34	4	96	.01	.12
Stimulus Type * Auditory Difference	2.81	8	192	< .01	.11
Vowel Type* Auditory Difference	20.93	2	48	< .01	.47
Stimulus * Vowel Type * Auditory Difference	7.15	8	192	< .01	.23

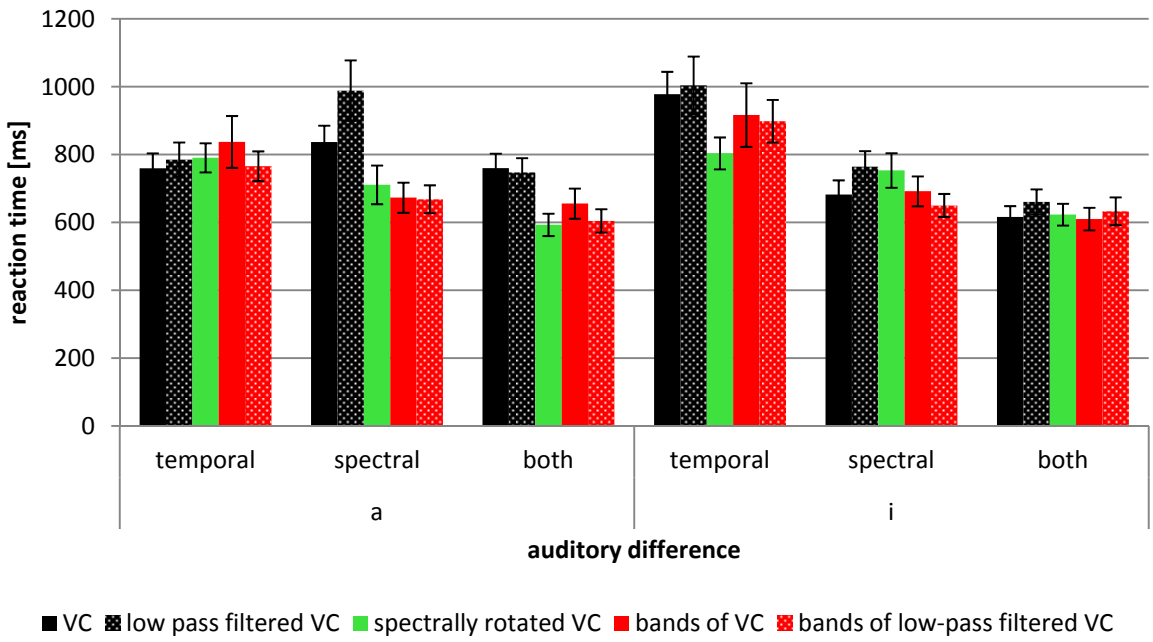

Figure 23: Means and standard errors of reaction times for each experimental condition of Experiment 1.

All interactions of this analysis of variance became significant. These interactions can be explained by the fact that the difference of reaction time for the speech and non-speech

stimuli was especially high for these contrasts, which are supposed to be difficult: the spectral condition for the vowel pair /a/ - /a:/ ($t(24) = 6.57$, $p < .01$, $d = 1.32$) and the temporal condition for the vowel pair /ɪ/ - /i:/ ($t(24) = 2.71$, $p = .01$, $d = 0.54$). Only the spectro-temporal contrast between speech and non-speech stimuli for the vowel pair /a/ - /a:/ reached significance as well ($t(24) = 6.84$, $p < .01$, $d = 1.37$).

To rule out any speed-accuracy trade off, the point-biserial correlation coefficient between the correctness of the response (0 = error, 1 = correct response) and the reaction time was calculated. The correlation was $r = -.16$ ($p < .01$).

Discussion

The major goal of this chapter was to extend the German vowel length discrimination paradigm by using spectrally rotated non-speech stimuli with the same complexity as the speech-like ones. In addition, a second non-speech version was created, including bands of formants with lower complexity, while maintaining the most important frequencies of the vowels.

The aim was to replicate the pattern of results of the speech stimuli reported by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation) in this extended version of the German vowel length discrimination paradigm and to compare it to the discrimination performance for the non-speech stimuli.

The first hypothesis dealt with the question of whether low pass filtering of the vowel center stimuli would influence the overall discrimination performance. There was no systematic difference between the two stimulus types. 4000Hz was chosen as the cut-off frequency for the low pass filtered vowel center stimuli, comparable to most studies dealing with spectrally rotated speech (e.g., Davids et al., 2011; Evans et al., 2013; Narain et al., 2003; Okada et al., 2010; Scott et al., 2000; Scott et al., 2006; Scott et al., 2009; Sörqvist et al., 2012; Vandermosten et al., 2010; Vandermosten et al., 2011). The most important frequencies of the speech signal are supposed to lie between 500 and 4000Hz (Wilmanns & Schmitt, 2002) and it has been shown that the first two formants are sufficient for the correct identification of vowels (Nawka & Wirth, 2008). In the light of these facts it is assumed that the intelligibility of the speech sound would not be impaired (e.g., Scott et al., 2000; Scott & Wise, 2004) and, indeed, discrimination performance in our study was actually not affected by the low pass filtering. However, the naturalness of these low pass filtered

sounds was rated much weaker compared to the vowel center stimuli with respect to the whole frequency spectrum. Some participants were even unable to identify the low pass filtered stimuli as vowels of the German language. Furthermore, reaction times tended to be longer for the low pass filtered stimuli, indicating that they were not perceived in the same way as the vowel center stimuli.

Hypothesis 2 was that the pattern of results found by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation) would be replicated in the current experiment, as the vowel center stimuli are based on the German vowel system. For the vowel center stimuli, discrimination scores should be dependent on the type of vowel and the auditory contrast. As expected, performance was less accurate for the temporal contrast of the vowel pair /ɪ/ – /i:/ and the spectral contrast of the vowel pair /a/ – /a:/, but not for the temporal contrast of the vowel pair /a/ – /a:/ and also not for the spectral contrast of the vowel pair /ɪ/ – /i:/. This pattern of results was found for both the vowel center stimuli and the low pass filtered vowel center stimuli (see Figure 20). These results confirm the Hypotheses 2a-d. The vowel center stimuli are based on natural spoken German vowels. Temporal differences are smaller in the tense-lax pair /i:/ - /ɪ/ compared to /a:/ - /a/. On the other hand, /a:/ and /a/ show a similar spectral pattern (Ungeheuer, 1969), whereas /i:/ and /ɪ/ can be easily distinguished on the basis of their spectral properties (Bennet, 1968; Strange & Bohn, 1998; Weiss, 1974).

These findings are comparable to the results reported by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation). In their experiments the vowels were embedded into a CVC syllable. In contrast, the vowel center stimuli were presented without frame in the current experiment. Nevertheless, the drop of performance for the spectral condition of the vowel pair /a/ - /a:/ and the temporal condition of the vowel pair /ɪ/ - /i:/ was still observed. This means that the replication of the results based on the German vowel length discrimination paradigm used by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation) was successful. It was also shown that difficult contrasts lead to longer reaction times.

The next hypothesis (Hypothesis 2e) addressed the role of the spectro-temporal condition. As two different auditory cues are provided in the spectro-temporal contrast of the vowel center stimuli, performance should be better in the spectro-temporal condition compared to the performance when only a temporal or a spectral cue is available. Indeed, performance in

the spectro-temporal condition was significantly better compared to the spectral or temporal condition alone, as indicated by higher discrimination indexes and shorter reaction times. This observation is in accordance with the results reported by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation).

In the current experiment, the set of stimuli also included non-speech stimuli with comparable complexity to the vowel center stimuli. Consequently, the same pattern of results should be observed for the spectrally rotated vowel center stimuli. Performance should drop in the spectral condition of the vowel pair /a/ - /a:/ and in the temporal condition of the vowel pair /ɪ/ - /i:/ (Hypotheses 3a-d). As expected, there was no decrease of performance for both the spectral condition of the vowel pair /ɪ/ - /i:/ and the temporal condition of the vowel pair /a/ – /a:/ for the spectrally rotated stimuli. A drop in performance was only observed for the temporal condition of the vowel pair /ɪ/ – /i:/. Interestingly, performance in the spectral condition of the vowel pair /a/ – /a:/ was not affected by the spectrally rotated vowel center stimuli. This means that although the vowels and the spectrally rotated vowels were matched with respect to complexity, the difficulty of the spectral condition before and after the spectral rotation was not comparable. It was already mentioned that the spectrally rotated vowels do not contain harmonic partials. Therefore, they evoke a completely different hearing impression compared to the vowels. This could be the reason why the difficulty of the spectral contrast is not preserved by the spectral rotation.

The next hypothesis (Hypothesis 3e) addressed the role of the spectro-temporal condition in the spectrally rotated stimuli. Two different auditory cues are provided in the spectro-temporal contrast of the spectrally rotated vowel center stimuli. This should make correct discrimination easier. Comparable to the speech stimuli, performance and reaction times were significantly better in the spectro-temporal compared to the spectral or temporal condition.

There were two versions of the bands of formants, one based on the vowel center stimuli, and the other one based on the low pass filtered vowel center stimuli. There should be no systematic differences between the performance for the bands of formants on the basis of vowels and the low pass filtered vowels (Hypothesis 4). This is what was actually observed. This pattern of results was expected because the two types of the bands of formants were created with the same procedure and with similar values (see Tables 4 and 5). Moreover,

participants reported that the hearing impression of the two different stimulus types was quite similar.

The next hypotheses concern the pattern of results when the bands of formants are presented. The spectral condition of the bands of formants should be easier to discriminate compared to the spectral condition of the vowel center stimuli, low pass filtered vowel center stimuli or spectrally rotated vowel center stimuli (Hypothesis 5a). Performance in the spectral condition of the bands of formants did not drop in the same manner as was observed in the vowel center stimuli. Although the bands of formants are based on the vowel center stimuli and the low pass filtered vowel center stimuli, they are less complex. There is one large difference in the spectral contrast compared to the vowel center stimuli: The vowel center stimuli have the same pitch and differ only with respect to timbre. In contrast, the two stimuli of a spectral contrast in the bands of formants differ with respect to pitch. It was already mentioned that the human ear is able to distinguish very small differences between the pitch of two sounds (Hellbrück & Ellermeier, 2004). Performance was probably enhanced as a result of additional information about pitch differences. This additional information is only provided in the spectral condition and not in the temporal condition leaving a segue to the next hypothesis: The temporal contrast for the vowel pair /ɪ/ – /i:/ should be harder to discriminate compared to the vowel pair /a/ – /a:/ (Hypothesis 5b). Performance in the temporal condition was significantly reduced for the vowel pair /ɪ/ – /i:/ for the bands of formants. This drop in performance is expected whenever the temporal contrast is kept low and independent of stimulus type, as spectral information is not needed to compare the length of two stimuli with the same spectral pattern.

The last hypothesis concerns the role of the spectro-temporal condition for the bands of formants. The same pattern of results as for the other stimulus types was expected. As two different auditory cues are provided in the spectro-temporal contrast, performance should improve in this condition compared to the other performance, in which only a temporal or spectral cue is available (Hypothesis 5d). As expected, performance was highest in the spectro-temporal condition.

Conclusion

Taken together, the German vowel length discrimination paradigm used by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation) was replicated. This means that the overall performance was unaffected by the absence of the frame of the CVC syllable. As only the steady state portion was used in this experiment. The usage of the steady state portion means that contrary to the stimuli used by Groth and colleagues (2011) and Steinbrink et al. (in preparation) there is no spectral change within each stimulus. Even so, the pattern of results remains the same. Blesser (1972) described that the perception of some consonants go unaffected by the spectral rotation of the signal and that the perception of spectrally rotated vowels is highly dependent upon the frame in which they are embedded. The aim of the current experiment was to create an equally complex non-speech analogue, and so only isolated vowels were used as speech sounds and consonants were omitted.

One crucial finding of the current experiment shows that the difficulty of the spectral contrast is incomparable for the speech-like and spectrally rotated speech stimuli. Although both stimulus types are matched with respect to complexity, the timbre of /a/ and /a:/ should prove to be more similar than in the spectrally rotated versions of these stimuli. Conversely, performance dropped for the spectrally rotated vowel pair /ɪ/ – /i:/ in the spectral condition compared to the vowel center stimuli (see Figure 20).

The comparison of the vowel center stimuli with full spectrum and the low pass filtered vowel center stimuli revealed that the former were perceived to be more speech like than the latter. However, the low pass filtering of the speech sound is a precondition for the creation of the spectrally rotated speech. To circumvent this short coming, a modification of the spectral rotation will be presented in the following chapter which enables to compare the original speech sound with an equally complex non-speech sound with a complete spectrum (comparable to vowels).

The extended version of the German vowel length discrimination paradigm will be used in the next chapter for the comparison of the auditory processing of speech and non-speech sounds in dyslexic adults and age matched controls. This is the first study in which the complexity of speech and non-speech stimuli is controlled for while the processing of temporal, spectral and spectro-temporal cues is investigated in dyslexic adults.

Chapter 3:

The processing of speech and non-speech in dyslexic adults

> "I want you to wonder, not only about
> what you read but at the miracle that
> you can read."
>
> Vladimir Nabokov

This chapter deals with the specific nature of auditory processing deficits in developmental dyslexia. It is commonly accepted that phonological deficits represent the core symptom of the specific reading disorder. What remains unclear, however, is the issue of whether these phonological deficits might be speech specific or whether they might be caused by more general auditory problems. Most studies which compared the auditory processing of speech and non-speech stimuli in dyslexia did not control for the complexity of both stimulus types, for their size of contrast and for task difficulty. The modified German vowel length discrimination paradigm, as introduced in Chapter 2, is used to investigate the impairment of sound processing in dyslexic adults. This approach enables to control for the complexity of the task and the stimuli, as the same discrimination task is used to investigate several types of stimuli (vowel center stimuli, spectrally rotated vowel center stimuli and bands of formants) in one sample of participants. In addition, multiple acoustical parameters are varied within each type of stimulus, while maintaining task complexity.

Developmental dyslexia

The term developmental dyslexia or specific reading disorder refers to specific difficulties in learning to read despite normal intelligence, unaffected sensory abilities, motivation and conventional instruction (American Psychiatric Association, 1994; Démonet, Taylor, & Chaix, 2004; Lyon, Shaywitz, & Shaywitz, 2003). These difficulties are supposed to be already evident during childhood and can be accompanied with poor spelling performance (ICD-10 of WHO, Dilling & Freyberger, 2012). An accurate diagnosis of developmental dyslexia can only be made after the first instructions of written language (Warnke, 2008).

German is a so-called “shallow” language with regular orthography and grapheme-phoneme correspondences (Brunswick, McDougall, & de Mornay Davies, 2010). This means that one letter is mostly only represented by one sound and vice versa (e.g., the phoneme /b/ is represented by the letter “b”). Contrary to this, one sound can be described by a whole set of different letters in deep languages like English and French. The regular orthography enables German dyslexics to achieve a high level of reading accuracy (Goswami, 1999). However, reading speed was found to be impaired in German dyslexic children (Landerl, Wimmer, & Frith, 1997; Wimmer, 1993; Wimmer, Landerl, & Frith, 1999; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003). The symptoms remain stable during school years and are detectable into adulthood (Groth et al., 2011; Kohn, Wyschkon, Ballaschk, Ihle, & Esser, 2013; Shaywitz & Shaywitz, 2005; Svensson & Jacobson, 2006).

Slow reading of non-words was interpreted as a sign of phonological deficits (Wimmer, 1996) and spelling deficits in German are associated with phonological deficits as well (Wimmer, Mayringer, & Landerl, 2000). However, the phonological impairment is supposed to be the same in German and English (Landerl et al, 1997; Wimmer, 1996; Ziegler et al., 2003).

The prevalence of the specific reading disorder is dependent on the respective criteria of dyslexia in each study (Rodgers, 1983; Shaywitz, Fletcher, Holahan, & Shaywitz, 1992). The values vary from 4-8% in Germany (Plume & Warnke, 2007).

This disorder is associated with a broad range of psychosocial impairments. During their school career, students with dyslexia regularly experience failure, humiliation and a lack of understanding (Hughes & Dawson, 1995; Undheim, 2003). There is also an enhanced rate of antisocial and behavioral problems during childhood and adulthood (Esser & Schmidt, 1994;

Frisk, 1999; Heiervang, Stevenson, Lund, & Hugdahl, 2001). Additional risk factors which are commonly found in dyslexic children include a low self-concept (Alexander-Passe, 2006; Boetsch, Green, & Pennington, 1996; Burden, 2005) and emotional problems, like anxiety (Casey, Levy, Brown, & Brooks-Gunn, 1992) and depression (Alexander-Passe, 2006; Boetsch et al., 1996). The suicidal tendency is also higher in dyslexic children and adolescents (Daniel et al., 2006).

Maughan and colleagues (1985) report early school leaving as a supplementary factor. Adolescents with specific reading disorder are less likely to earn a high-school diploma (Dummer-Smoch 2007; Esser & Schmidt, 1993) and are more likely to become unemployed (Kohn et al., 2013).

Considering the evidence, prevention and intervention are crucial to help alleviate some of the problematic symptoms. A precondition to developing effective prevention programs and various forms of therapy is to understand the etiology of developmental dyslexia.

Etiology

There is an overall agreement that there is a neurobiological cause of dyslexia (Denckla & Rudel, 1976). These neuronal deficits have been shown in a wide range of studies dealing with the functional and structural correlates of developmental dyslexia (Csépe, 2003; Démonet et al., 2004; Galaburda, LoTurco, Ramus, Fitch, & Rosen, 2006; Shaywitz, Mody, & Shaywitz, 2006; Shaywitz et al., 2007; see Habib, 2000 for an overview). Interestingly enough, the same unusual neurobiological characteristics are found for participants with different mother tongues (Paulesu et al., 2001).

The reading process is very complex and involves a variety of component skills (*Functional Coordination Deficit model*, Lachmann, 2002; Steinbrink & Lachmann, in press), and so disruptions can be found on different levels (Frith, 1999; Snowling, 2000; Tunmer & Hoover, 1992). Therefore, it is questionable, if a single cause might explain all cases of dyslexia (Lachmann, 2002)

With this information in mind, it is not surprising that numerous theories concerning the etiology of the specific reading disorder have been published. Some prominent theories are, for example the *phonological deficit hypothesis* (Snowling, 1981; Snowling, 2000; Stanovich, 1988), the *temporal deficit hypothesis* (Tallal, 1980), the *cerebellar deficit hypothesis* (Denckla, 1985), the *automatization deficit hypothesis* (Eckert et al., 2003; Fawcett &

Nicolson, 1992; Nicolson, Fawcett, & Dean, 2001), the *magnocellular deficit hypothesis* (Demb, Boynton, Best, & Heeger, 1998; Galaburda & Livingstone, 1993; Stein, 2001).

The focus of the current work lies in the auditory processing in dyslexia, and so only theories which enable predictions about the performance in auditory tasks will be taken into consideration. It is also important not to rule out the possibility that visual or motor aspects might still explain at least some cases of developmental dyslexia (Ramus, 2003).

Phonological deficit hypothesis

The ability to process, perceive and represent phonemes correctly is thought to be crucial for the development of reading and writing (Bryant, MacLean, Bradley, & Crossland, 1990; Elbro, 1996; Goswami & Bryant, 1990). Deficits in these so-called phonological skills may play a key role in the etiology of developmental dyslexia in accordance with the phonological deficit hypothesis (Snowling, 1981; Snowling, 2000; Stanovich, 1988). Phonological deficits are also found regularly in individuals with specific reading disorder (Ramus, 2003; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Wagner & Torgesen, 1987).

Phonological skills comprise several subtypes: phonological awareness, phonological short term memory and the perception of phonemes. Dyslexic children and adults were found to be impaired in all classes, even at pre-school age (Pennington & Lefly, 2001).

The first class of phonological deficits concerns deficits in the perception, discrimination, and identification of phonemes (Adlard & Hazan, 1998; Godfrey, Syrdal-Lasky, Millay, & Knox, 1981), and also speech perception in noise (Ziegler, Pech-Georgel, George, & Lorenzi, 2009). The correct perception of phonemes is a prerequisite for developing appropriate representations for each speech sound of the mother tongue's phoneme repertoire (Godfrey et al., 1981; Watson & Miller, 1993) and to be able to establish phoneme-grapheme correspondences. Therefore, deficits in the phonemic perception might hinder the acquisition of the alphabetic principle (Frith, 1985; Share, 1995; Snowling, 1995). In this context it was noted that phonological processing is necessary but not sufficient alone to understanding the alphabetic principle (Tunmer, Herriman, & Nesdale, 1988). Perceptual aspects of speech influence the development of phonological awareness as well (Yavas & Gogate, 1999).

Phonological awareness is the ability to detect (Liberman, Shankweiler, Fischer, & Carter, 1974) and manipulate the sounds (e.g., deleting or adding a sound, Bruce, 1964) within a

word and to synthesize words from constituent phonemes (Torgesen et al., 1989). Dyslexic children are frequently impaired in phonological awareness tasks (Bradley & Bryant, 1983; Elbro & Jensen, 2005) and these problems are detectable into adulthood (Bruck, 1992; Elbro, Nielsen, & Petersen, 1994; Ramus, 2003). Segmentation (the detection of the phonemes within a word) and blending (to synthesize words from constituent phonemes) play a crucial role in learning to read and write (Blevins, 1997). Training programs with focus on phonological awareness and letters can prevent subsequent problems in reading and writing (Bus & van IJzendoorn, 1999; Ehri et al., 2001) and improve the reading performance of dyslexic children (Ehri et al., 2001).

The last class of the phonological deficits concerns problems in phonological short-term memory tasks, like remembering a string of numbers, letters or the recall for pictures (Jeffries & Everatt, 2004; Nelson & Warrington, 1980; Steinbrink & Klatte, 2008). These problems also persist into adulthood (Smith-Spark & Fisk, 2007).

The issues above are thought to originate from the quality of the underlying phonological representations (Fowler, 1991; Wagner et al., 1993). The quality might be reduced as a result of an underspecification of the speech sounds representations (Adlard & Hazan, 1998; Manis et al., 1997; Mody, Studdert-Kennedy, & Brady, 1997; Swan & Goswami, 1997) or as a result of deficits in the access to the speech sound representations (Boets, Wouters, van Wieringen, & Ghesquière, 2007; Ramus & Szenkovits, 2008) in dyslexia. The latter explanation is supported by studies using rapid automized naming (RAN) of objects, number, colors etc. Dyslexic children and adults were frequently shown to be slower in these tasks (Denckla & Rudel, 1976; Fawcett & Nicolson, 1994; Swan & Goswami, 1997) which concern the retrieval of phonological representations from long-term memory.

General auditory deficits in dyslexia

The speech perception deficits in dyslexia are very subtle and, consequently, difficult to detect (Mody et al., 1997), e.g., in noise (Pennington, van Orden, Smith, Green, & Haith, 1990). Some authors suggest that these deficits are domain specific and phonological. To them, dyslexia is classified as a particularly linguistic problem, as the processing of speech and non-speech are thought to involve different mechanisms (e.g., Liberman, 1989; Ramus, 2003; Vellutino, 1987). They assume that auditory impairments might co-exist with speech perception problems, but they also refute the notion that a general auditory deficit could be

the source of the phonological problems (e.g., Breier, Fletcher, Foorman, Klaas, & Gray, 2003; Mody et al., 1997; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998b).

However, it was stated that a detailed representation of spectral and temporal properties enhances the forming of phonological representations out of acoustic signals (Ahissar, Protopapas, Reid, & Merzenich, 2000; Corriveau, Goswami, & Thomson, 2010) and the role of auditory deficits in dyslexia should not be ignored, as there is a broad range of theories, which favor their causal role.

The most prominent example is the *temporal deficit hypothesis* (Tallal, 1980; Tallal & Gaab, 2006). It was originally proposed for developmental aphasia (Tallal & Piercy, 1974; Tallal & Piercy, 1975) but the authors extended the theory to developmental dyslexia. According to the temporal deficit hypothesis, people with reading difficulties should show a specific deficit in the processing of brief or rapidly changing sounds and also, whenever the presentation rate is very high (Gaab et al., 2007; Nagarajan et al., 1999; Tallal, 1980; Tallal, 1984; Tallal, Merzenich, Miller, & Jenkins, 1998; Tallal, 2000; for a review see Farmer & Klein, 1995). As a result, speech, which is characterized by rapid changes of frequency and amplitude, cannot be integrated and the normal development of the phonological system should fail, which, in turn, hinders the ability to learn to read and write (Tallal, Miller, & Fitch, 1993). In accordance with this theory, reading and writing performances were shown to be influenced by the temporal auditory processing in healthy children (Boets, Wouters, van Wieringen, Smedt, & Ghesquière, 2008; Hood & Conlon, 2004) as a result of enhanced phonological awareness (Corriveau et al., 2010).

There is controversy surrounding the specificity of the temporal perception deficit (Beaton, 2004). It might be confined to stimuli in the linguistic domain (e.g. Schulte-Körne et al., 1998a) or to a general auditory temporal deficit (e.g., Steinbrink, Ackermann, Lachmann, & Riecker, 2009). Moreover, there is no consensus about the term “temporal”. Studdert-Kennedy and Mody (1995) claim that a short duration or a short inter-stimulus interval cannot be defined as temporal features, as such stimuli do not include any change in time. This idea is supported by studies in which amplitude (AM) and frequency (FM) modulations (Talcott & Witton, 2002; Witton, Stein, Stoodley, Rosner, & Talcott, 2002) were detected less frequently by dyslexic participants compared to the control group.

However, the processing of slower temporal modulations might also play a role, especially for the correct identification of syllables and the perception of the rhythm of speech and stress as postulated by the temporal sampling framework for dyslexia (Goswami, 2011).

Measures of frequency are also reported to be regularly impaired in dyslexia (Ahissar et al., 2000; Amitay, Ahissar, & Nelken, 2002; Cacace, McFarland, Ouimet, Schrieber, & Marro, 2000; Hari, Sääskilahti, Helenius, & Uutela, 1999; McAnally & Stein, 1996; King, Lombardino, Crandell, & Leonard, 2003; Lachmann et al., 2005; Montgomery, Morris, Sevcik, & Clarkson, 2005; Walker, Givens, Cranford, Holbert, & Walker, 2006).

Findings concerning auditory deficits in dyslexia were reviewed by Hämäläinen and colleagues (2012). They concluded that measures of frequency, rise time, duration discrimination, amplitude modulation, and frequency modulation are most often impaired in dyslexia.

Reasons for the contradicting results

The fact that auditory deficits were not found in all studies with dyslexic children and adults was taken as evidence against the causal role of a general auditory impairment in dyslexia (e.g. Hill, Bailey, Griffiths, & Snowling, 1999). However, the absence of a behavioral deficit does not mean that there are no abnormalities at the psychophysiological level (Stoodley, Hill, Stein, & Bishop, 2006).

There was also a considerable difference between the studies: There is a broad range of different tasks used to investigate auditory processing in dyslexia (Banai & Ahissar, 2006; France et al., 2002; Lachmann & van Leeuwen, 2007), e.g., temporal order judgments (e.g. Tallal, 1980), gap detection (van Ingelghem et al., 2001), same-different judgments with two (e.g., Groth et al., 2011) or more stimuli (e.g., Hill et al., 1999; Vandermosten et al., 2010), high-low discrimination (e.g., Banai & Ahissar, 2006) and the passive oddball task (Bishop, 2007). Banai and Ahhisar (2006) were able to show that the performance of dyslexic participants is dependent on the type of task. Their performance decreases with increasing working memory load. This is the reason why the authors recommend using the same-different task with two stimuli to estimate auditory discrimination in dyslexia to avoid confounding with working memory load. Another advantage of the same-different task is that it can be performed successfully without prior identification of the two stimuli; so, by extension, it is not dependent on the access to phonological representations (Ahissar, 2007;

Ramus & Szenkovits, 2008) and deficits in this task cannot be explained by under specifications of long-term phonological representations (Boada & Pennington, 2006; Swan & Goswami, 1997). Another possibility would be to use the mismatch negativity as an index of auditory discrimination, as it is also not dependent on attention (Bishop, 2007).

In addition to the burden on memory load, there is an additional reason to avoid using the temporal order judgment task. Low performance in this kind of task cannot be interpreted unambiguously, as it remains unclear whether the deficits lie in the temporal judgment itself or in the auditory discrimination (Ben-Artzi, Fostick, & Babkoff, 2005).

Another reason for the contradicting results concerns the choice of stimuli in the non-speech conditions (Breedin, Martin, & Jerger, 1989). Hardly any study (but see Vandermosten et al., 2010 and Vandermosten et al., 2011) controlled for the complexity of the non-speech stimuli (Parviainen, Helenius, & Salmelin, 2005). Furthermore, the size of contrasts was incomparable for the speech and non-speech condition in many studies but it might be the most influential variable in this context (Bishop, 2007). As already mentioned, the speech perception deficits in dyslexia are only very subtle and therefore hard to discover (Mody et al., 1997). Indeed, it is unsurprising that studies in which the non-speech contrasts were many times higher than those of the speech stimuli do not report auditory deficits for non-speech in dyslexia.

Vowel length perception and dyslexia

Problems of phonemic length discrimination are thought to be one additional risk factor for dyslexia (Pennala et al., 2010). Finnish newborns with genetic risk for dyslexia have a differential hemispheric preference for vowel duration changes, as indexed by the MMN. The MMN evoked by vowel duration changes was found to be processed more likely in the right hemisphere, whereas the left hemisphere was more active during this task in age matched newborns without genetic risk for dyslexia (Leppänen, Pihko, Eklund, & Lyytinen, 1999; Pihko et al., 1999). These perceptual deficits could result in underspecified phonological representations (see the *phonological deficit hypothesis*).

The correct discrimination of short and long vowels is crucial for the German orthography. Short vowels are often followed by a double consonant (e.g., *nett* [*engl. nice*]) (Warnke, Schulte-Körne, & Ise, 2012). Long vowels are often followed by a “silent h” (e.g., *Stahl* [*engl. steel*]) and the /i:/ is often written as “ie” (e.g., *Lied* [engl. *song*]) (Landerl, 2003).

The correct spelling of vowel length was found to be difficult even for normally developing children (Klicpera & Gasteiger-Klicpera, 1998). Therefore, it is not surprising that German poor spellers have considerable problems with vowel length categorization (Landerl, 2003). As a result, vowel length discrimination tasks have been included in the latest intervention programs (e.g. Marburger Rechtschreibtraining, Schulte-Körne & Mathwig, 2009; Lautarium, see Klatte, Steinbrink, Prölß, Estner, Christmann, & Lachmann, in press for an evaluation).

The German vowel length discrimination paradigm and dyslexia

The German vowel length discrimination paradigm (Groth et al., 2011; Steinbrink et al., 2012; Steinbrink et al., in preparation) was previously introduced in Chapter 2 of the present work (see vowel length discrimination in German). It was originally developed to investigate the processing of the temporal, spectral and spectro-temporal aspects of speech signals in developmental dyslexia. The advantage of this approach is that it minimizes methodological confounds, like task complexity, which could be the main reason why phonological deficits are found more frequently compared to auditory deficits. As phonological tasks like phoneme deletion, non-word repetition, and RAN show a higher working memory load compared to simple discrimination tasks, the latter should be easier for dyslexic children and adults. That is why a simple same-different task was used within the German vowel length discrimination paradigm to minimize effects of attention and short-term memory. Contrary to prior research in which the temporal aspects of speech signals in dyslexia were investigated by stretching or compressing whole syllables (McAnally, Hansen, Cornelissen, & Stein, 1997) or single phonemes (Rey, Martino, Espesser, & Habib, 2002) the current approach manipulates syllables within the phoneme boundaries of the German language (Groth et al., 2011). Moreover, the temporal difference between tense and lax vowels should be small enough to uncover temporal processing deficits, as they lie within the time window that was proposed by Tallal and Piercy (1975). Note that the spectro-temporal condition of the German vowel length discrimination paradigm is a phonological rather than an auditory task, as it involves the discrimination of original German phonemes. Contrary to this, the temporal and spectral conditions involve auditory processing, as the manipulated vowels are included.

The first study which used the vowel length paradigm (Groth et al., 2011) compared the discrimination performance of 20 dyslexic adolescents and adults in the spectro-temporal

and temporal condition to that of 20 aged-matched controls. All of the participants were German native speakers. All seven German vowel pairs were included and embedded within two non-words (nVp and fVp). Both groups showed no problems with same trials. As mentioned before in Chapter 2, both groups performed nearly perfect within the spectro-temporal condition for all seven vowel pairs. There was, however, a drop in performance in the temporal condition with increasing vowel height in both groups; but the dyslexic adolescents and adults showed consistently inferior performance compared to that of the control group for all seven vowel pairs. This finding supports the idea of a temporal processing deficit in dyslexia (Farmer & Klein, 1995; Tallal, 1980). However, consistent with prior research, this temporal deficit was not found for the whole sample, but only for 65% of the dyslexic participants.

The entire pattern of behavioral results was replicated in a following fMRI study (Steinbrink et al., 2012). The hemodynamic brain activation was recorded while the participants performed the same-different task. Low temporal discrimination scores were associated with decreased activation of the insular cortices and the left inferior frontal gyrus.

The spectral condition, as introduced in Chapter 2, was included in a following behavioral study with 8 to 10 year old children with and without the diagnosis of specific reading disorder (Steinbrink et al., in preparation). Three vowel pairs were used with increasing vowel height: /a/ - /a:/, /o:/ - /ɔ/ and /ɪ/ - /i:/. Performance was better in the spectro-temporal condition compared to the spectral or temporal one for both groups. The discrimination index d' dropped systematically with vowel height in the spectral and temporal condition in both groups. In the temporal condition, performance dropped with vowel height (from /a/ - /a:/ to /ɪ/ - /i:/), whereas the opposite pattern of results could be observed for the spectral condition. This finding is in accordance with the properties of the German vowel system (see Chapter 2 of this thesis). The dyslexic children showed a significantly lower discrimination index for all vowels and conditions except the temporal condition of the vowel pair /ɪ/ - /i:/ and the spectral condition of the vowel pair /a/ - /a:/. These differences probably did not reach significance due to the level of difficulty for both groups. In opposition to the dyslexic adults (Groth et al., 2011; Steinbrink et al., 2012), the dyslexic children were also impaired in the spectro-temporal condition. The explanation could go two different ways. First, dyslexic adults could have been able to compensate their deficit by using the redundant information of the spectro-temporal signal, whereas the

dyslexic children did not yet develop such a strategy. The second explanation concerns the fact that discrimination performance in the spectro-temporal condition was at ceiling level for both groups in the study by Groth and colleagues (2011). There is a possibility that the task was not difficult enough to uncover group differences.

Experiment 2

The main question of this experiment is the specific nature of auditory processing deficits in dyslexia. Most studies which compared auditory processing of speech and non-speech did not control for the complexity of both stimulus types, for their size of contrast and task difficulty. The modified German vowel length discrimination paradigm, as introduced in Chapter 2, is used to investigate the impairment of sound processing in dyslexic adults. This approach enables to control for the complexity of the task and the stimuli, as the same discrimination task is used to investigate several types of stimuli (vowel center stimuli, spectrally rotated vowel center stimuli, and bands of formants) in one sample of participants. As spectrally rotated speech shows the same spectro-temporal properties as the original speech signal, it is an equally complex non-speech analogue. Moreover, multiple acoustical parameters are varied within each type of stimulus, while maintaining task complexity.

Participants

42 German adolescents and adults, aged between 14 and 25 years, participated in this experiment. 21 of them were part of the dyslexic group. They reported having problems in reading and writing since primary school up to now. The control group (N=21) was matched to the dyslexic group with respect to age ($t(40) = 0.70$, $p = .49$), sex (χ^2 (1) = 0.10, $p = .76$) and non-verbal intelligence ($t(40) = -1.21$, $p = .23$) (see Table 11 for details). The Culture Fair Test (CFT 20-R, German version, Weiß, 2006) was used to measure the non-verbal intelligence of each participant. The criterion for inclusion in the study was a non-verbal IQ equal to or above 81. This value corresponds to one standard deviation (15 IQ points) below the mean, which was corrected by the confidence interval reported in the manual (4 IQ points). However, the two groups are not comparable concerning their school education (χ^2 (2) = 10.67, $p = < .01$), with higher education levels for the control group. No one reported a history of neurological diseases, psychiatric or attention disorders or hearing problems.

A German reading test for adults (Schulte-Körne, 2001) was used. The dependent measure was the time, which was required to read a list of real words and a list of non-words. The number of errors was also taken into consideration. In addition, all participants completed a standardized German spelling test for adolescents and adults (Rechtschreibungstest, (RT); Kersting & Althoff, 2004).

Table 11: Comparison of the two groups investigated in Experiment 2 in relation to age, sex and IQ.

		Dyslexics	Controls	Comparison of groups	*p*
Age [years]	Mean	19.10	18.48	$t(40) = 0.70$	.49
	Minimum	14	15		
	Maximum	25	22		
	SD	3.45	2.16	$F(20,20) = 2.55$	.02
Sex	N(male)	12	11	χ^2 (1) = 0.10	.76
	N(female)	9	10		
IQ	Mean	102.86	107.95	$t(40) = -1.21$	.23
	SD	13.23	14.03	$F(20,20) = 1.12$	.40

The dyslexic group's performance on the reading and writing test was significantly poorer compared to the that of the control group (see Table 12 for details), indicated by slower word ($t(40) = 3.30$, $p < .01$) and non-word reading ($t(40) = 4.96$, $p < .01$), less accurate reading of words ($t(40) = 4.96$, $p < .01$) and non-words ($t(40) = 2.78$, $p < .01$), more errors ($t(40) = 7.14$, $p < .01$) and poorer standard values ($t(40) = -6.93$, $p < .01$) in the spelling test.

Table 12: Comparison of the two groups investigated in Experiment 2 in relation to reading and writing skills as revealed by t-tests of independent samples.

			Dyslexics	Controls	Comparison of groups	*p*
Reading words	Speed [s]	mean	55.95	42.29	$t(40) = 3.30$	< .01
		SD	12.53	14.28	$F(20,20) = 1.30$	.28
	Errors	mean	2.90	0.62	$t(40) = 4.96$	< .01
		SD	2.23	1.16	$F(20,20) = 3.70$	< .01
Reading non-words	Speed [s]	mean	102.52	70.10	$t(40) = 4.96$	< .01
		SD	22.84	19.41	$F(20,20) = 1.38$	.24
	Errors	mean	9.19	4.43	$t(40) = 2.78$	< .01
		SD	6.42	4.53	$F(20,20) = 2.01$	.06
Writing (RT)	Raw value	mean	36.24	13.38	$t(40) = 7.14$	< .01
		SD	9.29	11.35	$F(20,20) = 1.49$	.19
	Standard value	mean	81.52	102.63	$t(40) = -6.93$	< .01
		SD	9.29	10.63	$F(20,20) = 1.31$	.28

Material

This experiment included a subset of the stimuli of Experiment 1. Three stimulus types (vowel center stimuli with full spectrum, a modified version of the spectrally rotated vowel center stimuli and the bands of formants based on the vowel center stimuli) and both vowel types (/a/ - /a:/ and /ɪ/ - /i:/) were included. The same auditory contrasts (temporal, spectral and spectro-temporal) like in Experiment 1 were used. The unfiltered version of the vowels was chosen, as they sound more natural (see Experiment 1). In like fashion, the spectrally rotated stimuli were modified to obtain non-speech stimuli with the same complexity and the full frequency spectrum of the vowels by adding all frequencies of the vowel above 4000Hz to the spectrally rotated stimulus. This means that only the lower part (below 4000Hz) was modified by the inversion. The upper frequencies were not affected (see Figure 24). The adding of frequencies above 4000Hz was performed in Audition (version CS5.5, Abobe). Importantly, this new approach enables to compare equally complex speech and non-speech stimuli without prior low pass filtering of the speech signal. The spectrograms of these spectrally rotated vowel center stimuli are shown in Figures 25 and 26.

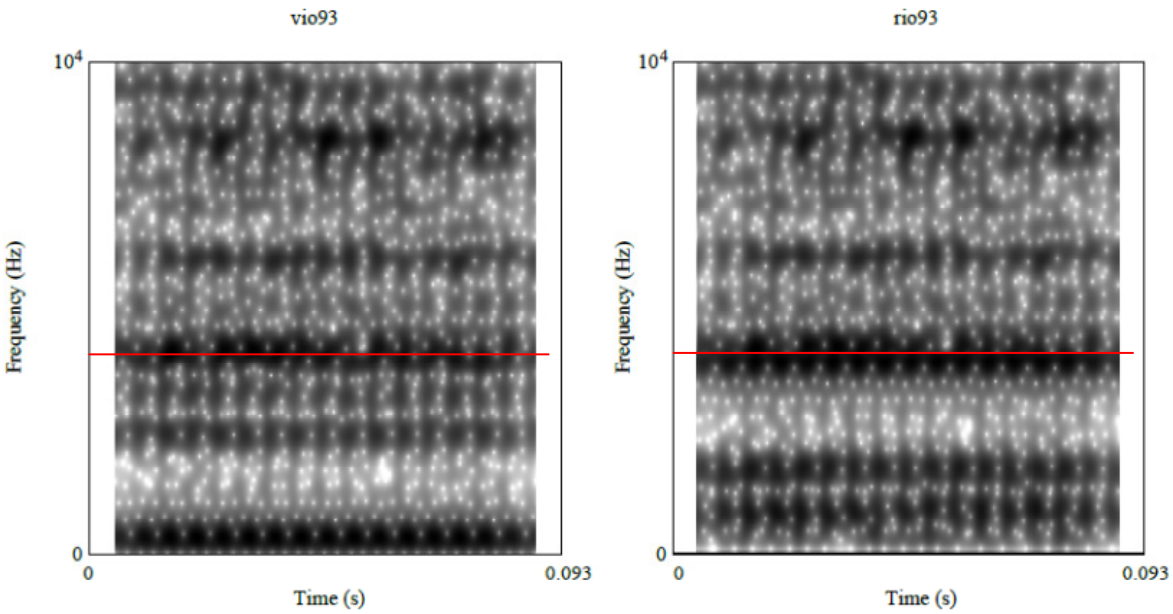

Figure 24: Spectrograms of the vowel center stimulus based on /i:/ and the modified spectrally rotated version of this stimulus. Only the lower part below 4000Hz, indicated by the red line, was modified by the inversion. The upper frequencies were not affected.

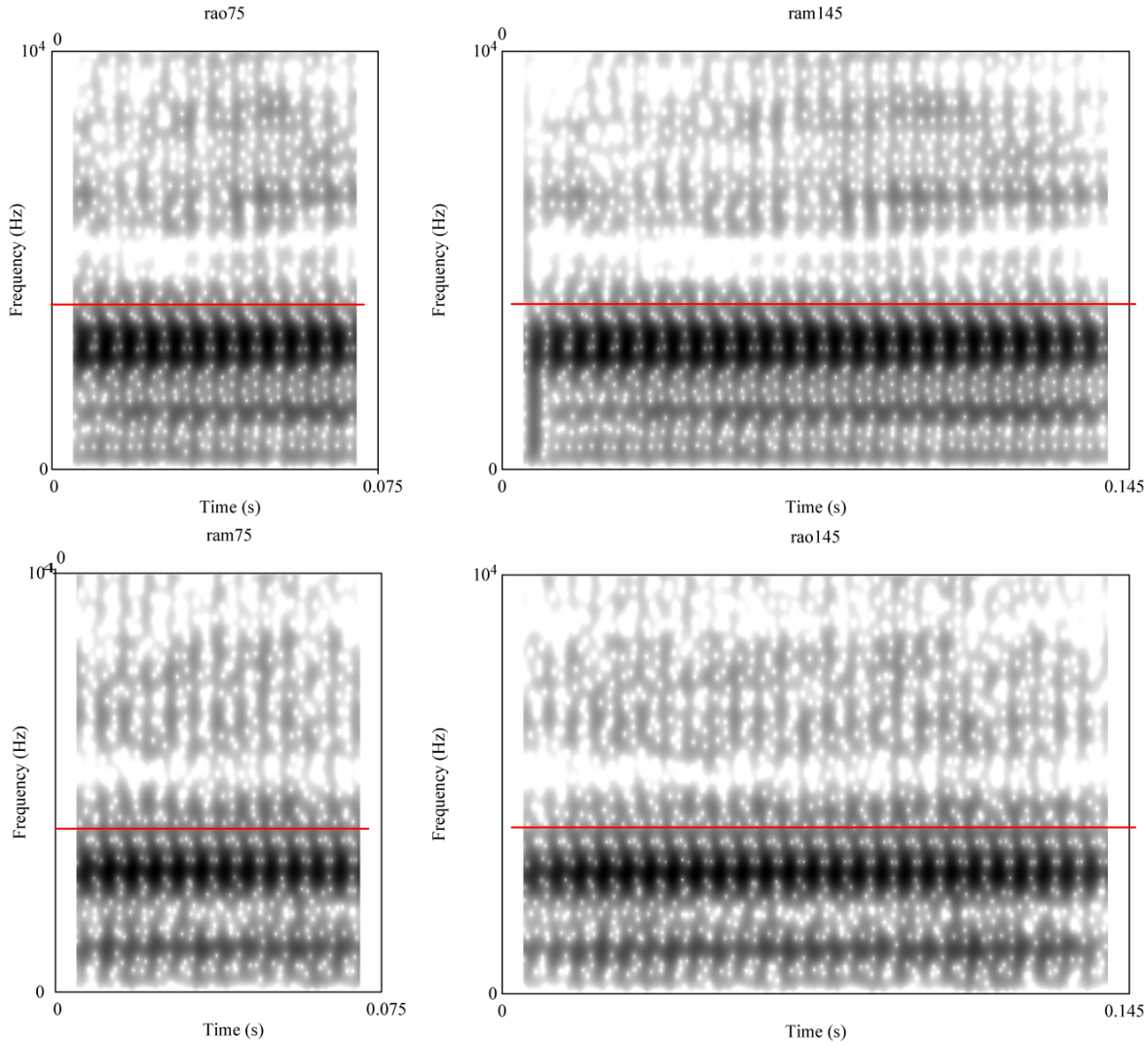

Figure 25: Spectrograms of the four spectrally rotated vowel center stimuli with complete spectrum based on the vowel pair /a/ - /a:/. Only the lower part, below 4000Hz, indicated by the red line, was modified by the inversion. The upper frequencies were not affected.

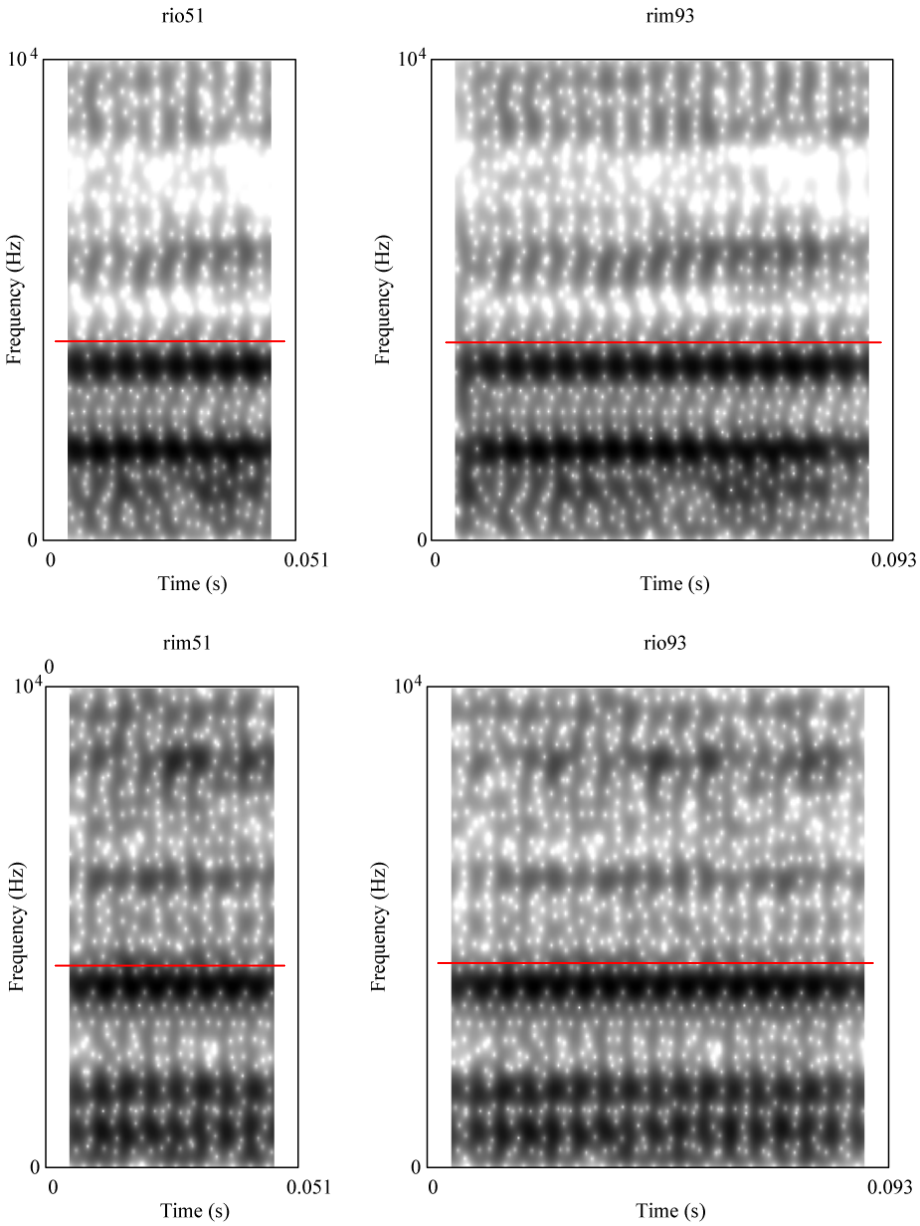

Figure 26: Spectrograms of the four spectrally rotated vowel center stimuli with complete spectrum based on the vowel pair /ɪ/ - /i:/. Only the lower part, below 4000Hz, indicated by the red line, was modified by the inversion. The upper frequencies were not affected.

Task and apparatus

The task was the same as in Experiment 1 and the same equipment with equal settings was used (see Chapter 2 for details). After having completed the same-different task, each participant listened again to the three stimulus types and was asked to rate each category as speech-like (7 points) or completely non-speech-like (1 point) or something in between.

Design

The complete design is illustrated in Table 13. In total, one block comprised 192 trials with one block for each stimulus type. Together there were *three stimulus types*: vowel center stimuli with full spectrum, spectrally rotated vowel center stimuli with full spectrum and bands of formants based on the vowel center stimuli with the full spectrum. The order of the blocks was counterbalanced between participants.

Within each block there were *two vowel types*: /a/ - /a:/ and /ɪ/ - /i:/. During one half of the trials one stimulus was presented twice (same condition), whereas two different stimuli could be distinguished during the second half of the trials (different condition). There were *three types of auditory difference*: temporal, spectral and both.

The order of the trials in each block was pseudo randomized in accordance with the following rules: there were a maximum of three trials in sequence which required the same response and, in addition, vowel identity changed at least after every third trial.

Dependent variables

d' was calculated as reported by Macmillan and Creelman (1991) for same-different designs (see Chapter 2 for details). The mean reaction times to correct responses were also calculated. Reaction times which exceeded three seconds were excluded from the analysis.

Table 13: Experimental design in Experiment 2.

	Type of stimulus	different condition			same condition
		temporal	spectral	both	
/a/ - /a:/	**vowel center**	Vao75 vs. Vam145 (8x) Vao145 vs. Vam75 (8x)	Vao75 vs. Vam75 (8x) Vao145 vs. Vam145 (8x)	Vao75 vs. Vao145 (16x)	Vao75 vs. Vao75 (12x) Vao145 vs. Vao145 (12x) Vam75 vs. Vam75 (12x) Vam145 vs. Vam145 (12x)
	spectrally rotated vowel center	Rao75 vs. Ram145 (8x) Rao145 vs. Ram75 (8x)	Rao75 vs. Ram75 (8x) Rao145 vs. Ram145 (8x)	Rao75 vs. Rao145 (16x)	Rao75 vs. Rao75 (12x) Rao145 vs. Rao145 (12x) Ram75 vs. Ram75 (12x) Ram145 vs. Ram145 (12x)
	bands of formants on the vowel center	Bao75 vs. Bam145 (8x) Bao145 vs. Bam75 (8x)	Bao75 vs. Bam75 (8x) Bao145 vs. Bam145 (8x)	Bao75 vs. Bao145 (16x)	Bao75 vs. Bao75 (12x) Bao145 vs. Bao145 (12x) Bam75 vs. Bam75 (12x) Bam145 vs. Bam145 (12x)
/ɪ/ - /i:/	**vowel center**	Vio51 vs. Vim93 (8x) Vio93 vs. Vim51 (8x)	Vio51 vs. Vim51 (8x) Vio93 vs. Vim93 (8x)	Vio51 vs. Vio93 (16x)	Vio51 vs. Vio51 (12x) Vio93 vs. Vio93 (12x) Vim51 vs. Vim51 (12x) Vim93 vs. Vim93 (12x)
	spectrally rotated vowel center	Rio51 vs. Rim93 (8x) Rio93 vs. Rim51 (8x)	Rio51 vs. Rim51 (8x) Rio93 vs. Rim93 (8x)	Rio51 vs. Rio93 (16x)	Rio51 vs. Rio51 (12x) Rio93 vs. Rio93 (12x) Rim51 vs. Rim51 (12x) Rim93 vs. Rim93 (12x)
	bands formants on the vowel center	Bio51 vs. Bim93 (8x) Bio93 vs. Bim51 (8x)	Bio51 vs. Bim51 (8x) Bio93 vs. Bim93 (8x)	Bio51 vs. Bio93 (16x)	Bio51 vs. Bio51 (12x) Bio93 vs. Bio93 (12x) Bim51 vs. Bim51 (12x) Bim93 vs. Bim93 (12x)

Hypotheses

(1) The spectro-temporal condition should be easier to discriminate compared to the temporal or spectral contrast for both groups (see Chapter 2 for details)

(2) The influences of the German vowel system should be observable in both groups (see Chapter 2 for details):

   a) For the vowel pair /a/ - /a:/, the spectral condition should be more difficult compared to the temporal condition in both groups

   b) For the vowel pair /ɪ/ - /i:/, the temporal condition should be more difficult compared to the spectral condition in both groups

   a) This interaction of vowel type and auditory contrast should be the most salient for the vowel center stimuli compared to the two non-speech stimulus types

(3) Concerning group differences, the same pattern of results as reported by Groth and colleagues (2011) should be observed for the vowel center stimuli, as a similar approach was chosen:

   a) Both groups should perform at ceiling level in the spectro-temporal condition

   b) The dyslexic group should be impaired in the temporal condition, indicated by smaller discrimination indexes

(4) As dyslexic children were severely impaired in the spectral condition of the German vowel length discrimination paradigm (Steinbrink et al., in preparation) and due to the fact that spectral deficits have also been found in dyslexic adults (Ahissar et al., 2000), the dyslexic adults should also be impaired in the spectral condition of this experiment

(5) If the auditory deficit can be generalized to the processing of non-speech stimuli, dyslexic adults should also be impaired in the spectral and temporal condition of the spectrally rotated vowel center stimuli, as these stimuli show the same complexity and a similar size of contrasts compared to the vowel center stimuli (see Chapter 2)

(6) If the auditory deficit can be generalized to the processing of non-speech stimuli even of lower complexity compared to speech stimuli, dyslexic adults should also be impaired in the spectral and temporal condition of the bands of formants (see Chapter 2)

Results

A 3*2*3*2 analysis of variance (ANOVA) with repeated measures was conducted, including the within-factors *Stimulus type* (3: vowel center stimuli vs. spectrally rotated vowel center stimuli vs. bands of formants based on the vowel center stimuli), *Vowel type* (2: /a/ - /a:/ vs. /ɪ/ - /i:/) and *Auditory contrast* (3: temporal vs. spectral vs. spectro-temporal) and the *Group* factor (2: dyslexic group vs. control group). An overall view of the data is given in Table 14 and Figure 27-29. Every time the assumption of sphericity was rejected as revealed by the Mauchly's test, *F* values were corrected according to Greenhouse-Geisser. The Bonferroni correction was used whenever multiple *t*-tests for independent and dependent samples were conducted.

The ANOVA based on the discrimination index d' revealed a significant main effect of *Stimulus type* ($F(2,80) = 8.18$, $p < .01$). The spectrally rotated vowel center stimuli were easier to discriminate compared to the vowel center stimuli ($t(41) = 4.62$, $p < .01$, $d = 0.72$) and the bands of formants ($t(41) = 3.03$, $p = .01$, $d = 0.48$). The difference between the vowel center stimuli and bands of formants did not reach significance ($t(41) = -0.98$, $p = .33$).

The main effect of *Auditory contrast* was also found to be significant ($F(2,80) = 87.49$, $p < .01$). The spectro-temporal condition was discriminated more accurately compared to the spectral ($t(41) = 7.68$, $p < .01$, $d = 1.19$) and temporal condition ($t(41) = 11.40$, $p < .01$, $d = 1.76$). The temporal condition was discriminated less accurately compared to the spectral one ($t(41) = -7.51$, $p < .01$, $d = 1.16$).

There was a significant main effect of *Vowel type* ($F(1,40) = 7.58$, $p < .01$). The vowel pair /ɪ/ - /i:/ was easier to discriminate compared to the vowel pair /a/ - /a:/ ($t(41) = 2.70$, $p = .01$, $d = 0.42$). However, there was a significant interaction between *Stimulus* and *Vowel type* ($F(2,80) = 14.63$, $p < .01$). The difference of performance between the two vowel pairs /a/ - /a:/ and /ɪ/ - /i:/ was only significant for the vowel center stimuli ($t(41) = -4.79$, $p < .01$, $d = 0.74$) and not for the two non-speech stimulus types ($t(41) = -0.90$, $p = .37$) for the rotated vowels and $t(41) = 1.51$, $p = .14$ for the bands of formants).

Moreover, the ANOVA revealed a significant interaction between *Stimulus type* and *Auditory contrast* ($F(4,160) = 2.13$, $p < .01$). For the temporal condition, no significant differences between the *Stimulus type* were found ($F(2,80) = -1.17$, $p = .31$), whereas discrimination performance varied systematically for the spectral ($F(2,80) = 18.94$, $p < .01$) and spectro-

temporal conditions ($F(2,80) = 7.11$, $p < .01$) for different stimulus types as revealed by three additional analyses of variance.

Table 14: Results of the analysis of variances based on d' in Experiment 2.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Stimulus type	8.18	2	80	< .01	.17
Vowel type	7.58	1	40	< .01	.16
Auditory difference	87.49	2	80	< .01	.69
Group	11.98	1	40	< .01	.23
Stimulus * Vowel type	14.63	2	80	< .01	.27
Stimulus type * Auditory contrast	2.13	4	160	< .01	.21
Vowel type * Auditory contrast	133.41	2	80	< .01	.77
Stimulus * Vowel type * Auditory contrast	34.53	4	160	.44	.02
Stimulus type * Group	0.08	2	80	.89	< .01
Vowel type * Group	7.58	1	40	.11	.06
Auditory contrast * Group	0.65	2	80	.53	.02
Stimulus * Vowel type * Group	0.183	2	80	.83	< .01
Stimulus type * Auditory contrast * Group	1.83	4	160	.13	.04
Vowel type * Auditory contrast * Group	0.41	2	80	.66	.01
Stimulus * Vowel type * Auditory contrast * Group	0.88	4	160	.48	.02

Moreover, there was a significant interaction between *Vowel type* and *Auditory contrast* ($F(2,80) = 133.41$, $p < .01$). The vowel pair /ɪ/ - /i:/ was harder to discriminate compared to the vowel pair /a/ - /a:/, but only in the temporal condition ($t(41) = -8.25$, $p < .01$, $d = 1.26$). The opposite pattern of results was found for the spectral ($t(41) = 10.08$, $p < .01$) and spectro-temporal condition ($t(41) = 4.54$, $p < .01$).

To prove the influences of the German vowel system, additional t-tests for dependent samples were conducted: For the vowel center stimuli, the temporal condition was easier compared to the spectral one for the vowel pair /a/ - /a:/ ($t(41) = 5.20$, $p < .01$, $d = 0.81$), whereas the spectral condition was found to be easier compared to the temporal one for the

vowel pair /ɪ/ - /i:/ ($t(41) = 11.80$, $p < .01$, $d = 1.82$). This pattern of results was only found for the vowel pair /ɪ/ - /i:/ in the two non-speech stimuli ($t(41) = 8.74$, $p < .01$, $d = 1.36$ for the rotated vowels and $t(41) = 9.87$, $p < .01$, $d = 1.52$ for the bands of formants). The temporal condition of the vowel pair /a/ - /a:/ was even harder to discriminate than the spectral one in the spectrally rotated vowel center stimuli ($t(41) = -4.72$, $p <.01$, $d = 0.73$). The same direction was observed for the bands of formants, but this difference did not reach significance ($t(41) = -2.31$, $p = .16$). The triplet interaction between type of stimulus, type of vowel and auditory difference did not reach significance ($F(4,160) = 34.53$, $p = .44$).

A significant main effect of group was found ($F(1,40) = 11.98$, $p < .01$, $f = 0.54$), explained by a significantly better performance of the control group ($t(40) = 3.46$, $p < .01$, $d = 1.08$) (see Figures 27-29 for group comparisons in the temporal, spectral and spectro-temporal condition). None of the interactions with the *Group* factor reached significance (see Table 14).

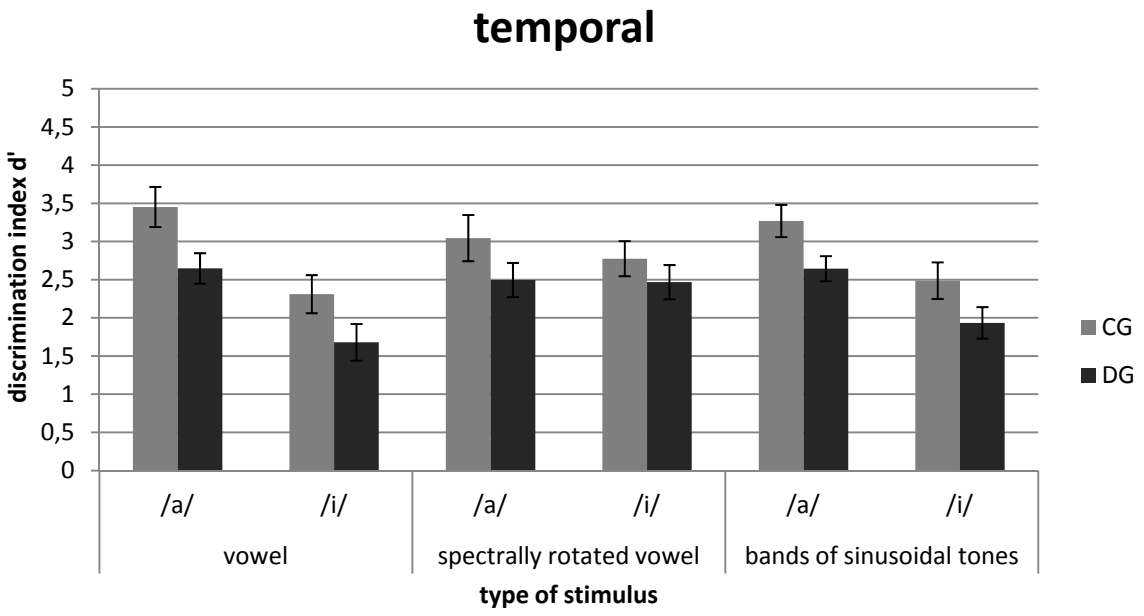

Figure 27: Means and standard errors of d' for the temporal condition of Experiment 2. The discrimination index d' is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the temporal condition.

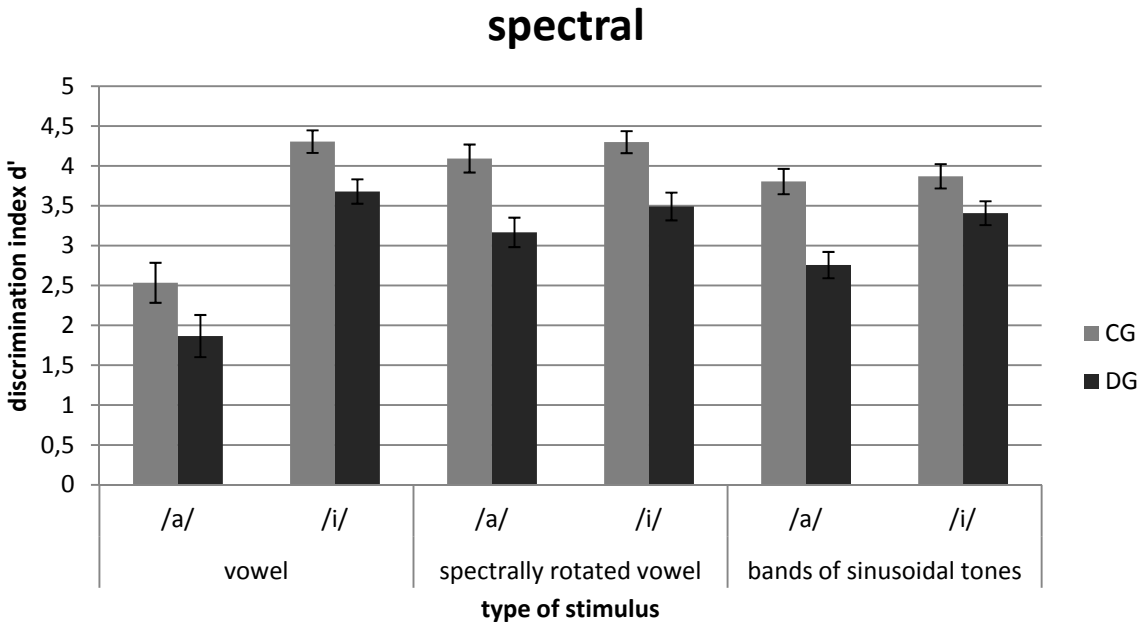

Figure 28: Means and standard errors of d' for the spectral condition of Experiment 2. The discrimination index d' is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the spectral condition.

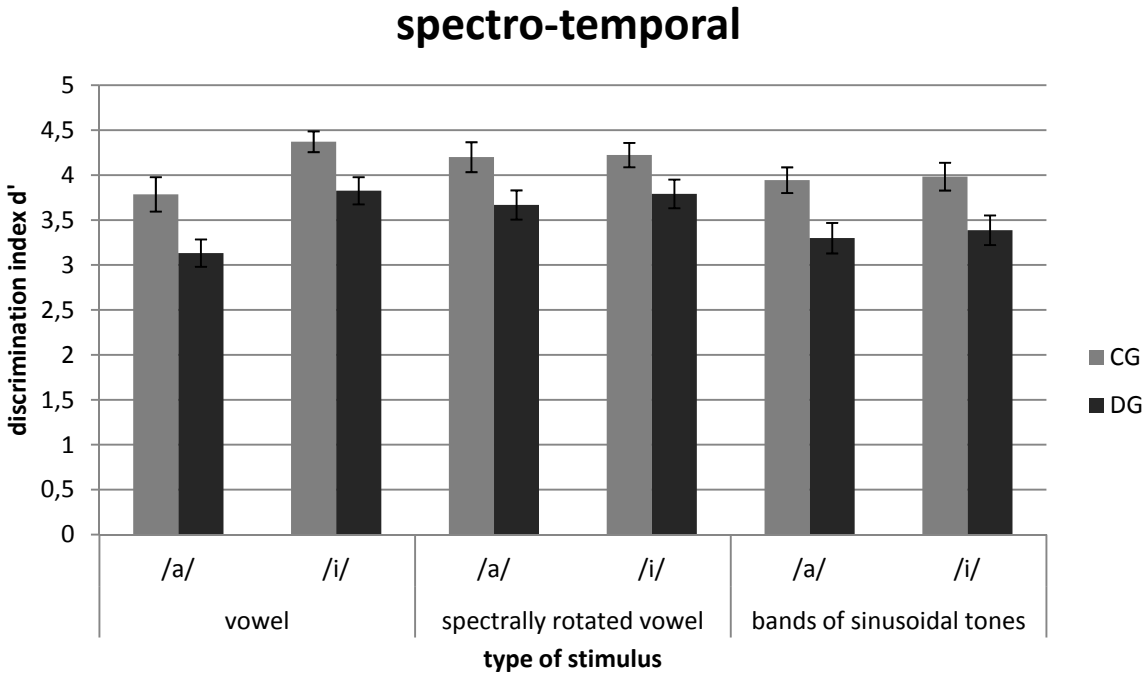

Figure 29: Means and standard errors of d' for the spectro-temporal condition of Experiment 2. The discrimination index d' is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the spectro-temporal condition.

A second ANOVA based on the reaction times was conducted. The main effects of *Stimulus type* ($F(2,80) = 2.4$, $p = .10$) and *Vowel type* ($F(1,40) = 1.19$, $p = .28$) did not reach significance. There was a significant main effect of *Auditory contrast* ($F(2,80) = 71.24$, $p < .01$). Responses to spectro-temporal contrasts were faster compared to spectral ($t(41) = 5.41$, $p < .01$, $d = 0.85$) or temporal ones ($t(41) = 10.00$, $p < .01$, $d = 1.56$) and responses to spectral contrasts were faster compared to those of the temporal ones ($t(41) = 7.41$, $p < .01$, $d = 1.12$).

Table 15: Results of the analysis of variance based on reaction times in Experiment 2.

Factor	F	df(factor)	df(error)	p	partial eta²
Stimulus type	2.40	2	80	.10	.06
Vowel type	1.19	1	40	.28	.03
Auditory difference	71.24	2	80	< .01	.64
Group	1.38	1	40	.25	.03
Stimulus * Vowel type	4.81	2	80	.01	.11
Stimulus type * Auditory contrast	0.73	4	160	.57	.02
Vowel type * Auditory contrast	17.55	2	80	< .01	.31
Stimulus * Vowel type * Auditory contrast	8.77	4	160	< .01	.18
Stimulus type * Group	1.25	2	80	.29	.03
Vowel type * Group	0.09	1	40	.77	< .01
Auditory contrast * Group	0.55	2	80	.58	.01
Stimulus * Vowel type * Group	0.27	2	80	.76	< .01
Stimulus type * Auditory contrast * Group	1.01	4	160	.40	.03
Vowel type * Auditory contrast * Group	0.12	2	80	.89	< .01
Stimulus * Vowel type * Auditory contrast * Group	3.42	4	160	.01	.08

There was a significant interaction between *Stimulus type* and *Vowel type* ($F(2,80) = 4.81$, $p = .01$). A significant difference between the vowel pair /a/ - /a:/ and /ɪ/ and /i:/ was found only for the vowel center stimuli ($t(41) = 2.93$, $p = .02$, $d = 0.44$) and not for the spectrally rotated stimuli ($t(41) = 0.92$, $p = .36$) or the bands of formants ($t(41) = -2.10$, $p = .12$). The interaction between *Stimulus type* and *Auditory contrast* was not statistically significant ($F(4,160) = 0.73$,

p = .57). In addition, the ANOVA revealed a significant interaction between *Vowel type* and *Auditory difference* ($F(2,80) = 17.55$, $p < .01$) and a significant triplet interaction between *Stimulus type, Vowel type*, and *Auditory difference* ($F(4,160) = 8.77$, $p < .01$). In the temporal condition, responses to the vowel pair /a/ - /a:/ were faster compared to the vowel pair /ɪ/ - /i:/ ($t(41) = -4.19$, $p < .01$, $d = 0.65$), whereas the opposite pattern of results was observed for the spectral ($t(41) = 3.41$, $p < .01$, $d = 0.53$) and spectro-temporal condition ($t(41) = 3.17$, $p < .01$, $d = 0.48$).

The main effect of *Group* did not reach significance ($F(1,40) = 1.19$, $p = .25$) and the interaction between *Stimulus type*, *Vowel type*, *Auditory contrast* and *Group* ($F(4,160) = 3.42$, $p = .01$) was the only one which proved to be statistically significant (see Table 15 and Figure 30-32 for details). However, none of the group comparisons for each sub-condition revealed significant group differences on reaction time after the Bonferroni correction.

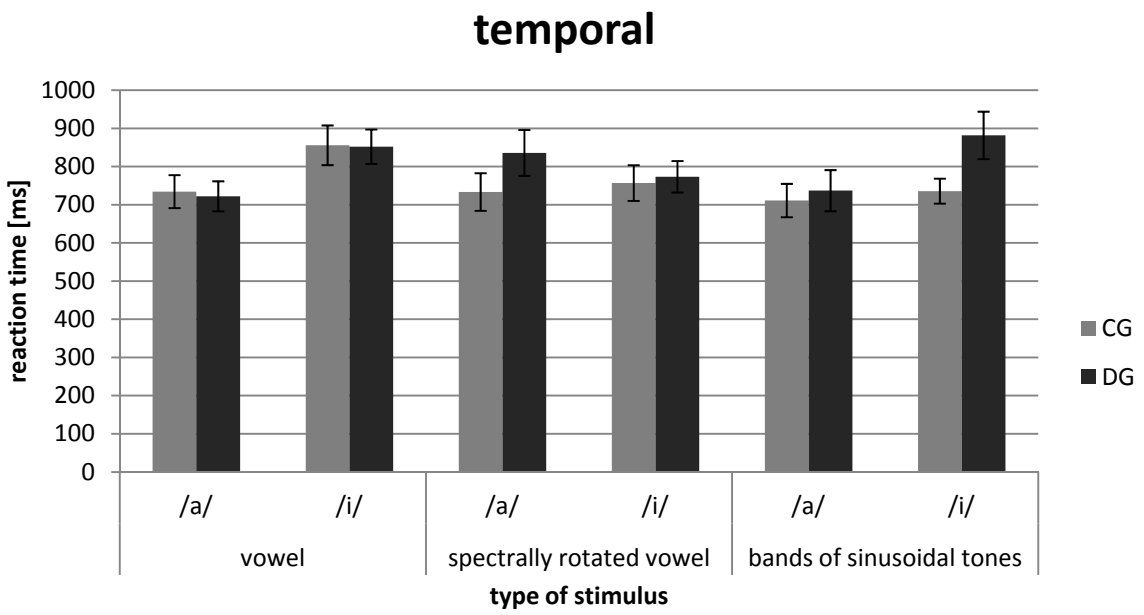

Figure 30: Means and standard errors of reaction times for the temporal condition of Experiment 2. The reaction time is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the temporal condition.

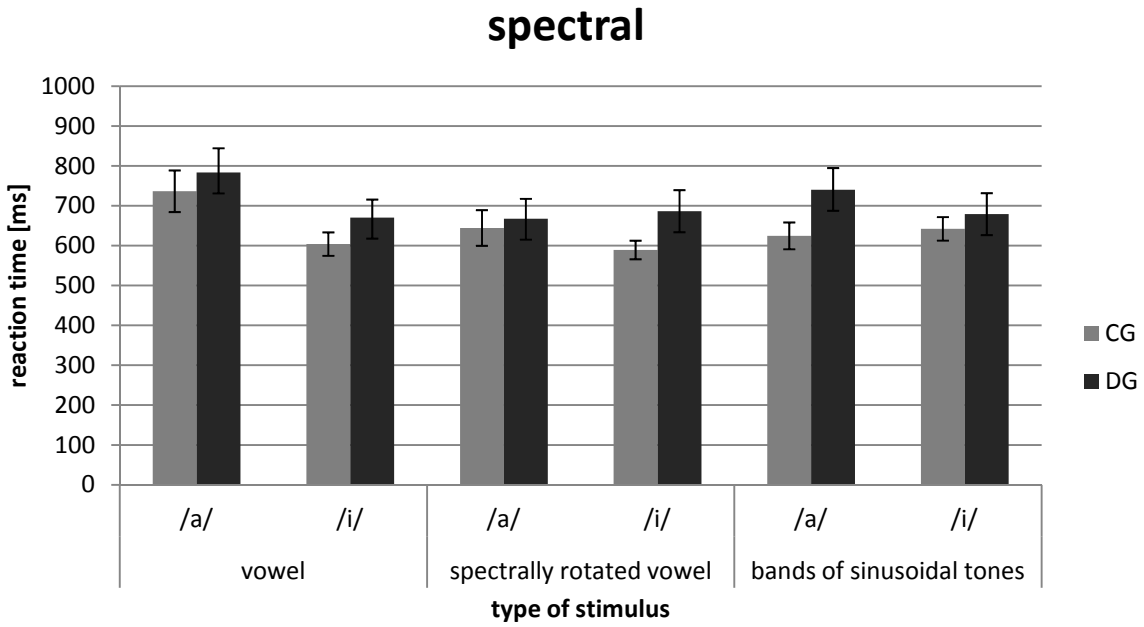

Figure 31: Means and standard errors of reaction times for the spectral condition of Experiment 2. The reaction time is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the spectral condition.

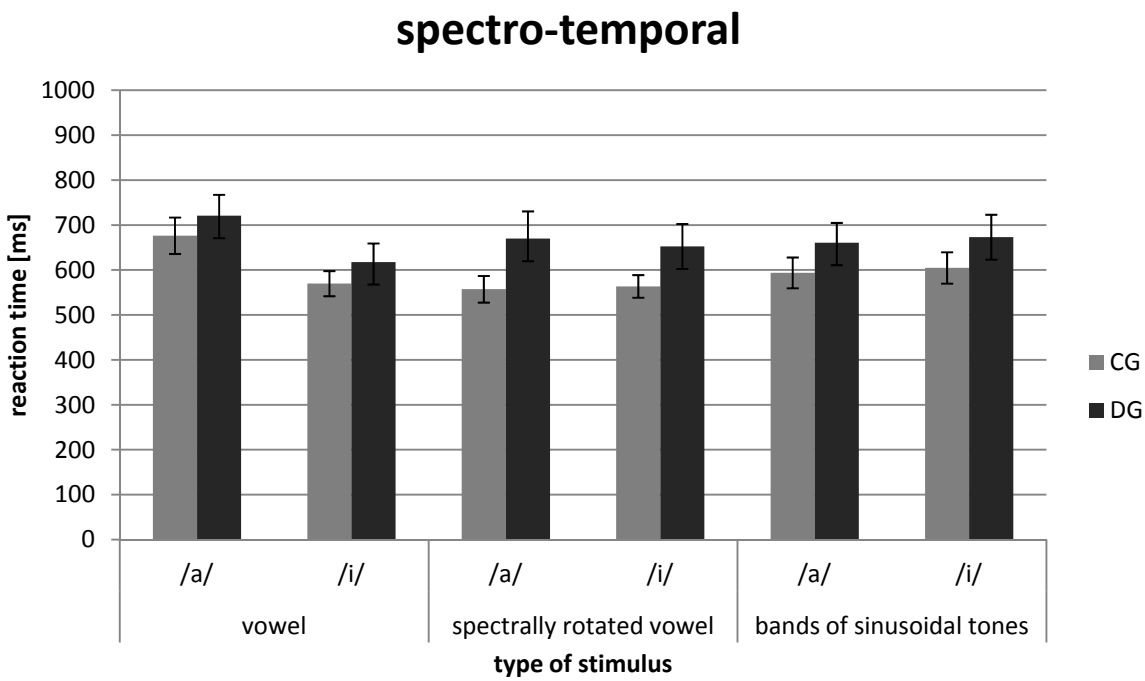

Figure 32: Means and standard errors of reaction times for the spectro-temporal condition of Experiment 2. The reaction time is displayed on the y-axis. Both groups (control group = grey, dyslexic group = black) are compared for each type of stimulus in the spectro-temporal condition.

To check whether the performance of the dyslexic groups for the vowel center stimuli were associated with their performance for the two non-speech stimuli, bivariate correlations according to Pearson were calculated separately for each auditory contrast on the basis of d' (see Table 16).

Table 16: Pearson correlations between the discrimination performance for vowel center stimuli and the two non-speech stimulus types for the dyslexic group.

N = 21 (dyslexic group)		spectrally rotated vowel center stimuli	bands of formants
vowel	temporal	.80**	.65**
center	spectral	.67**	.29
stimuli	spectro-temporal	.65**	.49*

Comparable to Ramus and colleagues (2003), each dyslexic adult was classified in accordance with his or her individual pattern of deficits. A deficit was defined as one standard deviation under the control group's performance. This procedure was performed twice - one time for the three stimulus types (see Figure 33) and one time for the three auditory contrasts (see Figure 34).

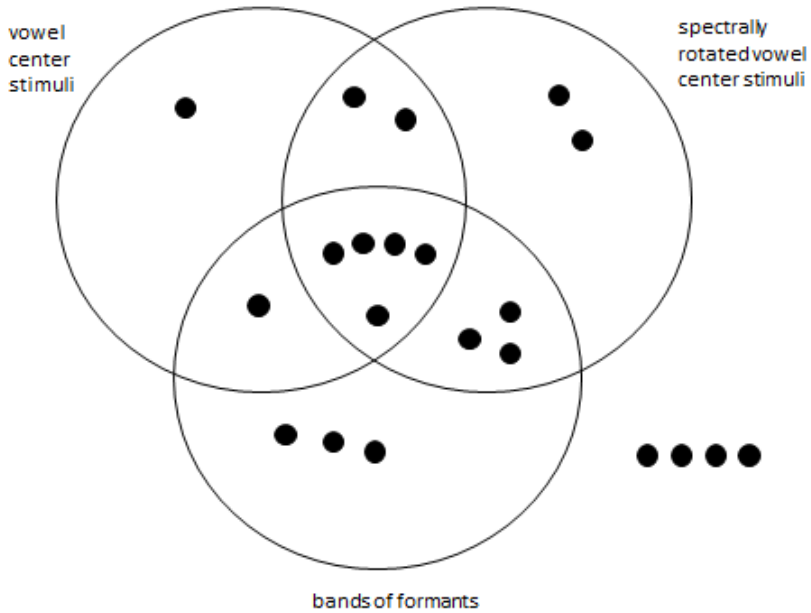

Figure 33: Classification of each dyslexic participant's deficit based on the three stimulus types. Each dot represents one person.

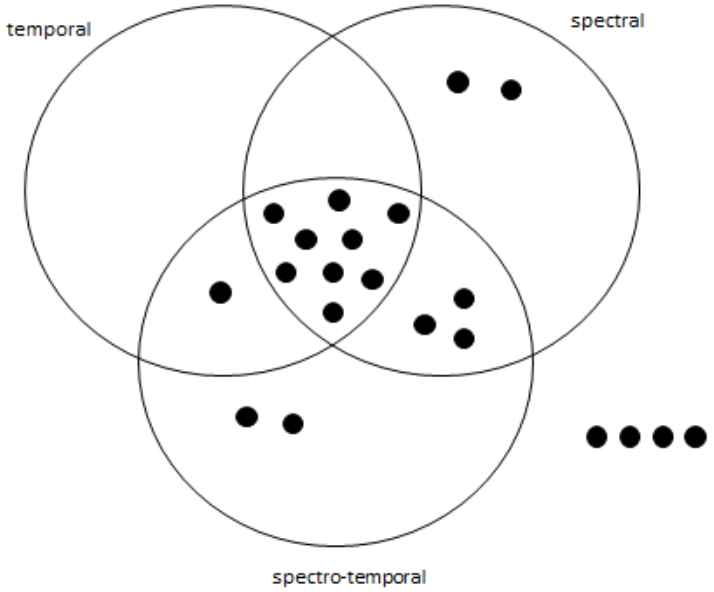

Figure 34: Classification of each dyslexic participant's deficit based on the three auditory contrasts. Each dot represents one person.

Four individuals did not show any deficit at all. There was only one dyslexic participant who was impaired specifically for the vowel center stimuli. Eight of the remaining sixteen dyslexic participants showed deficits in both the vowel center stimuli and in at least one non-speech type of stimulus. The other half was impaired only for non-speech stimuli (see Figure 33 for details).

There was not a single dyslexic participant with a specific temporal deficit. Nine participants showed deficits for all three auditory contrasts. Only two participants were specifically impaired in the spectral and spectro-temporal condition each time (see Figure 34 for details).

To rule out a speed-accuracy trade off, the correlation between the correctness of the answer (0 = error, 1 = correct response) and the reaction time was calculated for each group. The correlation was $r = -.12$ ($p < .01$) for the control group and $r = -.15$ ($p < .01$) for the dyslexic group.

To take into account the role of attention, the order of blocks was chosen as within factor in two additional ANOVAs, one based on the discrimination index d' and another one based on reaction time (see Figures 35 and 36). For the discrimination index d' there was a tendency for better performance in later blocks compared to the first one ($F(2,80) = 2.72$, $p = .07$). Performance in the last block was significantly better compared to the first one ($t(41) = 2.41$,

p = .02, d = 0.38). The interaction between group and position of the block was not found to reach significance (F(2,80) = 0.75, p = .48).

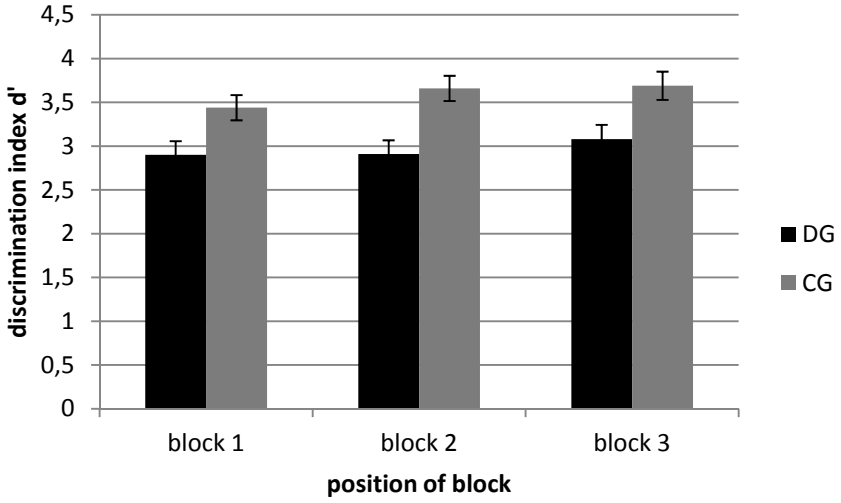

Figure 35: Means and standard errors of the discrimination index of each group for each experimental block.

There was no overall shift in reaction time for both groups (F(2,80) = 1.39, p = .26), but there was a significant interaction between group and position of block (F(2,80) = 4.70, p = .01). T-tests for dependent samples did not reveal any significant shifts of reaction time in the control group, whereas the dyslexic group's reaction time was significantly slower at the end of the experiment compared to the first block (t(20) = 2.61, p = .05, d = 0.57) (see Figure 36).

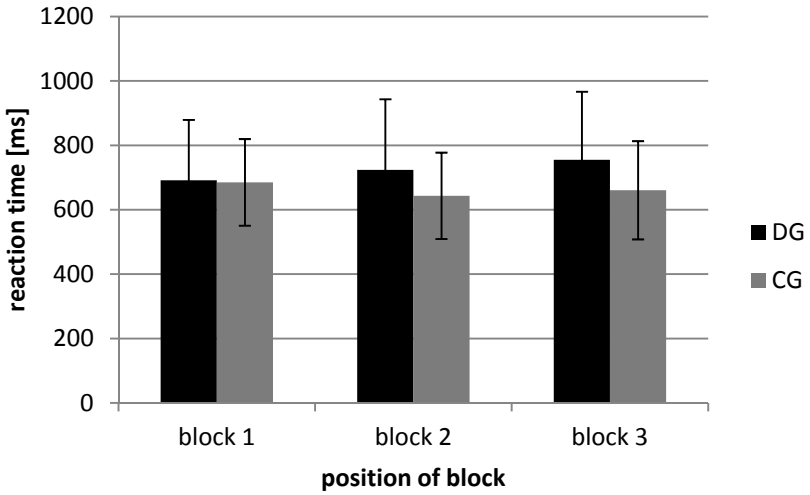

Figure 36: Means and standard errors of the reaction time of each group for each experimental block.

The last ANOVA concerned the stimulus ratings. A significant main effect of type of stimulus was found ($F(2,80) = 138.80$, $p < .01$). The vowel center stimuli were rated as to be more similar to speech compared to the spectrally rotated vowels ($t(41) = 13.93$, $p < .01$) or the bands of formants ($t(41) = 13.31$, $p < .01$), whereas no difference was found for the spectrally rotated vowel center stimuli and the bands of formants ($t(41) = 0.00$, $p = 1.00$) (see Figure 37). The main effect of group ($F(1,40) = 3.52$, $p = .07$) and the interaction between type of stimulus and group did not reach significance ($F(2,80) = 0.56$, $p = .57$).

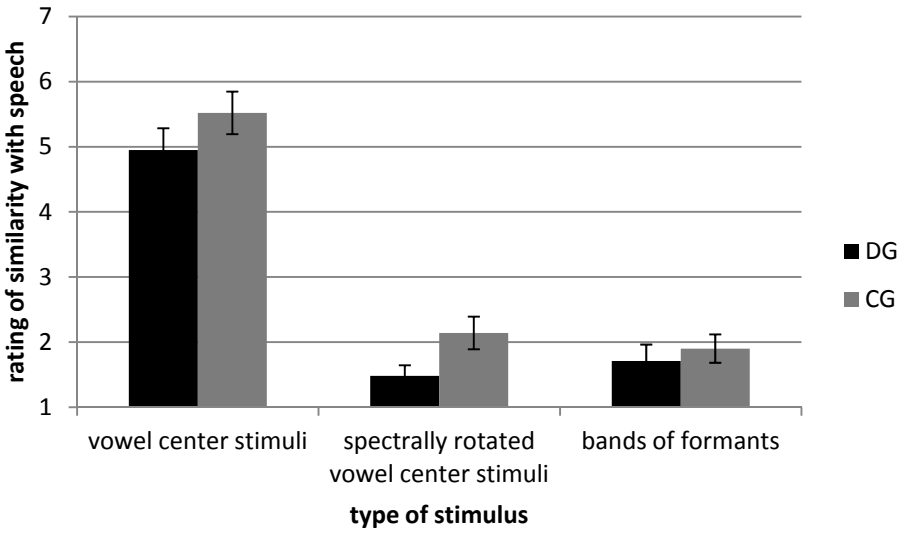

Figure 37: Means and standard errors of the stimulus rating seperately for each group (dyslexic group = grey, control group = black). Higher values mean that the stimulus was more likely to be rated as speech like.

Discussion

The modified German vowel length discrimination paradigm (see Experiment 1) was used to investigate dyslexic adults. Their performance was compared to that of an age and IQ matched control group. The discrimination performance will be compared to those reported by Groth and colleagues (2011), Steinbrink and colleagues (2012) and Steinbrink and colleagues (in preparation). Moreover, the role of the processing of non-speech stimuli in dyslexia will be discussed, followed by some comments on the role of attention and subgroups in dyslexia. In the end, imperfections regarding the choice of participants and stimuli as well as the outlooks on future trends in research will be offered.

Vowel length discrimination in adults with and without dyslexia

The same overall pattern of results as observed in Experiment 1 was found for both groups: The spectro-temporal condition was easier to discriminate compared to the temporal or spectral contrast for both groups (see Hypothesis 1). Furthermore, the influences of the German vowel system were observable in both groups (Hypothesis 2). Regarding the vowel pair /a/ - /a:/ the spectral condition was more difficult compared to the temporal condition (Hypothesis 2a), whereas for the vowel pair /ɪ/ - /i:/ the temporal condition was more difficult compared to the spectral condition (Hypothesis 2b). This interaction of vowel type and auditory contrast was only found for the vowel center stimuli and not for the two non-speech stimulus types (Hypothesis 2c).

As expected, the dyslexic group's performance was significantly worse compared to the control group in the temporal (Hypothesis 3b) and the spectral condition (Hypothesis 4) for the vowel center stimuli. These results correspond with those reported by Groth and colleagues (2011) and Steinbrink and colleagues (in preparation). However, the dyslexic adults of the current experiment were also impaired in the spectro-temporal condition. This result was unexpected, as no differences were found by Groth and colleagues (2011) and Steinbrink and colleagues (2012) for dyslexic adults (Hypothesis 3a). Both groups were, however, at ceiling level in these two studies. Perhaps the chosen stimuli were too easy to discriminate and therefore, not suitable to reveal group differences. In prior studies, the vowels were embedded within a syllable. Contrary to this, only the steady state part of the vowels was used in the current experiment. It could be that this contrast was difficult enough to circumvent any ceiling effects which might have obscured group differences.

Auditory deficits in dyslexia

The discrimination deficit was also found for both non-speech stimulus types in the dyslexic group (Hypotheses 5 and 6). There was only one person within the dyslexic group who showed a specific deficit in the processing of speech stimuli (see Figure 33). Conversely, eight persons had problems concerning the discrimination of the non-speech stimuli only. The other half was impaired for both stimulus classes. These findings favor the idea of a general auditory processing deficit in dyslexia. If the deficit would be speech specific (e.g., Liberman, 1989; Ramus, 2003; Vellutino, 1987), most persons of the dyslexic group should show reduced discrimination indexes for the vowel center stimuli only.

There is a point of contention, however, in which one could argue that auditory impairments might co-exist with speech perception deficits without being the source of the phonological problems (e.g., Breier et al., 2003; Mody et al., 1997; Schulte-Körne et al., 1998b) or they could possibly deteriorate the phonological deficit without being the core cause of dyslexia (Ramus, 2003).

The notion that phonological problems are caused in at least some cases of dyslexia by general auditory problems can be proven by using longitudinal study designs. This approach was recently chosen with a large group of children with genetic risk for dyslexia (Boets et al., 2008). Dynamic auditory processing was found to be associated with speech perception and phonological awareness, which were found to be predictive for future reading performance. However, the causal relation between auditory processing and speech perception and phonological awareness is not explained by this model.

The predictive nature of dynamic auditory processing was shown in a following longitudinal study (Corriveau et al., 2010). Additionally, impaired frequency discrimination for tones, as indexed by the MMN, was also reported for kindergarteners with genetic risk for dyslexia (Maurer, Bucher, Brem, & Brandeis, 2003). If these deficits are found prior to school entry, they can be regarded as predictors of the following reading behavior (Goswami, 2003). The causal link between auditory deficits and phonological discrimination skills in dyslexia is also provided by a training study (Schäffler, Sonntag, Hartnegg, & Fischer, 2004). After training for general auditory abilities (intensity and frequency discrimination, gap detection, time-order judgment, side-order judgment), the experimental group's performances in a phonological discrimination task and in a spelling test were significantly better compared to the waiting and placebo group. Furthermore, it has been shown in longitudinal designs that rapid auditory processing skills can be used as a predictor of later language abilities (Benasich & Tallal, 2002; Benasich, Thomas, Choudhury, & Leppänen, 2002; Choudhury, Leppänen, Leevers, & Benasich, 2007), reading (Hood & Conlon, 2004; Steinbrink, Zimmer, Lachmann, Dirichs, & Kammer, in press), and spelling outcomes (Steinbrink et al., in press).

In summary, these results support the idea that a general auditory impairment could be the cause of the phonological problems in at least some cases of dyslexia and this assumption can be explained by the fact that a precise representation of spectral and temporal features can facilitate the conversion of acoustical sounds into phonological representations (Ahissar et al., 2000).

Temporal and spectral auditory deficits in dyslexia

It was mentioned before that there is no consensus about the term "temporal": Studdert-Kennedy and Mody (1995) claimed that a short duration or a short inter-stimulus interval cannot be defined as temporal features, as such stimuli do not include any change in time. However, all stimuli in the current experiment did not include any change in time because they were based on the steady state part of German vowels. The spectrally rotated vowels and the bands of formants did not change over time either. Consequently, dyslexics show deficits in the discrimination of brief sounds and these deficits can be observed with steady state stimuli as well (Tallal, 1980). However, this finding does not rule out the possibility that dyslexics might show additional deficits for changing state stimuli (e.g., AM and FM stimuli, Hämäläinen et al., 2012; Talcott et al., 2000).

Comparable to prior research (Amitay, Ben-Yehudah, Banai, & Ahissar, 2002), the auditory deficits in this sample were not limited to the temporal domain, as they were also observable for the spectral and spectro-temporal contrasts. There was no participant with a specific deficit in the temporal domain at all. Temporal deficits always occurred together with spectro-temporal deficits (see Figure 34). Auditory deficits in the spectral domain have been reported frequently in dyslexic children and adults, but only for frequency changes that did not exceed 10% (see Hämäläinen et al., 2012 for a review). Of course, these results do not question that processing of other auditory features might also be affected in some dyslexic children and adults.

The role of attention

As dyslexia is often accompanied by attention deficits (Gilger, Pennington, & DeFries, 1992; Rüsseler, Kowalczuk, Johannes, Wieringa, & Münte, 2002; Willcutt & Pennington, 2000), the lower discrimination indexes in this experiment might be due to more clerical mistakes within the dyslexics group (Breier et al., 2003). In this case, the lower performance could not be a consequence of an auditory deficit, or it might be a consequence of both (Snowling, 2001). It should be noted that participants were preselected in order to exclude those who reported former attention problems and the overall performance in the dyslexic group did not drop in the course of the experiment. Furthermore, auditory deficits have also been reported frequently in studies in which the MMN was used as an index of auditory discrimination (Bishop, 2007). As the MMN is recorded without the participant's attention

(Näätänen et al., 2007) these findings prove that a lower performance in auditory tasks of dyslexic children and adults is not only due to attention problems.

Multicausal subgroups

The results presented in the Figures 30-32 might evoke the impression that all members of the dyslexic group performed worse compared to the control group. However, 19% of the dyslexic group did not show any auditory deficit at all (see Figures 33 and 34). The remaining participants showed a broad range of different patterns and only 43% were impaired in all three auditory contrasts (temporal, spectral and spectro-temporal) and only 24% underperformed the control group for all three types of stimuli (vowel center stimuli, spectrally rotated vowel center stimuli and bands of formants).

It is probable that developmental dyslexia cannot be explained by a single cause (Lachmann, 2002; Naidoo, 1972) and multicausal subgroups have been reported regularly in prior research (e.g. Bakker, 1992; Boder, 1973; Castles & Coltheart, 1993; Heim et al., 2008; Ingram, 1963; Johnson & Myklebust, 1967; see Watson & Willows, 1993 for an overview). It could also be that a child might have multiple deficits and not only one (Bishop, 2006; Snowling, 2008). This means that the results of each study are highly dependent on the composition of its respective sample, and so it is not surprising that the auditory deficit was not found for all participants in this experiment.

Choice of participants

It should be noted that there are some short comings regarding the chosen sample in this experiment. To begin with, none of our participants had an official diagnosis, although each member of the dyslexic group reported having reading problems since primary school. A second short coming concerns the matching of the control and dyslexic group. Although both groups were comparable with regard to age and sex, the IQ was found to be slightly higher in the control group. The members of the control group also had a higher level of school education. However, the significant main effect of group did not vanish ($F(1,39) = 4.98$, $p < .01$) when the IQ was added as covariate into the ANOVA. This finding proves that the auditory deficit of the dyslexic group is still observable after controlling for IQ differences.

Choice of stimuli

There were even three persons within the dyslexic group who showed lower discrimination indexes only for the bands of formants and not for the other two stimulus types with higher complexity. This means that the higher complexity of the speech signal compared to most non-speech stimuli should not be the reason for the absence of auditory deficits in prior research. There are also numerous studies which revealed significant group differences concerning frequency and duration discrimination using single sinusoidal tones (e.g., Banai & Ahissar, 2004; Heath, Bishop, Hogben, & Roach, 2006). The most important factor to reveal significant group differences, especially concerning spectral differences, seems to be the size of contrast, which should be at the most 10% (Hämäläinen et al., 2012). Another possibility would be to use threshold measures, as this procedure circumvents the problem of finding the optimal contrast between the experimental stimuli.

It has been proposed that phonological representations are used for the processing of spectrally rotated speech (Azadpour & Balaban, 2008). Therefore one might question whether spectrally rotated speech might be classified as a non-speech signal. The results of the stimulus ratings did not reveal any differences between the spectrally rotated vowel center stimuli and the bands of formants; both non-speech stimulus types were rated as less likely to be speech-like compared to the vowel center stimuli. These findings support the assumption that the spectrally rotated vowel center stimuli were actually perceived as non-speech. This finding does not rule out that spectrally rotated syllables might be classified as more speech like, as they contain spectrally rotated consonants, which are less affected by the inversion (Blesser, 1972, and see Chapter 2 for details).

Conclusion

Taken together, these findings show that the German vowel length discrimination paradigm is a suitable tool to establish proof of an auditory and phonological deficit in at least some German dyslexic adults and children. The auditory deficits were not limited to temporal features but became also obvious during the spectral processing. However, these results cannot be generalized to the processing of non-speech stimuli in dyslexia. This is the starting point of the final experiment of this thesis.

These results support the idea that the phonological deficits in at least some cases of dyslexia might be caused by a general auditory deficit in the temporal, spectral and spectro-

temporal dimension. The complexity of the auditory stimuli plays only a minor role, if even one at all. This finding is crucial for the comparability of prior studies as the heterogeneous data situation should be a result of other factors, like the kind of task (Banai & Ahissar, 2006), the size of contrasts (Hämäläinen et al., 2012), the composition of the sample and the criterion for being included into the study etc., rather than a result of the varying complexities of speech and non-speech stimuli.

Indeed, it will be a challenge to incorporate all of these factors in future research to explain the contradictory findings. Longitudinal designs should be used more often, as they extend correlational findings by providing the causal link between the factors. Causal relations can also be revealed by training studies which enable the transferring of theoretical knowledge about the etiology of developmental dyslexia to practical intervention methods.

The idea that speech sounds might be processed in a different way compared to non-speech is not restricted to the research field of developmental dyslexia. In fact, this idea forms the key assumption of the *domain specific models* of speech perception.

The aim of the following chapter (Chapter 4) is to test the *domain specific* and the *cue specific models* of speech perception by means of the extended German vowel length discrimination paradigm. An EEG component, called the mismatch negativity (MMN), will be used to investigate auditory discrimination of speech and non-speech stimuli at the pre-attentive level. On the one hand, if there are differences found in the processing of the speech and non-speech stimuli, this would support the *domain specific models* as the complexity of speech and non-speech stimuli is comparable and controlled for in this paradigm. On the other hand, if the size of the MMN is modulated by stimulus complexity only, this could be explained by the *cue specific models* of speech perception.

Chapter 4:

The role of stimulus complexity in the processing of speech and non-speech as revealed by the MMN

> "He who truly masters a language is also dominated by this language, as, when he speaks, he has to allow himself to be spoken by the language."
>
> Sigbert Latzel

The *domain specific models* assume that the differences in processing between speech and non-speech are observable even at early stages of auditory processing. One well-suited way to investigate the early processes of the auditory system is an EEG component, called the mismatch negativity (MMN; Näätänen et al., 1978). According to the *domain specific models*, the differences between speech and non-speech sounds can already be observed in this stage of processing. Contrary to this, no differences between the speech and non-speech sounds are expected following the *cue specific models*, if both stimulus types are identical concerning their physical properties. Additionally, *cue specific models* assume that non-speech sounds with lower complexity would not be processed in the same way as other stimulus types with higher complexity.

The German vowel length discrimination paradigm was extended successfully within a behavioral same-different task in Chapters 2 and 3 of this thesis. In this paradigm, vowel center stimuli can be compared to non-speech stimuli with either the same or lower stimulus complexity. For this reason, this stimulus set is appropriate for the testing of the *domain specific* and *cue specific models* by means of the MMN. To my knowledge, this is the first time that the MMN elicited by speech and non-speech sounds was compared while controlling for stimulus complexity.

One disadvantage of the spectral rotation is that the harmonic structure of the original stimulus is not preserved. With this in mind, the last goal of this chapter is to investigate the impact of harmony on the size of the MMN (see Experiment 4). Only if harmony has no

impact on the size of the MMN, a larger MMN of the vowel center stimuli compared to the spectrally rotated vowel center stimuli could be interpreted as an example of the specific processing of speech stimuli independent of their physical properties.

The mismatch negativity

The mismatch negativity (MMN) (Näätänen et al., 1978; Näätänen, 1979; Näätänen & Michie, 1979) is a change specific component of the event related potential that was originally found in the auditory domain. Using magnetoencephalography (MEG), its magnetic counterpart can be found (MMNm; Alho, 1995). The MMN can also be found in the visual (e.g., Alho, Woods, Algazi, & Näätänen, 1992; Heslenfeld, 2003; Pazo-Alvarez, Cadaveira, & Amenedo, 2003), olfactory (Akatsuka et al., 2005; Krauel, Schott, Sojka, Pause, & Ferstl, 1999) and somatosensory domain (Kekoni et al., 1997; Shinozaki, Yabe, Sutoh, Hiruma, & Kaneko, 1998).

For these studies, typically a so-called oddball paradigm is used. In this paradigm, one stimulus (or different stimuli which share a common dimension, e.g., intensity, pitch or duration) is repeated very often. This stimulus is called the standard. The frequent presentation of the standard results in a representation of the repetitive aspect (Horváth, Czigler, Sussman, & Winkler, 2001; Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996). During a small proportion of all trials the standard is replaced by the so called “deviant” which differs from the standard in at least one property. This difference is perceived as a mistake, as the standard stimulus is expected. The difference wave, which is calculated by the subtraction of the standard from the deviant, mostly shows a negative peak between 100 and 250 ms (Bishop, Hardiman, & Barry, 2011) at frontal and central electrodes (Lieder et al., 2013) and shows inverted polarity especially with nose reference (Deacon, Gomes, Nousak, Ritter, & Javitt, 2000). This negative component is called the MMN.

The MMN was found to changes of different kinds, e.g., to changes in pitch (Berti, Roeber, & Schröger, 2004; Hari et al., 1984; Jacobsen & Schröger, 2001; Näätänen et al., 1978; Schröger, 1996), intensity (Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1987; Schröger, 1996), timbre (Caclin et al., 2006; Goydke, Altenmüller, Möller, & Münte, 2004; Tervaniemi, Ilvonen, Karma, Alho, & Näätänen, 1997; Tervaniemi, Winkler, & Näätänen, 1997; Toiviainen et al., 1998), sound duration (Grimm, Widmann, & Schröger, 2004;

Joutsiniemi et al., 1998; Kaukoranta, Sams, Hari, Hämäläinen, & Näätänen, 1989; Roeber, Widmann, & Schröger, 2003), spatial location (Kaiser & Lutzenberger, 2001; Kujala, Alho, Paavilainen, Summala, & Näätänen, 1992; Paavilainen, Karlsson, Reinikainen, & Näätänen, 1989), rise time (Lyytinen, Blomberg, & Näätänen, 1992), inter-stimulus interval (Ford & Hillyard, 1981; Näätänen, Jiang, Lavikainen, Reinikainen, & Paavilainen, 1993; Nordby, Roth, & Pfefferbaum, 1988a), and stimulus order (Nordby, Roth, & Pfefferbaum, 1988b; Schröger, Tervaniemi, & Näätänen, 1995).

The amount of difference between the standard and the deviant stimulus influences the magnitude and the latency of the MMN: The magnitude increases with increasing difference between standard and deviant whereas the latency decreases (e.g., Berti et al., 2004; Sams, Paavilainen, Alho, & Näätänen, 1985, for frequency; Jaramillo, Paavilainen, & Näätänen, 2000; Näätänen, Syssoeva, & Takegata, 2004, for duration; Rinne, Särkkä, Degerman, Schröger, & Alho, 2006, for intensity).

The amplitude of the MMN is also modulated by the ratio between the probability of standard and deviant: The lower the probability of the deviant the higher the amplitude of the MMN (e.g., Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005; Ritter et al., 1992; Sabri & Campbell, 2001).

The MMN is considered to be an objective index of auditory discrimination (Näätänen, 2008). Significant correlations between the magnitude of the MMN and the performance in a behavioral discrimination task have been reported in some studies (e.g., Aaltonen, Eerola, Lang, Uusipaikka, & Tuomainen, 1994; Lang et al., 1990; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007). However, the linear relation between the magnitudes of the MMN and active discrimination tasks has not been observed in all studies (e.g., Alho & Sinervo, 1997; Allen, Kraus, & Bradlow, 2000; Dalebout & Stack, 1990; Paavilainen, Arajärvi, & Takegata, 2007). In these cases, the performance of the behavioral task was mostly at the ceiling (e.g., Colin et al., 2009) or at the bottom level (e.g. Allen et al., 2000).

The MMN is elicited independently from the participants' attention during the classical oddball paradigm (Sussman et al., 2003). Consequently, the MMN has been used in a variety of clinical fields, e.g. research of users of cochlear implants (e.g., Ponton & Don, 1995; Wable, van den Abbeele, Gallégo, & Frachet, 2000), schizophrenia (Michie et al., 2000; Umbricht & Krljes, 2005), dyslexia (e.g., Csépe, Gyurkocza, & Osman-Sagi, 1998; Kujala & Näätänen, 2001; Lachmann et al., 2005; see Bishop, 2007 for a review), specific language

impairment (e.g. Barry et al., 2008; Korpilahti, Krause, Holopainen, & Lang, 1998; Rinker et al., 2007; see Bishop, 2007 for a review) and coma (Naccache, Puybasset, Gaillard, Serve, & Willer, 2005; Wijnen, van Boxtel, Eilander, & Gelder, 2007). However, as the MMN can be replicated well only at the group level, predictions on the basis of individuals should not be made (e.g., Escera & Grau, 1996; Näätänen & Kreegipuu, 2012; Näätänen, Paavilainen, Rinne, & Alho, 2007).

The MMN is also found in children (Cheour, Leppänen, & Kraus, 2000), newborns (Alho, Sainio, Sajaniemi, Reinikainen, & Näätänen, 1990) and even in premature infants (Cheour-Luhtanen et al., 1996; Draganova et al., 2005; Huotilainen et al., 2005). Its latency decreases with age and its amplitude follows a u-shaped function with increasing age (Cheour et al., 2000).

The duration of the stimuli within the oddball paradigm should be at least 30ms, but longer stimulus durations do not influence the magnitude of the MMN (Paavilainen, Jiang, Lavikainen, & Näätänen, 1993; Tervaniemi, Schröger, Saher, & Näätänen, 2000b).

The MMNs elicited by different features are independent from each other (Deacon, Nousak, Pilotti, Ritter, & Yang, 1998) and can be added linearly (Schröger, 1995) when presented within one deviant. As deviants of different features do not influence each other, Näätänen and colleagues (2004) proposed the so-called multifeature paradigm in which five different deviant types are presented within one block having a probability of 10% for each deviant type. The probability of the standard is reduced to 50%, but the strength of the MMN is not decreased as four of the five deviants share the same feature with the standard. Therefore, each MMN shows the same strength as if the probability of the standard would have been 90%. This approach is very time-effective and has also been shown to be useful in the investigation of the processing of speech stimuli (Pakarinen et al., 2009; Partanen, Vainio, Kujala, & Huotilainen, 2011).

MMN of different stimulus types

In the first investigations dealing with the MMN, pure sinusoidal tones were used (e.g., Näätänen et al., 1978; Näätänen, 1979, Näätänen & Michie, 1979). The MMN can, however, also be elicited with non-speech stimuli of higher complexity (e.g., noise bursts, harmonic tones, chords) and even speech stimuli (e.g., vowels, syllables, words). There is much

ongoing research investigating the issue of how the properties of the different stimulus types might influence the magnitude, latency, and distribution on the scalp of the MMN.

Speech versus non-speech

There were some attempts to compare the MMN of speech and non-speech stimuli. Tervaniemi and colleagues (1999) for instance, compared the MMNm evoked by the vowels /e/ and /o/ to the MMNm of two chords: A major and A minor. The chords evoked a larger MMNm compared to the vowels. Additionally, the source of the MMNm to phonemes was found to be more superior compared to the chords. The authors concluded that speech and music are processed differently within the brain, as the MMNm was found in spatially distinctive areas.

A larger MMN to non-speech compared to speech stimuli was also reported by Wunderlich and colleagues (2001). They compared words (/bæd/ and /dæd/), syllables (/bæ/ and /dæ/) and single sinusoidal tones with a 10% pitch change (400 and 440Hz, 1500 and 1650Hz, 3000 and 3300Hz). The sinusoidal tones evoked a larger MMN compared to the words and syllables. No difference was reported for the words and syllables. However, it seems doubtful that the speech and non-speech stimuli in this experiment were matched with respect to the difficulty of contrast: Although discrimination performance was 100% for each stimulus type, /bæd/ and /dæd/ can only be distinguished on the basis of the place of articulation. In contrast to this minute spectral difference, there was an increase of pitch by 10% in the sinusoidal tones. The larger MMN could be a consequence of a larger contrast. This idea is supported by an additional experiment of the authors, presented in the same study. They produced non-speech stimuli with a higher complexity: They were composed of three sinusoidal tones: (1) 400Hz + 3000Hz + 1500Hz, (2) 400Hz + 3000Hz + 1650Hz. After 80 milliseconds 1500Hz was changed to 1650Hz and vice versa. The resulting stimuli differed with respect to the spectral drift after 80ms. The MMN to these more complex stimuli was comparable to the size of the speech stimuli MMN.

Nikjeh and colleagues (2009) compared the MMN of pure tones (1,5% and 6% pitch change), harmonic tones with three overtones (1,5% and 6% pitch change), and speech syllables (/ba/ and /da/), but found no differences.

There are also studies in which a larger MMN for speech stimuli compared to non-speech stimuli is reported. Jaramillo and colleagues (2001) compared vowels to harmonic tones. The vowel /e/ was changed with respect to identity (/o/), pitch (increment from 105 to 117Hz), or duration (decrement from 400 to 200ms). The tones were composed of a fundamental frequency of 105Hz and ten overtones. Comparable to the speech stimuli, the pitch of the fundamental frequency was increased to 117Hz in the spectral condition and duration was decreased from 400 to 200ms in the temporal condition. The authors reported a larger MMN for vowels for the durational contrast. Contrary to this, no differences between the speech and tone stimuli were reported for the spectral condition.

Čeponiene and colleagues (2002) did, however, find a higher MMN in children for the vowel /œ/ compared to a four partial sinusoidal tone and a single sinusoidal tone (458Hz) for both durational (decrement from 260 to 160ms) and spectral contrasts (increment of pitch of 10%). The four partial sinusoidal tones were composed of the first four formants of the vowel (458, 1370, 2054, and 3537Hz).

Sorokin and colleagues (2010) compared complex disharmonic non-speech stimuli to CV syllables. Five different deviants were presented within a multifeature paradigm (Näätänen et al., 2004) for both the speech and the non-speech condition: change of vowel (/i:/ vs. /e:/), consonant (/p/ vs. /k/), vowel duration (decrement of 50ms), frequency (+/-8%), and intensity (+/-6dB). Significant differences of the MMN amplitudes between the speech and non-speech stimuli were observed for the change of vowel and frequency, but not for the durational deviants.

Jaramillo and colleagues (1999) proposed that there might be an interaction between the direction of duration change (increment versus decrement) and type of stimulus (speech versus non-speech). They compared the vowel /a/, a low pass-filtered version of this vowel (cutoff of all frequencies beyond F2), noise, and a single sinusoidal tone (540Hz). For the durational decrement, the MMNs of the speech stimuli were larger compared to the MMNs of non-speech stimuli. Conversely, the non-speech stimuli evoked a larger MMN compared to the speech stimuli, when duration was increased. However, this observation failed to be replicated in some studies dealing with the durational MMN; Amenedo and Escera (2000) and Jaramillo and colleagues (2000) investigated the role of direction in the durational MMN for non-speech stimuli (sinusoidal tones and white noise). Neither of them reported any effects of direction. The first study which investigated the MMN for duration increments and

decrements (Näätänen, Paavilainen, & Reinikainen, 1989) did not report any effects of direction either. There is even one study in which a 50% decrease in tone duration evoked a greater MMN compared to an increment (Colin et al., 2009). How can these inconsistent results of the temporal MMN be explained?

The answer is provided by two studies in which the same stimulus calculation method was used (Peter, McArthur, & Thompson, 2010; Takegata, Alku, Ylinen, & Näätänen, 2008). Their approach controls for biases induced by the properties of the stimuli, as each stimulus is used as standard and deviant in a separate block. Takegata and colleagues (2008) compared the vowel /e/, the chord A major, white noise and a single sinusoidal tone (450Hz). The interaction between direction of duration change and stimulus type as reported by Jaramillo and colleagues (1999) was only found for the noise stimulus. The MMN evoked by the vowel was higher than the MMN of the sinusoidal tone for both duration increment and decrement. The MMN of the chord was comparable in size to the MMN of the vowel.

Peter and colleagues (2010) compared the traditional calculation method to the same stimulus calculation method used by Takegata and colleagues (2008). They compared the size of the MMN to duration increments (200 versus 300ms) and decrements (300 versus 200ms) in a single sinusoidal tone of 1000Hz. The size of the MMN was only increased for the duration increment in the traditional method. This means that the interaction reported by Jaramillo and colleagues (1999) might not have been found if they had used the same stimulus calculation method.

A summary of the MMN and MMNm studies which compared the amplitude of the MMN of speech to non-speech stimuli with lower complexity is given in Table 17. All results with higher MMN amplitudes for the speech stimuli compared to the non-speech ones are highlighted in bold.

Table 17: Summary of MMN/MMNm studies, which compared the amplitude of the MMN/MMNm of speech to non-speech stimuli with lower complexity. Results with higher MMN amplitudes for the speech stimuli compared to the non-speech ones are highlighted in bold.

	Difference of the amplitude of the MMN/MMNm between speech and non-speech stimuli			
Study	Vowel change (timbre)	Change of pitch	Change of duration	Change of consonant
Tervaniemi et al. (1999)	vowel (/e/ vs. /o/) < chord (A major vs. A minor)			
Wunderlich et al. (2001).				/b/ vs. /d/ < 10% pitch change in tones /b/ vs. /d/ = spectral drift after 80ms for 1 out of 3 sine waves
Nikjeh et al. (2009)				/b/ vs. /d/ = 1,5 and 6% pitch change in sine waves and harmonic tones
Jaramillo et al. (2001)	**vowel (/e/ vs. /o/) > harmonic tone (+8,5% pitch)**	vowel (+8,5%) = harmonic tone (+8,5%)	**vowel (-200ms) > harmonic tone (-200ms)**	
Čeponiene et al. (2002)		**vowel (+10%) > harmonic tone (+10%)** **vowel (+10%) > sine wave (+10%)**	**vowel (-100ms) > harmonic tone (-100ms)** **vowel (-100ms) > sine wave (-100ms)**	
Sorokin et al. (2010)	**vowel (/i:/ vs. /e:/) > complex non-speech analogue**	**Vowel (+/-8%) > complex non-speech analogue**	vowel (-50ms) = complex non-speech analogue (-50ms)	/p/ vs. /k/ = complex non-speech analogue
Jaramillo et al. (1999)			**Vowel (-80ms) > noise (-80ms)** **Vowel (-80ms) > sine wave (-80ms)** Vowel (+80ms) < noise (+80ms) Vowel (+80ms) < sine wave (+80ms)	
Takegata et al. (2008)			vowel = chord (+/- 80/160ms) **vowel > sine wave (+/- 80/160ms)** **vowel > noise (- 80/160ms)** vowel = noise (+80/160ms)	

The role of complexity in the MMN

Some studies dealt with the question of whether the complexity of a stimulus, defined as the number of different frequencies within the signal, might influence the magnitude of the MMN.

Tervaniemi and colleagues (2000a) were able to show that the size of the frequency MMN of two single sinusoidal tones is smaller compared to the frequency MMN of the same tones which were enriched with two overtones, but the size of the MMN does not grow with increase in number of overtones from two to four (Tervaniemi et al., 2000b).

Takegata and colleagues (2008) reported a higher MMN to duration changes in chords than in sinusoidal tones.

Zion-Golumbic and colleagues (2007) also investigated whether harmonically rich stimuli (two overtones) might evoke a bigger MMN compared to a single sinusoidal tone. In one block, the standard calculation method for the MMN was used. In another block, they controlled for differences of the N100. For the classical method, the harmonic stimuli evoked a higher MMN for both the pitch and duration MMN. The spectral MMN was comparable in size for both types of stimuli, when the N100 was controlled for. However, there was a difference in the temporal MMN, even when the N100 was controlled for.

Moreover, Alho and colleagues (1996) were able to show that the MMNm to pitch changes in single tones, chords, and patterns of tones do not share the same source.

These studies support the idea that the number of different frequencies within a stimulus could modify the properties of the MMN.

Spectrally rotated speech and the mismatch negativity

As previously mentioned, the confounding factor stimulus complexity can be controlled for by using spectrally rotated speech. To the best of my knowledge, there is only one MMN study in which spectrally rotated speech was used as non-speech analogue. Davids and colleagues (2011) compared children with and without specific language impairment. They used two words, /pan/ as standard and /kan/ as deviant, and their spectrally rotated counterparts in the non-speech condition. Both stimulus types evoked a significant MMN and their sizes of the MMN were not reported to be different in the control group. These results coincided in line with their pilot study which included 16 healthy adults. The findings

suggest that speech is not processed in a special way compared to non-speech. However, as already mentioned in the introduction of Experiment 1, plosives and nasals are still perceived as consonants after the spectral rotation (Blesser, 1972). So it could be that the spectrally rotated stimuli might not be perceived as completely non-speech like in this study.

The role of the native language in the MMN

The first evidence of a special role of the mother tongue in investigations dealing with the MMN was provided by Näätänen and colleagues (1997). One group of their participants was Estonian, the other Finnish. The vowel /e/ was used as standard stimulus and the vowels /œ/, /o/ and /õ/ served as deviants. /õ/ is an Estonian phoneme, which is not part of the Finnish phoneme inventory. /œ/ and /o/ are found in both languages. No group differences were found for /œ/ and /o/, but the MMN elicited by /õ/ was smaller in the Finnish group compared to the Estonian group. The authors concluded that the MMN is higher for speech stimuli which form part of the phoneme repertoire of the native language.

This phenomenon was also shown for other languages. Nenonen and colleagues (2003) compared Russian adults who spoke Finnish fluently as their second language (Nenonen, Shestakova, Huotilainen, & Näätänen, 2003) to adults with Finnish as their mother tongue. In Finnish, quantity is phonetically relevant. This means that the duration of a sound (vowel or consonant) can influence the meaning of the word. The temporal MMN (200 versus 150ms) of a syllable (/ka/) and of a harmonic rich tone (500 + 1000 + 1500Hz) was calculated in both groups. The MMN amplitude was lower for the second language speakers of Finnish compared to the native speakers for the syllables, whereas no difference was found between the groups for the harmonic tones. The phoneme representations appeared to be acquired during early childhood.

Kirmse and colleagues (2007) compared German and Finnish adults. Vowel quantity is only important for some tense-lax pairs in the German language, especially for the vowel pair /a/ - /a:/. Syllables (/sasa/) and tones were used. Within the syllable, only the duration of the vowels was changed whereas the consonants remained stable. The latency of the temporal MMN was shorter for the Finnish participants compared to the German ones for both the speech and non-speech stimuli. Contrary to this, there was no difference between the groups for the spectral condition of the tones. This pattern of results for non-speech stimuli

was also reported by Tervaniemi and colleagues (2006). Taken together, the participants' mother tongue should always be considered while comparing different MMN studies.

The role of harmony in the MMN

The MMN can also be evoked by musical stimuli, like tones (Meyer et al., 2011; Nikjeh et al., 2009) and chords (Bergelson & Idsardi, 2009; Tervaniemi et al., 1999). Tervaniemi and colleagues (1999) compared the MMNm of vowels and chords. The chords evoked a higher MMNm compared to the vowels. Furthermore, people who are highly familiar to musical stimuli show a larger amplitude and shorter latency of the MMN for musical stimuli (Nikjeh et al., 2009) and even speech stimuli (Kühnis, Elmer, Meyer, & Jäncke, 2013; Nikjeh et al., 2009).

Takegata and colleagues (2008) compared the temporal MMN for the vowel /e/, the chord A major, band-pass filtered white noise, and a single sinusoidal tone. The MMN of the chord was comparable in size to the MMN of the vowel for the duration increment and decrement. The noise evoked a significantly smaller MMN than the vowel and the chord for the duration decrement. For the duration increment however, the noise evoked a larger MMN compared to the two harmonic stimuli.

To the best of my knowledge, there is no study dealing with the question of whether there is a difference in the pre-attentive processing of harmonic and disharmonic stimuli, or not. If harmonic stimuli evoke a larger MMN compared to disharmonic stimuli, the larger MMN of the vowels compared to the disharmonic non-speech stimuli could be a consequence of the harmonic structure of the vowel and not due to differences in the processing of speech and non-speech.

Experiment 3

The aim of this experiment was to compare the magnitude of the MMN of speech and non-speech stimuli which are matched with respect to complexity and controlled for considering the difficulty of each contrast.

Participants

30 adults took part in this experiment. There was one group for each vowel type (/a/ – /a:/ vs. /ɪ/ - /i:/). The ratio of male and female participants was equal in both groups (10 females, 5 males). The mean age in the /a/ - /a:/ group was 23.87 years, with a standard deviation of 2.90 years. The range was 19 to 30 years. The mean age of the /ɪ/ - /i:/ group was 22.47 years, with a standard deviation of 2.92 years. The range was 18 to 30 years. A t-test of independent samples and the Levene-test did not reveal any differences of age between both groups in relation to the mean ($t(28) = 1.32$, $p = .20$) or the standard deviation ($F(1,28) = 0.03$, $p = .86$). All of participants were students of the University of Mainz, except two persons. No one reported impaired hearing. All were native speakers of German. Both groups were matched with respect to their former musical education.

Material

This experiment included a subset of the stimuli of Experiment 2. Three stimulus types (vowel center stimuli with full spectrum, spectrally rotated vowel center stimuli with complete spectrum and the bands of formants based on the vowel center stimuli) and both vowel types (/a/ - /a:/ and /ɪ/ - /i:/) were included. The shortened version of the originally long vowel was not included (vam75, vim51, ram75, rim51, bam75, bim51) (see Table 18 and 19 for an overview).

Task

All participants started with the EEG session, which was composed of three blocks. Afterwards, the same stimuli were presented within an active same-different task. A passive oddball task was used. Participants watched a silent movie and were asked to ignore all auditory stimuli, which were presented via headphones.

Comparable to the Experiments 1 and 2, all stimuli were presented within a same-different task during the behavioral task. The stimulus onset asynchrony (SOA) was the same as in the oddball task (500ms). To rule out any effects of handedness on reaction time, key

assignments were counterbalanced. There was a short practice block comprised of 8 trials to familiarize participants with the task. During these trials there was an acoustic feedback following incorrect button presses. No feedback was given during the experimental block. There was no time limit for the participants' response. The inter-trial interval (ITI) was 2000ms.

Apparatus

All stimuli were presented with an external soundcard (UGM96, ESI Audiotechnik GmbH, Leonberg, Germany) binaurally via closed headphones (Beyerdynamic DT 770) with an intensity of 66 dB (SPL) or 60dB(A), respectively. The intensity was measured with an artificial head (HSM III.0, HEAD acoustics, Aachen, Germany). The operating system on the laptop was Windows XP. Presentation (version 14.5, Neurobehavioral Systems, Albany, California) was used to control the experimental protocol. All sessions took place in an acoustically attenuated and electrically shielded chamber.

The electroencephalogram (EEG) was recorded continuously with a SynAmps amplifier (Neuroscan, Sterling, VA). The electrode impedance was kept under 5kOhm. Seven Ag/AgCl electrodes were attached according to the 10-20-system at the following positions: F3, Fz, F4, Cz, Pz and additionally upon the left and right mastoid (LM and RM) (see Figure 38). The reference electrode was placed on the tip of the nose. The vertical and horizontal electrooculogram (EOG) was recorded additionally to control for eye movements. The sampling rate was 500Hz and an online notch filter (50Hz) was applied.

Design

The design of the experiment was similar in the oddball and same-different task: There were *two Vowel types*: /a/ - /a:/ and /ɪ/ - /i:/. In this experiment, the *Vowel type* was a between subject factor. Taken together, there were *three Stimulus types*: vowel center stimuli with full spectrum, spectrally rotated vowel center stimuli with full spectrum and bands of formants based on the vowel center stimuli with the full spectrum. There was one block for each stimulus type and the order of the blocks was counterbalanced between participants in the oddball task. During the same-different task, all stimulus types were presented within

one block. In both the oddball and the same-different task, there were *two types of Auditory contrast*: temporal and spectral.

Within each oddball block, both types of auditory difference were presented (see Table 18). There was one block for each type of stimulus (vowel center stimuli, spectrally rotated vowel center stimuli, bands of formants). The sequence of the blocks was counterbalanced. Each block was comprised of 2000 stimuli: 1600 standard stimuli (p = 0.8) and 400 deviant stimuli, 200 for each type ($p_{temporal}$ = 0.1, $p_{spectral}$ = 0.1). The stimulus onset asynchrony was kept constant at 500ms during the experiment. At the beginning of each block, 14 standards were presented. During the entirety of the experiment there were at least 3 standard stimuli before each deviant.

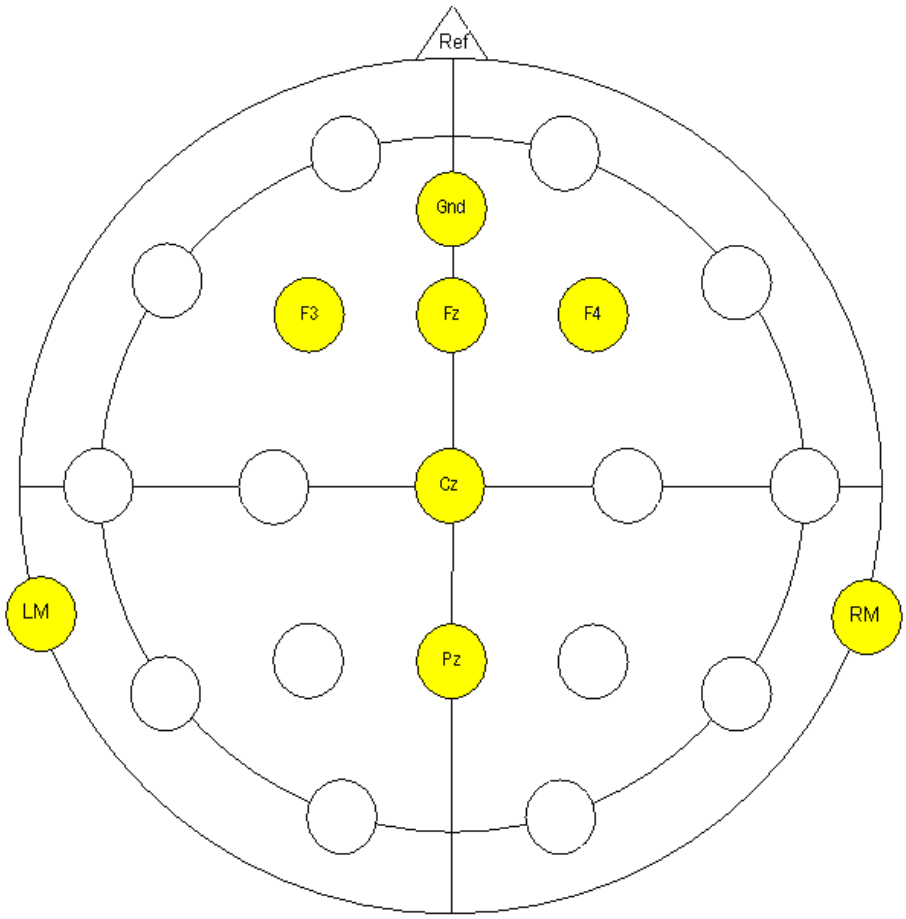

Figure 38: Positions of electrodes used in Experiment 3 and 4 according to the 10-20-system.

Table 18: Experimental design of all trials with stimuli based on /a/ - /a:/ and /ɪ/ - /i:/ in the oddball paradigm.

	Deviant		**Standard**
	Temporal N = 200 p = 0.1	Spectral N = 200 p = 0.1	N = 1600 p = 0.8
Vowel center stimuli (VC)	Vao75/Vio51	Vao145/Vio93	Vam145/Vim93
Spectrally rotated vowel center stimuli (RVC)	Rao75/Rio51	Rao145/Rio93	Ram145/Rim93
Bands of formants (BFVC)	Bao75/Bio51	Bao145/Bio93	Bam145/Bim93

The active discrimination task comprised 216 trials. The complete design is illustrated in Table 19. The order of the trials in each block was pseudo randomized in accordance with the following rules: There were no more than three trials in sequence which required the same response. In addition, the type of stimulus was changed at least after every third trial.

Table 19: Experimental design of all trials with stimuli based on /a/ - /a:/ and /ɪ/ - /i:/ in the same-different task.

	Different condition		**Same condition**
	Temporal	Spectral	
VC	Vao75 vs. Vam145/ Vio51 vs. Vim93 (18x)	Vao145 vs. Vam145/ Vio93 vs. Vim93 (18x)	Vao75 vs. Vao75 Vao145 vs. Vao145 Vam145 vs. Vam145/ Vio51 vs. Vio51 Vio93 vs. Vio93 Vim93 vs. Vim93 (12x)
RVC	Rao75 vs. Ram145/ Rio51 vs. Rim93 (18x)	Rao145 vs. Ram145/ Rio93 vs. Rim93 (18x)	Rao75 vs. Rao75 Rao145 vs. Rao145 Ram145 vs. Ram145/ Rio51 vs. Rio51 Rio93 vs. Rio93 Rim93 vs. Rim93 (12x)
BFVC	Bao75 vs. Bam145/ Bio51 vs. Bim93 (18x)	Bao145 vs. Bam145/ Bio93 vs. Bim93 (18x)	Bao75 vs. Bao75 Bao145 vs. Bao145 Bam145 vs. Bam145/ Bio51 vs. Bio51 Bio93 vs. Bio93 Bim93 vs. Bim93 (x12)

Dependent variables

All EEG analyses were performed with the Matlab (version R2011A; Mathworks) toolbox ERPLAB (Luck & Lopez-Calderon, 2013), which is integrated in the EEGLAB toolbox (Delorme & Makeig, 2004). First, an offline band-pass filter ranging from 1 to 30Hz was used. The event related potentials (ERPs) were computed separately for the three types of stimuli, standards and deviants. The time window ranged from 200ms before to 500ms after stimulus onset. The first 200ms served as the baseline for the averaged signal. The first 10 standards of each block and all epochs containing eye movements greater than 75µV were excluded. The dependent value was the area under the difference curve within a time window of 50ms around the peak latency (Beauchemin & Beaumont, 2005). First, the difference curve for each type of stimulus was formed by subtracting the ERP of the standard from the ERP of the deviant: $ERP_{deviant}$ - $ERP_{standard}$ (see Figure 39).

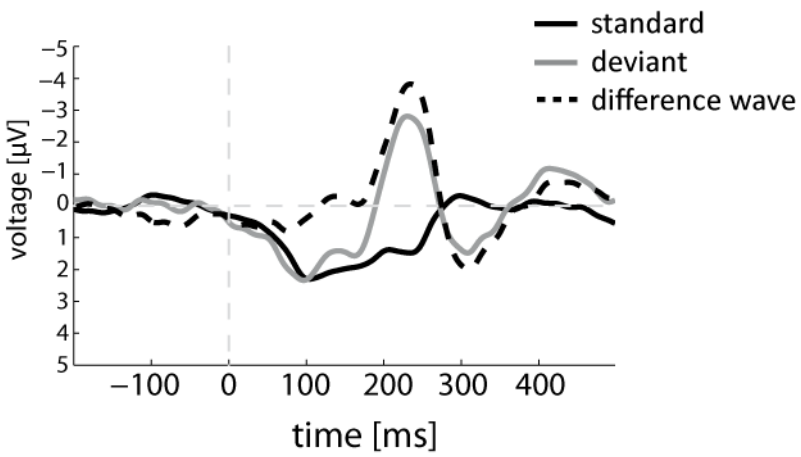

Figure 39: ERP curve evoked by the standard (black) and deviant (grey) stimulus. The difference curve (black dashed) is calculated by the following formula: $ERP_{deviant}$ - $ERP_{standard}$.

The peak latency of each difference curve was established within a time window between 100 and 300ms. The size of the MMN was estimated as the area under the difference curve ranging from 25ms before to 25ms after the peak latency (see Figure 40).

For the behavioral data the discrimination index d' (see Experiment 1) and reaction times of correct responses (see Experiment 1) were calculated.

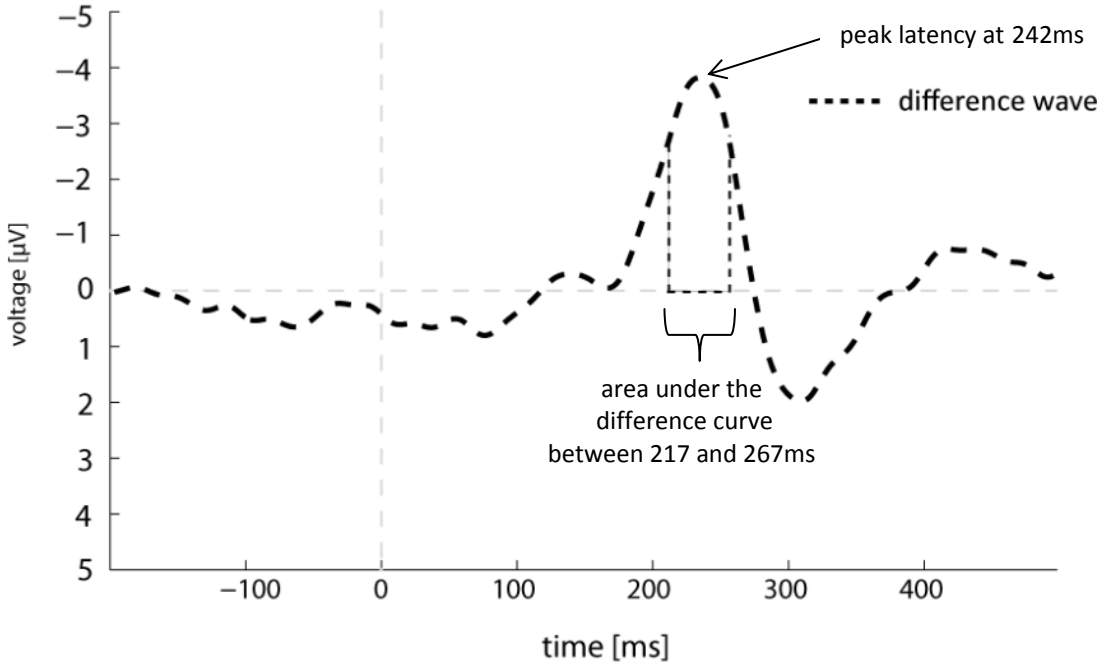

Figure 40: Example of a difference curve. The dependent value is the area under this curve from 25ms before to 25ms after the peak latency.

Hypotheses

(1) If speech is processed differently and more efficiently compared to non-speech as revealed by the MMN independently of complexity of the stimuli (see the *domain specific models*), the following pattern of results should occur:

   b) The magnitude of the MMN should be larger for the vowel center stimuli compared to the MMN of the spectrally rotated vowel center stimuli and the bands of formants

   c) There should be no difference between the spectrally rotated vowels and the bands of formants concerning the magnitude of the MMN

(2) If the size of the MMN is dependent on the complexity (see the *cue specific models*) and the "speechness" of the stimuli (see the *domain specific models*), the following pattern of results should occur:

   a) The magnitude of the MMN should be larger for the vowel center stimuli compared to the MMN of the spectrally rotated vowel center stimuli and the bands of formants

   b) There should be a difference between the spectrally rotated vowels and the bands of formants concerning the magnitude of the MMN

(3) If the size of the MMN is only dependent on the complexity of each stimulus (see the *cue specific models*), the following pattern of results should occur:

   a) The magnitude of the MMN should be equal for the vowel center stimuli and the spectrally rotated vowel center stimuli

   b) There should be a difference between the spectrally rotated vowels and the bands of formants concerning the magnitude of the MMN

(4) As the speech stimuli are based on the German vowel system (see Chapter 1), a different pattern of results is expected for the vowel pairs /a/ - /a:/ and /ɪ/ - /i:/:

   a) For the vowel pair /a/ - /a:/, the temporal MMN should be larger compared to the spectral MMN

   b) For the vowel pair /ɪ/ - /i:/, the spectral MMN should be larger compared to the temporal MMN

(5) Concerning the behavioral data, the same pattern of results as in Experiment 1 is expected (see Chapter 2)

Results

The average of rejected trials is depicted in Table 20. All ERPs and difference waves are depicted in Figures 41 and 42. T-tests for one sample based on the area of MMN revealed that the MMN was observed for every condition (see Table 21).

Table 20: Mean, standard error and maximum of rejected trials for the vowel pair /a/ - /a:/ and /ɪ/ - /i:/ in Experiment 3.

Vowel type		/a/ - /a:/			/ɪ/ - /i:/		
		Mean	Standard error	Maximum	Mean	Standard error	Maximum
VC	Temporal deviant	50.61	7.16	120	34.80	5.70	77
	Spectral deviant	47.60	6.70	115	39.00	5.07	76
	Standard	390.83	54.81	917	308.16	44.29	582
RVC	Temporal deviant	46.53	7.09	112	39.93	6.16	85
	Spectral deviant	47.87	6.77	111	43.80	6.80	98
	Standard	387.63	52.51	885	310.27	50.92	704
BFVC	Temporal deviant	43.4	6.28	98	41.00	5.43	81
	Spectral deviant	38.4	5.59	92	39.60	5.88	79
	Standard	343.57	47.54	818	318.61	46.89	651

Table 21: T-tests for one sample based on the area of the MMN for each stimulus type, vowel type and auditory contrast.

		/a/ - /a:/		/ɪ/ - /i:/	
Stimulus type	Condition	*t*(14)	*p*	*t*(14)	*p*
Vowel center	Temporal	11.45	< .01	8.47	< .01
Vowel center	Spectral	8.84	< .01	7.93	< .01
Spectrally rotated vowel	Temporal	5.97	< .01	5.07	< .01
Spectrally rotated vowel	Spectral	8.77	< .01	7.55	< .01
Bands of formants	Temporal	6.89	< .01	6.41	< .01
Bands of formants	Spectral	7.08	< .01	5.76	< .01

Three mixed model ANOVAs were conducted for the within factors, *Stimulus type* (vowel center stimuli, spectrally rotated vowel center stimuli with full spectrum, bands of formants) and type of *Auditory contrast* (temporal, spectral) and the between factor *Vowel type* (/a/ -

/a:/, /ɪ/ - /i:/). There was one ANOVA for each dependent variable (area of MMN, d', reaction time). Every time the assumption of sphericity was rejected, as revealed by the Mauchly's test, degrees of freedom were corrected according to Greenhouse Geisser. The alpha value of the post hoc t-tests was always adjusted following the Bonferroni correction.

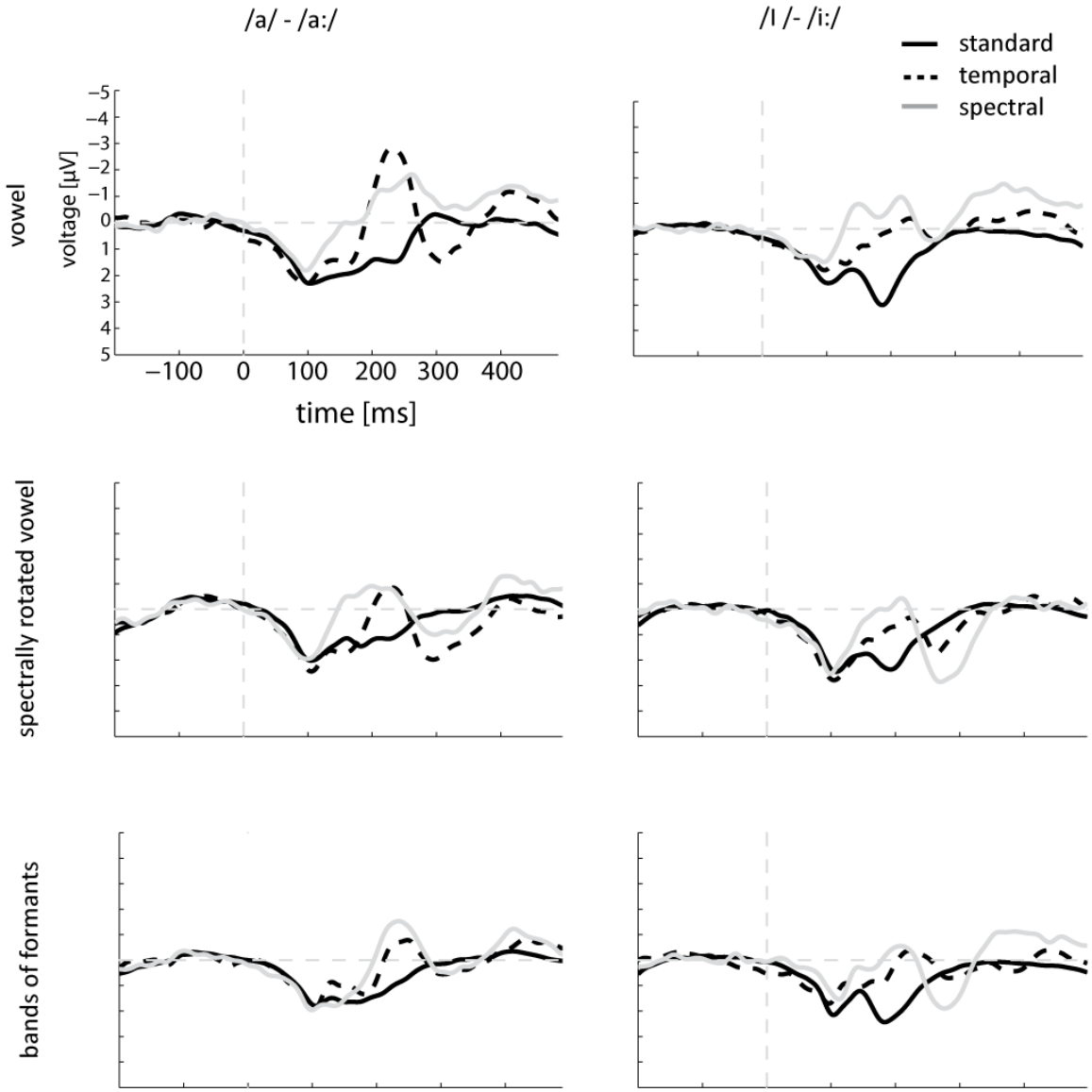

Figure 41: ERPs at Fz for each stimulus type (vowel center stimuli at the top, spectrally rotated vowel center stimuli in the middle, bands of formants at the bottom), vowel type (/a/ - /a:/ on the left side, /ɪ/ - /i:/ on the right side). The ERPs of the standards are represented by the black solid line. The ERPs of the temporal deviants are represented by the black dashed line and spectral deviants by the grey solid line.

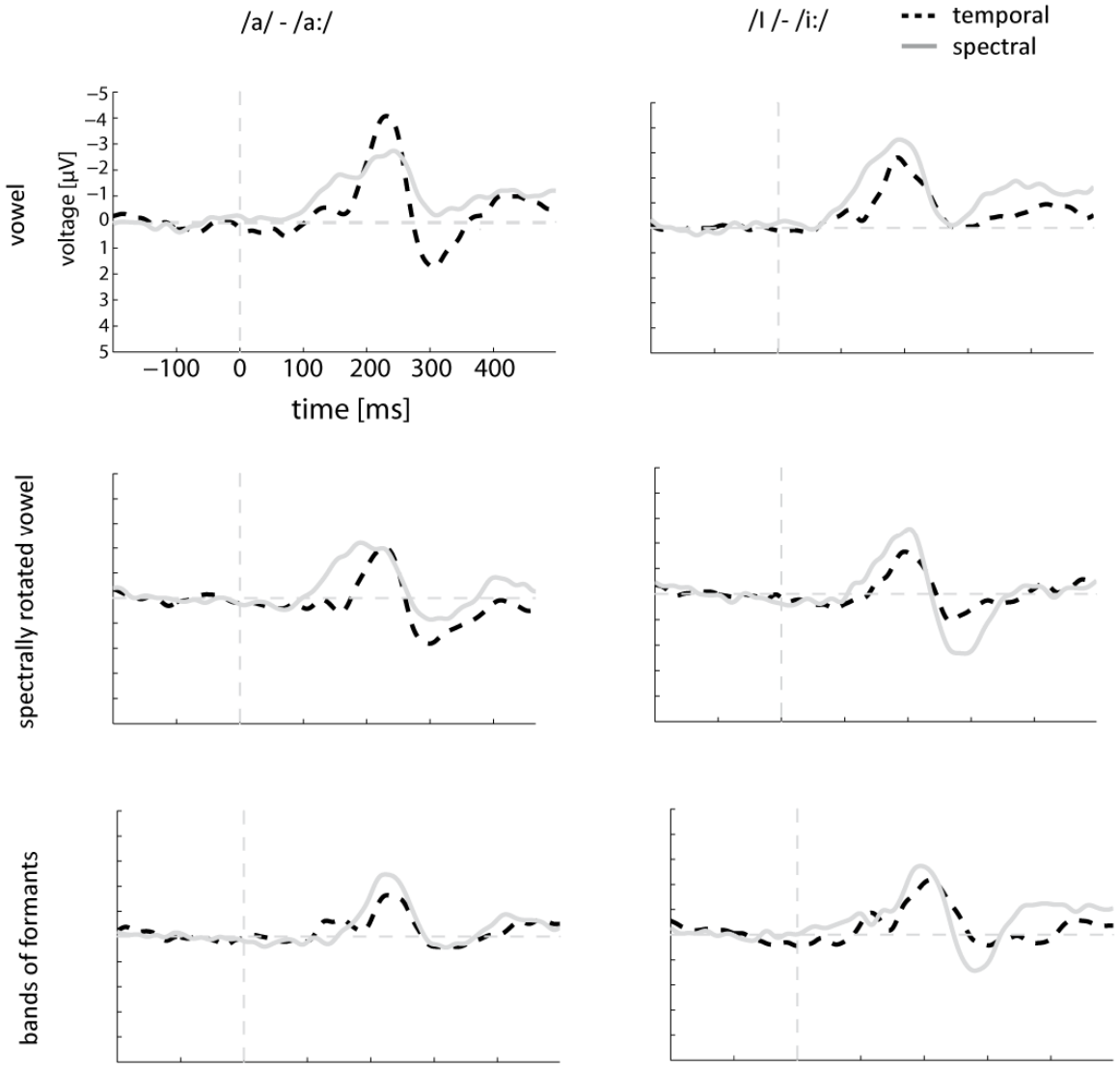

Figure 42: Difference waves at Fz for each stimulus type (vowel center stimuli at the top, spectrally rotated vowel center stimuli in the middle, bands of formants at the bottom), vowel type (/a/ - /a:/ on the left side, /ɪ/ - /i:/ on the right side). The difference waves of the temporal deviants are represented by the dashed black line and spectral ones by the solid grey line.

The results of the ANOVA for the area of the MMN at Fz are reported in Table 22. The means and standard errors of each condition are shown in Figure 43. There was a significant main effect of *Stimulus type* ($F(2,56) = 14.26$, $p < .01$). Vowel center stimuli evoked a significantly larger MMN area compared to spectrally rotated vowel center stimuli ($t(29) = 5.35$, $p < .01$, $d = 1.00$) and compared to the bands of formants ($t(29) = 3.67$, $p < .01$, $d = 0.87$). There was no difference between the area of the MMN of the spectrally rotated vowels and the bands of formants ($t(29) = -0.48$, $p = .63$). There was no significant main effect of *Vowel type* ($F(1,28) =$

0.04, p = .85). A significant difference between the spectral and temporal deviant (F(1,28) = 6.06, p = .02) was found. The spectral deviant evoked a larger area of MMN compared to the temporal deviant (t(29) = 2.35, p = .03, d = 0.44).

Table 22: Results of the analysis of variances based on the area of the MMN in Experiment 3.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Type of stimulus	14.26	2	56	< .01	.34
Type of vowel	0.04	1	28	.85	< .01
Auditory contrast	6.06	1	28	.02	.18
Type of stimulus * vowel	1.50	2	56	.23	.05
Type of stimulus * auditory contrast	2.37	2	56	.10	.08
Vowel * auditory contrast	3.93	1	28	.06	.12
Type of stimulus * vowel * auditory contrast	4.32	2	56	.02	.13

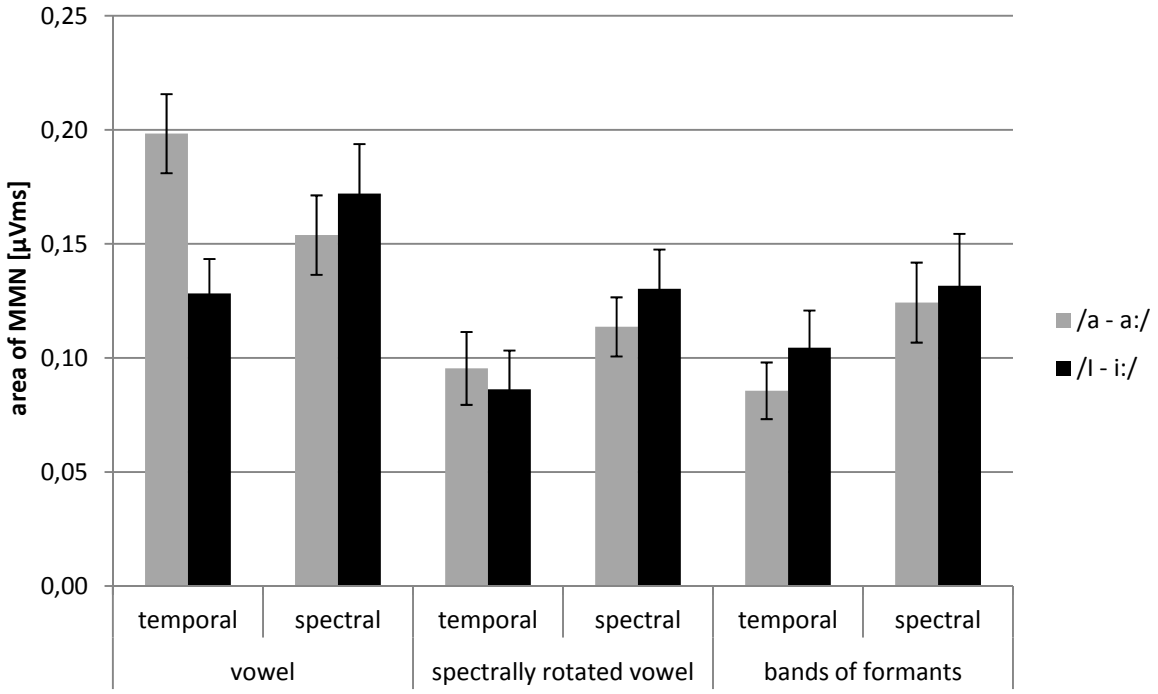

Figure 43: Means and standard errors of the area of MMN for each experimental condition of Experiment 3.

The interactions between *Stimulus* and *Vowel type* (F(2,56) = 1.50, p = .23), *Stimulus type* and *Auditory contrast* (F(2,56) = 2.37, p = .10) and *Auditory contrast* and *Vowel type* (F(1,28)

= 3.93, p = .06) did not reach significance. There was a significant interaction between *Stimulus type*, *Vowel type* and *Auditory contrast* ($F(2,56) = 4.32$, $p = .02$). This triplet interaction seems to be based on an interaction between the *Vowel type* and the *Auditory contrast* in the vowel center stimuli. To test this idea, two additional analyses of variance were conducted: One for the vowel center stimuli and one for the non-speech stimuli. For the vowel center stimuli, there was a significant interaction between *Stimulus type* and the *Auditory contrast* ($F(1,28) = 16.89$, $p < .01$). For the vowel pair /a/ - /a:/, the temporal deviant evoked a larger area of MMN compared to the spectral one ($t(14) = 3.03$, $p = .02$, $d = 0.66$). For the vowel pair /ɪ/ - /i:/, the spectral deviant evoked a larger MMN compared to the temporal one ($t(14) = -2.79$, $p = .03$, $d = 0.60$). This pattern of results is illustrated in the upper part of Figure 42. Contrary to this, the interaction between type of vowel and auditory contrast did not reach significance for the two non-speech stimulus types ($F(1,28) = 0.90$, $p = .35$ for the spectrally rotated vowels and $F(1,28) = 0.15$, $p = .70$ for the bands of formants).

The results of the analysis of variance with d' as dependent value are illustrated in Table 23 and Figure 44.

Table 23: Results of the analysis of variances based on d' in Experiment 3.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Stimulus type	4.03	2	56	.02	.13
Vowel type	0.89	1	28	.35	.03
Auditory contrast	11.45	1	28	< .01	.29
Stimulus * Vowel type	0.07	2	56	.93	< .01
Stimulus type * Auditory contrast	2.08	2	56	.14	.07
Vowel type * Auditory contrast	11.03	1	28	< .01	.28
Stimulus * Vowel type * Auditory difference	3.60	2	56	.03	.11

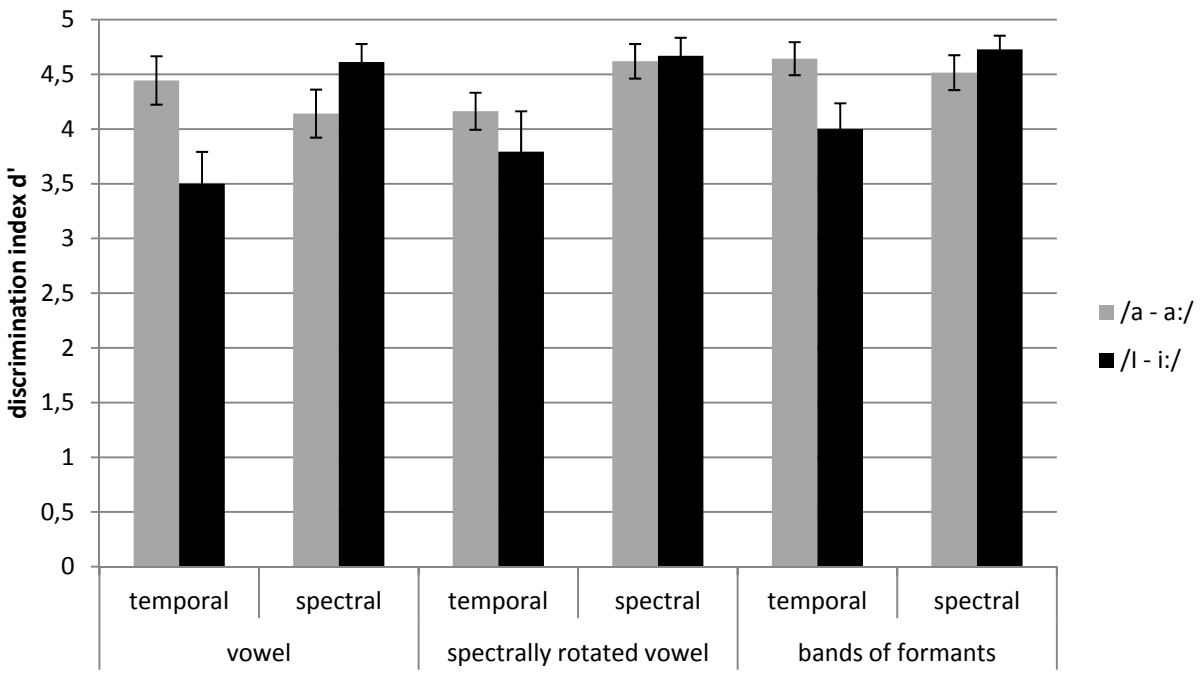

Figure 44: Means and standard errors of d' for each experimental condition of Experiment 3.

The ANOVA revealed a significant main effect of *Stimulus type* ($F(2,56) = 4.03$, $p = .02$) and *Auditory contrast* ($F(1,28) = 11.45$, $p < .01$). Discrimination of the vowel center stimuli was more difficult compared to the bands of formants ($t(29) = -2.69$, $p = .04$, $d = 0.49$). The spectrally rotated stimuli were slightly easier to discriminate compared to the vowel center stimuli, but this difference did not reach significance ($t(29) = 1.42$, $p = .16$). The difference between the spectrally rotated vowels and the bands of formants was not statistically significant either ($t(29) = -1.56$, $p = .13$).

The spectral difference was easier to discriminate compared to the temporal one ($t(29) = 2.29$, $p < .01$, $d = 0.53$). The interactions between *Stimulus* and *Vowel type* ($F(2,56) = 0.07$, $p = .93$) and between *Stimulus type* and *Auditory contrast* ($F(2,56) = 2.08$, $p = .14$) did not reach significance. However, there was a significant interaction between *Vowel type* and *Auditory contrast* ($F(1,28) = 11.03$, $p < .01$). The temporal condition tended to be more difficult for the vowel pair /a/ - /a:/ compared to the vowel pair /ɪ/ - /i:/ ($t(29) = 2.20$, $p = .07$). For the spectral condition, no difference between the two vowel types was found ($t(29) = -1.19$, $p = .24$).

There was a significant interaction between *Stimulus type*, *Vowel type,* and *Auditory difference* ($F(2,56) = 3.6$, $p = .03$). This triplet interaction might be explained by the fact that

the interaction between *Vowel type* and *Auditory contrast* seems to be more salient for the vowel center stimuli compared to the two non-speech conditions (see Figure 44).

As the spectrally rotated vowel center stimuli used in this experiment included frequencies beyond 4000Hz, performance for these stimuli was compared to the performance for the conventional spectrally rotated stimuli used in Experiment 1 with t-tests for independent samples. No systematic difference concerning the discrimination performance of the two experiments was found (see Table 24 and Figure 45).

Table 24: Results of the t-tests for independent samples comparing the discrimination performance for the spectrally rotated vowel center stimuli of Experiments 1 and 3.

Vowel type	Auditory contrast	*t*	df	*p*
/a/ - /a:/	Temporal	0.08	38	.94
	Spectral	0.16	38	.88
/ɪ/ - /i:/	Temporal	-1.1	38	.27
	Spectral	-0.80	38	.43

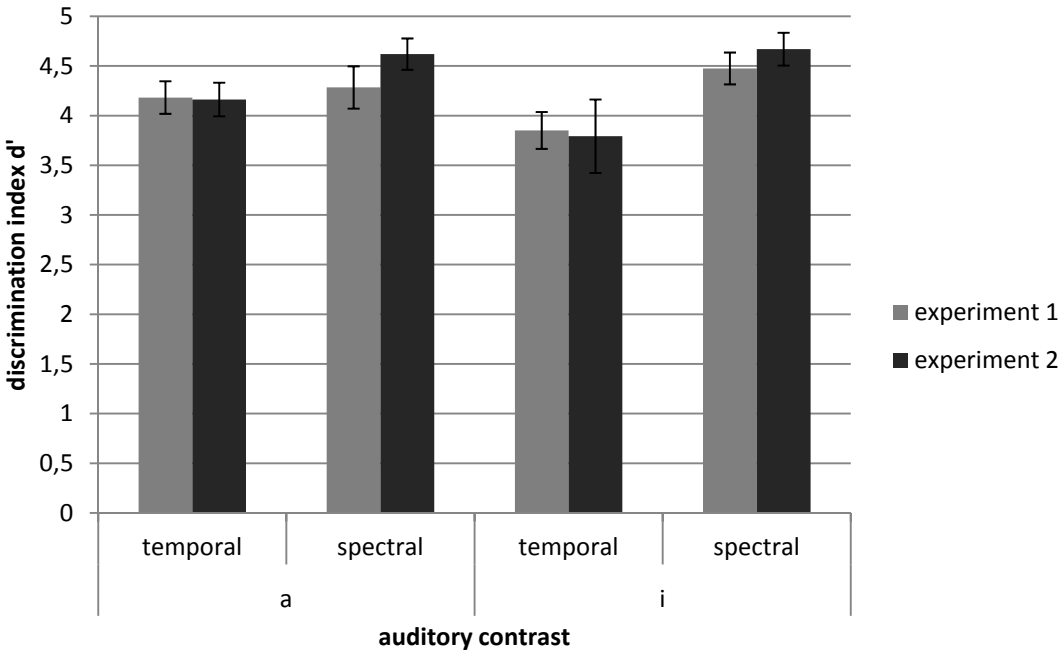

Figure 45: Comparison of the discrimination performance for the spectrally rotated vowel center stimuli of Experiments 1 and 3.

The last analysis of variance was based on the mean reaction times of correct responses and the results are depicted in Table 25 and Figure 46.

Table 25: Results of the analysis of variance based on reaction times in Experiment 3.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Stimulus type	0.54	2	56	.58	.02
Vowel type	1.36	1	28	.25	.05
Auditory contrast	54.89	1	28	< .01	.66
Stimulus * Vowel type	0.23	2	56	.80	< .01
Stimulus type * Auditory contrast	7.88	2	56	< .01	.26
Vowel type * Auditory contrast	7.28	1	28	.01	.21
Stimulus * Vowel type * Auditory difference	9.71	2	56	<.01	.26

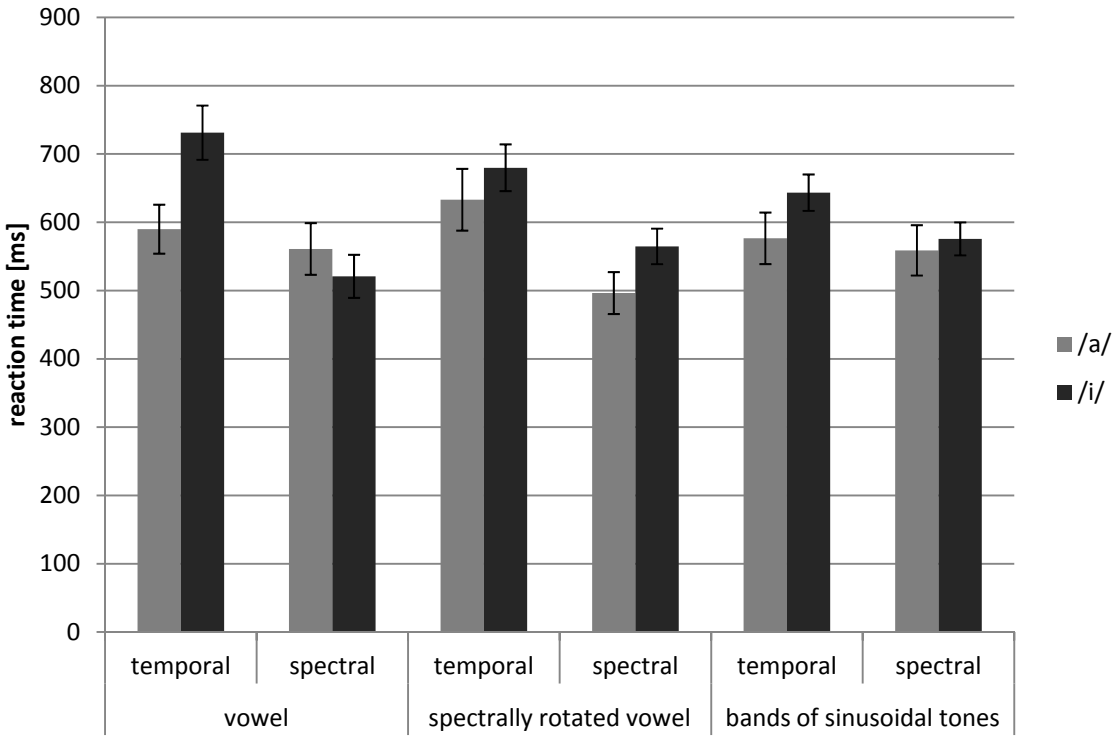

Figure 46: Means and standard errors of the reaction times for each experimental condition of Experiment 3.

The main effects of *Stimulus type* ($F(2,56) = 0.54$, $p = .58$) and *Vowel type* ($F(1,28) = 1.36$, $p = .25$) did not reach significance. There was a significant main effect of *Auditory Difference* ($F(1,28) = 54.89$, $p = < .01$). The spectral condition was discriminated faster compared to the

temporal condition ($t(29) = -6.72$, $p < .01$, $d = -1.24$). The interaction between S*timulus type* and *Vowel type* did not reach significance ($F(2,56) = 0.23$, $p = .80$). However, the interactions between S*timulus type* and *Auditory contrast* ($F(2,56) = 7.88$, $p < .01$), between *Vowel type* and *Auditory contrast* ($F(1,28) = 7.28$, $p = .01$) and between *Stimulus type*, *Vowel type*, and *Auditory contrast* ($F(2,56) = 9.71$, $p < .01$) reached significance. To examine these interactions, two additional analyses of variance were conducted, one for the vowel center stimuli and one for the two non-speech conditions. There was a significant interaction between the *Vowel type* and *Auditory contrast* in the vowel center stimuli ($F(1,28) = 17.70$, $p < .01$). This interaction was not significant for the non-speech stimuli ($F(1,28) = 0.30$, $p = .59$). To rule out any speed-accuracy trade off, the correlation between the accuracy of the answer (0 = error, 1 = correct response) and the reaction time was calculated. The correlation was $r = -.03$ ($p = .01$).

Discussion

The aim of this experiment was to find out whether speech and non-speech stimuli with the same complexity might be processed differently by the human brain. The MMN was used as index of pre-attentive auditory discrimination and compared to the discrimination performance in an active same-different task with the same stimulus set.

Role of "speechness" and complexity in the MMN

The analysis of variance based on the area of the MMN revealed a main effect of stimulus type. Vowel center stimuli evoked a larger MMN compared to the spectrally rotated vowel center stimuli and the bands of formants. No difference was found between the two non-speech stimulus types. This pattern of results goes in line with Hypothesis 1 and the *domain specific models*. The vowel center stimuli and the spectrally rotated vowel center stimuli are matched with respect to complexity. This means that the same number of different frequencies is included in both signals at each time point. If the different size of the MMN of speech and non-speech stimuli in previous experiments was mediated by the complexity of the stimuli (Hypothesis 3, *cue specific models*), there should have been no differences between the vowel center stimuli and the spectrally rotated vowel center stimuli. Nevertheless, it might be possible that speech stimuli are processed more efficiently

compared to non-speech, however, this relation does not rule out any additional effect of complexity (Hypothesis 2, a combination of the *domain specific and cue specific models*). In this scenario, one would expect a larger MMN for the vowel center stimuli compared to the spectrally rotated vowel center stimuli. There should also be an additional difference between the two non-speech stimulus types, as the bands of formants show a lower complexity than the spectrally rotated vowel center stimuli. The difference between the two stimulus types was far from reaching significance. As such, an additional effect from stimulus complexity seems to be unlikely. In summary, stimulus complexity does not explain differences in the size of the MMN of speech and non-speech stimuli in this experiment.

Although the vowel center stimuli were harder to discriminate compared to the non-speech stimuli, they were processed more efficiently by the brain, as they evoked a larger area of MMN. This finding coincides with the concept of language specific phoneme representations (Näätänen et al., 1997). It was shown that vowels evoke a larger MMN when they are part of the participants' mother tongue.

Influences of the German vowel system

The vowel center stimuli are based on the German vowel system (see Chapter 2). This is the reason why there should be a larger MMN for the salient conditions: For the vowel pair /a/ - /a:/, the temporal MMN should be larger than the one for the spectral condition and for the vowel pair /ɪ/ - /i:/, the spectral MMN should be larger than the temporal one (Hypothesis 4). There was a significant interaction between type of vowel and auditory contrast for the vowel center stimuli. As illustrated in Figures 42 and 43, the temporal contrast evoked a larger area of MMN compared to the spectral one for the vowel pair /a/ - /a:/, whereas the opposite pattern of results was found for the vowel pair /ɪ/ - /i:/. The difference between the vowel center stimuli and the two non-speech stimulus types was therefore largest in the salient conditions, namely for the temporal condition of the vowel pair /a/ - /a:/ and the spectral condition of the vowel pair /ɪ/ - /i:/.

The relation between the MMN and the active discrimination performance

The analysis of variance based on the discrimination index d' also revealed a significant main effect of stimulus type. However, performance for the vowel center stimuli was worse compared to the two non-speech stimulus types. This means that the expected positive, linear relation between the discrimination performance and the size of the MMN (see e.g., Aaltonen et al., 1994; Lang et al., 1990; Pakarinen et al., 2007) was not found between different stimulus types. Contrary to this, within one type of stimulus, discrimination performance and the area of the MMN went into the same direction: The interaction between type of vowel and auditory contrast was found for the behavioral data and the MMN. Discrimination for the vowel pair /a/ - /a:/ was better and faster for the temporal compared to the spectral condition and the temporal MMN was higher compared to the spectral one. For the vowel pair /ɪ/ - /i:/ the opposite pattern of results was found: discrimination was better and faster for the spectral condition compared to the temporal one and the area of the MMN was also higher for the spectral contrast. The interaction between type of vowel and auditory contrast was only found for the vowel center stimuli and not for the two non-speech stimulus types. To sum up, within one type of stimulus, conditions which were discriminated more easily and more rapidly, also lead to a higher area of MMN. However, this relation between discrimination performance and the size of the MMN was not found between the speech and non-speech stimuli.

The role of the size of contrast in the MMN

Based on the assumption that speech stimuli are always processed more efficiently compared to non-speech, one would have expected to find a larger MMN for speech stimuli in all studies that deal with speech and non-speech stimuli. Results however are mixed. Only a few studies report a higher MMN for speech stimuli compared to non-speech stimuli (see Table 17 for a summary). These experiments share one property: The difficulty of the contrasts was kept constant in these experiments (e.g., 10% increment of pitch or a duration decrement from 400ms to 200ms). This implies that the same size of contrast was used for speech and non-speech stimuli.

Contrary to this, there are some studies in which the difficulty of the contrast was not matched: In the experiment of Tervaniemi and colleagues (1999), /e/ and /o/ were used in the speech condition and A major and A minor in the non-speech condition. The chords

evoked a higher MMNm compared to the vowels. It would appear doubtful that the amount of spectral change is the same for both conditions. Another example is the study of Wunderlich and colleagues (2001). They used /bæd/, /dæd/, /bæ/, and /dæ/ in the speech condition. For the non-speech condition, they used tones with a 10% pitch increment. It is not surprising that such a large contrast evoked a larger MMN compared to the small contrast between /d/ and /b/ in the speech condition. The same approach was also chosen in a study by Nikjeh and colleagues (2009). They compared the MMN of pure tones (1.5% and 6% pitch change) and harmonic tones with three overtones (1.5% and 6% pitch change) to the MMN evoked by the speech syllables /ba/ and /da/. In this case, the spectral contrast of the non-speech stimuli was much smaller compared to those in the study of Wunderlich and colleagues (2001). As a consequence, they did not find any differences between the speech and non-speech stimuli. Nevertheless, it seems doubtful that a pitch change of 6% represents the same difficulty as the contrast between /b/ and /d/. This could be why they did not find a larger MMN for the speech stimuli. All things considered, the mixed pattern of results appears to be a consequence of the different contrasts for speech and non-speech stimuli. Most studies in which the contrasts were equally difficult on the physical level for the speech and non-speech stimuli reported enhanced processing of speech stimuli.

It would be useful to provide additional discrimination indexes based on active discrimination tasks to control for the difficulty of contrasts. However, the vowel center stimuli and the spectrally rotated vowel center stimuli are not matched concerning one property, as only the vowel center stimuli show a harmonic structure.

The role of harmony

To my knowledge, there is no MMN study dealing with the question of whether harmonic stimuli might be processed more efficiently compared to disharmonic ones. According to the *cue specific models* of speech perception, it could be possible that the difference of the size of the MMN between the vowel center stimuli and the spectrally rotated vowel center stimuli could be mediated by the harmonic structure of the vowel. Experiment 4 was conducted to investigate this question.

Experiment 4

The vowel center stimuli and spectrally rotated vowel center stimuli used in Experiment 3 were matched with respect to complexity. Nonetheless, only the vowel center stimuli showed a harmonic structure. The aim of Experiment 3 was to find out whether the difference between the size of the MMN of the vowel center stimuli and the spectrally rotated vowel center stimuli could be explained by the fact that only the vowel center stimuli are harmonic. To achieve this goal, two non-speech stimulus types with the same complexity were compared: one with a harmonic, the second one with a disharmonic structure.

Participants

Fourteen adults (5 male) took part in the experiment. All of them had previously participated in Experiment 3. Ten were members of the /ɪ/ - /i:/ group. The mean age was 23.14 years with a standard deviation of 3.42 years. The range was 18 to 30 years.

Material, Task, and Apparatus

During the harmonic condition, two different tones were used. These tones had the same pitch as the vowel center stimuli in Experiment 3 (186Hz). The tones were generated by a clarinet and a saxophone. As a result, the two tones differed only with respect to timbre. The duration was matched to the vowel /a:/ of Experiment 3 (145ms). To create two disharmonic stimuli with the same complexity, both tones were spectrally rotated (see Chapter 2). The modified version as described in Chapter 3 was chosen to receive spectrally rotated stimuli with a complete spectrum. The spectrograms of the tones and the spectrally rotated tones are illustrated in Figure 47.

A classical oddball paradigm was used. Participants were seated in a comfortable chair and asked to ignore all auditory stimuli while watching a silenced film. The apparatus was the same as in Experiment 3.

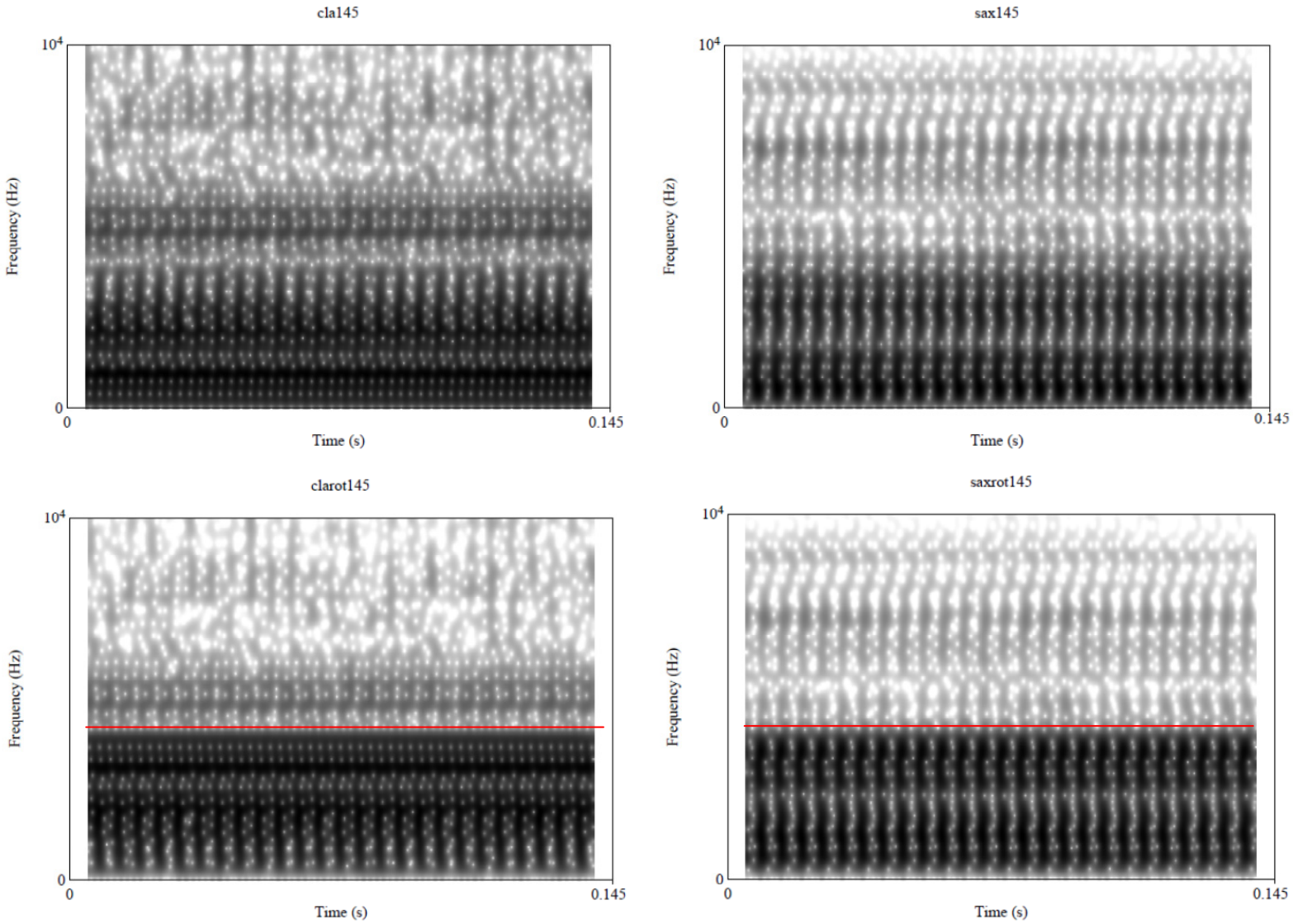

Figure 47: Spectrograms of the tones and spectrally rotated tones used in Experiment 4.

Design

There were separate blocks for the tones and the spectrally rotated tones. Additionally, every stimulus was presented as standard in one block and as deviant in another. In total, four blocks were presented to each participant. The sequence of the blocks was mixed for each subject. Each block contained 1050 standard stimuli (p = 84%) and 200 deviant stimuli (16%). The SOA was 500ms.

Dependent variables

First, an offline band-pass filter ranging from 1 to 30Hz was used. The ERPs were computed separately for each standard and deviant. The time window ranged from 200ms before to 500ms after stimulus onset. The first 200ms served as baseline for the averaged signal. The first 10 standards of each block and all epochs containing eye movement larger than 75µV were excluded.

The dependent value was the area under the difference curve within a 50ms time window around the peak latency (see Experiment 3). The peak latency was estimated at the fronto-central electrode. As each stimulus was presented once as standard and once as deviant, the difference curve was calculated with the following formula: deviant (cla) + deviant (sax) - standard (cla) - standard (sax). This procedure eliminates any possibility that the MMN might be distorted by different stimulus properties of the two tones. All seven positions of electrode (F3, Fz, F4, Cz, Pz, LM and RM) were included for the ANOVA.

Hypothesis

(1) If harmonic stimuli evoke a larger MMN compared to disharmonic stimuli, the tones should show a larger MMN area compared to the spectrally rotated tones.

Results

The ERPs of the tones and the spectrally rotated tones are illustrated in Figures 48 and 49. The difference waves are illustrated in Figure 50. The average of rejected trials is depicted in Table 26 and 27.

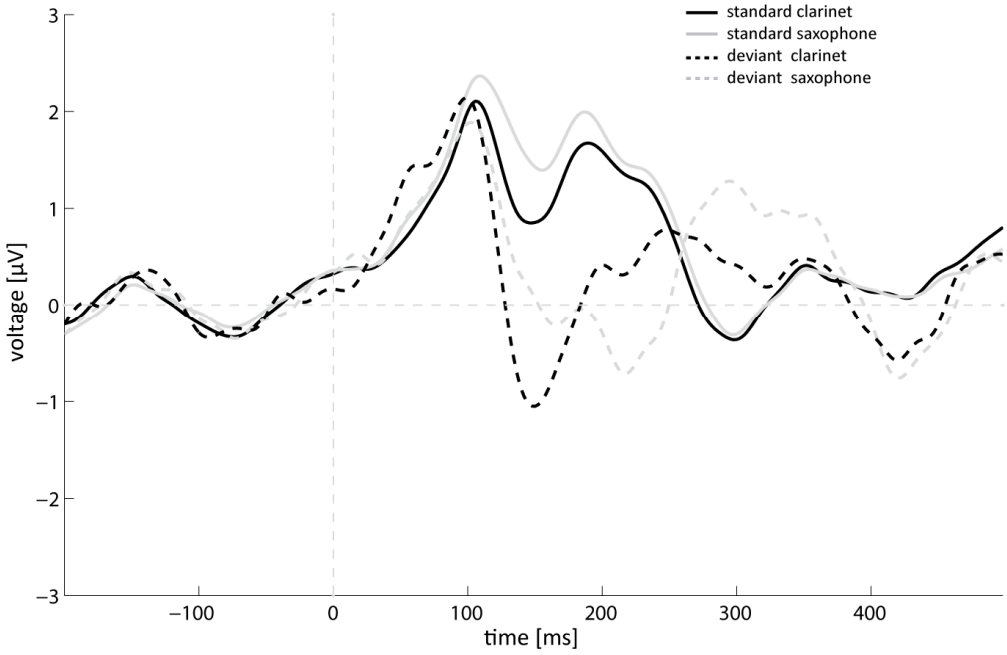

Figure 48: ERPs for each tone at Fz in Experiment 4. The ERPs of the standards are represented by the solid lines (clarinet = black, saxophone = grey). The ERPs of the deviants are represented by the dashed lines (clarinet = black, saxophone = grey). Time (in ms) is displayed on the x-axis, voltage (in µV) on the y-axis.

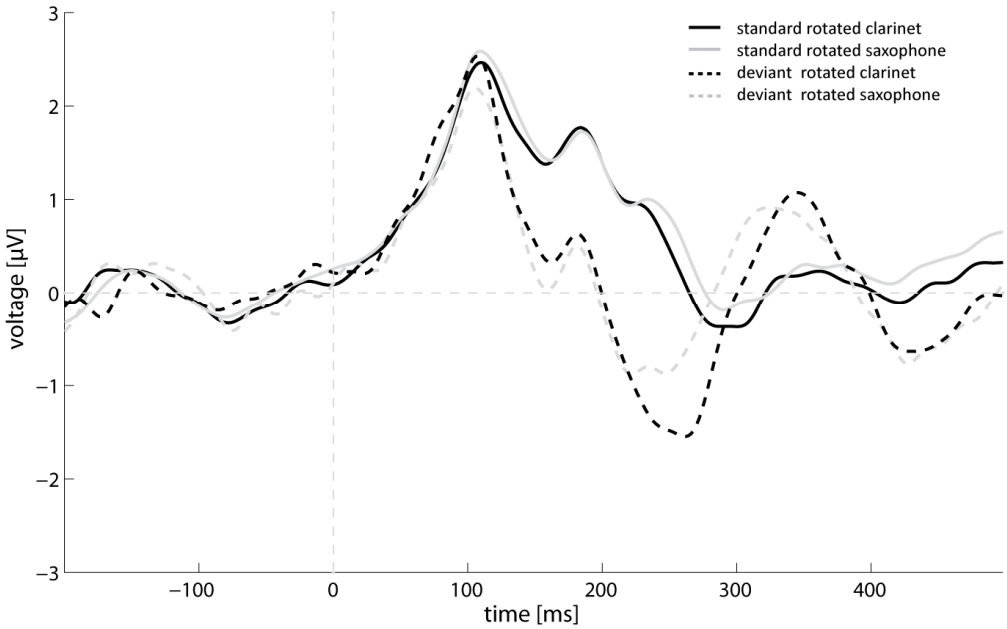

Figure 49: ERPs for each spectrally rotated tone at Fz in Experiment 4. See Figure 48 for details.

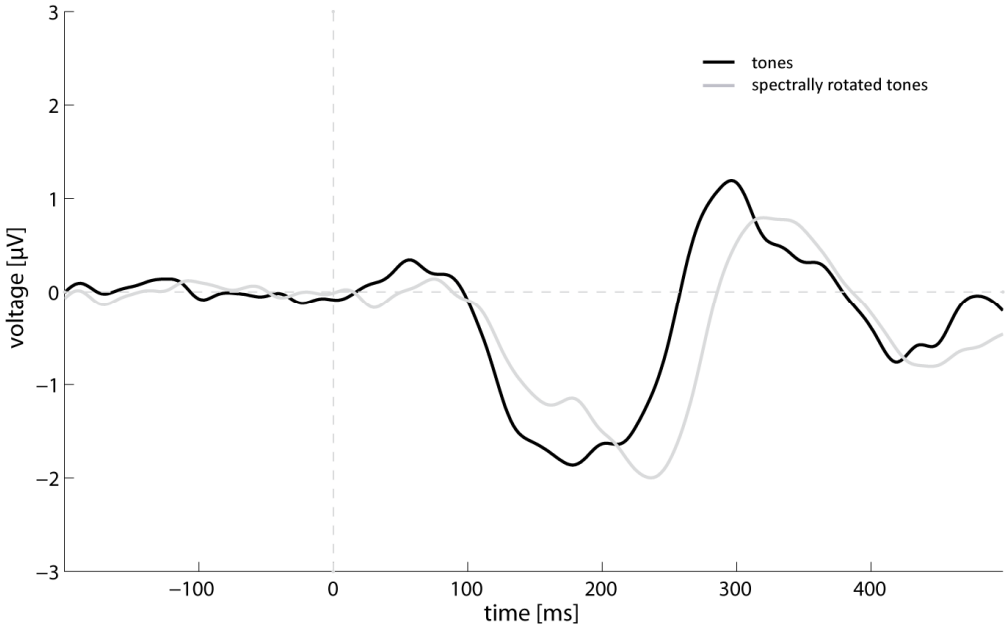

Figure 50: Difference waves for the tones (black) and spectrally rotated tones (grey) at the Fz in Experiment 4. Time (in ms) is displayed on the x-axis, voltage (in µV) on the y-axis.

Table 26: Mean, standard error, and maximum of rejected trials for the tones ("cla" = clarinet, "sax" = saxophone) in Experiment 4.

Tones		Mean	Standard error	Maximum
Cla	Standard	184.28	30.51	342
	Deviant	30.07	4,68	60
Sax	Standard	155.55	26.83	344
	Deviant	34.86	6.01	78

Table 27: Mean, standard error, and maximum of rejected trials for the spectrally rotated tones ("cla" = clarinet, "sax" = saxophone) in Experiment 4.

Spectrally rotated tones		Mean	Standard error	Maximum
Cla	Standard	151.73	23,84	309
	Deviant	32.50	5.94	70
Sax	Standard	172.95	29.62	396
	Deviant	26.50	4.64	60

A significant MMN was found for each type of stimulus at each electrode (see Table 28).

Table 28: T-tests for one sample on the basis of the area of MMN for each position of electrode**.**

Stimulus type	Location	*t*(13)	*p*
Tones	F3	6.75	< .01
	Fz	7.62	< .01
	F4	8.22	< .01
	Cz	7.04	< .01
	Pz	5.50	< .01
	LM	4.91	< .01
	RM	6.50	< .01
Spectrally rotated tones	F3	9.22	< .01
	Fz	11.04	< .01
	F4	11.20	< .01
	Cz	8.50	< .01
	Pz	8.83	< .01
	LM	6.85	< .01
	RM	5.29	< .01

An ANOVA for the two within factors *Harmony* and *Electrode* was conducted. There was a significant main effect of *Electrode* ($F(6,78) = 20.45$, $p < .01$) (see Table 29). As expected, the area of MMN was larger at frontal electrodes (see Figure 51). The main effect of *Harmony* ($F(1,13) < 0.01$, $p = .94$) and the interaction between *Harmony* and *Electrode* ($F(6,78) = 0.15$, $p = .69$) did not reach significance.

Table 29: Results of the analysis of variances based on the area of the MMN in Experiment 4.

Factor	*F*	df(factor)	df(error)	*p*	partial eta²
Harmony	< 0.01	1	13	.94	< .01
Electrode	20.45	6	78	< .01	.61
Harmony * Electrode	0.27	6	78	.69	.02

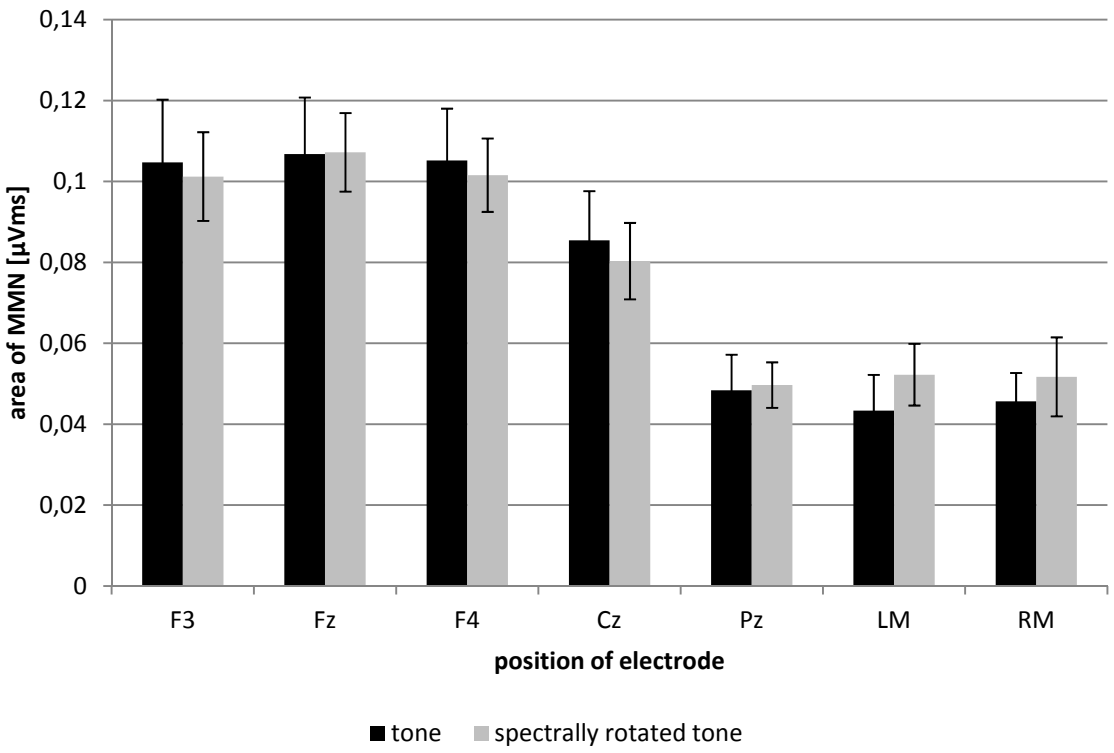

Figure 51: Means and standard errors of the area of the MMN for the tones and spectrally rotated tones in Experiment 4.

Discussion

The area of the MMN for both the tones and spectrally rotated tones decreased systematically with the position of electrode from frontal to occipital (see Figure 51). This pattern of results was expected, as the MMN was regularly found to be larger for frontal and central electrodes (Näätänen et al., 2007). As there was no systematical difference concerning the area of the MMN between the tones and the spectrally rotated tones, these data do not support the idea that harmonic stimuli might be processed more efficiently by the human brain (see Figure 51). Nevertheless, a main effect that did not reach significance is no proof of similarity as the test power in such a small sample is not sufficient. However, the partial eta² was smaller than .01 in this experiment, implying that although we cannot prove that there is no difference between the tones and spectrally rotated tones, it is clear that the impact of harmony on the area of the MMN was negligible in this experiment.

Conclusion

In Experiment 3, it was shown that vowel center stimuli evoke a larger MMN compared to non-speech stimuli, independently of stimulus complexity. Contrary to the two non-speech stimulus types, the vowel stimuli were harmonic. To my knowledge, there is no study dealing with the MMN in which harmonic stimuli were compared to disharmonic ones. In most experiments, vowels are compared to single sinusoidal tones or harmonic tones (e.g., Čeponiené et al., 2002; Jaramillo et al., 2001). Consequently, the difference between speech and non-speech stimuli could not be moderated by harmony in these experiments. However, in Experiment 3 of this work, the vowel center stimuli were compared to their spectrally rotated counterparts. Only the former ones show a harmonic structure. It is therefore important to prove that the difference of the area of the MMN is due to the "speechness" of the vowels and not due to the harmonic structure. As harmony seems to play only a negligible role on the area of the MMN, as shown in Experiment 4, it is unlikely that the significant difference between the vowels and the non-speech stimuli is only due to the influence of harmony. Taken together, these data support the *domain specific models* of speech perception.

One deficiency of the current experiment is that it does not take lateralization into consideration. There are many studies dealing with the issue of whether or not there might be a hemispherical specialization for speech and non-speech stimuli in the brain (e.g., Rinne et al., 1999; Shtyrov, Kujala, Palva, Ilmoniemi, & Näätänen, 2000). Only seven electrodes were used in the current experiment. Therefore, it was inappropriate to incorporate the issue of lateralization to our design. There are numerous functional imaging studies in which spectrally rotated speech is used. It is also conceivable to use these stimuli in EEG experiments with dipole analysis.

Another deficiency concerns the control for the influences of the properties of each stimulus on the MMN - This potential confounding factor was only controlled for in Experiment 4, as each stimulus was once presented as standard and once as deviant. However, the pattern of results in Experiment 3 seems unlikely to be distorted to a great extent, as the size of the MMN is comparable to the active discrimination performance: The MMN decreased significantly for all contrasts which are thought to be difficult (the temporal condition of the vowel pair /ɪ/ - /i:/ and the spectral condition of the vowel pair /a/ - /a:/). These findings support the validity of this experiment.

Chapter 5:

General discussion

As mentioned in the beginning, speech is one of the most complex sounds in our daily environment. Nevertheless, the role of stimulus complexity has hardly been taken into consideration in prior auditory research, especially during the comparison of speech and non-speech processing. To make progress in filling this gap, the focus of this thesis has been put on controlling for stimulus complexity (see Chapter 2). This approach enabled to reveal auditory deficits in developmental dyslexia (see Chapter 3) and to test the *domain specific* and *cue specific models* of speech perception (see Chapters 1 and 4).

The aim of Experiment 1 (see Chapter 2) was to create and evaluate non-speech and speech stimuli with the same and lower complexity than German vowels. The vowel center stimuli were created on the basis of a German vowel length discrimination paradigm (Groth et al., 2011; Steinbrink et al., 2012; Steinbrink et al., in preparation). Spectrally rotated speech (Blesser, 1972; Scott et al., 2000) was chosen as non-speech analogue with comparable physical complexity. However, the procedure of the spectral rotation was modified to produce spectrally rotated vowels with a complete spectrum which is not limited by the band pass filter (see Chapters 3 and 4). Importantly, this new approach circumvents the disadvantage of the low pass filtering of the speech stimuli which was found to impair the perceived naturalness of the speech signal.

Additionally, a completely new non-speech stimulus class was developed in Experiment 1, i.e. bands of formants. These stimuli are not only based on single frequencies of the formants, but instead they include the formants' bandwidth. Therefore, they are more complex and more comparable to the physical structure of vowels than single sine waves which represent the formant frequency only and which were commonly used in prior research.

In Experiment 2, these newly developed stimuli were used to investigate general auditory and speech processing in dyslexic adults compared to age matched controls (see Chapter 3). The dyslexic group was found to be impaired for all stimulus types (vowel center stimuli,

spectrally rotated vowel center stimuli, and bands of formants) and all auditory contrasts (temporal, spectral, spectro-temporal). This result goes in line with the assumption of a general auditory impairment which is not restricted to temporal features as one of the causes in developmental dyslexia.

The aim of the following chapter (Chapter 4) was to test the *domain specific* and the *cue specific models* of speech perception with the same stimulus set (see Experiment 3). An EEG component, called the MMN (Näätänen et al., 1978; Näätänen, 1979; Näätänen & Michie, 1979) was used to investigate the auditory discrimination of speech and non-speech stimuli at the pre-attentive level. According to the domain specific models of speech perception, differences in the processing of the speech and non-speech stimuli should be found independently from stimulus complexity, whereas no differences are expected between the vowel center stimuli and the spectrally rotated vowel center stimuli according to the cue specific models. Vowel center stimuli evoked a larger MMN compared to the spectrally rotated vowel center stimuli and the bands of formants, indicating that speech was processed more efficiently compared to non-speech, independently from stimulus complexity.

However, the vowel center stimuli and the spectrally rotated vowel center stimuli were not matched with respect to one feature: Harmony. According to the *cue specific models* of speech perception, this feature might be an additional confounding factor, which might be the reason for the differences in the processing of the speech and non-speech stimuli in Experiment 3. Therefore, the role of harmony was investigated in Experiment 4. No difference was found between the size of the MMN of the tones and the spectrally rotated tones. This finding shows that the difference between the size of the MMN of the speech and non-speech stimuli in Experiment 3 was not due to the harmonic and disharmonic structure of the different stimulus types. Both experiments (Experiments 3 and 4) could be explained on the basis of the *domain specific models*.

One advantage of the newly developed stimulus set is that it was not just controlled for physical features. The stimulus rating (see Chapter 3) proved that the vowel center stimuli were actually perceived as speech and the spectrally rotated vowel center stimuli were perceived as non-speech. Such control ratings should be included in future studies dealing with spectrally rotated speech to make sure that no phonological representations are used

for its processing (Azadpour & Balaban, 2008), especially when consonants are used, as these are hardly impaired by the inversion through spectral rotation (Blesser, 1972). Moreover, as the complete spectrum of the vowel center stimuli was used instead of the low pass filtered version, the perceived naturalness of the vowel center stimuli was increased.

Another advantage of the developed stimulus set is that the difficulty of the contrasts was controlled for to avoid bottom and ceiling effects that might have covered differences between dyslexics and controls and between speech and non-speech stimuli. This factor seems to be even more important than stimulus complexity, as differences between dyslexic groups and age matched controls were also reported for single sinusoidal tones (e.g., Banai & Ahissar, 2004; Heath et al., 2006), but only for contrasts smaller than 10% (Häämälainen et al., 2012). Our results support the assumption that stimulus complexity only plays a minor role, if even one at all, as the dyslexic group was also impaired for non-speech stimuli with lower complexity (bands of formants). This finding is crucial for the comparability of prior studies because it tells us that the heterogeneous data situation is likely a result of other factors, like the kind of task (Banai & Ahissar, 2006), the size of contrasts (Hämälainen et al., 2012), the composition of the sample and the criterion for being included into the study etc., rather than a result of the varying complexities of speech and non-speech stimuli.

The final advantage of the chosen stimulus set concerns the different auditory features (temporal, spectral and spectro-temporal) which were taken into consideration. Most studies focused on either temporal, spectral or spectro-temporal processing, while neglecting the other ones. Contrary to this short coming, all three auditory features were investigated in Experiment 2.

However, this work also shows some room for improvements. The results of Experiment 2 go in line with the assumption of a general auditory impairment as a cause in developmental dyslexia, as the dyslexic group was found to be impaired for all stimulus types and auditory features. However, the deficit was not found consistently in the dyslexic group. This finding is not surprising as it is likely that developmental dyslexia cannot be explained by a single cause (Lachmann, 2002; Naidoo, 1972) and multicausal subgroups have been reported regularly in prior research (see Chapter 3 for details). It could also be that a dyslexic child or adult might have multiple deficits and not only one (Bishop, 2006; Snowling, 2008). This means that the results of each study are highly dependent on the composition of its

respective sample, and therefore, it is not surprising that the auditory deficit was not found for all participants in this experiment. One solution for this problem would be to use larger samples to be able to estimate the prevalence of auditory deficits in dyslexic children and adults more reliably.

It should also be noted that there are some short comings regarding the chosen sample in this experiment. To begin with, none of the dyslexic participants had an official diagnosis, although each one reported having reading problems since primary school. A second short coming concerns the matching of the control and dyslexic group. The members of the control group had a higher level of school education. However, the auditory deficit of the dyslexic group was still observable after controlling for IQ differences.

Although participants with a diagnosis of attention disorder were excluded and the discrimination accuracy of the dyslexic group did not drop during the course of the experiment, it could be argued that the lower discrimination performance of the dyslexic group was not due to auditory impairments but due to attention deficits (Breier et al., 2003) or that it might be a consequence of both (Snowling, 2001). This was also indicated by the observation that reaction times slowed down in the dyslexic group only.

A MMN experiment, as introduced in Chapter 4, might be the solution to solve this dilemma. The MMN, as an objective index of auditory discrimination, without any requirement on attention, would be suitable in this context and was already used in many studies dealing with auditory processing in dyslexia (see Bishop, 2007 for a review). As the deficits can be located on different levels (Frith, 1985), it is also possible to reveal group differences on the neurophysiological level that might not be apparent on the behavioral level (Stoodley et al., 2006). It was shown in Experiment 3 that the MMN can be reliably evoked by each stimulus class which was used to investigate the dyslexic sample. The multifeature design (Näätänen et al., 2004) was used in Experiment 3 which is more time efficient compared to the classical oddball paradigm and therefore, especially suitable for clinical samples and children. It could also be used in the context of longitudinal studies which enable, in contrast to cross-sectional studies, to reveal causal relations between auditory deficits, phonological impairments and dyslexia.

The aim of Experiment 3 was to investigate the role of stimulus complexity during the auditory processing in the healthy human brain. The MMN was chosen as index of auditory

discrimination performance. According to the *cue specific models* of speech perception, differences between the processing of speech and non-speech sounds should be moderated by the different physical properties of the two stimulus classes. Contrary to this assumption, according to the *domain specific models* of speech perception, differences between speech and non-speech stimuli should persist even after controlling for the physical properties and the difficulty of the contrasts. The size of the MMN was found to be independent of stimulus complexity. The vowel center stimuli evoked a larger MMN compared to the two non-speech stimulus types, although the vowels were harder to discriminate compared to the latter ones. This difference was highest for these contrasts which are salient for the German vowel system: the temporal contrast of the vowel pair /a/ - /a:/ and the spectral contrast of the vowel pair /ɪ/ - /i:/. This finding supports the concept of language specific phoneme representations (Näätänen et al., 1997). It was shown that vowels evoke a larger MMN when they are part of the participants' mother tongue. These influences of the mother tongue underpin the idea of the *domain specific models* of speech perception.

The results of Experiment 3 and 4 are explainable with the *domain specific models*. However, these findings are no proof of one class of models or a counterevidence for the other one, as both classes of models can be appropriate depending on the context (Zatorre & Gandour, 2008). Especially by using imaging techniques, there is evidence that simple acoustic features, like temporal and spectral resolution, can explain patterns of hemispheric specialization (e.g., Nicholls, 1996; Zatorre & Belin, 2001; Zatorre, Belin, & Penhune, 2002). The combination of EEG measures and imaging techniques, such as simultaneous EEG-fMRI recording (Ritter & Villringer, 2006) could be one approach towards the understanding and integrating of the contradicting findings. Sure, the research on the processing of speech and non-speech sounds has not yet been finished. However, this work contributs to this debate as it revealed some crucial factors that should be taken into consideration in this research field: the detailed control of the stimulus features (e.g., complexity and harmony) and the control for the size of contrast between speech and non-speech stimuli.

General conclusion

It has been shown in the past that both the *domain specific* and the *cue specific models* of speech perception can account for findings in auditory research depending on context (see Zatorre & Gandour, 2008). This is why the essential conclusion of this thesis is that, although all experiments in this thesis speak against stimulus complexity or harmony as moderating factors for observed differences between speech and non-speech processing in auditory research, these features should be taken into consideration in future auditory research, as the *cue specific models* of speech perception are far from being disproved.

References

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

Aaltonen, O., Tuomainen, J., Laine, M., & Niemi, P. (1993). Cortical Differences in Tonal versus Vowel Processing as Revealed by an ERP Component Called Mismatch Negativity (MMN). *Brain and Language*, *44*(2), 139–152.

Abrams, D. A., Ryali, S., Chen, T., Balaban, E., Levitin, D. J., & Menon, V. (2012). Multivariate Activation and Connectivity Patterns Discriminate Speech Intelligibility in Wernicke's, Broca's, and Geschwind's Areas. *Cerebral Cortex*, *23*(7), 1703-1714.

Adlard, V., & Hazan, A. (1998). Speech Perception in Children with Specific Reading Difficulties (Dyslexia). *The Quarterly Journal of Experimental Psychology Section A*, *51*(1), 153–177.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in cognitive sciences*, *11*(11), 458–465.

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

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.

Alexander-Passe, N. (2006). How dyslexic teenagers cope: an investigation of self-esteem, coping and depression. *Dyslexia (Chichester, England)*, *12*(4), 256–275.

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., Sainio, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *77*(2), 151–155.

Alho, K., & Sinervo, N. (1997). Preattentive processing of complex sounds in the human brain. *Neuroscience Letters*, *233*(1), 33–36.

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

Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, *82*(5), 356–368.

Allen, J., Kraus, N., & Bradlow, A. (2000). Neural representation of consciously imperceptible speech sound differences. *Perception & Psychophysics*, *62*(7), 1383–1393.

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

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual on Mental Disorders: (DSM-IV)* (4th ed.). Washington, DC: American Psychiatric Press.

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory Processing Deficits in Reading Disabled Adults. *Journal of the Association for Research in Otolaryngology*, *3*(3), 302–320.

Amitay, S., Ben-Yehudah, G., Banai, K., & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain*, *125*(10), 2272–2285.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2007). A Common System for the Comprehension and Production of Narrative Speech. *Journal of Neuroscience*, *27*(43), 11455–11464.

Azadpour, M., & Balaban, E. (2008). Phonological Representations Are Unconsciously Used when Processing Complex, Non-Speech Signals. *PLoS ONE*, *3*(4), e1966.

Bakker, D. J. (1992). Neuropsychological classification and treatment of dyslexia. *Journal of learning disabilities*, *25*(2), 102–109.

Banai, K., & Ahissar, M. (2004). Poor Frequency Discrimination Probes Dyslexics with Particularly Impaired Working Memory. *Audiology and Neuro-Otology*, *9*(6), 328–340.

Banai, K., & Ahissar, M. (2006). Auditory Processing Deficits in Dyslexia: Task or Stimulus Related? *Cerebral Cortex*, *16*(12), 1718–1728.

Barry, J. G., Hardiman, M. J., Line, E., White, K. B., Yasin, I., & Bishop, D. V. M. (2008). Duration of auditory sensory memory in parents of children with SLI: A mismatch negativity study. *Brain and Language*, *104*(1), 75–88.

Baumann, R. (2010). *Physiologie.* (Klinke, R., Ed.). Stuttgart, New York: Thieme.

Bayerdörfer, H.-P. (2002). *Stimmen, Klänge, Töne: Synergien im szenischen Spiel*. Tübingen: Narr.

Beaton, A. (2004). *Dyslexia, reading, and the brain: A sourcebook of psychological and biological research* (1st ed.). Hove, East Sussex, New York: Psychology Press.

Beauchemin, M., & Beaumont, L. de. (2005). Statistical analysis of the mismatch negativity: To a dilemma, an answer. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 18–24.

Becker, T. (1998). *Das Vokalsystem der deutschen Standardsprache*. Frankfurt am Main: P. Lang.

Behrends, J. C., Bischofberger, J., Deutzmann, R., Ehmke, H., Frings, S., Grissmer, S., et al. (2010). *Duale Reihe Physiologie*. *Duale Reihe*. Stuttgart: Thieme.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*(1), 17–26.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., & Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, 10, 536-540.

Belvins, W. (1997). *Phonemic awareness activities for early reading success*. New York: Scholastic Professional Books.

Ben-Artzi, E., Fostick, L., & Babkoff, H. (2005). Deficits in temporal-order judgments in dyslexia: evidence from diotic stimuli differing spectrally and from dichotic stimuli differing only by perceived location. *Neuropsychologia*, *43*(5), 714–723.

Benasich, A. A., & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural brain research*, *136*(1), 31–49.

Benasich, A. A., Thomas, J. J., Choudhury, N., & Leppänen, P. H. T. (2002). The importance of rapid auditory processing abilities to early language development: evidence from converging methodologies. *Developmental psychobiology*, *40*(3), 278–292.

Bennet, D. (1968). Spectral form and duration as cues in the recognition of English and German Vowels. *Language and Speech*, (11), 65–85.

Benson, R. R., Richardson, M., Whalen, D. H., & Lai, S. (2006). Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. *NeuroImage*, *31*(1), 342–353.

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically Dissociating Speech and Nonspeech Perception in the Brain Using fMRI. *Brain and Language*, *78*(3), 364–396.

Bergelson, E., & Idsardi, W. J. (2009). A neurophysiological study into the foundations of tonal harmony. *NeuroReport*, *20*(3), 239–244.

Berti, S., Roeber, U., & Schröger, E. (2004). Bottom-Up Influences on Working Memory: Behavioral and Electrophysiological Distraction Varies with Distractor Strength. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *51*(4), 249–257.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by funtional magnetic resonance imaging. *Journal of Neuroscience*, *17*(1), 353–362.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human Temporal Lobe Activation by Speech and Nonspeech Sounds. *Cerebral Cortex*, *10*(5), 512–528.

Bishop, D. V. M. (2006). Dyslexia: what's the problem? *Developmental Science*, *9*(3), 256–257.

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., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402–416.

Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *Journal of Speech and Hearing Research*, *15*, 5–41.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of experimental child psychology*, *95*(3), 153–193.

Boder, E. (1973). Developmental Dyslexia: a Diagnostic Approach Based on Three Atypical Reading-spelling Patterns. *Developmental Medicine & Child Neurology*, *15*(5), 663–687.

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

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer. Retrieved from http://www.praat.org/

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.

Boets, B., Wouters, J., van Wieringen, A., Smedt, B. de, & Ghesquière, P. (2008). Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain and language*, *106*(1), 29–40.

Boetsch, E. A., Green, P. A., & Pennington, B. F. (1996). Psychosocial correlates of dyslexia across the life span. *Development and Psychopathology*, *8*(03), 539.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature*, *301*(5899), 419–421.

Breedin, S. D., Martin, R. C., & Jerger, S. (1989). Distinguishing auditory and speech-specific perceptual deficits. *Ear and hearing*, *10*(5), 311–317.

Bregman, A. S. (1995). *Auditory scene analysis: Perceptual organization of sound* (New ed.): Bradford Books.

Breier, J. I., Fletcher, J. M., Foorman, B. R., Klaas, P., & Gray, L. C. (2003). Auditory temporal processing in children with specific reading disability with and without attention deficit/hyperactivity disorder. *Journal of speech, language, and hearing research : JSLHR*, *46*(1), 31–42.

Bruce, D. J. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology*, *34*(2), 158–170.

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

Brunswick, N., McDougall, S., & de Mornay Davies, P. (2010). *Reading and dyslexia in different orthographies*. Hove, East Sussex, New York: Psychology Press.

Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology*, *26*(3), 429–438.

Budinger, E., Heil, P., König, R., & Scheich, H. (2005). *The auditory cortex: A synthesis of human and animal research*. Mahwah, NJ [u.a.]: Erlbaum.

Burden, R. L. (2005). *Dyslexia and self-concept: Seeking a dyslexic identity*. London, Philadelphia: Whurr.

Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., et al. (1994). An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America*, *96*(4), 2108.

Cacace, A. T., McFarland, D. J., Ouimet, J. R., Schrieber, E. J., & Marro, P. (2000). Temporal processing deficits in remediation-resistant reading-impaired children. *Audiology & neuro-otology*, *5*(2), 83–97.

Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M.-H., & McAdams, S. (2006). Seperate neural processing of timbre dimensions in auditory sensory memory. *Journal of Cognitive Neuroscience*, *18*(12), 1959–1972.

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

Carstensen, K.-U. (2004). *Computerlinguistik und Sprachtechnologie: Eine Einführung* (2nd ed.). München: Elsevier, Spektrum, Akad. Verl.

Casey, R., Levy, S. E., Brown, K., & Brooks-Gunn, J. (1992). Impaired emotional health in children with mild reading disability. *Journal of developmental and behavioral pediatrics : JDBP*, *13*(4), 256–260.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149–180.

Čeponiené, R., Yaguchi, K., Shestakova, A., Alku, P., Suominen, K., & Näätänen, R. (2002). Sound complexity and 'speechness' effects on pre-attentive auditory discrimination in children. *International Journal of Psychophysiology*, *43*(3), 199–211.

Cheour-Luhtanen, M., Alho, K., Sainio, K., Rinne, T., Reinikainen, K., Pohjavuori, M., et al. (1996). The ontogenetically earliest discriminative response of the human brain. *Psychophysiology*, *33*(4), 478–481.

Cheour, M. H. T., Leppänen, P., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology*, *111*(1), 4–16.

Choudhury, N., Leppanen, P. H., Leevers, H. J., & Benasich, A. A. (2007). Infant information processing and family history of specific language impairment: converging evidence for RAP deficits from two paradigms. *Developmental Science*, *10*(2), 213–236.

Colin, C., Hoonhorst, I., Markessis, E., Radeau, M., Tourtchaninoff, M. de, Foucher, A., et al. (2009). Mismatch Negativity (MMN) evoked by sound duration contrasts: An unexpected major effect of deviance direction on amplitudes. *Clinical Neurophysiology*, *120*(1), 51–59.

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.

Csépe, V. (2003). *Dyslexia: Different brain, different behavior*. *Neuropsychology and cognition: Vol. 23*. New York: Kluwer Academic/Plenum Publishers.

Csépe, V., Gyurkocza, E. E., & Osman-Sagi, J. (1998). Normal and disturbed phoneme perception as reflected by the mismatch negativity: do event-related potentials help to understand dyslexia? *Pathophysiology*, *5*(1), 202.

Dalebout, S. D., & Stack, J. W. (1990). Mismatch negativity to acoustic differences not differentiated behaviorally. *Journal of the American Academy of Audiology*, *10*, 388–399.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, *39*(6), 507–514.

Davids, N., Segers, E., van den Brink, D., Mitterer, H., van Balkom, H., Hagoort, P., & Verhoeven, L. (2011). The nature of auditory discrimination problems in children with specific language impairment: An MMN study. *Neuropsychologia*, *49*(1), 19–28.

Deacon, D., Gomes, H., Nousak, J. M., Ritter, W., & Javitt, D. (2000). Effect of frequency separation and stimulus rate on the mismatch negativity: an examination of the issue of refractoriness in humans. *Neuroscience Letters*, *287*(3), 167–170.

Deacon, D., Nousak, J. M., Pilotti, M., Ritter, W., & Yang, C.-M. (1998). Automatic change detection: Does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology*, *35*(4), 413–419.

Dehaene-Lambertz, G. (2000). Cerebral Specialization for Speech and Non-Speech Stimuli in Infants. *Journal of Cognitive Neuroscience*, *12*(3), 449–460.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

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.

Démonet, J.-F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *The Lancet*, *363*(9419), 1451–1460.

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

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.

Derry, R. (2006). *PC audio editing with Adobe Audition 2.0: Broadcast, desktop and CD audio production* (3rd ed.). Oxford, Burlington, MA: Focal Press.

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

Dilling, H., & Freyberger, H. J. (Eds.). (2012). *Taschenführer zur ICD-10-Klassifikation psychischer Störungen: DSM-IV-TR; … unter Berücksichtigung der Änderungen entsprechend ICD-10-GM (German Modifikation)* (6th ed.). Bern: Huber.

Draganova, R., Eswaran, H., Murphy, P., Huotilainen, M., Lowery, C., & Preissl, H. (2005). Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *NeuroImage*, *28*(2), 354–361.

Dummer-Smoch, L. (2007). Theoretische und schulpraktische Argumente für die Vereinbarkeit der beiden kontrovers diskutierten Konzepte Legasthenie / Allgemeine LRS (pp. 23–36). In G. Schulte-Körne (Ed.), *Legasthenie und Dyskalkulie. Aktuelle Enwicklungen in Wissenschaft, Schule und Gesellschaft*. Bochum: Dr. Dieter Winkler.

Eckert, M. A., Leonard, C. M., Richards, T. L., Aylward, E. H., Thomson, J., & Berninger, V. W. (2003). Anatomical correlates of dyslexia: frontal and cerebellar findings. *Brain : a journal of neurology*, *126*(Pt 2), 482–494.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read: Evidence From the National Reading Panel's Meta-Analysis. *Reading Research Quarterly*, *36*(3), 250–287.

Eichler, J. (2011). *Physik: Für das Ingenieurstudium - prägnant mit knapp 300 Beispielaufgaben* (4th ed.). Wiesbaden: Vieweg + Teubner.

Eimas, P D. (1963). The relation between identification and discrimination along speech and non-speech continua. *Language and Speech*, 6, 206-217.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing*, *8*(6), 453–485.

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.

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, *44*(1), 203–226.

Escera, C., & Grau, C. (1996). Short-term replicability of the mismatch negativity. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *100*(6), 549–554.

Esser, G., & Schmidt, M. H. (1993). Die langfristige Entwicklung von Kindern mit Lese-Rechtschreibschwäche. *Zeitschrift für Klinische Psychologie*, *22*, 100–116.

Esser, G., & Schmidt, M. H. (1994). Children with specific reading retardation - early determinants and long-term outcome. *Acta paedopsychiatrica*, *56*(3), 229–237.

Evans, S., Kyong, J. S., Rosen, S., Golestani, N., Warren, J. E., McGettigan, C., et al. (2013). The Pathways for Intelligible Speech: Multivariate and Univariate Perspectives. *Cerebral Cortex*,

Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton de Gruyter.

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.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: Facts and models* (3rd ed.). *Springer series in information sciences: Vol. 22*. Berlin, New York: Springer.

Fawcett, A. J., & Nicolson, R. I. (1992). Automatisation deficits in balance for dyslexic children. *Perceptual and motor skills*, *75*(2), 507–529.

Fawcett, A. J., & Nicolson, R. I. (1994). Naming Speed in Children with Dyslexia. *Journal of Learning Disabilities*, *27*(10), 641–646.

Flanagan, J. L. (1958). Pitch Discrimination for Synthetic Vowels. *The Journal of the Acoustical Society of America*, *30*(5), 435–442.

Ford, J. M., & Hillyard, S. A. (1981). Event-Related Potentials (ERP s) to Interruptions of a Steady Rhythm. *Psychophysiology*, *18*(3), 322–330.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: a tribute to Isabelle Y. Liberman* (pp. 97–117). Hillsdale, New York: Lawrence Erlbaum Accociates Ltd.

France, S. J., Rosner, B. S., Hansen, P. C., Calvin, C., Talcott, J. B., Richardson, A. J., & Stein, J. F. (2002). Auditory frequency discrimination in adult developmental dyslexics. *Perception & psychophysics*, *64*(2), 169–179.

Friesecke, A. (2007). *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker; mit … 145 Tabellen*. München: Saur.

Frisk, M. (1999). A complex background in children and adolescents with psychiatric disorders: developmental delay, dyslexia, heredity, slow cognitive processing and adverse social factors in a multifactorial entirety. *European child & adolescent psychiatry*, *8*(3), 225–236.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–322). London: Lawrence Erlbaum Accociates Ltd.

Fujisaki, H., Nakamura, K., & Imoto, T. (1975). Auditory perception of duration of speech and non-speech stimuli. In G. Fant & M. A. A. Tatham (Eds.), *Auditory Analysis and Perception of Speech* (pp. 197–220). New York: Academic Press.

Gaab, N., Gabrieli, J. D. E., Deutsch, D. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology and Neuroscience*, *25*, 295–310.

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

Galaburda, A. M., LoTurco, J., Ramus, F., Fitch, R. H., & Rosen, G. D. (2006). From genes to behavior in developmental dyslexia. *Nature Neuroscience*, *9*(10), 1213–1217.

Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). A twin study of the etiology of comorbidity: attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(2), 343–348.

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.

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

Goswami, U. (2003). Why theories about developmental dyslexia require developmental designs. *Trends in Cognitive Sciences*, *7*(12), 534–540.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. *Essays in developmental psychology*. Hove: Lawrence Erlbaum.

Goydke, K. N., Altenmüller, E., Möller, J., & Münte, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research*, *21*(3), 351–359.

Grimm, S., Widmann, A., & Schröger, E. (2004). Differential processing of duration changes within short and long sounds in humans. *Neuroscience Letters*, *356*(2), 83–86.

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.

Habib, M. (2000). The neurological basis of developmental dyslexia: An overview and working hypothesis. *Brain*, *123*(12), 2373–2399.

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-Related Brain Potential Correlates of Human Auditory Sensory Memory-Trace Formation. *Journal of Neuroscience*, *25*(45), 10494–10501.

Hagendorf, H., Krummenacher, J., Müller, H.-J., & Schubert, T. (2011). *Wahrnehmung und Aufmerksamkeit: Allgemeine Psychologie für Bachelor*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hämäläinen, J. A., Salminen, H. K., & Leppänen, P. H. T. (2012). Basic Auditory Processing Deficits in Dyslexia: Systematic Review of the Behavioral and Event-Related Potential/Field Evidence. *Journal of learning disabilities*, 46(5), 413-427.

Hari, R., Hämäläinen, M., Ilmoniemi, R. J., Kaukoranta, E., Reinikainen, K., Salminen, J., et al. (1984). Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: Neuromagnetic recordings in man. *Neuroscience Letters*, *50*(1-3), 127–132.

Hari, R., Sääskilahti, A., Helenius, P., & Uutela, K. (1999). Non-impaired auditory phase locking in dyslexic adults. *Neuroreport*, *10*(11), 2347–2348.

Heath, S. M., Bishop, D. V. M., Hogben, J. H., & Roach, N. W. (2006). Psychophysical indices of perceptual functioning in dyslexia: A psychometric analysis. *Cognitive neuropsychology*, *23*(6), 905–929.

Heiervang, E., Stevenson, J., Lund, A., & Hugdahl, K. (2001). Behaviour problems in children with dyslexia. *Nordic journal of psychiatry*, *55*(4), 251–256.

Heike, G. (1970). Lautdauer als Merkmal der wahrgenommenen Quantität. Qualität und Betongung im Deutschen. *Proceedings of the 6th International Congress of Phonetic Sciences*, 433–437.

Heike, G. (1971). Quantitative und qualitative Differenzen von /a(:)/-Realisationen im Deutschen. *Proceedings of the 7th International Congress of Phonetic Sciences*, 725–729.

Heim, S., Tschierse, J., Amunts, K., Wilms, M., Vossel, S., Willmes, K., et al. (2008). Cognitive subtypes of dyslexia. *Acta neurobiologiae experimentalis*, *68*(1), 73–82.

Hellbrück, J., & Ellermeier, W. (2004). *Hören: Physiologie, Psychologie und Pathologie* (2nd ed.). Göttingen [u.a.]: Hogrefe, Verl. für Psychologie.

Heslenfeld, D. J. (2003). Visual mismatch negativity. In J. Polich (Ed.), *Detection of change: event-related potential and fMRI findings.* Boston: Kluver Academics Publishers.

Hickok, G., Love, T., Swinney, D., Wong, E. C., & Buxton, R. B. (1997). Functional MR Imaging during Auditory Word Perception: A Single-Trial Presentation Paradigm. *Brain and Language*, *58*(1), 197–201.

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.

Hood, M., & Conlon, E. (2004). Visual and auditory temporal processing and early reading development. *Dyslexia (Chichester, England)*, *10*(3), 234–252.

Horváth, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Cognitive Brain Research*, *12*(1), 131–144.

Howard, D., Patterson, K., Wise, R. J. S., Brown, W. D., Friston, K., Weiler, C., & Frackowiak, R. (1992). The cortical localization of the lexicons. *Brain*, *115*(6), 1769–1782.

Hughes, W., & Dawson, R. O. N. (1995). Memories of school: Adult dyslexics recall their school days. *Support for Learning*, *10*(4), 181–184.

Huotilainen, M., Hotakainen, M., Parkkonen, L., Taulu, S., Simola, J., & et al. (2005). Short-term memory functions of the human fetus recorded with magnetoencephalography. *NeuroReport*, *16*, 81–84.

Ingram, T. T. S. (1963). Deleyed development of speech with special reference to dyslexia. *Proceedings of the Royal Society of Medicine*, *56*, 199–203.

Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, *38*(4), 723–727.

Jaramillo, M., Alku, P., & Paavilainen, P. (1999). An event-related potenital (ERP) study of duration changes in speech and non-speech sounds. *NeuroReport*, *10*(16), 3301–3305.

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.

Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, *290*(2), 101–104.

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

Johnson, D. J., & Myklebust, H. R. (1967). *Learning disabilities; educational principles and practices*. New York: Grune & Stratton.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 369–381.

Joutsiniemi, S.-L., Ilvonen, T., Sinkkonen, J., Huotilainen, M., Tervaniemi, M., Lehtokoski, A., et al. (1998). The mismatch negativity for duration decrement of auditory stimuli in healthy subjects. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(2), 154–159.

Kaiser, J., & Lutzenberger, W. (2001). Location changes enhance hemispheric asymmetry of magnetic fields evoked by lateralized sounds in humans. *Neuroscience Letters*, *314*(1-2), 17–20.

Kaukoranta, E., Sams, M., Hari, R., Hämäläinen, M., & Näätänen, R. (1989). Reactions of human auditory cortex to a change in tone duration. *Hearing Research*, *41*(1), 15–21.

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.

Kersting, M., & Althoff, K. (2004). *Rechtschreibungstest (RT)*. Göttingen: Hogrefe.

King, W. M., Lombardino, L. J., Crandell, C. C., & Leonard, C. M. (2003). Comorbid auditory processing disorder in developmental dyslexia. *Ear and hearing*, *24*(5), 448–456.

Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *International Journal of Psychophysiology*, 67(2), 131 -143.

Klatte, M., Steinbrink, C., Prölß, A., Estner, B., Christmann, C. A., & Lachmann, T. (in press). Effekte des computerbasierten Trainingsprogramms „Lautarium“ auf die phonologische Verarbeitung und die Lese-Restschreibleistungen bei Grundschulkindern.

Klicpera, C., & Gasteiger-Klicpera, B. (1998). *Psychologie der Lese- und Schreibschwierigkeiten*. Weinheim: Beltz, PVU.

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

Kohn, J., Wyschkon, A., Ballaschk, K., Ihle, W., & Esser, G. (2013). Verlauf von Umschriebenen Entwicklungsstörungen: Eine 30-Monats-Follow-up-Studie. *Lernen und Lernstörungen*, *2*(2), 77–89.

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, H. A. (1998). The late MMN wave in normal and language impaired children. *Pathophysiology*, *5*, 202.

Krauel, K., Schott, P., Sojka, B., Pause, B. M., & Ferstl, R. (1999). Is There a Mismatch Negativity Analogue in the Olfactory Event-Related Potential? *Journal of Psychophysiology*, *13*(1), 49–55.

Kühnis, J., Elmer, S., Meyer, M., & Jäncke, L. (2013). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: An EEG study. *Neuropsychologia*,

Kujala, T., Alho, K., Paavilainen, P., Summala, H., & Näätänen, R. (1992). Neural plasticity in processing of sound location by the early blind: an event-related potential study. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *84*(5), 469–472.

Kujala, T., & Näätänen, R. (2001). The mismatch negativity in evaluating central auditory dysfunction in dyslexia. *Neuroscience & Biobehavioral Reviews*, *25*(6), 535–543.

Lachmann, T. (2002). Reading disability as a deficit in functional coordination and information integration (pp. 165-198). In E. Witruk, A. D., Friederici, & Lachmann, T.

(Eds.). *Basis functions of language, reading and reading disability*. Boston: Kluwer/Springer.

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*, *56*(2), 105–120.

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

Lachs, L., & Pisoni, D. B. (2004). Cross-Modal Source Information and Spoken Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 378–396.

Landerl, K. (2003). Categorization of vowel length in German poor spellers: An orthographically relevant phonological distinction. *Applied Psycholinguistics*, *24*(04).

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

Lang, H. A., Nyrke, T., Ek, M., Aaltonen, O., Raimo, I., & Näätänen, R. (1990). Pitch discrimination performance and auditory event-related potentials. In C. H. M. Brunia, A. W. K. Gaillard, A. Kok, G. Mulder, & M. N. Verbaten (Eds.), *Psychophysiological brain research. vol. 1* (pp. 294–298). Tilburg, the Netherlands: Tilburg University Press.

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

Liberman, I. Y. (1989). Phonology and beginning reading revisited. In C. v. Euler, I. Lundberg, & G. Lennerstrand (Eds.), *Brain and reading* (pp. 207–220). New York: Stockton.

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

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*(2), 201–212.

Liddell, H., & Scott, R. (1996). *A Greek-English lexicon: With a revised supplement*. Oxford: Clarendon press.

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., Stephan, K. E., & Sporns, O. (2013). Modelling Trial-by-Trial Changes in the Mismatch Negativity. *PLoS Computational Biology*, *9*(2), e1002911.

Luck, S. J. & Lopez-Calderon, J. (2013). ERPLAB toolbox: a toolbox for ERP data analysis. Retrieved from http://erpinfo.org/erplab.

Lühr, R. (1993). *Neuhochdeutsch: Eine Einführung in die Sprachwissenschaft* (4th ed.). München: W. Fink.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14.

Lyytinen, H., Blomberg, A., & Näätänen, R. (1992). Event-Related Potentials and Autonomic Responses to a Change in Unattended Auditory Stimuli. *Psychophysiology*, *29*(5), 523–534.

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

Mangold, M. (2005). *Duden Band 6: Das Aussprachewörterbuch*. Mannheim: Dudenverlag.

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.

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant Rule Learning Facilitated by Speech. *Psychological Science*, *18*(5), 387–391.

Maughan, B., Gray, G., & Rutter, M. (1985). Reading retardation and antisocial behaviour: a follow-up into employment. *Journal of child psychology and psychiatry, and allied disciplines*, *26*(5), 741–758.

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.

McAnally, K. I., Hansen, P. C., Cornelissen, P. L., & Stein, J. F. (1997). Effect of time and frequency manipulation on syllable perception in developmental dyslexics. *Journal of speech, language, and hearing research : JSLHR*, *40*(4), 912–924.

McAnally, K. I., & Stein, J. F. (1996). Auditory temporal coding in dyslexia. *Proceedings. Biological sciences / The Royal Society*, *263*(1373), 961–965.

McArthur, G. M., Bishop, D. V. (2005). Speech and non-speech processing in people with specific language impairment: a behavioral and electrophysiological study. *Brain and Language*, 94(3), 260-273.

Meyer-Eppler, W. (1950). Reversed speech and repetition systems as means of phonetic research. *The Journal of the Acoustical Society of America*, *22*(6), 804–806.

Meyer, M., Elmer, S., Ringli, M., Oechslin, M. S., Baumann, S., & Jancke, L. (2011). Long-term exposure to music enhances the sensitivity of the auditory system in children. *European Journal of Neuroscience*, *34*(5), 755–765.

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.

Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J., & Doolong, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: an example of categorical perception. *The Journal of the Acoustical Society of America*, 60, 410-417.

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., Freeman, R. B., & Palermo, D. S. (1975). The ontogeny of brain lateralization for speech and non-speech stimuli. *Brain and Language*, *2*, 356–368.

Montgomery, C. R., Morris, R. D., Sevcik, R. A., & Clarkson, M. G. (2005). Auditory backward masking deficits in children with reading disabilities. *Brain and language*, *95*(3), 450–456.

Moore, B. C. J., & Tan C.-T. (2003). Perceived naturalness of spectrally distorted speech and music. *The Journal of the Acoustical Society of America*, *114*(1), 408–419.

Möser, M. (2007). *Messtechnik der Akustik* (1st ed.). Berlin: Springer.

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

Näätänen, R. (1979). Orienting and Evoked Potentials. In H. D. Kimmel, E. H. van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in huamns* (pp. 61–75). New Jersey: Erlbaum.

Näätänen, R. (2008). Mismatch negativity (MMN) as an index of central auditory system plasticity. *International Journal of Audiology*, *47*(s2), S16.

Näätänen, R., Gaillard, A. W. K., & Mäntsyalo, S. (1978). Early selelctive-attention effect on evoked potenital reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, R., Jiang, D., Lavikainen, J., Reinikainen, K., & Paavilainen, P. (1993). Event-related potentials reveal a menory trace for temporal features. *NeuroReport*, *5*(3), 310–312.

Näätänen, R., & Kreegipuu, K. (2012). The mismatch negativity (MMN). In S. J. Luck & E. S. Kappenman (Eds.), *Oxford library of psychology. The Oxford handbook of event-related potential components* (pp. 143–157). Oxford: Oxford University Press.

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

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

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1987). The mismatch negativity to intensity changes in an auditory stimulus sequence. *Electroencephalography and Clinical Neurophysiology*, *40*, 125–131.

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, *107*(1-3), 347–352.

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.

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

Näätänen, R., Syssoeva, O., & Takegata, R. (2004). Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology*, *41*(4), 660–663.

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.

Naidoo, S. (1972). *Specific dyslexia: the research report of the ICAA Word Blind Centre for Dyslexic Children*. London: Pitman.

Narain, C., Scott, S. K., Wise, R. J. S., Rosen, S., Leff, A., Iversen, S. D., & Matthews, P. (2003). Defining a Left-lateralized Response Specific to Intelligible Speech Using fMRI. *Cerebral Cortex*, *13*(12), 1362–1368.

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

Neath, I., Surprenant, A. M., & Crowder, R. G. (1993). The context-dependent stimulus suffix effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(3), 698–703.

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

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research*, *16*(3), 492–495.

Nicholls, M. E. (1996). Temporal processing asymmetries between the cerebral hemispheres: evidence and implications. *Laterality*, 1(2), 97-137.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in neurosciences*, *24*(9), 508–511.

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.

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988a). Event-Related Potentials to Breaks in Sequences of Alternating Pitches or Interstimulus Intervals. *Psychophysiology*, *25*(3), 262–268.

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988b). Event-Related Potentials to Time-Deviant and Pitch-Deviant Tones. *Psychophysiology*, *25*(3), 249–261.

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. *Journal of Neuroscience*, *27*(9), 2283–2289.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.-H., Saberi, K., et al. (2010). Hierarchical Organization of Human Auditory Cortex: Evidence from Acoustic Invariance in the Response to Intelligible Speech. *Cerebral Cortex*, *20*(10), 2486–2495.

Paavilainen, P., Arajärvi, P., & Takegata, R. (2007). Preattentive detection of nonsalient contingencies between auditory features. *NeuroReport*, *18*, 159–163.

Paavilainen, P., Jiang, D., Lavikainen, J., & Näätänen, R. (1993). Stimulus duration and the sensory memory trace: An event-related potential study. *Biological Psychology*, *35*(2), 139–152.

Paavilainen, P., Karlsson, M.-L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology*, *73*(2), 129–141.

Pahn, J. (2000). *Sprache und Musik: Beiträge der 71. Jahrestagung der Deutschen Gesellschaft für Sprach- und Stimmheilkunde e.V., Berlin, 12.-13. März 1999*. *Zeitschrift für Dialektologie und Linguistik. Beihefte: Vol. 107*. Stuttgart: Franz Steiner.

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

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

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

Parviainen, T., Helenius, P., & Salmelin, R. (2005). Cortical differentiation of speech and nonspeech sounds at 100 ms: implications for dyslexia. *Cerebral cortex (New York, N.Y. : 1991)*, *15*(7), 1054–1063.

Paulesu, E., Démonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, et al. (2001). Dyslexia: Cultural Diversity and Biological Unity. *Science*, *291*(5511), 2165–2167.

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

Peelle, J. E., Gross, J., & Davis, M. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, *23*(6), 1378–1387.

Pennala, R., Eklund, K., Hamalainen, J., Richardson, U., Martin, M., Leiwo, M., et al. (2010). Perception of Phonemic Length and Its Relation to Reading and Spelling Skills in Children With Family Risk for Dyslexia in the First Three Grades of School. *Journal of Speech, Language, and Hearing Research*, *53*(3), 710–724.

Pennington, B. F., & Lefly, D. L. (2001). Early Reading Development in Children at Family Risk for Dyslexia. *Child Development*, *72*(3), 816–833.

Pennington, B. F., van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child development*, *61*(6), 1753–1778.

Peter, V., McArthur, G., & Thompson, W. F. (2010). Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neuroscience Letters*, *482*(1), 71–75.

Pfister, B., & Kaufmann, T. (2008). *Sprachverarbeitung: Grundlagen und Methoden der Sprachsynthese und Spracherkennung*. Berlin, Heidelberg: Springer.

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

Plume, E., & Warnke, A. (2007). Definition, Symptomatik, Prävalenz und Diagnostik der Lese-Rechtschreib-Störung. *Monatsschrift Kinderheilkunde*, *155*(4), 322–327.

Pollmann, S. (2008). *Allgemeine Psychologie*. München: Reinhardt.

Ponton, C. W., & Don, M. (1995). The mismatch negativity in cochlear implant users. *Ear and Hearing*, *16*(1), 131–146.

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.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, (212), 947–950.

Rey, V., Martino, S. d., 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.

Riggenbach, P. (2000). *Funktionen von Musik in der modernen Industriegesellschaft: Eine Untersuchung zwischen Empirie und Theorie*. Marburg: Tectum.

Rinker, T., Kohls, G., Richter, C., Maas, V., Schulz, E., & Schecker, M. (2007). Abnormal frequency discrimination in children with SLI as indexed by mismatch negativity (MMN). *Neuroscience Letters*, *413*(2), 99–104.

Rinne, T., Alho, K., Alku, P., Holi, M., Sinkkonen, J., Virtanen, J., & et al. (1999). Analysis of speech sounds is left-hemisphere predominant at 100-150 ms after sound onset. *NeuroReport*, *10*(5), 1113–1117.

Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, *1077*(1), 135–143.

Ritter, P., & Villringer, A. (2006). Simultaneous EEG-fMRI recording. *Neuroscience & Behavioral Reviews*, 30(6), 823-838.

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., & Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, *83*(5), 306–321.

Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology*, *53*(3), 369–373.

Roeber, U., Widmann, A., & Schröger, E. (2003). Auditory distraction by duration and location deviants: a behavioral and event-related potential study. *Cognitive Brain Research*, *17*(2), 347–357.

Rosen, S., & Iverson, P. (2007). Constructing adequate non-speech analogues: what is special about speech anyway? *Developmental Science*, *10*(2), 165–168.

Rüsseler, J., Kowalczuk, J., Johannes, S., Wieringa, B. M., & Münte, T. F. (2002). Cognitive brain potentials to novel acoustic stimuli in adult dyslexic readers. *Dyslexia (Chichester, England)*, *8*(3), 125–142.

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, *398*(6730), 760.

Sabri, M., Binder, J., Desai, R., Medler, D. A., Leitl, M. D., & Liebenthal, E. (2008). Attentional and linguistic interactions in speech perception. *NeuroImage*, *39*(3), 1444–1456.

Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*, *12*(1), 171–180.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology /Evoked Potentials Section*, *62*(6), 437–448.

Sauter, D., & Eimer, M. (2010). Rapid Detection of Emotion from Human Vocalizations. *Journal of Cognitive Neuroscience*, *22*(3), 474–481.

Schäffler, T., Sonntag, J., Hartnegg, K., & Fischer, B. (2004). The effect of practice on low-level auditory discrimination, phonological skills, and spelling in dyslexia. *Dyslexia (Chichester, England)*, *10*(2), 119–130.

Schröder, M. R. (1968). Reference signal for signal quality studies. *The Journal of the Acoustical Society of America*, *44*, 1735–1736.

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

Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *100*(6), 517–526.

Schröger, E., Tervaniemi, M., & Näätänen, R. (1995). Time course of loudness in tone patterns is automatically represented by the human brain. *Neuroscience Letters*, *202*(1-2), 117–120.

Schulte-Körne, G. (2001). *Lese-Rechtschreibstörung und Sprachwahrnehmnung: Psychometrische und neurophysiologische Untersuchungen zur Legasthenie*. *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 14*. Münster: Waxmann.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998a). 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. (1998b). Role of auditory temporal processing for reading and spelling disability. *Perceptual and motor skills*, *86*(3 Pt 1), 1043–1047.

Schulte-Körne, G., & Mathwig, F. (2009). *Das Marburger Rechtschreibtraining: Ein regelgeleitetes Förderprogramm für rechtschreibschwache Kinder* (4th ed.). Bochum: Winkler.

Scott, S. K., Blank, C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400–2406.

Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. S. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *The Journal of the Acoustical Society of America*, *125*(3), 1737-1743.

Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. S. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech—A positron emission tomography study. *The Journal of the Acoustical Society of America*, *120*(2), 1075-1083.

Scott, S. K., & Wise, R. J. S. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition*, *92*(1-2), 13–45.

Sendlmeier, W. F. (1981). Der Einfluß von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen. *Phonetica*, *38*(5-6), 291–308.

Sendlmeier, W. F. & Seebode, J. (2006). *Formantkarte des deutschen Vokalsystems*. Retrieved from http://www.kgw.tu-berlin.de/forschung/Formantkarten

Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Demonet, J. (2001). Perceptual Discrimination of Speech Sounds in Developmental Dyslexia. *Journal of Speech, Language, and Hearing Research*, *44*(2), 384–399.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science*, *270*(5234), 303–304.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151-226.

Shaywitz, B. A., Fletcher, J. M., Holahan, J. M., & Shaywitz, S. E. (1992). Discrepancy compared to low achievement definitions of reading disability: results from the Connecticut Longitudinal Study. *Journal of Learning Disabilities*, *25*(10), 639–648.

Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., ConsTable, R. T., Fulbright, R. K., et al. (2007). Age-related changes in reading systems of dyslexic children. *Annals of neurology*, *61*(4), 363–370.

Shaywitz, S. E., Mody, M., & Shaywitz, B. A. (2006). Neural Mechanisms in Dyslexia. *Current Directions in Psychological Science*, *15*(6), 278–281.

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.

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.

Sidtis, J. J. (1980). On the nature of the cortical function underlying right hemisphere auditory perception. *Neuropsychologia*, *18*(3), 321–330.

Sjerps, M. J., Mitterer, H., & McQueen, J. M. (2011). Constraints on the processes responsible for the extrinsic normalization of vowels. *Attention, Perception, & Psychophysics*, *73*(4), 1195–1215.

Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory*, *15*(1), 34–56.

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

Snowling, M. J. (1995). Phonological processing and developmental dyslexia. *Journal of Research in Reading*, *18*(2), 132–138.

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

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*(1), 37–46.

Snowling, M. J. (2008). Specific disorders and broader phenotypes: the case of dyslexia. *Quarterly journal of experimental psychology (2006)*, *61*(1), 142–156.

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

Sörqvist, P., Nöstl, A., & Halin, N. (2012). Disruption of writing processes by the semanticity of background speech. *Scandinavian Journal of Psychology*, *53*(2), 97–102.

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

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2006). Converging Language Streams in the Human Temporal Lobe. *Journal of Neuroscience*, *26*(28), 7328–7336.

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

Steinbrink, C., Ackermann, H., Lachmann, T., & Riecker, A. (2009). Contribution of the anterior insula to temporal auditory processing deficits in developmental dyslexia. *Human Brain Mapping*, *30*(8), 2401–2411.

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. (2008). Phonological working memory in German children with poor reading and spelling abilities. *Dyslexia*, *14*(4), 271–290.

Steinbrink, C., Klatte, M., Lachmann, T. (in preparation). Phonological, temporal and spectral processing in vowel length discrimination is impaired in German primary school children with developmental dyslexia.

Steinbrink, C., & Lachmann, T. (in press). *Lese-Rechtschreibstörung: Grundlagen – Diagnostik – Intervention*. Heidelberg: Springer.

Steinbrink, C., Zimmer, K., Lachmann, T., Dirichs, M., & Kammer, T. (in press). Development of rapid temporal processing and its impact on literacy skills in primary school children. *Child Development*.

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.

Stoppelman, N., Harpaz, T., & Ben-Shachar, M. (2013). Do not throw out the baby with the bath water: choosing an effective baseline for a functional localizer of speech processing. *Brain and Behavior*, 3(3), 211-222.

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

Studdert-Kennedy, M., & Mody, M. (1995). Auditory temporal perception deficits in the reading-impaired: a critical review of the evidence. *Psychonomic Bulletin and Review*, *2*, 508–514.

Sussman, E., Winkler, I., & Wang, W. (2003). MMN and attention: Competition for deviance detection. *Psychophysiology*, *40*(3), 430–435.

Svensson, I., & Jacobson, C. (2006). How persistent are phonological difficulties? A longitudinal study of reading retarded children. *Dyslexia*, *12*(1), 3–20.

Swan, D., & Goswami, U. (1997). Picture Naming Deficits in Developmental Dyslexia: The Phonological Representations Hypothesis. *Brain and Language*, *56*(3), 334–353.

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.

Talcott, J. B., & Witton, C. (2002). A sensory-linguistic approach to normal and impaired reading development. In R. M. Joshi, E. Witruk, A. D. Friederici, & T. Lachmann (Eds.), *Neuropsychology and cognition. Basic Functions of Language, Reading and Reading Disability* (pp. 213–240). Boston, MA: Springer US.

Talcott, J. B., Witton, C., McLean, M. F., Hansen, P. C., Rees, A., Green, G. G. R., & Stein, J. F. (2000). From the Cover: Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Sciences*, *97*(6), 2952–2957.

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

Tallal, P. (1984). Temporal or phonetic processing deficit in dyslexia? That is the question. *Applied Psycholinguistics*, *5*(02), 167.

Tallal, P. (2000). Experimental studies of language learning impairments: From research to remediation. In D. B. Bishop & L. Leonard (Eds.), *Speech and language impairments in children* (pp. 131–156). Hove: Psychology Press.

Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in neurosciences*, *29*(7), 382–390.

Tallal, P., Merzenich, M. M., Miller, S., & Jenkins, W. (1998). Language learning impairments: integrating basic science, technology, and remediation. *Experimental Brain Research*, *123*(1-2), 210–219.

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*, 27–47.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*(1), 83–93.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*, *13*(1), 69–74.

Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, *226*(1), 1–4.

Tervaniemi, M., Ilvonen, T., Sinkkonen, J., Kujala, A., Alho, K., Huotilainen, M., & Näätänen, R. (2000a). 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., et al. (2006). Selective tuning of cortical sound-feature processing by language experience. *European Journal of Neuroscience*, *23*(9), 2538–2541.

Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Functional Specialization of the Human Auditory Cortex in Processing Phonetic and Musical Sounds: A Magnetoencephalographic (MEG) Study. *NeuroImage*, *9*(3), 330–336.

Tervaniemi, M., Schröger, E., Saher, M., & Näätänen, R. (2000b). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans – a mismatch negativity study. *Neuroscience Letters*, *290*(1), 66–70.

Tervaniemi, M., Winkler, I., & Näätänen, R. (1997). Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. *NeuroReport*, *8*, 2571–2574.

Toiviainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M., & Näätänen, R. (1998). Timbre similarity: convergence of neural, behavioral, and computational approaches. *Music Perception*, *16*, 223–241.

Torgesen, J. K., Wagner, R. K., Balthazar, M., Davis, C., Morgan, S., Simmons, K., … (1989). Developmental and individual differences in performance on phonological synthesis tasks. *Journal of experimental child psychology*, *47*(3), 491–505.

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. M. (2000). The irrelevant sound effect: Does speech play a special role? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1750–1754.

Tunmer, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic Abilities and Beginning Reading. *Reading Research Quarterly*, *23*(2), 134.

Tunmer, W. E., & Hoover, W. A. (1992). Cogntive and linguistic factors in learning to read. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 175–214). Hillsdale, N.J: Lawrence Erlbaum Accociates Ltd.

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

Undheim, A. M. (2003). Dyslexia and psychosocial factors. A follow-up study of young Norwegian adults with a history of dyslexia in childhood. *Nordic journal of psychiatry*, *57*(3), 221–226.

Ungeheuer, G. (1969). Das Phonemsystem der deutschen Hochlautung. In T. Siebs (Ed.), *Deutsche Aussprache. Reine und gemässigte Hochlautung mit Aussprachewörterbuch* (19th ed.). Berlin: VMA-Verl.

Uwer, R., Albrecht, R., & Suchodoletz, W. von. (2002). Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine & Child Neurology*, *44*(08).

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.

Vandermosten, M., Boets, B., Luts, H., 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.

Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Wouters, J., & Ghesquière, P. (2011). Impairments in speech and nonspeech sound categorization in children with dyslexia are

driven by temporal processing difficulties. *Research in Developmental Disabilities*, *32*(2), 593–603.

Vellutino, F. R. (1987). Dyslexia. *Scientific American*, *256*(3), 34–41.

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.

Wable, J., van den Abbeele, T., Gallégo, S., & Frachet, B. (2000). Mismatch negativity: a tool for the assessment of stimuli discrimination in cochlear implant subjects. *Clinical Neurophysiology*, *111*(4), 743–751.

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.

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & et al. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology*, *85*(1), 83–103.

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.

Whalen, D. H., Liberman, A. L. (1987). Speech perception takes precedence over non-speech perception. *Science*, 237, 169-171.

Warnke, A. (2008). Umschriebene Entwicklungsstörungen. In H.-J. Möller, G. Laux, & H.-P. Kapfhammer (Eds.), *Psychiatrie und Psychotherapie* (pp. 1120–1150). Berlin, Heidelberg: Springer Berlin Heidelberg.

Warnke, A., Schulte-Körne, G., & Ise, E. (2012). Developmental Dyslexia. In M. E. Garralda & J.-P. Raynaud (Eds.), *IACAPAP book series. The working with children and adolescents series: Vol. 19. Brain, mind, and developmental psychopathology in childhood* (pp. 173–198). Lanham, Md: Jason Aronson.

Watson, B. U., & Miller, T. K. (1993). Auditory perception, phonological processing, and reading ability/disability. *Journal of speech and hearing research*, *36*(4), 850–863.

Watson, C., & Willows, D. M. (1993). Evidence for a visual-processing-deficit subtype among disabled readers. In D. M. Willows, R. S. Kruk, & E. Corcos (Eds.), *Visual processes in reading and reading disabilities* (pp. 287–309). Hillsdale, New York: Erlbaum.

Weinzierl, S. (Ed.). (2008). *VDI*. *Handbuch der Audiotechnik*. Berlin, Heidelberg: Springer.

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

Weiß, R. H. (2006). *Grundintelligenztest Skala 2 Revision (CFT 20-R)*. Göttingen: Hogrefe.

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.

Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. *Journal of learning disabilities*, *33*(2), 179–191.

Wilmanns, J., & Schmitt, G. (2002). *Die Medizin und ihre Sprache: Lehrbuch und Atlas der Medizinischen Terminologie nach Organsystemen*. Landsberg/Lech: Ecomed.

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

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: evidence from children learning to read German. *Journal of experimental child psychology*, *61*(1), 80–90.

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

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology*, *92*(4), 668–680.

Winkler, I., Cowan, N., Csépe, V., Czigler, I., & Näätänen, R. (1996). Interactions between Transient and Long-Term Auditory Memory as Reflected by the Mismatch Negativity. *Journal of Cognitive Neuroscience*, *8*(5), 403–415.

Wirth, G., Ptok, M., & Schönweiler, R. (2000). *Sprachstörungen, Sprechstörungen, kindliche Hörstörungen: Lehrbuch für Ärzte, Logopäden und Sprachheilpädagogen* (5th ed.). Köln: Dt. Ärzte-Verl.

Witton, C., Stein, J. F., Stoodley, C. J., Rosner, B. S., & Talcott, J. B. (2002). Separate influences of acoustic AM and FM sensitivity on the phonological decoding skills of impaired and normal readers. *Journal of cognitive neuroscience*, *14*(6), 866–874.

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.

Yavas, M. S., & Gogate, L. J. (1999). Phoneme awareness in children: a function of sonority. *Journal of psycholinguistic research*, *28*(3), 245–260.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946-953.

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., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 5058, 846–849.

Zatorre, R. J., Gandour, J. T.. (2008). Neural specializations for speech and pitch: moving beyond that dichotomies. *Philosophical Transactions of the Royal Society B*, 363, 1087-1104.

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

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.

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.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, mich bei all den Personen zu bedanken, die mir meine Dissertation ermöglicht und mich dabei unterstützt haben.

Mein besonderer Dank gilt **Thomas Lachmann**, der mir die Möglichkeit zur Promotion gab und mir stets den Rücken gestärkt hat.
Danke auch an **Claudia Steinbrink**, für ihre intensive Betreuung und für ihr immer offenes Ohr bei Problemen.
Ein weiterer Dank geht an **Stefan Berti**, für die Unterstützung bei den EEG Experimenten.

Weiterhin möchte ich **Bernhard Schaaf-Christmann** für die Programmierung des Matlab Skripts für die Herstellung der Formantenbänder danken, sowie **Martin Dirichs** für seine Unterstützung bei der Programmierung der Experimente.

Außerdem möchte ich **Petra Linner** für ihre Unterstützung bei der Datenerhebung von Experiment 2 danken.

Vielen Dank auch an **Joanne Hall** und **Tina Weiß** für das Korrekturlesen der Arbeit, sowie an meine Kolleginnen **Andrea Prölß**, **Barbara Estner** und **Kirstin Bergström** für ihre zahlreichen kleinen Anregungen und Hilfestellungen.

Curriculum Vitae

Name:	Corinna Anna Christmann
06/2011 – present	Research assistant, Department of Cogntive and Developmental Psychology, University of Kaiserslautern
10/2006 – 04/2011	Study of Psychology at Johannes Gutenberg University Mainz
	Diploma Thesis: 'No influence of identity in the gaze direction aftereffect', Department of Psychology, Methods Section, Johannes Gutenberg University Mainz