Alexander Werth, Marie Josephine Rocholl, Karen Henrich, Manuela Lanwermeyer, Hanni Th. Schnell, Ulrike Domahs, Joachim Herrgen, and Jürgen Erich Schmidt

The interaction of vowel quantity and tonal cues in cognitive processing: An MMN-study concerning dialectal and standard varieties

Abstract: In this study, the influence of two dialectal prosodic features on the processing of lexical meaning during spoken word recognition was investigated in German dialect and non-dialect speakers. Previous studies in the field of German dialectology investigated differences between dialectal varieties and the Standard German variety by using mainly offline production and perception studies. The present study concentrates on brain responses to the phonological contrast of vowel quantity combined with tone accents, which occur in Germany exclusively in the Middle-Franconian dialect area (Moselle-Franconian, Ripuarian and southern Low Franconian dialects) but not in Standard German. In an event-related potential-study using a classic oddball paradigm, two groups of participants (dialect and Standard German speakers) were presented with two

Alexander Werth (Corresponding author): Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, alexander.werth@uni-marburg.de
Marie Josephine Rocholl: Amazon Development Center, Aureliusstraße 2, D-52064 Aachen, mjroc@amazon.com
Karen Henrich: Max-Planck-Institut für empirische Ästhetik, Grüneburgweg 14, D-60322 Frankfurt am Main, karen.henrich@aesthetics.mpg.de
Manuela Lanwermeyer: Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, manuela.lanwermeyer@uni-marburg.de
Hanni Th. Schnell: Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, hanni.schnell@uni-marburg.de
Ulrike Domahs: Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, ulrike.domahs@uni-marburg.de
Joachim Herrgen: Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, herrgen@uni-marburg.de
Jürgen Erich Schmidt: Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Pilgrimstein 16, D-35032 Marburg, schmidtj@uni-marburg.de

DOI 10.1515/9783110542899-008

words of a minimal pair ([ʃa̠ː²l] 'stale' vs. [ʃa̠lˑ²] 'acoustic noise') which have inverted lengths for the vowel and the lateral but both bear Tone Accent 2. Late mismatch negativity effects resulting from pre-attentive processing differ in amplitude and latency between the two groups of participants indicating varying phonological relevance of prosodic cues in these two varieties. Although both participant groups perceive Tone Accent 2 as a high tone, only the dialect group uses rules of tone-text-association within the minimal pair for lexical access.

Keywords: vowel quantity, tone accent, dialect, event-related potentials (ERP), auditory evoked responses, late mismatch negativity (MMN), contrast enhancement

1 Introduction

1.1 General Introduction

The Standard German vowel system contains 15 monophthongs and three diphthongs and can therefore be regarded as a relatively rich system of vowel phonemes compared to most other languages.[1] In some German dialects such as Moselle-Franconian, phoneme inventories are even more extensive. In Mayen, which is part of the Moselle-Franconian dialect area, 20 monophthongs and three diphthongs exist in the vowel inventory. While tenseness and long quantity[2] and then laxness and short quantity appear for most of the vowel phonemes in complementary distribution in the Standard German variety, all tense vowels can also appear with short quantity in stressed positions in this dialect (e.g., [ʃtuf] 'living room' and [ʃtof] 'textile fabric'). This means that there are absolutely parallel rows of short and long vowels concerning quality, and therefore vowel quantity has a larger phonological load than in the Standard German variety. Thus, an examination of quantity seems very worthwhile for this particular dialect variety.

In addition, in the Middle-Franconian dialects including Moselle-Franconian, there are tone accents associated with vowels or other sonorants, which can be used for lexical access. But the question still remains how these phonological

1 Some authors such as Wiese (2011) also consider the two vowels [ə] und [ɐ] to be phonemes of Standard German, both of which can only occur in an unstressed position. Also the vowel [ɛː] is discussed concerning its phoneme status. It is considered here, but it is not used by all German speakers and is distinctive only in some minimal pairs.

2 The term quantity is used here only for phonological distinction, whereas length refers to the units of perception and duration to the acoustic units.

differences between the different German varieties influence vowel processing. This question will be examined by means of electrophysiological responses to vowel perception.

Research findings already exist for neural processing of prosodic cues like length (e.g., Näätänen, Paavilainen, and Reinikainen 1989; Jaramillo, Alku, and Paavilainen 1999; Jaramillo et al. 2001; Amenedo and Escera 2000; Menning et al. 2002; Nenonen et al. 2003; Ylinen, Huotilainen, and Näätänen 2005; Ylinen et al. 2005, 2006; Kirmse et al. 2008; Chládková, Escudero, and Lipski 2013) and tone (e.g., Gandour et al. 2000, 2004; Gandour 2006, 2007; Chandrasekaran, Krishnan, and Gandour 2007, 2009; Kaan et al. 2008; Fournier et al. 2010), but there is hardly any research on the interaction of these cues for word processing. Only a few studies explored the processing of fundamental frequency (F0) and length, but only on the basis of isolated sounds (e.g., Czigler and Winkler 1996; Levänen et al. 1993; Wolff and Schröger 2001; Jaramillo et al. 2001) or vowels (e.g., Jaramillo et al. 2001) and therefore without direct relevance to word access. Moreover, a consensus on the existence of an additive processing of different cues (cf. Wolff and Schröger 2001) or a separated processing (cf. Jaramillo et al. 2001) remains to be found.

A number of results from the studies concerning cognitive processing of length or acoustic duration differences are summarised in the following: i) Depending on the direction in which the various stimuli are tested (long to short or short to long), MMN effects diverge to a different degree (cf. Jaramillo, Alku, and Paavilainen 1999; Takegata et al. 2008; Colin et al. 2009). Similarly, ii) the stimulus material used (e.g., harmonic tones vs. vowels) is considered to be another factor which influences the strength of the effect (cf. Jaramillo, Alku, and Paavilainen 1999; Jaramillo et al. 2001; Takegata et al. 2008; Christmann et al. 2014), and also iii) the phonological relevance of length in native languages seems to be a relevant factor (cf. Nenonen et al. 2003; Ylinen et al. 2006; Kirmse et al. 2008). It can be concluded that differences in length represent an important value for neural processing and depend on varying factors. But the stimuli used only range from harmonic tones to individual vowels to pseudowords. So the question remains as to whether the identified neural signatures can also be transferred to real words because differences in phonological quantity have an important function for lexical access unlike pure duration differences between harmonic tones.

The results mentioned above are valid only for simple length contrasts, but in some phonological systems they appear in conjunction with other prosodic features. Therefore, the question arises how quantity differences are processed when additional cues like tone accents can be used for lexical access.

Varieties in which such an issue can be investigated very well are the Moselle-Franconian dialects. These dialects spoken in the western part of Germany are of particular interest to researchers as they have developed a form-meaning mapping in which lexical and morphosyntactic distinctions between words are often derived from prosodic, i.e., tonal cues alone (as referred to tone accents like Tone Accent 1 and Tone Accent 2; cf. Schmidt 1986; Gussenhoven and Peters 2004; Werth 2011; Köhnlein 2011). In the Moselle-Franconian dialects "the contrast is acoustically manifested by a complex phenomenon consisting of a constant length opposition (Tone Accent 1 is always shorter) and a robust pitch difference" (Werth 2012: 187–188). Werth (2012: 192) describes the pitch difference as follows: "Accent 2 is marked with a lexical high tone [H_{lex}] on the second mora which Accent 1 lacks". For example, in the dialect of Mayen, the dialect words [d̥a̠͡ʊ^{1}f] 'pigeon' and [d̥a̠͡ʊ^{2}f] 'baptism' are distinguished by prosodic tone accent features (characterised by superscripts in the transcription) but not by features at a segmental level, while in the Standard German variety the phonemic segmental distinction /b/ versus /f/ is used to differentiate between the same lexemes: [ta̠͡ʊbə] 'pigeon' and [ta̠͡ʊfə] 'baptism'. Roughly speaking, compared to Standard German, Moselle-Franconian dialects show a preferred tendency to distinguish word meaning by prosody in that tone accents occur in all words that include heavy nuclei, i.e., on syllables containing a long vowel, a diphthong or a short vowel followed by a sonorant coda (cf. Schmidt 1986; Werth 2011; Köhnlein 2011).

Taken together, the previous considerations may mean that the combination of certain kinds of prosodic features like tone accents and vowel quantity are more relevant for lexical access in a tone-accent dialect than in the Standard German variety. To address this issue, we have investigated the interaction of two types of prosodic features in natural speech: i) tone-text-association, which is the association of tones with prosodic domains like syllables, moras or words, and ii) vowel quantity. To do so, we tested two participant groups: i) speakers of the Moselle-Franconian dialect who use tone accents as a cue for lexical access and ii) speakers of the Standard German variety who do not.

1.2 The interaction of tone and length in the Moselle-Franconian dialect of Mayen

Length is a feature in the Moselle-Franconian dialects which can be complementarily distributed between vowels and the following sonorant. Its auditory and acoustic salience is clearly evident in words with Tone Accent 2 (cf. Schmidt 1986: 185–191). Thus, if there is a long vowel followed by a sonorant in the rhyme of words with Tone Accent 2, the sonorant is short (e.g., [ʃa̠ː^{2}l] ‘stale’, see Table 1, condition b.). On the other hand, if the vowel is short, the sonorant is prolonged (e.g., [ʃa̠lˑ^{2}] ‘acoustic noise’, see Table 1, condition a. In CV phonology (cf. e.g., Clements and Keyser 1983; Lass 1984; Hayes 1999), these time units can be represented with the skeletal positions VCC (= short vowel + long sonorant) and VVC (= long vowel + short sonorant).[3] Phonetically, the occurrence of long sonorants can be explained by the fact that the realization of Tone Accent 2 requires more time than available in two short consecutive segments (short vowel + short sonorant).[4] By prolonging the sonorant, a sufficiently large time interval for full realization of the tonal pattern is provided (cf. Schmidt 1986: 185–191).

In some phonological theories, this large time interval can be analyzed through an underlying bimoraic domain (μμ), if we follow the definition of moras as tone-bearing units. Often, moras are defined units of weight and length (cf. e.g., Trubetzkoy 1939; Féry 2001; Hyman 2003), but in most tone and tonal accent languages, moras also operate as tone-bearing units (cf. e.g., Zec 1994; Hyman 2003; Zhang 2002; Werth 2011, 2012), which is their crucial function within the Moselle-Franconian dialects. Both moras are of decisive importance; intonational tone associated with the first mora expresses the communicative meaning, while the pitch level on the second mora (presence of a lexical high tone for Tone Accent 2 and absence for Tone Accent 1) expresses the lexical meaning (cf. Werth 2011, 2012). Although sonorant consonants are lengthened due to the association of the tone accent with the second mora, consonantal length alone could not be seen here as a phonologically distinctive feature. Schmidt (1986) assumes that

---

3 Other authors like Wiese (1996), Ramers (1998) or Ramers and Vater (1995) associate also long vowels with two skeletal positions, but with the symbols VC. To reflect the difference between the two mentioned conditions (long vowel + short sonorant consonant and short vowel + long sonorant consonant), we follow authors like Lass (1984) and Hayes (1999), using the symbols VV for long vowels in the phonological representation.

4 This larger time frame required for the realization of high tones in opposition to low tones is due to the fact that an increase in fundamental frequency in speech production universally requires more muscle activity and therefore more production time than for an F0-decrease (cf. Ohala 1972, 1978; Ohala and Ewan 1973; Sundberg 1973, 1979).

phonetic length of the sonorant consonant is to be regarded solely as a suprasegmental feature expression of Tone Accent 2. Thus, when preceded by a short vowel, the sonorant is lengthened to provide enough space for the realization of tonal information on the second mora. Therefore, length differences on the sonorant consonants are an epiphenomenon of the association of Tone Accent 2-lexical high tone with the consonant.

Another study that deals with Moselle-Franconian tone accents, and thus also constitutes direct relevance for the present study, is by Werth (2011, cf. 2012). Acoustically, tone accents are based on a phonetic-prosodic feature complex of fundamental frequency, duration and intensity, as has been demonstrated in numerous studies (cf. Heike 1962; Schmidt 1986; Gussenhoven and Peters 2004; Peters 2006; Werth 2011). However, Werth (2011) has shown that for the identification of tone accents, the tonal features represent the crucial cue in perception. In his perception tests with native speakers, he manipulated all of the three acoustic components belonging to tone accents and came to the result that "native listeners identified the opposite tone accent (Tone Accent 2 in place of Tone Accent 1 and vice versa) significantly often when F0 was manipulated, but almost never when length [...] or intensity [...] were" (Werth 2012: 190). Thus, length can be considered a redundant feature, whereas the lexical tone represents the relevant information for the identification process.

Another important result of the identification task was that listeners always responded to the F0-movement at the end of the word. For this reason, it seems sensible to associate the lexical tone to the second mora. Furthermore, this lexical tone is always associated with a long segment with an underlying bimoraic structure; in monosyllabic words with a VCC-structure (condition a.: short vowel + long sonorant consonant), the second mora and therefore the lexical high tone (H_{lex}) is located on the consonant. Thus, a tone accent cannot be identified until the sonorant has been perceived. In monosyllabic words with a VVC-structure (long vowel + short sonorant consonant), the second mora is already located on the vowel and can be identified as soon as the vowel has been perceived.

Thus, tone accent speakers use a combination of several cues (vowel quantity, tone-text-association and length of the sonorant consonant) for lexical access; however, according to the results of Werth (2011), length appears to be redundant while tone is relevant. However, there are minimal pairs with contrasts in vowel quantity in the Moselle-Franconian dialects, as well (e.g., [ʃa̱lˑ^{2}] 'acoustic noise' – [ʃa̱ː^{2}l] 'stale'). Thus, the question remains whether tonal cues are redundant and less relevant in online word recognition than vowel quantity in these word pairs.

In sum, the phonetic and phonological conditions of monosyllabic words contrasting in vowel quantity, tone-text-association and lateral length and the relevance of these cues for the perception of Moselle-Franconian dialect speakers can be represented as follows:

Tab. 1: Phonological and perceptual representation of monosyllabic words with inverted length contrasts combined with Tone Accent 2 in the Moselle-Franconian dialect of Mayen

Condition	**a.**			**b.**		
Phonology	μ | V | short	+	μ | CC | long^{Hlex}	μ | VV | long^{Hlex}	+	μ | C | short
Perception	ø		2 cues (H_{lex}+ length)	H_{lex}	+	length

1.3 The interaction of tone and length in the Standard German variety

As tone accents do not occur in the Standard German variety, the interaction of tone and length information must be different from that of the Moselle-Franconian. Tonal information is not used to differentiate lexical meanings, but is processed in another way. Monosyllabic stimuli realised with Tone Accent 2 by a Moselle-Franconian speaker (see Table 2, condition c. and d.) can be processed by the Standard German listeners, as well. But without dialectal competence they are only able to perceive a simple high tone (H*) without further lexical meaning. In addition to the variation of loudness in Standard German, changes in the temporal structure as well as pitch changes are used to realise stress (cf. Stock and Zacharias 1982; Jessen et al. 1995; Dogil 1999; Pompino-Marschall 2003 etc.). Therefore, there may be a certain sensitivity towards the perception of these prosodic parameters. This means that the length of vowels and consonants as well as the high tone are perceived, but only vowel quantity is used distinctively for lexical retrieval with the beginning of the lateral.

Vowel length is a phonological cue in the Standard German variety with low functional load. Unlike the Moselle-Franconian dialects or real quantity languages such as the Finno-Ugric languages Finnish, Hungarian or Estonian, most quantity contrasts in Standard German are coupled with differences in vowel quality, e.g., /oː/ – /ɔ/ or /iː/ – /ɪ/. The only exceptions are the vowel pairs /ɛ/ –

/ɛː/ and /a̠/ – /a̠ː/, which differ only in the feature [± long] (cf. Fuhrhop and Peters 2013). Long consonants with sense-discriminative function like the Finno-Ugric languages or long sonorant consonants due to lexical tones like the Moselle-Franconian dialects do not exist in the Standard German variety.

In sum, it can be stated that there is a difference in the phonological prosodic systems between Moselle-Franconian dialect speakers and speakers of the Standard German variety concerning the relevance of tone accents and of contrasts in length for lexical access. Vowel quantity has a low functional load in the Standard German variety as well, but there is no phonological relevance of tonal information and consonantal length.

The present study will examine the question how important the various cues are in the neural processing of the two mentioned participant groups. According to the results of the perception study by Werth (2011), it can be deduced that there must be a hierarchy for the individual cues in processing, i.e., primarily and secondarily relevant cues. Commonly occurring cues like the "duration ratio" (vowel + closure), formant transition and the length of voicing are used (with varying relevance) for the identification of fortis and lenis in the context of nasal plosion in the Standard German variety (cf. Kohler 1979). Gussenhoven (2004) describes this phenomenon as contrast enhancement that is defined for phonology in Werth (2012: 196) as "the fact that linguistic function (in a broader sense) is encoded in several different formal dimensions". In our study, we expect visible differences in the hierarchy of the various cues because of their differing phonological relevance.

Thus, the phonetic and phonological conditions of monosyllabic words contrasting in vowel quantity, tone-text-association and lateral length as well as the relevance of these cues for the perception of Standard German speakers can be represented as follows:

Tab. 2: Phonological and perceptual representation of monosyllabic words with inverted length contrasts combined with a high tone in non-dialectal listening competence

Condition	**c.**			**d.**		
Phonology	μ		μ	μ		μ
	V		CC	VV		C
	short	+	long^{H*}	long^{H*}	+	short
Perception	ø		2 cues (H*+ length)	H*	+	length

2 Methods

To test our hypotheses, we conducted an event-related potential (ERP) study with a categorical oddball paradigm using the electroencephalograpy (EEG) technique. This design was chosen in order to examine the so-called mismatch negativity (MMN), a fronto-central negative component, usually peaking at 150 to 250 ms from change onset. This component is elicited when an infrequent deviation ('deviant') occurs among frequently repeated sound patterns ('standard'). The repetitive presentation of standards creates a short-term memory trace in the auditory cortex for this pattern. If this pattern is violated by an infrequent deviant, an MMN is elicited, as it reflects an automatic, pre-attentive response to any change in auditory stimulation, regardless of the participants' attention. It thus indicates that this stimulus deviates from the memory representation of the preceding series of standards (cf. Näätänen et al. 2007).

2.1 Stimuli

Two monosyllabic words from the Moselle-Franconian tone accent dialect spoken in Mayen were selected for this study: [ʃa̠ː²l] 'stale' – [ʃa̠l·²] 'acoustic noise'.[5] These words differ in vowel quantity in one phoneme only ([a̠ː] vs. [a̠]), but they have an additional difference in lateral length ([l] vs. [l·], see Figure 1), which is caused by the tone-mora-association. Furthermore, there is a difference in the tone-text-association which means that the lexical high tone of Tone Accent 2 is associated with the lateral in [ʃa̠l·²] while in [ʃa̠ː²l] the lexical high tone is already associated with the vowel (cf. Schmidt 1986; Werth 2011 and Figure 1). Together with the minimal pair [ʃa̠l¹] '(apple) skin' and [ʃa̠·¹l] 'shell, shawl', a total of four lexemes can be distinguished by length differences in combination with tone accent differences (cf. Schmidt 1986). Notice that in the present study, these words with Tone Accent 1 were not tested, only the two lexemes with Tone Accent 2.

5 The tested word pair was part of an experiment in which two other minimal pairs with other phonological contrasts (Tone Accent 1 vs. Tone Accent 2 and vowel quality /o/ vs. /u/) were also examined (cf. Schmidt 2017).

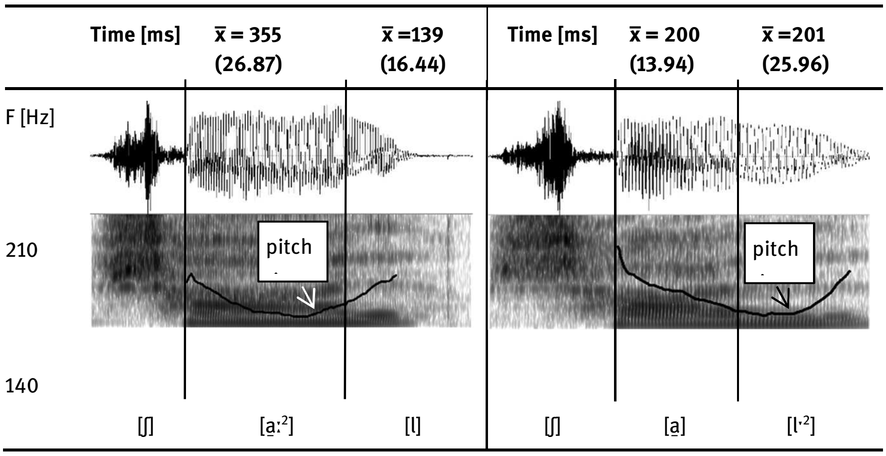


Fig. 1: Average vowel and lateral segment durations and pitch curves of Accent 2 in [ʃa̠ː²l] and [ʃa̠l·²]

The critical word pairs were recorded several times to obtain different tokens with a natural variation for each word. Finally, eight different natural tokens per word served as frequently presented stimuli (= standard[6]) and as infrequently presented stimuli that are incongruent with the memory representation of the preceding stimuli, cf. Näätänen et al. (2007) (= deviant). This acoustic variability was chosen to create a more natural speech perception condition as well as a memory trace for the presented standard condition, since it has been shown that a higher and hence more natural variability in standard items leads to a more reliable abstraction or trace form of the different acoustic stimuli presented (Phillips et al. 2000). Moreover, the standard and the deviant representations of one critical item were phonetically identical, i.e., the standards as well as the deviants presented comprised the same eight different tokens. As a result, purely acoustic effects between standards and deviants could be distinguished from phonetic and phonological effects (cf. Scharinger, Monahan, and Idsardi 2011).

All stimuli were digitally recorded with a sampling rate of 44.1 kHz and a 16 bit (mono) sample size, using an electret microphone (Beyerdynamic MC 930) and the software Adobe Audition 2.0. The stimuli were spoken naturally by a female native speaker of the Moselle-Franconian dialect from Mayen at a normal speech rate.

6 If we refer to the Standard German variety in the present study, we always use the term Standard German.

For each word, eight acoustically variant tokens were selected. Table 3 depicts mean values of the parameters pitch, and frequencies of the first two vowel formants for tokens of the two word types. Statistical analyses utilizing Mann-Whitney-U tests revealed that the distribution of pitch values for the two word types differed significantly ($U = 64$, $p < .001$). Despite these statistical differences, it is questioned here that the acoustic differences exceed the perceptional threshold. According to Nooteboom (1997: 645) pitch differences between stimuli can be reliably discriminated, if the pitch difference exceeds a difference of three semitones (one semitone corresponds to a frequency difference of approximately 6%). Consequently, a perceptible difference should exhibit a difference of roughly 25 Hz. The pitch difference observed in our stimulus material is clearly smaller.

Tab. 3: Acoustic properties of conditions with mean values (and standard deviation)

Stimulus	Pitch [Hz]	1. Vowel Formant	2. Vowel Formant
[ʃa̠ː²l]	158 (5.66)	710 (79.55)	1352 (79.72)
[ʃa̠lˑ²]	169 (8.57)	753 (28.59)	1365 (80.46)

Finally, all items were controlled for and normalised in intensity to ~ 80 dB SPL (Table 3). This adjustment was carried out using the sound recording and analysis software PRAAT (version 5.3.08, Boersma and Weenink 2012).

2.2 Procedure

The experiment was carried out twice: first with participants from Mayen raised in the Moselle-Franconian dialect area and in a second run with Standard German speakers in Marburg. In both group sessions, exactly the same conditions and the identical set of stimuli were used. All stimuli of the minimal pair [ʃa̠ː²l] – [ʃa̠lˑ²] were presented in two experimental blocks. In one block, stimuli with a long vowel served as standard and stimuli with a short vowel as deviant. The opposite was the case in another block in which a short vowel served as standard and a long vowel as deviant. In total, together with two further conditions not reported here,[7] six blocks consisting of 1000 items each (15% deviants) were presented with each requiring approximately 25 minutes.

In order to avoid sequence effects, the block order was varied across participants. Moreover, two blocks containing the same lexical material were never presented directly one after the other. Each block started with ten standards which were not included in data analysis. Next, the standards and deviants were presented in a classic passive oddball paradigm, i.e., in a pseudo-randomised order in which a deviant was presented after two up to eight standards. The inter-stimulus interval (offset-to-onset) was 900 ms. Stimuli were presented via two loudspeakers at a comfortable listening level (~ 65 dB SPL).

During the experiment, participants were comfortably seated in front of a computer screen in a dimly lit and quiet room. They were instructed to watch a silent movie and to disregard the auditory presentation. After the first block, all participants reported that they were able to ignore the auditory signal and to concentrate entirely on the movie presented. Between blocks, participants were offered a break to rest their eyes. All procedures were performed in compliance with relevant laws and institutional guidelines.

2.3 Participants

25 native speakers of the Moselle-Franconian dialect from Mayen (16 women; mean age 50.0,[8] age range 31 to 62) and 21 speakers of Standard German not born or raised in the Moselle-Franconian dialect area (16 women, mean age 52.8, age range 42 to 61) participated in the two distinct experimental sessions. Both groups were closely matched in their mean age. All subjects were right-handed, monolingual and had normal or corrected-to-normal vision. All participants with a dialectal background were born and raised in Mayen and still live there. Their dialect competence was tested and verified via a dialect pre-test. None of the participants had hearing deficits, which was verified by an online-hearing-test.[9]

7 Two other minimal pairs ([ʃtuf] ‘living room’ – [ʃtof] ‘textile fabric’ and [d̥a̠͡ʊ¹f] ‘pigeon’ – [d̥a̠͡ʊ²f] ‘baptism’) were tested in the same experiment beside the pair with contrast in vowel quantity presented in this study (cf. Schmidt 2017).

8 The mean age range for ERP studies is typically between 18 and 35 years. The relatively high mean age in both participant groups in the present study is due to the factor of dialect competence in the dialect group. Dialect competence is generally more stable in older peer groups so that the age range had to be extended for the purposes of the present study.

9 URL: http://www.powerone-batteries.com/de/wissen/hoertest/power-one-hoertest/?no_cache=1 (accessed 20 May 2016).

All participants gave their informed consent to this study and privacy rights were thoroughly obeyed. Each participant received monetary compensation for taking part in the study.

2.4 ERP recording and data processing

An electroencephalogram (EEG) was recorded from 26 Ag/AgCl electrodes, mounted on an elastic cap (EasyCap), according to the 10-20 system (F7, F3, Fz, F4, F8, FC5, FC1, FCz, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CPz, CP2, CP6, P7, P3, Pz, P4, P8, POz) with a “BrainVision” (Brain Products GmbH) amplifier. The C2 electrode served as the ground electrode. The reference electrode was placed at the tip of the nose. Four electrodes measured the **electrooculogram (EOG)**, i.e., horizontal and vertical eye movements to control for eye movements and blinks. Two electrodes were placed at the left and right mastoid. EEG and EOG were recorded continuously with a sampling rate of 500 Hz and filtered offline with a 0.16 to 30 Hz bandpass filter. All electrode impedances were kept below 5 kΩ. EEG recordings were re-referenced off-line to the linked mastoids to decrease the signal-to-noise ratio and hence to increase the MMN amplitude (cf. Schröger 1998). Averaged data were baseline corrected over 100 ms before vowel onset.

For the data analysis, all standard and deviant epochs starting at a baseline of 100 ms before the divergence point up to 900 ms after the vowel onset were automatically scanned for artifacts produced by eye or body movements. All epochs that included artifacts with an amplitude exceeding 75 microvolt were removed from the data set. Subsequently, all single-trial waveforms were individually screened for further artifacts. As a result of these observations, the data sets gathered from ten participants (eight women) in the dialect group and six participants (six women) in the Standard German group had to be excluded from the analysis because of the high number of artifacts (mainly eye blinks) per condition epoch (more than 50%). Thus, the data sets of 15 participants[10] were analysed per group (Moselle-Franconian dialect group: eight women, mean age 50.7 years; Standard German group: ten women, mean age 53.1 years). ERP responses to the first ten standard stimuli of each experimental block as well as to standard epochs immediately preceded by a deviant were excluded from data analysis.

10 Due to the given high mean age and its known influence on ERP components it would have been interesting to test two age groups per participant group to control for this factor. However, due to the small groups, a division into two age groups within each group of participants is not possible. We therefore refer to research on age effects on ERP patterns in the following discussion.

2.5 Data analyses

For the statistical analysis, an omnibus multifactorial repeated measures ANOVA was calculated with the factor GROUP (Moselle-Franconian dialect participants vs. Standard German participants), REGION (frontal [F3, Fz, F4], central [C3, Cz, C4], parietal [P3, Pz, P4]) and VOWEL QUANTITY (long vowel vs. short vowel). Averages were calculated from the onset of the divergence point up to 900 ms thereafter, with a baseline of 100 ms. For the statistical analysis, consecutive epochs of 50 ms were investigated in both groups in the classical MMN time window (100 to 200 ms).

Furthermore, additional time windows were calculated following visual inspection of the grand-average curves within each group. For these additional time windows, an omnibus ANOVA with the factors described above as well as a multifactorial ANOVA within each group was conducted. For the latter analyses, time windows were adjusted on the basis of visual inspection of the grand average curves. This was necessary as the effects' latency differed between the two groups. This was expected because of the varying dialect background between the two groups. For effects with more than one degree of freedom, Huynh-Feldt (1976) corrections were applied and corrected p-values are reported here.

3 Results

Data analyses aim at finding mean voltage differences between certain experimental conditions that manifest differently in the two experimental groups. In particular, the oddball paradigm leads to the expectation to find a significant negative mismatch response for deviants indexing pre-attentive perception of duration contrasts. Table 2 shows that negativity effects were found in both participant groups. Note that all presented results are derived from comparisons across oddball blocks, i.e., deviant related effects were examined by comparing phonetically identical stimuli. For example, [ʃa̠ː²l], which was presented as deviant in one block, was compared with [ʃa̠ː²l], which was presented as standard in another block. This procedure ensured that potential effects between standard and deviant which are purely elicited by acoustic differences could be excluded (cf. Eulitz and Lahiri 2004; Scharinger, Monahan, and Idsardi 2011).

For the comparison of amplitudes of the standard condition [ʃa̠ː²l] and the deviant condition [ʃa̠ː²l] in the time window between 100 and 200 ms, the omnibus ANOVA showed significant main effects for all three factors GROUP [$F(1, 28) = 4.81$, $p = .037$, $\eta^2p = .11$], REGION [$F(2, 56) = 6.53$, $p = .010$, $\eta^2p = .02$] and VOWEL QUANTITY

[$F(1, 28) = 4.94$, $p = .034$, $\eta^2 p = .02$]. There was no significant three way interaction, but a significant interaction between the factors GROUP and REGION [$F(2, 56) = 8.80$, $p = .003$, $\eta^2 p = .03$]. This significant interaction is in line with the expectation to find a more pronounced negativity in the frontal region as the MMN is typically distributed frontally. In fact, the post-hoc analysis of this interaction by REGION only revealed a significant result in the frontal region [frontal: factor GROUP: $F(1, 28) = 8.28$, $p < .05$, $\eta^2 p = .20$; factor VOWEL QUANTITY: $F(1, 28) = 5.71$, $p < .05$, $\eta^2 p = .03$].

In order to also resolve this significant interaction by GROUP and because of the significant main effects for all three factors, a multifactorial repeated measure ANOVA was also conducted for this time window within each group. This should have revealed possible between-group differences in the early MMN time window. However, these within-group calculations showed no significant main effects for the main factor VOWEL QUANTITY.

The comparison of the short vowel contrast between the standard condition [ʃa̠l·^{2}] and the deviant condition [ʃa̠l·^{2}] in the time window from 100 to 200 ms only elicited a main effect for REGION [$F(2, 56) = 29.12$, $p = .000$, $\eta^2 p = .01$]. However, in contrast to the typical early MMN time window, significant mismatch effects were found in later time windows for both contrast pairs. The comparison of the long vowel contrast between the standard condition [ʃa̠ː^{2}l] and the deviant condition [ʃa̠ː^{2}l] in a time window from 300 to 450 ms revealed significant main effects for the factors GROUP [$F(1, 28) = 9.13$, $p = .005$, $\eta^2 p = .19$], REGION [$F(2, 56) = 14.91$, $p = .000$, $\eta^2 p = .07$] and VOWEL QUANTITY [$F(1, 28) = 11.32$, $p = .002$, $\eta^2 p = .05$].

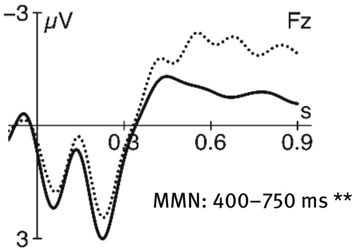
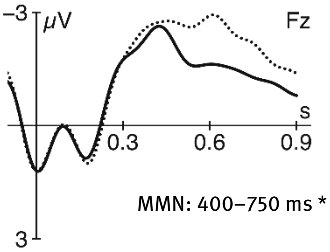
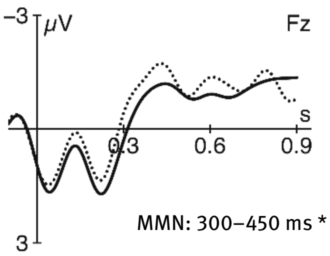
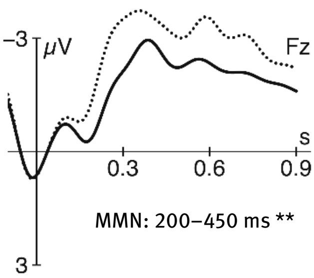
There was no significant three way interaction, but a significant interaction between the factors GROUP and REGION [$F(2, 56) = 8.06$, $p = .006$, $\eta^2 p = .04$]. The post-hoc analysis of this interaction by REGION revealed a more pronounced effect in the fronto-central region [frontal: factor GROUP: $F(1, 28) = 10.11$, $p < .01$, $\eta^2 p = .24$; factor VOWEL QUANTITY: $F(1, 28) = 9.53$, $p < .01$, $\eta^2 p = .05$; central: factor GROUP: $F(1, 28) = 8.78$, $p < .01$, $\eta^2 p = .21$; factor VOWEL QUANTITY: $F(1, 28) = 10.29$, $p < .01$, $\eta^2 p = .05$]. A post-hoc within-group comparison revealed a significant latency difference for the negativity effect between the two groups: it starts 100 ms earlier in the Standard German group than in the dialect group.

Due to this latency difference, post-hoc analyses were calculated for each group with the factors VOWEL QUANTITY and REGION in a time window from 300 to 450 ms for the dialect group and from 200 to 450 ms for the Standard German group. These analyses show significant main effects in both groups for the factor VOWEL QUANTITY [dialect group: $F(1, 14) = 6.04$, $p = .028$, $\eta^2 p = .04$ / Standard German group: $F(1, 14) = 9.50$, $p = .008$, $\eta^2 p = .09$] but no significant interaction between the factors VOWEL QUANTITY and REGION. The measurements of the deviant peak amplitude and peak latency also show that the negativity effect is more pronounced and peaks earlier in the Standard German group in comparison to the peak latency and peak amplitude of the deviant in the dialect group (see Table 5 in the appendix for peak latencies and amplitudes of the late MMN at Fz for all experimental conditions).

For the comparison of the short vowel contrast between the standard condition [ʃa̠lˑ^2] and the deviant condition [ʃa̠lˑ^2], similar results were found. For the omnibus ANOVA, a time window between 400 and 750 ms was investigated. It revealed significant main effects for the factors REGION [$F(2, 56) = 20.35$, $p = .000$, $\eta^2p = .08$] and VOWEL QUANTITY [$F(1, 28) = 22.58$, $p = .000$, $\eta^2p = .10$] but no significant interactions. However, hypothesis guided post-hoc analyses within each group were calculated from 400 to 750 ms for both groups. For the dialect group, a significant main effect was found for the factor VOWEL QUANTITY [$F(1, 14) = 14.24$, $p = .002$, $\eta^2p = .13$] without a significant interaction between the two factors VOWEL QUANTITY and REGION. Within the Standard German group, significant main effects were elicited for the factors REGION [$F(2, 28) = 21.14$, $p = .000$, $\eta^2p = .13$] and VOWEL QUANTITY [$F(1, 14) = 8.72$, $p = .010$, $\eta^2p = .08$], but again no significant interaction between these factors was found. For this contrast in vowel quantity, no significant latency differences were found between the two participant groups.

Table 4 displays an overview of the relevant time windows and the statistically significant results in both groups of participants for the critical word pair [ʃa̠ː^{2}l] – [ʃa̠lˑ^2].

Tab. 4: Grand average ERP responses for deviants (dotted lines) and standards (solid lines) at Fz of the two experimental conditions [ʃa̠l·²] (a.+c.) and [ʃa̠ː²l] (b.+d.) in different time windows for both groups of participants, measured from 100 ms prior the vowel onset up to 900 ms. Statistical significance is indicated by * (p < .05), ** (p < .01), *** (p < .001).

	Dialect speakers	**Standard German speakers**
Condition	a.	b.
Phonology	μ μ V CC short + long^Hlex	μ μ V CC short + long^H*
Perception	ø 2 cues (H_{lex}+ length)	ø 2 cues (H*+ length)
EEG-Responses Legend: ······· deviant —— standard		
Condition	c.	d.
Phonology	μ μ VV C long^Hlex + short	μ μ VV C long^H* + short
Perception	H_{lex} + length	H* + length
EEG-Responses Legend: ······· deviant —— standard		

4 Discussion

4.1 ERP component discussion

While differences in tone-text-association have not yet been in the focus of neurolinguistic research interest, length differences in pre-attentive processing regarding isolated vowels, pseudowords and tones have already been well examined (cf. Kirmse et al. 2008; Ylinen et al. 2006; Nenonen et al. 2003; Jaramillo, Alku, and Paavilainen 1999; Jaramillo et al. 2001).

In the present study, we used real words contrasting in vowel quantity and found comparable negativities, albeit with differences in terms of their respective latency. The component that was observed in the former studies was classified as an MMN (= Mismatch Negativity). It usually occurs if a mismatch between a memory trace of the currently valid auditory representation and an appearing stimulus takes place (cf. Schröger, SanMiguel, and Bendixen 2013). The latency for this component is usually reported between 100 to 250 ms after the beginning of an alteration (cf. Schröger, SanMiguel, and Bendixen 2013). If we look at our results, the latency of the detected negative-going deflection occurs at a later time window in most cases. Nevertheless, it seems unlikely that this component could be interpreted differently from previous similar studies. Although the latency is similar to the N400 component, the negativity effect found is not distributed largest over centro-parietal sites, which is usually described for the N400 (cf. Lau, Phillips, and Poeppel 2008; Kutas and Federmeier 2011), but has the fronto-central topographic distribution of the MMN.

Furthermore, the experimental design containing an oddball paradigm without an active task leads us to interpret the observed component as an expression of pre-attentive processing. From other studies, it is already known that the MMN may sometimes appear in a later time window, if complex auditory or even linguistic stimuli were used (cf. Korpilahti, Lang, and Aaltonen 1995; Korpilahti et al. 2001; Cheour et al. 2001), as is the case in the present study. Korpilahti et al. (2001) associate this late MMN (lMMN) with the automatic detection of lexical differences. In our study, we investigated phonological quantity differences using the words [ʃa̠ː²l] ‘stale’ and [ʃa̠l·²] ‘acoustic noise’, which also result in lexical differences. Thus, this could be an explanation for the difference in latency compared with length contrasts in isolated vowels, pseudowords or tones.[11]

11 Furthermore, it must be taken into consideration that the two groups of participants in our study are older than the usually younger subjects in electrophysiological experiments (mean age in our study: dialect group = 50.0, Standard German group = 52.8). Age effects are also found in

As an initial result of our study, we conclude that phonological quantity contrasts embedded in real words lead to a similar negative component in pre-attentive processing for dialect as well as for non-dialect speakers, similar to length contrasts without lexical access, but with a delay in latency due to the higher complexity of the stimuli.

4.2 Participant group 1 (Moselle-Franconian dialect speakers)

As is known from other studies, MMN effects are highly dependent on the direction of stimulus presentation in their amplitude and latency (cf. Jaramillo, Alku, and Paavilainen 1999; Takegata et al. 2008; Kirmse et al. 2008; Colin et al. 2009). In the present study we find differences as well: a decreased latency of the MMN component (a. vs. b.) and an earlier peak in the deviant condition [ʃa̠ː²l] (see Table 4, cf. Table 5 in Appendix). A closer look at the structure of the stimuli and the position of the high tone may deliver an explanation. While the words in both conditions [ʃa̠ː²l] and [ʃa̠l·²] bear Accent 2, the tone-text-association is different. Accent 2 bears the high tone on the second mora, thus the tone accent is associated with the vowel in [ʃa̠ː²l] as described in Section 2.1. In [ʃa̠l·²], the vowel bears no tonal information because the second mora is associated with the lateral. Consequently, the final lateral bears the tonal information. Thus, listeners are already able to perceive the difference through the tone-text-association when they hear the long vowel in the condition [ʃa̠ː²l]. Since the vowel quantity can be identified only after the offset of the vowel (cf. Czigler and Winkler 1996; Levänen et al. 1993), the processing of this cue can only start at the beginning of the lateral.

In the condition [ʃa̠l·²], the determination of **both** cues can only be accomplished by the perception of the lateral. Because the vowel has to be completed for vowel quantification and the tonal information is tied to the lateral, the listener has to wait for the lateral to gain lexical access. Thus, the vowel itself carries no information for the condition [ʃa̠l·²] that could be perceived as different from the standard stimuli. In the condition [ʃa̠ː²l], a negative component with an earlier latency seems to be triggered through the association of the high tone with the vowel. The results found suggest that lexical retrieval already starts with the perception of the high tone when this information is available before the quantity

pre-attentive processing with regard to an ERP's latency and distribution. This is especially reflected in smaller amplitudes (cf. Verleger et al. 1991; Pekkonen et al. 1996; Bertoli, Smurzynski, and Probst 2002; Cooper et al. 2006; Schiff et al. 2008; Rimmele et al. 2012), but also in a later appearance of the MMN in older participant groups (cf. Verleger et al. 1991; Bertoli, Smurzynski, and Probst 2002; Cooper et al. 2006).

cue. This suggests that vowel quantity is only a minor cue for lexical access in the Moselle-Franconian dialect.

Furthermore, the higher amplitude in the condition a. [ʃa̲lˑ^2] (see Table 4, cf. Table 5 in Appendix) indicates that the short vowel evokes an increased neural activity in pre-attentive processing. This result is in line with intrinsic stimulus effects presented in studies by Jaramillo, Alku, and Paavilainen (1999), Colin et al. (2009) and Sittiprapaporn (2012). Based on synthetically generated sound patterns, Colin et al. (2009) found a major effect in MMN responses in deviant sound patterns with shorter durations which evoked higher amplitudes than those of longer duration. These results suggest a general effect of vowel length on speech processing. The length effect on MMN amplitudes can be explained by the delay between the moment of deviance detection and the end of the deviant quantification process. In short deviants, deviance detection and quantification take place simultaneously, giving rise to the MMN amplitude. By contrast, in long deviants, deviance detection and quantification occur separately one after another. However, the effects observed here cannot be explained solely by intrinsic effects since the amplitude effects for the quantity contrast are highest for the tested quantity contrast in comparison to other contrasts (vowel quality and tone contrasts) tested in the same experiment. While the tested contrasts in tone and vowel quality were single deviants, the examined word pair in the present study is characterised through three differing features (tone-text-association, vowel quantity and lateral length). This difference might suggest that all these features are processed additively in the condition [ʃa̲lˑ^2], in which all occur during a short period of time, more precisely during the lateral perception. While vowel quantity could only be identified with the beginning of the lateral and lateral length with the end of this sonorant consonant, even the third cue, the high tone, is associated with the lateral. Thus, all cues are perceived during the 201 ms (mean value) of the lateral, which leads to an additive processing.

According to Ylinen, Huotilainen, and Näätänen (2005), additive responses indicate an independent analysis of the different features. The authors examined in their pseudoword MMN study the processing of stimuli that differ in terms of quality ([ip:i]), quantity ([iti]) or a combination of both ([ipi]) from a standard ([it:i]). By comparing the MMN results for single and combined deviant forms, they found that both features are processed additively and therefore the analysis of these features is carried out independently.

In the present analysis, the nature of the stimuli is slightly different; nevertheless, in the condition [ʃa̲ː^{2}l], deviating features occur one after another in an average time slot of 316.5 ms (beginning of the high tone at the middle of the vowel $\bar{x}$ = 177.5 ms (see Figure 1), vowel quantity with the start of the lateral and lateral length with the end of the lateral $\bar{x}$ = 139 ms), while they are perceptible within one segment ([l·²]) in the condition [ʃa̱l·²]. The higher amplitude in condition [ʃa̱l·²] could thus be an indication that not only phoneme quality and quantity, as described by Ylinen, Huotilainen, and Näätänen (2005) is processed independently, but also the two suprasegmental features of quantity and tone accents, if they both occur at the same time or in a very short period of time.

4.3 Participant group 2 (Standard German speakers)

As mentioned before, the main difference between the two participant groups of dialect and Standard German speakers is the phonological relevance of the tone accent features like the tone-text-association. However, there are some similarities in processing the investigated word pair with regard to latency and amplitude differences between the MMNs for both deviants.

Both deviants elicited significant late MMN effects in the Standard German speaker group, as well. Moreover, an earlier latency in the deviant condition [ʃa̱ː²l] contrasting with [ʃa̱l·²] (a. + c. vs. b. + d., see Table 4) can be observed in both participant groups. Thus, although high tones are not associated with lexical relevance in Standard German, Standard German speakers perceive this tonal information at different temporal points in both presented words. We therefore conclude that the high tone, regardless of whether it is used for lexical access (dialect group) or only as a salient marker (Standard German group), seems highly relevant in pre-attentive processing of deviance.

Furthermore, the higher amplitude in the condition c. [ʃa̱l·²] vs. d. [ʃa̱ː²l] (see Table 4, cf. Table 5 in Appendix) indicates that the short vowel with Accent 2 associated with the lateral evokes an increased neural activity in pre-attentive processing. Even without phonological relevance of the tone-text-association, additive processing of the two suprasegmental features quantity and tone accent takes place nearly at the same time. It seems like Standard German speakers perceive the high tone as a salient cue, and add this information to vowel quantity resulting in increased processing costs in comparison to the condition where these two cues are processed successively (condition d. [ʃa̱ː²l], see Table 4, cf. Table 5 in Appendix). Thus, we conclude that tonal information is processed independently from quantity regardless of phonological function.

Although the negativity effects occur in both groups, there are some relevant differences between the groups regarding the latency of the effects (cf. b. vs. d., see Table 4). For deviants with a long vowel, the effect starts 100 ms later in the dialect group. We assume that one reason for this difference in latency is the tonal information of Accent 2, as it represents an extra processing effort for the participants who speak the Moselle-Franconian dialect. Accent 2, characterised by a final rising pitch or the presence of a high tone on the second mora (see Introduction), plays an important role as a cue for lexical access during vowel perception, but only for speakers of the Moselle-Franconian dialect. While in Standard German no lexical meaning is carried by tone accents, Moselle-Franconian dialect speakers use this information to disambiguate lexemes and, according to our results, to distinguish vowel quantity, if the high tone precedes the vowel offset. With the high tone on the second mora, Moselle-Franconian dialect speakers may predict that the similar lexemes [ʃa̠ˑ^{1}l] 'shell, shawl' and [ʃa̠l^{1}] '(apple) skin', which lack this high tone on the second mora due to Accent 1 (cf. Werth 2011), can be excluded from the retrieval process. Since the second mora is associated with the vowel, the lexeme [ʃa̠lˑ^{2}] 'acoustic noise' can also be excluded because here the high tone is located on the lateral. Thus, the tone accent initiates lexical access while the length of the vowel ensures the identification for the group of dialect speakers. Since only the group of dialect speakers are able to use the tonal information for lexical processing and word recognition, their cognitive effort is even higher compared to Standard German speakers who have to process the quantity cues only for lexical retrieval. For Standard German speakers, the recognition of the high tone results only in the perception of differing phonetic information or the marking of salience or word stress without any further distinction. Thus, the high tone is an important lexical cue for the dialect group only. This higher complexity in processing might delay the negativity onset compared to the Standard German group.

Moreover, due to the later onset of the negativity effect, its latency is generally shorter than the effect elicited in Standard German speakers. Shorter MMN latencies have often been associated with higher discrimination sensitivity (cf. Kirmse et al. 2008; Partanen et al. 2011). Since tonal cues are only lexically relevant for the dialect speakers and the particular phonetic form of the stimuli presented is only known by the Moselle Franconian dialect speakers and not by the Standard German speakers, these advantages might explain the higher discrimination sensitivity of dialect speakers. Because they are used to discriminate tone accent contrasts in their own dialect, they acquire a more pronounced discrimination sensitivity and accuracy. In contrast, the group of Standard German speakers might process the given tonal cue as word stress information, leading to prolonged processing.[12] This process seems more time-consuming because a more complex analysis for the whole word is needed (cf. Kirmse et al. 2008).

Our results thus demonstrate that the quantity contrast combined with tonal information elicits a late MMN in both groups of participants. However, significant latency differences of this component between the two groups reflect their difference in processing a tone accent cue which contains phonologically relevant information only for the participants with competence in the Moselle-Franconian dialect, but not for the Standard German speakers.

5 Conclusion

In the study presented here, we investigated the importance of high tone information associated with different sound segments during the processing of real words with a phonological contrast in vowel quantity (Moselle-Franconian Accent 2 in words with long and short vowels). Our results show a late MMN effect for both conditions (long vs. short vowel) and in both participant groups (Standard German and dialect speakers). Latency differences reflect clear differences between the two groups and their differing phonological systems. Since the high tone carries no lexical meaning for the Standard German speakers, it can be processed as a simple salient intonation contour or a stress cue. In contrast, the dialect speakers use the same high tone as an important lexical cue. The differing processing of tonal cues is reflected in a shorter MMN latency for the long vowel in the dialect group, illustrating higher sensitivity and accuracy in word discrimination due to the possibility to use the high tone in combination with the quantity cue for lexical retrieval.

Thus, we can show that a contrast which is phonologically relevant in two different varieties elicits similar electrophysiological effects in form of an MMN. However, this signature is further modulated by cues that are only relevant for one of the systems due to the regional linguistic background.

Acknowledgements: The research presented here was supported by the Hessian State Ministry of Higher Education, Research and the Arts, Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Excellenz (LOEWE), *Research Focus:*

12 Because of the overlong vowels and salient pitch movements due to Tone Accent 2, monosyllables might be perceived as bisyllabic words by non-native speakers, which entails an evaluation of the stressed syllable in the perceived word.

Exploring fundamental linguistic categories grant, research project *Phonological word – Constituents of the Phonological Word* granted to Joachim Herrgen, Jürgen Erich Schmidt, and Alexander Werth. The authors thank mayor Wolfgang Treis and the employees of the municipal administration in Mayen for their generous support in the field and Lydia Riedl and Corinna Schwarz for their help in preparing the stimuli.

References

Amenedo, Elena & Carles Escera. 2000. The accuracy of sound duration representation in the human brain determines the accuracy of behavioral perception. European Journal of Neuroscience 12. 2570–2574.

Bertoli, Sibylle, Jacek Smurzynski & Rudolf Probst. 2002. Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task. Clinical Neurophysiology 113. 396–406.

Boersma, Paul & David Weenink. 2012. PRAAT: Doing Phonetics by Computer (ver. 5.3.08) [Computer program]. Amsterdam: Institut for Phonetic Sciences.

Chandrasekaran, Bharath, Ananthanarayan Krishnan & Jackson T. Gandour. 2007. Mismatch negativity to pitch contours is influenced by language experience. Brain Research 1128. 148–156.

Chandrasekaran, Bharath, Ananthanarayan Krishnan & Jackson T. Gandour. 2009. Sensory Processing of Linguistic Pitch as Reflected by the Mismatch Negativity. Ear & Hearing 30(5). 552–558.

Cheour, Marie, Pirjo Korpilahti, Olga Martynova & A. Heikki Lang. 2001. Mismatch Negativity and Late Discriminative Negativity in Investigating Speech Perception and Learning in Children and Infants. Audiology & Neuro-Otology 6. 2–11.

Chládková, Kateřina, Paola Escudero & Silvia C. Lipski. 2013. Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. Brain and Language 126. 243–252.

Christmann, Corinna, Stefan Berti, Claudia Steinbrink & Thomas Lachmann. 2014. Differences in sensory processing of German vowels and physically matched non-speech sounds as revealed by the mismatch negativity (MMN) of the human event-related brain potential (ERP). Brain and Language 136. 8–18.

Clements, George N. & Samuel Jay Keyser. 1983. CV phonology: A Generative Theory of the Syllable. Cambridge: MIT Press.

Colin, Cécile, Ingrid Hoonhorst, Emily Markessis, Monique Radeau, Marianne de Tourtchaninoff, Aurélie Foucher, Gregory Collet & Paul Deltenre. 2009. Mismatch Negativity (MMN) evoked by sound duration contrasts: An unexpected major effect of deviance direction on amplitudes. Clinical Neurophysiology 120, 51–59.

Cooper, Rowena J., Juanita Todd, Katherine McGill & Patricia T. Michie. 2006. Auditory sensory memory and the aging brain: A mismatch negativity study. Neurobiology of Aging 27. 752–762.

Czigler, István & István Winkler. 1996. Preattentive auditory change detection relies on unitary sensory memory representations. NeuroReport 7. 2413–2417.

Dogil, Grzegorz. 1999. The phonetic manifestation of word stress in Lithuanian, Polish, German and Spanish. In Harry van der Hulst (ed.), Word prosodic systems in the languages of Europe, 273–311. Berlin & New York: Mouton de Gruyter.

Eulitz, Carsten & Aditi Lahiri. 2004. Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition. Journal of Cognitive Neuroscience 16(4). 577–583.

Féry, Caroline. 2001. Phonologie des Deutschen: Eine optimalitätstheoretische Einführung, Teil II (Linguistics in Potsdam 11). Potsdam: Univ.-Bibliothek, Publ.-Stelle.

Fournier, Rachel, Carlos Gussenhoven, Ole Jensen & Peter Hagoort. 2010. Lateralization of tonal and intonational pitch processing: An MEG study. Brain Research 1328. 79–88.

Fuhrhop, Nanna & Jörg Peters. 2013. Einführung in die Phonologie und Graphematik. Stuttgart: Metzler.

Gandour, Jackson T. 2006. Tone: Neurophonetics. In Keith Brown (ed.), Encyclopedia of language and linguistics, vol. 12, 2nd edn., 751–760. Oxford: Elsevier.

Gandour, Jackson T. 2007. Neural substrates underlying the perception of linguistic prosody. In Carlos Gussenhoven & Tomas Riad (eds.), Tones and Tunes, Volume 2: Experimental Studies in Word and Sentence Prosody (Phonology and Phonetics 12.2), 3–25. Berlin & New York: Mouton de Gruyter.

Gandour, Jackson T., Yunxia Tong, Donald Wong, Thomas Talavage, Mario Dzemidzic, Yisheng Xu, Xiaojian Li & Mark Lowe. 2004. Hemispheric roles in the perception of speech prosody. NeuroImage 23. 344–357.

Gandour, Jack, Donald Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker & Gary D. Hutchins. 2000. A Crosslinguistic PET Study of Tone Perception. Journal of Cognitive Neuroscience 12(1). 207–222.

Gussenhoven, Carlos. 2004. The phonology of tone and intonation (Research surveys in linguistics). Cambridge: Cambridge University Press.

Gussenhoven, Carlos & Jörg Peters. 2004. A tonal analysis of Cologne Schärfung. Phonology 21. 251–285.

Hayes, Bruce. 1999. Inalterability in CV Phonology. In John A. Goldsmith (ed.), Phonological Theory: The Essential Readings. Malden, MA & Oxford: Blackwell.

Heike, Georg. 1962. Suprasegmentale Merkmale der Stadtkölner Mundart. Phonetica 8. 147–165.

Huynh, Huynh & Leonard S. Feldt. 1976. Estimation of the Box Correction for Degrees of Freedom From Sample Data in Randomized Block and Split-Plot Designs. Journal of Educational Statistics 1. 69–82.

Hyman, Larry M. 2003. A Theory of Phonological Weight, 2nd edn. (The David Hume series). Stanford: CSLI Publications.

Jaramillo, Maria, Paavo Alku & Petri Paavilainen. 1999. An event-related potential (ERP) study of duration changes in speech and non-speech sounds. NeuroReport 10. 3301–3305.

Jaramillo, Maria, Titta Ilvonen, Teija Kujala, Paavo Alku, Mari Tervaniemi & Kimmo Alho. 2001. Are different kinds of acoustic features processed differently for speech and non-speech sounds? Cognitive Brain Research 12. 459–466.

Jessen, Michael, Krzysztof Marasek, Katrin Schneider & Kathrin Clahßen. 1995. Acoustic correlates of word stress. Proceedings of the 13th international congress of phonetic sciences 1995 (ICPhS), Stockholm. 428–431.

Kaan, Edith, Christopher M. Barkley, Mingzhen Bao & Ratree Wayland. 2008. Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. BMC Neuroscience 9. 53.

Kirmse, Ursula, Sari Ylinen, Mari Tervanieme, Martti Vainio, Erich Schröger & Thomas Jacobsen. 2008. Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. International Journal of Psychophysiology 67. 131–143.

Kohler, Klaus J. 1979. Dimensions in the perception of fortis and lenis plosives. Phonetica 36. 332–343.

Köhnlein, Björn. 2011. Rule reversal revisited: synchrony and diachrony of tone and prosodic structure in the Franconian dialect of Arzbach. Utrecht: LOT.

Korpilahti, Pirjo, Christina M. Krause, Irma Holopainen & A. Heikki Lang. 2001. Early and Late Mismatch Negativity Elicited by Words and Speech-Like Stimuli in Children. Brain and Language 76. 332–339.

Korpilahti, Pirjo, A. Heikki Lang & Olli Aaltonen. 1995. Is there a late-latency mismatch negativity (MMN) component? Electroencephalography and Clinical Neurophysiology 95. P96.

Kutas, Marta & Kara D. Federmeier. 2011. Thirty years and counting: finding meaning of the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62. 621–647.

Lass, Roger. 1984. Phonology: An introduction to basic concepts (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Lau, Ellen F., Colin Phillips & David Poeppel. 2008. A cortical network for semantics: (de)constructing the N400. Nature Reviews Neuroscience 9(12), 920–933.

Levänen, Sari, Riitta Hari, Linda McEvoy & Mikko Sams. 1993. Responses of the human auditory cortex to changes in one versus two stimulus features. Experimental Brain Research 97. 177–183.

Menning, Hans, Satoshi Imaizumi, Pienie Zwitserlood & Christo Pantev. 2002. Plasticity of the human auditory cortex induced by discrimination learning of non-native, moratimed contrasts of the Japanese language. Learning Memory 9(5). 253–267.

Näätänen Risto, Petri Paavilainen & Kalevi Reinikainen. 1989. Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? Neuroscience Letters 107. 347–352.

Näätänen, Risto, Petri Paavilainen, Teemu Rinne & Kimmo Alho. 2007. The mismatch negativity (MMN) in basic research of central auditory processing: A review. Clinical Neurophysiology 118, 2544–2590.

Nenonen, Sari, Anna Shestakova, Minna Huotilainen & Risto Näätänen. 2003. Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. Cognitive Brain Research 16. 492–495.

Nooteboom, Sieb. 1997. Prosody of speech: Melody and rhythm. In William J. Hardcastle & John Laver (eds.), The handbook of phonetic sciences, 640–673. Oxford: Blackwell.

Ohala, John J. 1972. How is pitch lowered? Journal of the Acoustical Society of America 52. 124 (A).

Ohala, John J. 1978. Production of Tone. In Victoria A. Fromkin (ed.), Tone: a linguistic survey, 5–39. New York et al.: Academic Press.

Ohala, John J. & William G. Ewan. 1973. Speed of pitch change. Journal of the Acoustical Society of America 53. 345 (A).

Partanen, Eino, Martti Vainio, Teija Kujala & Minna Huotilainen. 2011. Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. Psychophysiology 48(10). 1372–1380.

Pekkonen, Eero, Teemu Rinne, Kalevi Reinikainen, Teija Kujala, Kimmo Alho & Risto Näätänen. 1996. Aging Effects on Auditory Processing: An Event-Related Potential Study. Experimental Aging Research: An International Journal Devoted to the Scientific Study of the Aging Process 22(2). 171–184.

Peters, Jörg. 2006. The Cologne word accent revisited. In Michiel de Vaan (ed.), Germanic Tone Accents (Zeitschrift für Dialektologie und Linguistik. Beihefte 131), 107–133. Stuttgart: Franz Steiner.

Phillips, Colin, Thomas Pellathy, Alec Marantz, Elron Yellin, Kenneth Wexler, David Poeppel, Martha McGinnis & Timothy Roberts. 2000. Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study. Journal of Cognitive Neuroscience 12(6). 1038–1055.

Pompino-Marschall, Bernd. 2003. Einführung in die Phonetik, 2nd edn. (de Gruyter Studienbuch). Berlin & New York: Walter de Gruyter.

Ramers, Karl-Heinz. 1998. Einführung in die Phonologie (UTB für Wissenschaft: Uni-Taschenbücher). München: Fink.

Ramers, Karl-Heinz & Heinz Vater. 1995. Einführung in die Phonologie (Kölner linguistische Arbeiten [KLAGE] – Germanistik 16). Hürth: Gabel.

Rimmele, Johanna, Elyse Sussman, Christian Keitel & Thomas Jacobsen. 2012. Electrophysiological Evidence for Age Effects on Sensory Memory Processing of Tonal Patterns. Psychology and Aging 27(2). 384–398.

Scharinger, Mathias, Philip J. Monahan & William J. Idsardi. 2011. You had me at "Hello": Rapid extraction of dialect information from spoken words. NeuroImage 56(4). 2329–2338.

Schiff, Sami, Pietro Valenti, Andrea Pellegrini, Maria Lot, Patrizia Bisiacchi, Angelo Gatta & Piero Amodio. 2008. The effect of aging on auditory components of event-related brain potentials. Clinical Neurophysiology 119. 1795–1802.

Schmidt, Jürgen Erich. 1986. Die mittelfränkischen Tonakzente (Mainzer Studien zur Sprach- und Volksforschung 8). Stuttgart: Franz Steiner.

Schmidt, Jürgen Erich. 2017. Neurodialektologie. Zeitschrift für Dialektologie und Linguistik 83(1). 56–91.

Schröger, Erich. 1998. Measurement and interpretation of the mismatch negativity. Behavior Research Methods, Instruments, & Computers 30(1). 131–145.

Schröger, Erich, Iria SanMiguel & Alexandra Bendixen. 2013. Prädiktive Modellierung in der auditiven Wahrnehmung. In Erich Schröger & Stefan Koelsch (eds.), Affektive und Kognitive Neurowissenschaft (Enzyklopädie der Psychologie 5), 11–45. Göttingen et al.: Hogrefe Verlag für Psychologie.

Sittiprapaporn, Wichian. 2012. Source Localization of Preattentive Processing for Different Vowel Duration Changes with Contour Tones in Monosyllabic Thai Words. Journal of Applied Sciences 12(15). 1580–1587.

Stock, Eberhard & Christina Zacharias. 1982. Deutsche Satzintonation, 3rd edn. Leipzig: VEB Verlag Enzyklopädie.

Sundberg, Johan. 1973. Data on maximum speed of pitch changes. Speech Transmission Laboratory – quarterly progress and status report 14. 39–47.

Sundberg, Johan. 1979. Maximum speed of pitch changes in singers and untrained subjects. Journal of Phonetics 7. 71–79.

Takegata, Rika, Mari Tervaniemi, Paavo Alku, Sari Ylinen & Risto Näätänen. 2008. Parameter-specific modulation of the mismatch negativity to duration decrement and increment: Evidence for asymmetric processes. Clinical Neurophysiology 119. 1515–1523.

Trubetzkoy, Nikolaj S. 1939. Grundzüge der Phonologie. Prag.

Verleger, Rolf, Winfried Neukäter, Detlef Kömpf & Peter Vieregge. 1991. On the reasons for the delay of P3 latency in healthy elderly subjects. Electroencephalography and clinical Neurophysiology 79. 488–502.

Werth, Alexander. 2011. Perzeptionsphonologische Grundlagen der Prosodie: Eine Analyse der mittelfränkischen Tonakzentdistinktion (Zeitschrift für Dialektologie und Linguistik. Beihefte 143). Stuttgart: Franz Steiner.

Werth, Alexander. 2012. Perceptual evidence for contrast enhancement in tone-intonation interaction. In Oliver Niebuhr (ed.), Understanding prosody, 187–200. Berlin & New York: Walter de Gruyter.

Wiese, Richard. 1996. The Phonology of German. Oxford: Clarendon Press.

Wiese, Richard. 2011. Phonetik und Phonologie (LIBAC. Linguistik für Bachelor). Paderborn: Fink.

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

Ylinen, Sari, Minna Huotilainen & Risto Näätänen. 2005. Phoneme quality and quantity are processed independently in the human brain. NeuroReport 16. 1857–1860.

Ylinen, Sari, Anna Shestakova, Paavo Alku & Minna Huotilainen. 2005. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. Language and Speech 48(3). 313–338.

Ylinen, Sari, Anna Shestakova, Minna Huotilainen, Paavo Alku & Risto Näätänen. 2006. Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. Brain Research 1072. 175–185.

Zec, Draga. 1994. Sonority constraints on prosodic structure. New York & London: Garland.

Zhang, Jie. 2002. The effects of duration and sonority on contour tone distribution: A typological survey and formal analysis (Outstanding dissertations in linguistics). New York & London: Routledge.

Appendix

Tab. 5: Peak latencies and amplitudes of the late MMN at Fz for all experimental conditions

deviant$_{/standard}$	**group**	**peak latency at Fz**	**amplitude at Fz deviant-standard**
[ʃa̠ː^{2}l]$_{/ʃa̠lˑ2/}$	Dialect	430 ms	−0.535 µV
[ʃa̠ː^{2}l]$_{/ʃa̠lˑ2/}$	Standard German	356 ms	−0.780 µV
[ʃa̠lˑ^{2}]$_{/ʃa̠ː2l/}$	Dialect	554 ms	−1.162 µV
[ʃa̠lˑ^{2}]$_{/ʃa̠ː2l/}$	Standard German	612 ms	−1.307 µV