Preattentive phonotactic processing

DISSERTATION

zur Erlangung des Grades eines Doktors der Philosophie der

Fakultät für Geistes- und Sozialwissenschaften der Helmut-Schmidt-Universität /

Universität der Bundeswehr Hamburg

vorgelegt von

Thomas Konstantin Jacobsen

aus Uelzen

Hamburg, 2015

Content

Part III

1 Introduction

This dissertation is based on a cumulus of publications that consists of three journal publications which deal with the empirical investigation of speech processing. Three experiments are presented and discussed, in which the preattentive processing of phonological rules in German is examined with event-related brain potentials (ERPs) as the dependent variable. Therefore the cognitive electrophysiology is introduced at first, because this method was used throughout these studies.

In the next chapters phonotactic knowledge is introduced and some Mismatch Negativity studies which deal with it are described, which is followed by the goal of the experiments.

The empirical section starts with an experiment examining the German phonological rule of g-deletion. The second one deals with final devoicing. Finally the goal of the last experiment was to find out whether it is possible to detect two Mismatch Negativities (MMN) with two predictable successive deviants.

In the last chapter a generalized summary of the present work is given with some thoughts of possible further experiments.

Part I

2 Theoretical Section

2.1 Cognitive Electrophysiology[1]

"In cognitive neurosciences, the mismatch negativity (MMN) component of the auditory event-related brain potentials (ERP) has become an well-established electrophysiological measure to investigate the phonological knowledge and its involvement in early, automatic stages of speech processing (for reviews, see Näätänen, 2001; Näätänen et al., 2007; Pulvermüller and Shtyrov, 2006; Shtyrov and Pulvermüller, 2007)."[1] To measure the electromagnetic activity of the brain, special electrodes are used, which are placed along the scalp (for further information, see Luck, 2005). With this method the fluctuation of voltage can be recorded, which is the electrical activity of neural sources in that region. ERPs are the brain responses to a specific event, which can be seen in this electroencephalography (EEG) signal. They are signal changes in the ongoing neuroelectromagnetic activity of the brain that are time-locked to an event (e.g. Rohrbaugh et al., 1990). These effects are very small and usually not visible in the signal, so that many trials of the same event have to be averaged.

The MMN, which is one of the components, was firstly described by Näätänen et al. (1978) and "reflects a preattentively operating memory-based mechanism that detects violations of an expectancy created by the regularity of the preceding acoustic stimulation (Näätänen, 1992; Näätänen and Winkler, 1999; Winkler, 2007; Näätänen et al., 2011, Schröger et al., 2013)"[1]. That means that during the presentation of a frequent event, the so-called standard, an infrequent event (the deviant) occurs, which violates the expectation of the system, which already extracted regularities from the standards. The standard could for example be the vowel 'i', which is repeated several times, followed by a deviant 'u' [i i i i u…]. A standard experimental design, called reversed oddball design, will also test it the

[1] Derived from Paper 1

other way round, both with a high number of repetitions. A negative deflection in the deviant-minus-standard difference wave calculated from the ERPs elicited by the standard and the deviant stimuli is defined as the MMN. It typically occurs between 100 and 250 ms after deviation onset.

These signals can be measured at a millisecond level, so the resolution is superior to behavioral measures (such as response time or accuracy of the response) which usually give information after the process of interest (e.g. auditory perception) has been terminated. The recording is noninvasive; Most importantly, the recording of ERPs does not require a behavioural response by the participant. Thus, it is possible to record brain responses even when the participant is not able to behaviorally respond (e.g. with infants or comatose patients) or when the stimulation can be ignored. Thus the recording of ERPs and ERFs represents a unique tool to study processing of task-irrelevant information (e.g. pre-attentive processing).

If two or more features of a standard change in the deviant, the MMN is as large as the sum of the MMNs to each of the individual deviants differing in only one feature (e.g., Paavilainen et al., 2001, Takegata et al., 1999, Wolff and Schröger, 2001; see experiment 3). The MMN is also larger, when the detection of an ill-formed stimulus coincide with the detection of for example a change of the duration. Ill-formed stimuli do produce a distinct MMN (Steinberg et al., 2010a, 2010b, 2011; see the next chapter), but it is difficult to recognize it, when this MMN coincides with a feature change, resulting only in one peak, which is higher than in control groups.

It is also important for the interpretation of oddball experiments, that the MMN often appears asymmetrically. That means that for example an ill-formed stimulus only causes an MMN when it occurs as a deviant. As a standard, the system would have already extracted regularities and the well-formed deviant would not produce a second peak.

The MMN can reflect the activation of linguistic knowledge because it is sensitive to higher order cognitive processes. Therefore it is a good tool for investigating phonological processing. The other advantage of this technique is that it can be used to investigate automatic processes, because its protocols are passive. That means that the participants of these experiments do not pay attention to the stimulation but try to ignore it.

The MMN component of the human ERP has been demonstrated to be a good tool for the investigation of language comprehension outside the focus of attention (e.g., Näätänen, 2001; Pulvermüller & Shtyrov, 2006). It is sensitive for processes at the segmental level (e.g., Näätanen et al., 1997; Winkler et al., 1999; Kirmse et al., 2008) as well as the lexical level (Jacobsen et al., 2004). It therefore appears worthwhile to also test processing phonotactic rules outside the focus of attention with this technique, although there are only a few studies to this at the moment.

2.2 Phonological/Phonotactic knowledge[2]

The phonology as such examines the emergence of the speech sounds in one language and how they can be combined (Hall 2000, Durand 1990). There are a certain number of sounds in every language and there are rules for legal and illegal combinations. In some languages, there are for example rules for the number of consonants in one cluster. In Japanese, (nonnasal) coda consonants are not allowed (Dehaene-Lambertz 2000).

Phonological knowledge is represented in long-term memory and does include language-specific speech sound inventories as well as abstract principles of the co-occurrences of phonemes and their restrictions, namely the phonotactic knowledge. It is activated automatically in speech processing, during sub-lexical processing stages, independently from the mental lexicon (e.g., Kenstowicz, 1994; de Lacy, 2007). Phonological rules are processed outside the focus of attention in language comprehension (e.g., Bonte et al., 2005; Steinberg et al., 2011).

The term phonotactic knowledge is used for the knowledge a speaker has about how often different speech sound combinations are actually used in a given language and how limited the possibilities of sound combinations in one language are (e.g. Coleman & Pierrehumbert, 1997; Hayes & Wilson, 2008; Frisch, Pierrehumbert & Broe, 2004), which is also part of electrophysiological research (e.g. Bonte et al. 2005, 2007). These studies compare the processing of frequently used sound sequences and combinations that are rarely used.

While investigating the processing effects of phonotactically illegal speech material, for example pseudowords with a combination of sounds which are illegal

[2] Derived from Paper 1

in that language, two contrary result types can be found depending on the properties of the experiment. First, a phonotactic repair transforms an illegal stimulus in a phonetically similar legal percept. It has been observed in the perception of phonotactically illegal obstruent clusters, either by modifying one of the obstruents or by inserting an additional vowel (for example: Cuetos, Hallé, Dominguez, & Segui, 2011; Davidson, 2011; Dehaene-Lambertz, Dupoux, & Gout, 2000; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Dupoux, Parlato, Frota, Hirose, & Peperkamp, 2011; Hallé & Best, 2007; Hallé, Segui, Frauenfelder, & Meunier, 1998; Segui, Frauenfelder, & Hallé, 2001; Wagner, Shafer, Martin, & Steinschneider, 2012).[3]

Second, the violation detection occurs if the illegal stimulus could not have been perceptually repaired but had to be perceived as illegal. With this technique it is also possible to give arguments for discussions, in which people debate about whether a phonotactic rule is really existent in one language. Phonotactic illegality has been observed to cause additional processing effort, for example in the perception of illegal vowel-nasal or vowel-fricative sequences or illegal offset-clusters (e.g. Cho & McQueen, 2006; Flagg, Cardy, & Roberts, 2006; Hwang, Monahan, & Idsardi, 2010; Otake, Yoneyama, Cutler, & van der Lugt, 1996; Poeppel & Monahan, 2011; Steinberg, Truckenbrodt, & Jacobsen, 2010a, 2010b, 2011; Weber, 2001).[3]

These two different outcomes can maybe be explained by the explanation of Hallé et al. (1998), who said that illegal input would be analyzed and represented veridically or it would be assimilated to the native phonotactic system. So it would be either interpreted as a nonnative sound combination or it is modified in a way that guarantees legality. Another factor could be the salience of the violated phonological rule as triggered by the phonological context which the illegal sequence is embedded in.[3]

---

[3] Derived from Paper 1

2.3 Previous MMN-studies on phonotactic restrictions[4]

Dehaene-Lambertz et al. (2000) showed in a cross-linguistic experiment using an active protocol that the phonotactic properties of the native language influence the behavioral and electrophysiological results. In Japanese, a coda consonant is not allowed, that is why their language-trained auditory system repairs stimuli in which a coda vowel is missing. They hear no difference between /igmo/ and /igumo/, whereas the French participants did. French native speakers showed brain responses similar to MMN, indicating the detection of a difference between the two stimuli presented, but no MMN was observed for Japanese speaking participants. These results indicate that Japanese speakers automatically compensated for the phonotactically illegal sequence *[gm] (illegality marked by the asterisk) by inserting a vowel, thereby turning [igmo] into [igumo]. It is therefore a good proof for a phonotactic repair.

A similar study was made by Wagner et al. (2012) who investigated the /pt/ and the /st/ cluster at the beginning of a syllable in native-English and native-Polish listeners. Using pseudowords, the task for the participants was to identify the number of syllables of the stimulus. Only the native-English listeners had problems with stimuli containing the /pt/ cluster, because at the initial position of a syllable it is an illegal combination of phonemes in English. Therefore nearly all listeners heard an additional syllable with an insertion of the vowel /ə/ between the clusters, they were not able to discriminate between /pt/ and /pət/ pseudowords.

Kharlamov et al. (2011) investigated the optional but productive Russian t-deletion process in the middle of a word-internal three-consonant cluster. Pseudowords [asna] and [astna] were used as stimuli. Russian participants had a higher error rate and slower response times in the identification and discrimination

[4] Derived from Paper 1 and Paper 2

tasks compared to English speakers. Conversely, only in the ERP-signals of the English speakers an MMN was observed, whereas in the Russian group no effect was found for the contrast of [asna] and [astna].

Mitterer and Blomert (2003) tested the processing of nasal place assimilation in Dutch with a passive oddball protocol. They used the following stimulus pairs: ‘tuinbank’ (“garden bank”) and the assimilated compound ‘tuimbank’ as well as ‘tuinstoel’ (“garden chair”) and ‘tuimstoel’. The MMN was only elicited in the second pair, because the change of the nasal in the deviant was not motivated by the assimilation rule.

Bonte et al. (2005) investigated the processing of phoneme clusters, with clusters that occur with high or low frequency in Dutch. The MMN responses were stronger for the deviants consisting of frequently occurring phoneme cluster which follow infrequently occurring cluster, than vice versa.

Steinberg et al (2010a, 2010b, 2011) investigated a German allophonic alternation related to two dorsal fricative allophones both represented orthographically as "ch". The palatal allophone of this fricative occurs after front vowels ([ɛç]) and the velar allophone after back vowels ([ɔx]). This alternation is also known as Dorsal Fricative Assimilation (DFA). They made a range of different experiments that all provide evidence for the effect of DFA in pre-attentive sublexical processing. Here are some of their findings:

The ill-formed pseudoword *[ɛx] combines a velar fricative with a front vowel. Employed as a deviant with the standard [ɔx], there is a mismatch with respect of the different vowels. The fricatives, however, are segmentally the same, so that comparison conditions that employ analogically structured well-formed pseudowords do not show an MMN for the fricative. However, the ill-formed deviant *[ɛx] showed an additional MMN response attributable to the fricative. This response is temporally separated from the MMN elicited by the distinct vowels and attributed to the abstract phonotactic ill-formedness of the deviant.

The same holds true with an experiment, where [ɛç] and [*ɔç] were contrasted, the ill-formed pseudoword now combines a back vowel with a front fricative. The ill-formed deviant showed again an additional MMN response, which was not there in the control condition, where [ɛʃ] and [ɔʃ] were contrasted.

Finally an experiment was made in which four ill-formed ([*ʏx], [*œx], [*ɔç] and [*ʊç]) and four well-formed pseudowords ([ʏç], [œç], [ɔx] and [ʊx]) were contrasted. All ill-formed stimuli taken together showed an MMN effect, which was missing at the well-formed stimuli taken together. This was an effect because of the detection of the phonotactic violation again.

2.4 Goal

With the experiments in this dissertation, new data points shall be added to the literature that establishes the relevance of phonotactic violations in pre-attentive processing. Therefore the MMN is used as a tool, because it is sensitive to higher order cognitive processes and can reflect the activation of linguistic knowledge. It also provides a temporally highly resolved on-line measure of speech processing and it can be used to investigate fairly automatic processes in passive oddball protocols.

As described, Dehaene-Lambertz, Dupoux and Gout (2000), who found out cross-language differences, did not use a clearly pre-attentive protocol. Their protocol was active, because they asked the participants whether they heard a difference, so the stimuli have not been ignored. The passive protocol used in the experiments of this dissertation shows the automatic processing from the long term memory.

Flagg et al. (2006) investigated an assimilatory nasalization process in English with a different method, a Magnetoencephalography (MEG) study. In a pseudoword like /ama/ the first vowel can optionally be produced nasalized as in [ãma]. That is why they took [ãma] and [aba] and combined (and recombined) the first vowel with the second and third sound of both pseudowords to get four stimuli. They found out that the prediction of a nasal consonant is stronger when a nasalized vowel appears. So the nasalization process is also relevant to pre-attentive processing.

Mitterer and Blomert (2003) used postlexical assimilation in contrast to this lexical phonological process. Steinberg et al. (2010a, 2010b, 2011) investigated a German phonotactic rule with the passive oddball paradigm. The experiments in this dissertation tie in with their results and have a look at two other German phonotactic rules. The first rule is the g-deletion rule. A voiced velar plosive is deleted when appearing syllable-final following an alveolar nasal. The second rule is the final devoicing rule. All voiced obstruents are devoiced when appearing at the end of a syllable. These two rules are investigated, which is described in the empirical section.

The last experiment deals with the question whether it is possible to detect two Mismatch Negativities (MMN) with two predictable successive deviants.

On the one hand, the processing of other German phonotactic rules shall be compared to the findings of Steinberg et al (…) to generalize their findings. But there is also the question whether under certain conditions illegal stimuli can be automatically repaired by the auditory system. And it is finally interesting how two successive deviants are processed when they are fully predictable.

Part II

3 Empirical Section

3.1 Publications

The empirical section is based on three journal publications that are listed below. In the first two papers two different phonological rules were examined. The first one is the g-deletion rule, the second one is the final devoicing rule, both are German phonological rules. The third experiment dealt with the question whether it is possible to get two MMNs with two successive deviants.

Paper 1: Steinberg, J., Jacobsen, T.K., Truckenbrodt, H., and Jacobsen, T. (2015). Repair or violation detection? Pre-attentive processing strategies of phonotactic illegality demonstrated on the constraint of g-deletion in German. *Brain and Language*, accepted for publication.

Paper 2: Truckenbrodt, H., Steinberg, J., Jacobsen, T.K., and Jacobsen, T. (2014). Evidence for the role of German final devoicing in pre-attentive speech processing: A mismatch negativity study. *Frontiers in Psychology*, **5**, 1-11.

Paper 3: Jacobsen, T.K., Steinberg, J., Truckenbrodt,, H. and Jacobsen, T. (2013). Mismatch Negativity (MMN) to successive deviants within one hierarchically structured auditory object. *International Journal of Psychophysiology*, **87**, 1–7.

3.2 G-Deletion

G-deletion pertains to distributional constraints concerning the dorsal nasal [ŋ] in German. In current phonological frameworks, [ŋ] is considered not to be a phoneme in the German sound inventory, but rather an allophone because of its limited distribution (e.g., Hall 1992, Ramers & Vater 1992, Wurzel 1980). Depending on the segmental context, [ŋ] either alternates with the coronal nasal /n/ or with the sequence /n+g/. Two phonological mechanisms are assumed to cause these alternations:

(1) Preceding /g/ or /k/, [ŋ] is considered to result from a regressive place assimilation of /n/ to the following dorsal plosive operating obligatory within a syllable.

(2) Elsewhere, [g] deleted after the assimilated nasal if it is not followed by a full vowel

This latter phenomenon is referred to as “g-deletion” in the phonological literature (Féry, Hohmann, & Stähle, 2009; Hall, 1992; Issatschenko, 1963; Kohrt, 1980; Wiese, 1996). G-deletion occurs in syllable final position and before any following syllable that is non-moraic, i.e. that does not contain a full vowel. Accordingly, sequences like “Menge” [mɛŋə] (crowd) undergo g-deletion. However, in the presence of a following full vowel, like in “Mango” [mɑŋgo] (mango), /g/ always remains present. Thus, sequences of syllable-final [ŋ] followed by any full vowel like *[mɑŋo] are phonotactically illegal as they display a misapplication of g-deletion. To summarize, [ŋ] is only allowed to occur either before homorganic plosives (i.e. /g,k/) or before something else than a full vowel, where g-deletion has applied.

In this study, we pursued three goals:

(1) As stated above, the body of electrophysiological research focusing on categorical phonotactic knowledge and its role in pre-attentive speech processing is rather small. By expanding the focus of interest on a different phonotactic restriction in German grammar we aimed to strengthen the existing evidence about the impact of categorical phonotactic knowledge on pre-attentive speech processing.

(2) So far, disparate findings have been reported on how phonotactically illegal speech input can be dealt with in speech processing. Here, we aimed to disentangle two different processing strategies on phonotactically violated speech input that have been reported in the literature. More specifically, we manipulated the phonological context of the illegal stimulus to induce either phonotactic repair or phonotactic violation detection.

(3) Different methodological objections have been raised concerning the use of naturally produced illegal speech material (for a detailed discussion, see Steinberg, 2014; Steinberg, et al., 2012). Among the proposed alternatives (such as using synthetic or heavily manipulated natural stimulus material or stimuli that have been articulated by a nonnative speaker in whose first language the sound sequence of interest would not be illegal), the application of moderately controlled utterances produced by a native-language speaker seemed the most appropriate compromise to us. Nevertheless, in our previous studies we employed professional speakers to avoid any potential artefacts due to the phonotactic violation. Here, we included two different stimulus sets realized by one non-professional and one professional speaker of German (from now on referred to as Experiment 1 vs. Experiment 2). By this we intended to estimate whether the performance of the illegal stimulus as mediated by the level of speech proficiency would influence the effects on phonotactic illegality processing.

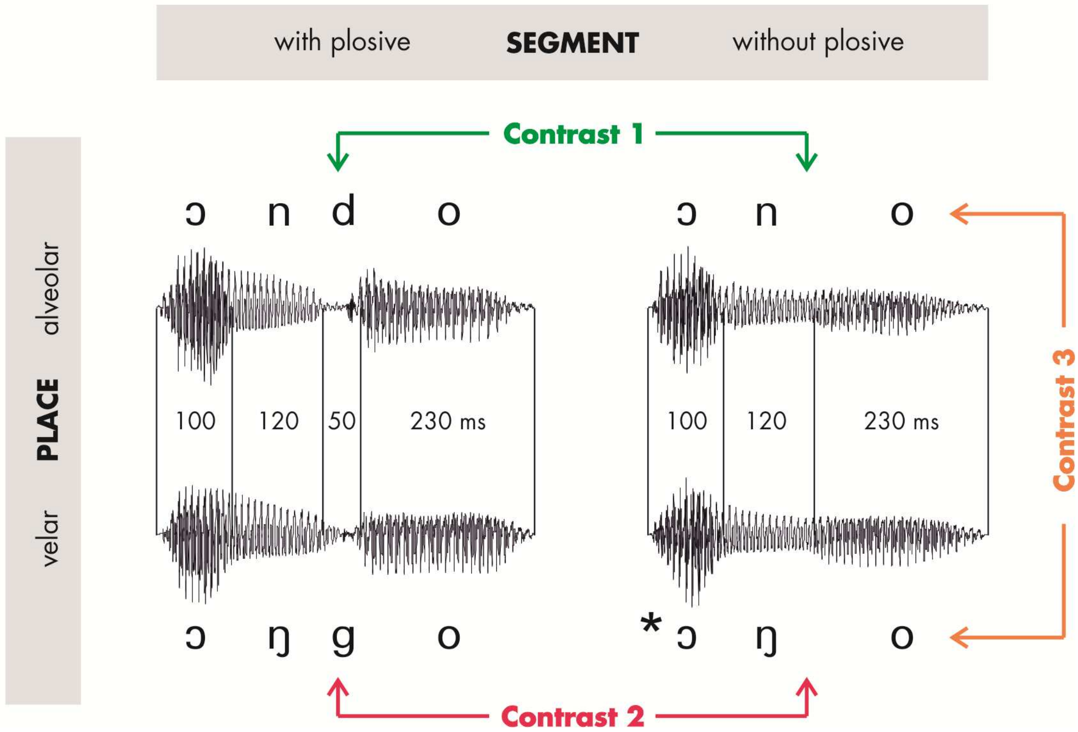

Figure 1. Experimental design of both experimental parts. The stimuli used differ in two features: segment (whether they contain a plosive or not) and place (whether the plosive and/or nasal are velar or alveolar)

We used four non-lexical disyllabic stimuli that were pair-wise contrasted in three experimental conditions as depicted in figure 1. Those stimuli entailing coronal consonants, i.e. consonants with an alveolar place of articulation, ([ɔndo] and [ɔno]), serve as phonotactically legal control condition in Contrast 1. The phonotactically violated stimulus, [ɔŋo], is presented twice, first among [ɔŋgo] in Contrast 2 and second among [ɔno] in Contrast 3. The segmental structure of the four stimuli is reflected by the phonological factors PLACE of articulation (alveolar: [ɔndo/ɔno] vs velar: [ɔŋgo/*ɔŋo]) and SEGMENT (with plosive: [ɔndo/ɔŋgo] vs without plosive [ɔno/*ɔŋo]). Finally, the factor CONTEXT determines whether the deviants [ɔno/*ɔŋo] were presented among phonologically analog ([ɔno/*ɔŋo] from Contrast 3) or different standards ([ɔndo/ɔŋgo] from contrasts 1 or 2, respectively).

Furthermore, the between-subjects factor EXPERIMENT differentiates the two data sets run with the stimuli of the naïve and professional speaker.

Based on this experimental design we set up the following hypotheses.

Hypothesis 1: Phonotactic illegality due to a misapplication of g-deletion affects pre-attentive speech processing. This impact is expected to modulate the amplitude of the MMN responses elicited by illegal deviants.

Hypothesis 2: this impact depends on the phonological context of the illegal deviant, i.e. the segmental structure standard stimulus. If, on the one hand, the context of the standard stimulation per se activates phonotactic knowledge on g-deletion, the nasal [ŋ] in the deviant will accordingly be interpreted as allophonic representation of /N+g/. Thus, a perceptual repair by inserting the missing /g/ in the input seems likely, reflected by the absence of any violation-related MMN response (henceforth: phonotactic repair hypothesis). If, on the other hand, g-deletion is not motivated by the context of the standard stimulation, we expect [ŋ] to be parsed as singular nasal as encoded in the acoustic signal. Consequently, the phonotactic illegality will become manifest with the onset of the following full vowel. In this case, a violation-related MMN response is expected to be elicited by the illegal deviant (henceforth: phonotactic evaluation hypothesis).

3.2.1 Materials and methods

Participants

Sixteen volunteers participated in each part of the study (Experiment 1: mean age of 24.4 years, range from 21 to 28, ten male; Experiment 2: mean age of 23.6 years, range from 20 to 37, nine male), all of them right-handed and monolingual native speakers of German. Handedness was assessed using an inventory adopted from Oldfield (1971). All participants reported normal auditory and normal or corrected-to-normal visual acuity and no neurological, psychiatric, or other medical problems. They gave informed written consent. The study conformed to The Code of Ethics of the World Medical Association (2013).

Materials

Four pseudowords were used as stimuli: [ɔno], [ɔndo], [ɔŋgo], and *[ɔŋo]. None of these have lexical meaning in German. Three of the stimuli are phonotactically well-formed in German, however, the nonword *[ɔŋo] inappropriately applied a g-deletion and therefore violates phonotactic requirements in German.

Stimulus material was digitally recorded in a sound-attenuated room with a 16 bit resolution and a 48-kHz sampling rate using a using a RME Fireface 800 recording device (Audio AG, Haimhausen, Germany) and a Neumann U87 Microphone (Georg Neumann GmbH, Berlin, Germany). The pseudowords were articulated numerous times both by a non-professional speaker with a background in phonetics (TKJ; Experiment 1) and an externally-hired professional speaker who was non-professional with respect to the purpose of the experiment (Experiment 2), both males with fundamental frequencies of about 100 Hz. Naturally spoken speech material was chosen in favor of fully synthesized or spliced material in order to avoid brain responses to artifacts in signal processing caused by mismatching or misleading phonetic gestures.

To include acoustic variability into the stimulus material, we selected 10 exemplars of each pseudoword were selected, resulting in a set of 40 tokens in total. Duration manipulations of each token were performed using the time-domain PSOLA tool of Praat software (Boersma & Weenink, 2010). Segmental durations of each stimulus were equated by setting the first vowel to 100 ms, the nasal to 120 ms, the plosive to 50 ms and the second vowel to 230 ms as depicted in Figure 1. Original segmental durations averaged across the ten token of each stimulus along with standard deviations (SD) were as follows: /ɔno/ 125 (SD 9.7) + 106 (6.3) + 228 (11.7) ms, /ɔndo/ 124 (3.3) + 166 (16.2) + 48 (5.5) + 248 (16.3) ms, /ɔŋgo/ 108 (9.2)

+ 137 (11.2) + 46 (3.2) + 227 (21.7) ms and */ɔŋo/ 110 (8.8) + 126 (12.2) + 226 (7) ms (Experiment 1); /ɔno/ 127 (9.8) + 117 (10.6) + 185 (14.7) ms, /ɔndo/ 122 (7.6) + 166 (8.9) + 38 (5) + 176 (23.7) ms, /ɔŋgo/ 110 (9.9) + 160 (8.4) + 45 (7.3) + 175 (15.4) ms and */ɔŋo/ 117 (7.7) + 138 (13.8) + 173 (15.3) ms (Experiment 2). Intensities were normalized using the root mean square (RMS) of the whole sound file by means of Praat.

Experimental design and procedure

The pseudowords were contrasted pairwise in orthogonal oddball conditions. In each oddball sequence, one pseudoword served as standard (p = .85 = 1360 trials) and the other was used as deviant. Six experimental conditions were employed. Conditions were based on three different pair-wise stimulus contrasts each of which being run twice with reversed stimulus probabilities: /ɔndo/ vs /ɔno/ (contrast 1), /ɔŋgo/ vs */ɔŋo/ (contrast 2) and /ɔno/ vs */ɔŋo/ (contrast 3).

Contrasts 1 and 2 base upon the differing segmental complexity of the employed stimuli, one of them containing a /plosive + nasal/ sequence while the other containing a single /nasal/. Consequently, an MMN is expected to reflect the additional respective missing plosive in the deviant in both conditions. Contrast 1 serves as a baseline condition as its stimuli contain coronal consonants and in that are neutral with respect to g-deletion. Contrast 2 presents the dorsal counterparts of the Contrast 1, however, in this contrast both of the stimuli are to be evaluated with respect of g-deletion. Finally, in Contrast 3 stimuli differ in the quality of the nasal (caused by different places of articulation: one coronal and the other dorsal) while segmental complexity is kept constant. In this contrast, only the ill-formed stimulus triggers phonotactic evaluation with regard to g-deletion. In Contrast 3, an MMN is expected to be elicited due to the change of the nasal.

Stimulus sequences of 1600 trials in total were generated per condition, using all ten tokens of each pseudoword equally. Standard and deviant stimuli were delivered in pseudo-randomized order forcing at least two standards to be presented between successive deviants. Sequences were split into two parts each, resulting in a total of twelve blocks per participant (3 stimulus contrasts x 2 directions of presentation x 2 parts per condition). The first and second half of each sequence was run on different days, i.e. the experimental sessions were split into two parts. Stimuli were presented with a stimulus onset asynchrony (SOA) randomly varying from 550 to 900 ms in units of 10ms. The order of the experimental blocks was counterbalanced between participants for both sessions. Participants were seated comfortably in a sound-attenuated and electrically shielded experimental chamber. They were instructed to ignore the auditory stimulation while watching a self-selected silent subtitled movie. Stimuli were presented binaurally at 53 dB SPL (Experiment 1) and 54 dB SPL (Experiment 2) through headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). Sound intensity was measured by means of an artificial head (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany). Informal questioning of the participants revealed that they had been able to ignore the auditory stimulation and that they had perceived all stimulus types as speech sounds. A whole experimental session lasted approximately 150 minutes (plus additional time for electrode application and removal) including ten short breaks of about two minutes each.

Electrophysiological recordings

The EEG (Ag/AgCl electrodes, Falk Minow Services, V-Amp EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from 26 standard scalp locations according to the extended 10–20 system (American Encephalographic Society, 1994, FP1, FP2, F7, F3, FZ, F4, F8, FC5, FC1, FCZ, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P7, P3, PZ, P4, P8, O1, O2) and from the left and right mastoids. The reference electrode was placed on the tip of the nose, and an additional electrode placed at AFz was used as ground during recording. Electroocular activity was recorded with two bipolar electrode pairs. The vertical electrooculogram (EOG) was obtained from the right eye by one supraorbital and one infraorbital electrode, and the horizontal EOG was obtained from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 10 kΩ. On-line filtering of the EEG and EOG signals was carried out using a 0.011 Hz high-pass and a 100 Hz low-pass filter. The signal was digitized with a 16 bit resolution at a sampling rate of 500 Hz.

Data analysis

Off-line signal processing was carried out using EEP 3.0 (ANT Neuro, Enschede, Netherlands). EEG-data were band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 800 ms, time-locked to the onset of the stimuli, and including a 100 ms pre-stimulus baseline, were extracted and averaged separately for stimulus probability (standard; deviant), each pseudoword per contrast (contrast 1: [ɔndo1], [ɔno1]; contrast2: [ɔŋgo2], [*ɔŋo2]; contrast 3: [ɔno3], [*ɔŋo3]), and participant. The ERP responses to the first five stimuli of each block as well as to each standard stimulus immediately following a deviant were excluded from the analysis. Epochs showing an amplitude change exceeding 100 μV at any of the recording channels were rejected. Grand-averages were subsequently computed from the individual-subject averages. To quantify the full MMN amplitude, the scalp ERPs were re-referenced to the averaged signal recorded from the left and right mastoids. This computation results in an integrated measure of the total neural activity underlying the auditory MMN (Schröger, 2005).

Identity MMN (iMMN) was calculated for each pseudoword per contrast by subtracting the ERPs elicited by the standard point by point from the ERPs elicited by the physically identical deviant obtained from the reversed stimulation (Pulvermüller & Shtyrov, 2006). Deviance-related effects were quantified by measuring the MMN amplitudes as mean voltages in fixed analysis windows of 40 ms. These windows were aligned a posteriori at the first peak latency of the grand-average difference waves of the pseudowords that could possibly be attributed to the effects of interest. Peak latencies were then averaged across C3, CZ, C4, F3, FZ and F4 electrode sites. Explained in more detail, the segmental MMN due to the change of the nasals in Contrast 3 was not be found before from 200 ms because the deviation occurred at 100 ms after stimulus onset and an MMN takes at least 100 ms to build up. Thus the analysis window was centered over the first negative-going deviation of the difference wave occurring after this point in time. Accordingly, the MMN relating to the phonotactic (il)legality in all contrasts was not to be expected earlier than 320ms because any misapplication of g-deletion could not be detected earlier than 220 ms after stimulus onset, i.e. the offset of the nasal. Differing from these criteria, the pseudoword [*ɔŋo2] in the data set of experiment 2 was not included in the respective latency averaging because there was no discernible peak in the respective time window to be considered. Due to the differing segmental structure of the stimuli, separate windows were assigned for those stimuli containing a plosive ([ɔŋgo2], [ɔndo1] and for those without plosives [*ɔŋo2], [*ɔŋo3], [ɔno1], [ɔno3]. Differing from these criteria, the pseudoword [*ɔŋo2] in the data set of Experiment 2 was not included in the respective latency averaging because there was no discernible peak in the respective time window to be considered.

Statistical analysis

Statistical analyses of the present data were performed with SPSS (IBM SPSS Statistics 21). For reasons of straightforwardness, only effects based on CZ electrode position are reported in the main article (unless stated otherwise). Full topographical analyses using the extra within-subject factors POSITION (F-/C-/P-line) and LATERALITY (3-/z-/4-line) are provided as supplementary material, as well as separate analyses for each of the data sets obtained from Experiments 1 and 2.

Effects attributable to the change of the nasals in Contrast 3 were analyzed by means of a univariate two-way mixed design ANOVA with the within-subject factor PLACE (place of articulation: coronal/dorsal) and the between-subjects factor EXPERIMENT (1/2). The level of type 1 error was set up to p <.05.

For analyzing possible effects due to the phonotactic (ill)legality of the stimuli, two univariate three-way mixed-design ANOVAs were run. (1) effects elicited by the pseudowords [ɔndo1], [ɔno1], [ɔŋgo2],and [*ɔŋo2] were analyzed using the within-subject factors PLACE and SEGMENT (segmental structure: with/without plosive), and the between-subjects factor EXPERIMENT; (2) deviance-related effects elicited by the pseudowords [ɔno1], [*ɔŋo2], [ɔno3], [*ɔŋo3] were analyzed using the within-subjects factors PLACE and CONTEXT (original oddball contrast: same/different), and with the between-subjects factor EXPERIMENT Since both ANOVAs were partially run with the same data ([ɔno1] and [*ɔŋo2]), the original level of type 1 error was adjusted to p* <.025 according to Bonferroni. The level of type 1 error was set to p <.05.

3.2.2 Results

In Experiment 1, an average of 13.4 % (SD 7.7 %) of the trials per participant was rejected prior to ERP computation, in Experiment 2, mean rejection rate was 14.3 % (SD 11.3 %). The grand-average deviant-minus-standard difference waves showed negative-going deflections (such as MMN) in several time ranges. First, in Contrast 3, deviance-related effects relating to the onset of the nasal at 100 ms after stimulus onset were observed and analyzed. The analysis window was set to 216-256 ms (Experiment 1) and 208-248 ms (Experiment 2) according to the criteria stated above. Furthermore, effects that could temporally be attributed to the phonotactic violation in [*ɔŋo], i.e. that relate to the onset of the second vowel at 220 ms after stimulus onset were analyzed. The appropriate analysis windows were set as follows: (1) stimuli containing a plosive ([ɔndo]/contrast 1; [ɔŋgo]/contrast 2): 376 to 416ms (Experiment 1) and 356 to 396 ms (Experiment 2); (2) stimuli without a plosive ([ɔno]/contrasts1 and 3; *[ɔŋo]/contrasts 2 and 3): 342 to 382 ms (Experiment 1) and 334 to 374 ms (Experiment 2). Later negativities possibly reflecting stimulus offset effects have not been analyzed.

- Nasal change: segmental iMMN in contrast 3

The two-way mixed-design ANOVA at CZ with the within-subjects factor PLACE (coronal/dorsal) and the between-subjects factor EXPERIMENT (1/2) revealed significant main effect of the factors PLACE ($F_{1,30} = 54.71$; $p < .001$; $\eta_p^2 = .646$) indicating a stronger MMN response for [ɔno3] compared to [*ɔŋo3], and EXPERIMENT ($F_{1,30} = 4.48$; $p = .043$; $\eta_p^2 = .130$), reflecting generally stronger MMN responses in Experiment 1 compared to Experiment 2. Descriptive statistics are as follows: experiment 1 – [ɔno3] mean MMN amplitude -2.52 μV, (SD 1.17); [*ɔŋo3] -.52 μV (.91); experiment 2 – [ɔno3] -1.56 μV (1.11); [*ɔŋo3] -.29 μV (.91).

A significant main effect PLACE was also found in the separate analyses of experiments 1 and 2 (see also the detailed statistical results that are provided as supplementary material).

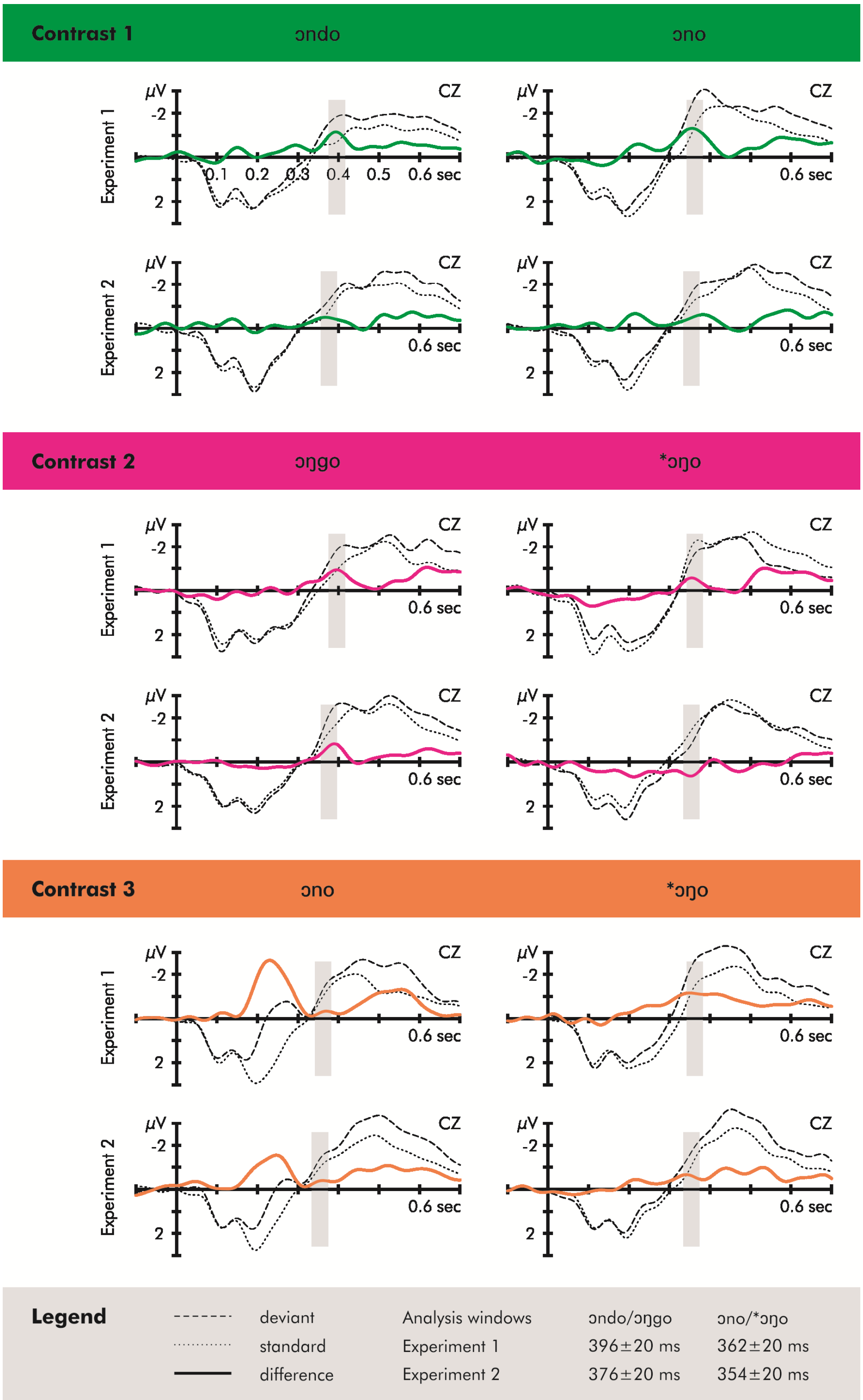

Figure 2. Grand-averaged, re-referenced ERP responses are represented separately for the pseudowords in every contrast the electrode CZ. Shown are ERPs elicited by the deviants (dashed lines), ERPs elicited by the standards (dotted lines), and Deviant-minus-Standard difference waves (solid lines). The bars mark the statistically analyzed time windows of the experimental parts. Scales are in milliseconds and microvolt.

- Illegality effects in all contrasts

a) Phonotactic Repair Hypothesis

The three-way mixed-design ANOVA at CZ with the within-subjects factor PLACE (coronal/dorsal) and the between-subjects factor EXPERIMENT (1/2) revealed a significant main effect of the between-subjects factor EXPERIMENT ($F_{1,30} = 13.24$; $p = .001$; $\eta p^2 = .306$) indicating generally stronger MMN responses in the data set of Experiment 1 compared to that from Experiment 2. However, there were no significant interactions between the linguistic factors and EXPERIMENT. The crucial interaction of PLACE*SEGMENT did not reach significance at CZ ($F_{2,30} = 4.16$; $p = .05$; $\eta p^2 = .122$) but at C3 where the effect was strongest ($F_{2,30} = 6.88$; $p = .014$; $\eta p^2 = .186$). Breaking down this interaction by the factor PLACE, two two-way mixed design ANOVAS were run comparing the MMNs separately for coronal and dorsal stimuli. The MMNs elicited by the coronal pseudowords [ɔndo1] and [ɔno1] (Contrast 1) did not differ significantly. Crucially, in Contrast 2, [ɔŋgo2] elicited a stronger MMN response than the illegal pseudoword [*ɔŋo2] (main effect SEGMENT $F_{1,30} = 11.52$; $p = .002$; $\eta p^2 = .278$), additionally, the factor EXPERIMENT was significant $F_{1,30} = 4.97$; $p = .003$; $\eta p^2 = .142$) as well. There were no significant interactions with SEGMENT in both analyses. Descriptive statistics were as follows: experiment 1 – [ɔndo1] mean MMN amplitude -1.05 μV, (SD 1.38); [ɔno1] -1.24 μV (1.36); [ɔŋgo2] -.87 μV (1.02); [*ɔŋo2] -.49 μV (.88); experiment 2 – [ɔndo1] -.47 μV (1.04); [ɔno1] -.56 μV (1.18); [ɔŋgo2] -.73 μV (.78); [*ɔŋo2] .45 μV (1.07).

b) Phonotactic Evaluation Hypothesis

The three-way mixed-design ANOVA at CZ with the within-subjects factor PLACE (coronal/dorsal) and the between-subjects factor EXPERIMENT (1/2) revealed the crucial interaction CONTEXT*PLACE to be highly significant ($F_{1,30} = 16.67$; $p < .001$; $\eta p^2 = .357$). Breaking down this interaction by the factor PLACE, two two-way mixed design ANOVAS were run comparing the MMNs separately for the (legal) coronal and the (illegal) dorsal stimuli. Looking at the coronal pseudowords, the main effect CONTEXT ($F_{1,30} = 5.27$, $p = .029$; $\eta p2 = .149$) did not reach significance. However, it reveals a statistical trend indicating that [ɔno] elicited stronger MMN responses when contrasted with [ɔndo] (Contrast 1) than with [*ɔŋo] (Contrast 3). There were no significant differences between the first and second data set. The analysis of the illegal dorsal pseudoword [*ɔŋo] from the contrasts 2 and 3, however, revealed a significant main effect of CONTEXT ($F_{1,30} = 12.71$; $p = .001$; $\eta p2 = .298$) based on a contrary effect: the illegal pseudoword elicited stronger MMN responses when being contrasted with [ɔno] (Contrast 3) than in the oddball context of [ɔŋgo] (Contrast 2). Furthermore, the factor EXPERIMENT was found to be significant ($F_{1,30} = 9.05$; $p = .005$; $\eta p2 = .232$) proving stronger effects in the first compared to the second data set. Descriptive statistics were as follows: experiment 1 – [ɔno1] mean MMN amplitude -1.24 µV, (SD 1.36); [ɔno3] -.28 µV (1.75); [*ɔŋo2] -.49 µV (.88); [*ɔŋo3] -1.14 µV (1.18); experiment 2 – [ɔno1] -.56 µV (1.18); [ɔno3] -.38 µV (1.17); [*ɔŋo2] .45 µV (1.07); [*ɔŋo3] -.52 µV (.68).

Separately analyzed, experiment 1 revealed a significant interaction CONTEXT*PLACE at CZ ($F_{1,15} = 7.972$; $p = .013$; $\eta p2 = .347$). Broken down, however, no main effect CONTEXT was found neither for ɔno nor for *ɔŋo. In the data set of experiment 2, the interaction CONTEXT*PLACE was found to be significant at CZ ($F_{1,15} = 10.001$; $p = .006$; $\eta p2 = .400$). Broken down analyses at CZ revealed a significant main effect of CONTEXT for *ɔŋo ($F_{1,15} = 14.700$; $p = .002$; $\eta p2 = .495$) but not for ɔno.

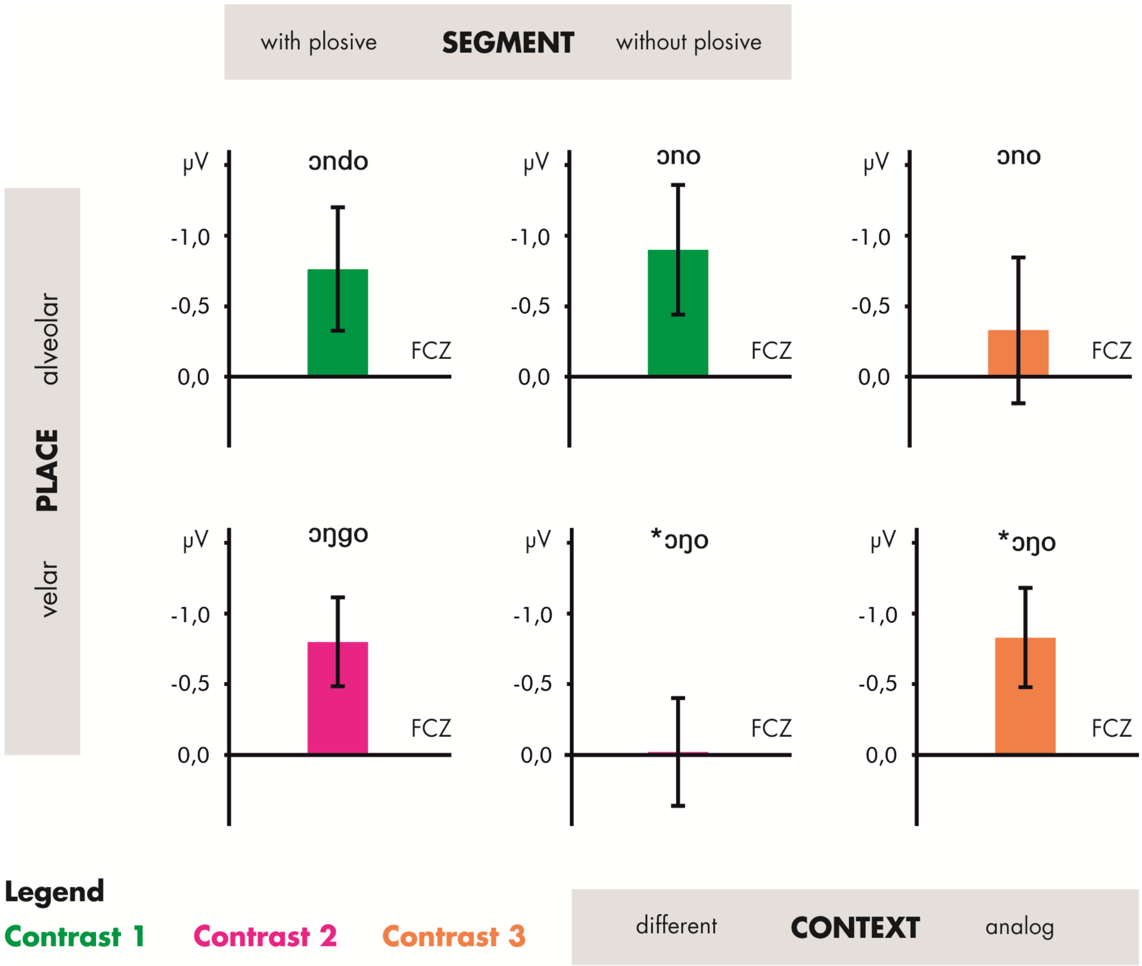

Figure 3. Mean MMN amplitudes in microvolt, based on re-referenced data at FCZ electrode, represented separately for the pseudowords in every contrast. Data are collapsed across experiments 1 and 2. Error bars indicate standard error of mean. The three between-subjects factors are depicted by grey fields.

3.2.3 Discussion

In this MMN study, we aimed to understand in more depth how the human speech processing system deals with phonotactically illegal speech input. According to our first goal, we intended to provide further evidence for the impact of a phonotactic grammar, i.e. categorical phonotactic knowledge, in sublexical and pre-attentive speech processing stages. For this purpose, we extended the phenomenal scope by focusing on g-deletion, a suitable restriction in German.

Moreover, according to our second goal, we aimed to disentangle two different processing strategies that seem to be effective in processing illegal nonsense sound sequences, termed “repair” and “violation detection” throughout this article. In this scope, we investigated whether processing of phonotactically illegal stimuli can be modulated by the immediate phonological context. By focusing on g-deletion, a suitable phonotactic restriction in German for this purpose, we further aimed to extend the phenomenal scope of pre-attentive processing effects of phonotactic grammar. Concretely, we expected repair to occur if the violated phonological constraint was made salient by the immediate context of the illegal stimulus. In contexts neutral with respect to the violated phonological constraint, however, illegality was expected to be processed explicitly. To test these hypotheses, we embedded one and the same illegal pseudoword as deviant into different legal standard stimulations while obtaining the MMN.

Finally, according to our third goal, we employed two analog stimulus sets produced by a naïve and a trained speaker, respectively. Thereby we intended to further validate our findings by exploratively estimating the impact of the speaker’s level of proficiency in producing illegal sound sequences.

As expected from previous MMN studies, the illegal stimulus led to specific effects on the MMN when presented as deviant, thus confirming Hypothesis 1. Moreover, these effects were different depending on the phonological context as established by the standard stimulus in the oddball conditions, confirming Hypothesis 2: Where standard and deviant could be related by the phonological rule of g-deletion (as in Contrast 2), a reduced MMN response was obtained to the illegal deviant stimulus as compared to the legal control condition (Contrast 1). We take this as indicator of a pre-attentive phonotactic repair mechanism, thus corroborating the phonotactic repair hypothesis. By contrast, where the standard stimulus did not set up any specific phonological context (as in Contrast 3), an enhanced MMN response was observed for the illegal deviant. This result is interpreted as an index for a post-hoc evaluation mechanism resulting in the detection of the phonotactic violation, supporting the phonotactic evaluation hypothesis. Although we observed significant overall differences in MMN amplitudes between the two data sets (obtained from the stimuli produced by the naïve and the professional speaker), this effect did not interact with any phonological factor of interest. So far, we take these observations as evidence in favor to our hypotheses.

In the following paragraphs, we aim to further explicate the cognitive processes potentially responsible for our findings. By this, we focus on different sources of predictions in order to differentiate several evaluative contexts regarding the illegal deviant stimulus within the oddball stimulation.

- Repair or evaluation – competing top-down influences

In any classic oddball stimulation, an ad-hoc model is set up based on the repetitive standard stimulation that, on the one hand, forms strong expectations about new incoming sensory information, and on the other hand, provides the template against which any next stimulus will be compared. This predictive model, however, is commonly seen as the outcome of a profound analysis integrating both sensory information and, if appropriate, information accessed from the long-term memory. Applied to (nonsense) linguistic stimulation material, such a relation between the mental representation of the standard and abstract phonological information accessible from the long term memory has been stated by Eulitz and Lahiri (2004). In other words, the language processing system presumably initiates a phonological evaluation of the standard stimulus by activating any relevant phonological information from LTM and integrating it into the standard model. From this perspective, abstract phonological information is accessible for deviant prediction and evaluation by means of the specific standard within the oddball stimulation.

However, despite from this local standard representation, there are other sources of information available to generate expectations and to provide evaluative criteria regarding new sensory information. Any linguistic stimulus (aside from the artificial situation of isolated speech sounds) establishes its own on-line phonological context providing probabilistic and categorical information about possible and probable following phonemes. These sources of information are located in and are directly accessible from LTM, and they do, in our view, not necessarily need to be mediated by means of the predictive standard model. Applied to the present study, the velar nasal /ŋ/, in German, per se limits the set of possible successors to (syllable offset, non-moric vowel, velar plosives), while explicitly prohibiting the sequence of */ŋ+full vowel/. Thus, with respect to the static phonological knowledge and irrespective of the specific standard it is embedded in, the deviant stimulus *ɔŋo violates both specific phonological expectations about possible successors and abstract well-formedness criteria.

According to the above assumptions, the processing of the incoming deviant is influenced by predictions from different sources. We suppose that the impact of these predictions on deviant processing depend on how strong these predictions cooperate. In contrast 3, the illegal deviant *ɔŋo was processed under the strong expectation of the standard sequence ɔno. Phonetic/phonological deviancy was detectable with nasal onset, i.e. from 120ms post stimulus on, and accordingly accompanied by an MMN response. At this point in time, however, the segmental context of the deviant (so far just the vowel) was still too small to set up any specific phonotactic expectation about any forthcoming segments. Thus, the following dorsal nasal had to be analyzed entirely bottom-up and treated as an informative phonetic cue indicating a significantly limited set of possible successors. These inherent expectations, however, were then violated with the onset of the unexpected final (full) vowel and were reflected by a post-hoc phonotactic error response (taken as the outcome of a phonotactic evaluation mechanism).

By contrast, no error signal was found in response to the illegal deviant in contrast 2, even thou, in this contrast, both the segmental deviancy from the standard model and the inherent phonotactic violation coincided at the absence of the plosive at 220 ms post onset. Here, both expectations seem to interact in a way that perceptually evened out the inherent discrepancy. We argue that, in this contrast, the phonological knowledge about g-deletion had already been activated during the initial stage of the standard stimulation. Moreover, integrated into the predictive model of the local standard it had been kept activated during the whole stimulation. Under these strong expectations, the velar nasal in the deviant had not to be analyzed veridically as a phonetic cue. Instead, it could be automatically transformed into the pre-activated abstract sequence /ŋ+g/ as predicted by the phonotactic grammar in German (taken as the outcome of a phonotactic repair mechanism). This interpretation is supported by contrast 1. Serving as a phonotactically legal baseline this contrast proved that the simple segmental deviancy of an additional/missing voiced plosive is phonetically sufficient to elicit a classic segmental MMN response without any interactions with inherent phonotactic limitations.

- No evidence for a “legalized” standard representation of *ɔNo

Apart from its role as deviant among the standard ɔŋgo (context strongly related to g-deletion), the illegal stimulus was presented as deviant among the standard ɔno (context neutral with respect to g-deletion) and, furthermore, it served as standard itself in the two reversed oddball conditions (fairly context free).Apart from its role as deviant among the standards ɔŋgo (context strongly related to g-deletion) and ɔno (context neutral with respect to g-deletion) the illegal stimulus served as standard itself in the two reversed oddball conditions of contrasts 2 and 3 (fairly context free). Thus, the question remains how the phonotactic illegality has been processed when being presented repetitively during stimulation and how it possibly affects the predicitve model that is build based on the standard. As we have discussed this issue in detail in previous publications (e.g., Steinberg, et al., 2010a, 2011), we will shortly line out our main point of views here: Phonotactic illegality, if not automatically repaired prior to further phonological processing, has been found to be reflected by a genuine mismatch response in the ERP of the illegal deviant. However, illegal standards did not elicit any measurable error signal. Two hypothetical explanations have been provided by us: First, the lack of error signal might result simply from neural adaptation in the repetitive standard stimulation. Second, according to the proposal of Eulitz and Lahiri (2004), the standard might be represented in a phonologically canonical format, thus neutralizing phonological mismatches. According to German phonology as outlined in the introduction, prevocalic /ŋ/ can always be traced back to underlying |ŋ+g|.

Given this, the illegal standard *ɔŋo would consequently be represented /ɔŋgo/, i.e. underspecified with regard to the place of articulation of the nasal and including the underlying plosive. However, this explanation clearly does not hold with our data: Applied to contrast 2, the deviant /ɔngo/ would be fully compatible with this hypothetical standard representation; by contrast, a distinct MMN is observable due to the additional plosive in the deviant. This indicates that the standard was, in fact, represented without any plosive. As for contrast 3, the data rule this explanation out twice: The nasal in the standard representation must have been specified for velar place of articulation given the large MMN response elicited by the nasal [n] in the deviant. Furthermore, a reconstructed underlying /g/ in the standard representation would have set up another deviancy, even if this fell within the temporal window of integration, thus not necessarily causing a second mismatch response. In sum, there is no evidence in our data supporting a canonical sparse standard representation of the illegal stimulus, leaving the adaptation hypothesis as explanation for the missing error signal in the illegal standards.

However, we wish to emphasize that, with respect to contrast 3, the data are supportive for the featurally underspecified lexicon (FUL)[5] model as proposed by Lahiri and Reetz (2002; 2010). The asymmetry in the nasal-related MMN in this contrast conforms with the predictions of FUL and are in line with a wide range of empirical support (e.g., Cornell, Lahiri, & Eulitz, 2011; Cornell, Lahiri, & Eulitz,

[5] The FUL model assumes that the underlying featural representation of a speech sound (as it is stored in the mental lexicon) does not contain feature specifications that are linguistically redundant. In FUL, all coronal consonants are assumed to be underspecified for their place of articulation. For speech recognition, FUL proposes a three-way matching mechanism between the actual speech input and the underlying representation: match (input and underlying representation have the same featural specification), mismatch (input and underlying representation have different feature representations), and no-mismatch (input is mappped against an underspecified, ‘empty’, featural slot).

2013; Eulitz & Lahiri, 2004; Scharinger, Bendixen, Trujillo-Barreto, & Obleser, 2012): a velar nasal in the deviant (ɔŋo) is subsumable under an underspecified nasal in the standard (ɔŋo) ("no mismatch" condition), indicated by a weak MMN response. By contrast, the alveolar nasal in the deviant (ɔno) is not compatible with a fully specified velar nasal in the standard (ɔŋo) ("mismatch" condition), reflected by a strong MMN amplitude.

- Methodological and conceptual considerations

Designing phonotactically illegal but nonetheless naturally sounding speech material whose articulatory implementation conform native phonetic knowledge is always a challenge (Steinberg, 2014; Steinberg, et al., 2012). In the present study, we included two analog stimulus sets produced by one non-professional (Experiment 1) and one professional (Experiment 2) speaker of German. Even though an overall main effect in the ERP amplitudes was found between the data sets (ERPs obtained in Experiment 1 being stronger than those from Experiment 2), this between-subjects factor did not interact with any phonological factors of interest. However, when focusing on the separate data sets (results are provided in the supplementary material), some interactions (even if they numerically point in the intended direction) appear to lack sufficient statistical power. As this limitation mainly pertains to experiment 1, the stimuli of which were articulated by the untrained speaker, we assume that a minimum level of speech proficiency is mandatory for a speaker to guarantee native-like articulatory integrity in producing illegal sound combinations.

So far, we investigated and discussed the processing effects of phonotactic illegality within the framework of the MMN, mainly focusing on a pre-attentive stage of processing. This approach imposed us on certain methodological pre-settings and theoretical assumptions regarding the underlying cognitive mechanisms of MMN elicitation. However, for the sake of converging validity, the phenomenon of interest is to be investigated and discussed by means of more than one methodology. In this regard, an interesting complement to our rationale comes from a study of Breen et al (2013) who used an active priming paradigm to investigate the perception of illegal onset-clusters in English (e.g. *dla), both behaviorally (similarity ratings) and by means of ERPs (elicited by the target). Their results provide behavioral evidence for a perceptual repair of the illegal input, however, this pattern was not reflected in the electrophysiological data: the target ERPs subsequent to an illegal (and likely repaired) prime did not facilitate target processing to the same extent as a prime identical to the target. This finding indicates that phonotactically illegal input even if it triggers active perceptual legalization (based on later attentional processes) may cause increased processing effort on prior, presumably automatic processing stages. In this respect, their findings contradict the results of Dehaene-Lamberts et al. (2000) who reported electrophysiological evidence of successful perceptual repair by means of an active MMN paradigm.

Several methodological parameters, however, vary across the mentioned studies, potentially affecting the results and thus limiting their validity (Breen, et al., 2013; Davidson & Shaw, 2012): From the linguistic point of view, there are the phonological and phonetic properties of the particular phonotactic constraint under investigation that may have had a specific impact on processing, depending on whether the illegality affects the featural, segmental or even prosodic level of representation. Also, it might be relevant whether the illegality becomes manifest with stimulus onset or later during the crucial item. With respect to the experimental design, several differences have to be kept in mind when comparing results: Dehaene-Lamberts et al. (2000) and Breen et al. (2013) used active protocols demanding participants attention focused on the stimulation while we used passive settings. Also, the particular protocol probably taps on other cognitive functions such as concrete predictions that notably affect stimulus processing. Differently from the study of Breen et al. and from our own studies, Dehaene-Lamberts et al (2000) used a cross-linguistic between-subjects design that was possibly affected by differing language-specific phonetic experiences between the samples. Additional work integrating several methodological approaches will be necessary to extract the essential phonological aspects of illegality processing from methodological side effects.

3.3 Final Devoicing

Final devoicing operates on what has classically been analyzed as a voicing contrast (see e.g. Rubach 1990, Hall 1992). Jessen and Ringen (2002) have argued that the contrast instead involves the feature [spread glottis] for the plosives, and Beckman et al. (2009), building on this, have argued that the German fricatives are specified for both [spread glottis] and [voiced] (see also Vaux 1998 for arguments that voiceless fricatives are specified [+spread glottis] across languages). In the present experiment, we employed a voicing distinction in fricatives. Assuming such a dual specification, we expect no effects of lexical underspecification, which have been argued to affect MMN by Eulitz and Lahiri (2004), Cornell et al. (2011, 2013), and Scharinger et al. (2012). Instead, voiced fricatives would be specified [+voiced] and voiceless ones would be specified [+spread glottis] in the mental lexical entries.

The German plosives and fricatives that allow such a laryngeal contrast, here transcribed in terms of voicing, are [p/b, t/d, k/g, f/v, s/z]. Both members of each pair can occur in the onset of a syllable before a vowel. In the classical analysis, the voiced values become voiceless in a syllable coda (Rubach 1990, Hall 1992). Thus, the two genitive forms [ra.d-əs] (Rades 'wheel-GEN') and [ra.t-əs] (Rates 'advice-GEN') distinguish [d] and [t] in the syllable onset before a vowel. However, in the nominative form, without the genitive suffix [-əs], the forms are identically pronounced [raːt] ('wheel'/'advice'). Here, /d/ and /t/ are in the syllable coda and only the voiceless pronunciation [t] occurs. The change from /d/ in the mental lexical entry /rad/ to [t] in the pronunciation in coda position is called final devoicing. There are different suggestions about the best way of describing and capturing the correct environment (see e.g. Lombardi 1991, 1999, Steriade 1997, Beckman et al. 2009). There is also a debate about whether the voicing neutralization is phonetically complete (see e.g. Port and O'Dell 1981, Beckman et al 2009). However, it is clear that the change takes place obligatorily in a set of core environments that include the word-final position, that there are no lexically marked exceptions, and that the change is bounded by the word.

Hwang et al. (2010) showed that English voicing agreement in consonant clusters as in their pseudoword stimuli [ʊts] and [ʊdz] lead to processing difficulties in non-agreeing clusters like *[ʊds], which were not found in non-agreeing clusters like *[ʊtz]. Poeppel and Monahan (2011; p. 947f.) refer to results of a related MEG experiment in which a distinction between [ʊts] and *[ʊds] was found around 150 ms after the onset of the fricative. Hwang et al. (2010) interpret their results in terms of the underspecification for voicing of voiceless plosives in English postulated by Lombardi (1991, 1999): Speakers predict a following voiced sound after [d] in *[ʊds] but do not predict a following voiceless sound after [t] in *[ʊtz] because [t] is underspecified for voicing. We think that this explanation may apply to a phonological surface structure in which the position preceding the fricative is conceivably one of laryngeal neutralization (see Steriade 1997). It is conceivable that the voicelessness of [t] preceding a fricative may be accounted for by laryngeal neutralization by the processing system, while the voicing of [d], if followed by a fricative, can only be licensed by agreement with the fricative. Our assumptions about the underlying featural specifications in German are thus not in conflict with these interesting results.

The voicing distinction between obstruents in German is phonetically implemented in several ways depending both on the manner of articulation and on the relative position of the sound. As we will focus on fricatives in intervocalic and final position in our study (see section 1.5. Experimental design for details), we will limit the following overview over phonetic voicing cues to these instances. Because in German, the phonetic implementations of the voicing contrast are – at least partly – neutralized in final positions, we also attend to voicing parameters obtained in languages like English in which FD is not operative. As shown by various phonetic studies (for an overview of the literature on the voicing distinction in German fricatives, see Jessen, 1998, pp. 65-66, 96; phonetic evidence on English fricatives is reviewed for instance by Stevens et al. 1992, and Maniwa and Jongman 2009) the voicing distinction between fricatives is mainly coded by three kinds of parameters: First, the duration of the fricative (as reflected by the duration of friction noise in the acoustic signal) is shorter in voiced compared to voiceless fricatives. Preceding full vowels show – to some degree – the reversed durational pattern. Second, there are several spectral indicators for fricative voicing, most importantly the presence of periodic low-frequency energy during the fricative (reflecting vocal fold vibrations). Additionally, voiced fricatives are characterized by a lower Center of Gravity (COG) and higher variance compared to voiceless fricatives (cf. Maniwa and Jongman 2009). Third, fricative voicing is indicated by a greater extent of the F1 transitions of preceding or following adjacent vowels (e.g. Stevens et al 1992). Accordingly, vowels following a voiced fricative have been shown to begin with lower F0 than vowels after voiceless ones (Jessen, 1998).

There is an interesting asymmetry in the roles played by standard and deviant in processing the oddball stimulation. Since the standard is repeated a number of times, and the pauses between the repetitions give sufficient time for it to be recognized as a particular phonological sound or sound sequence, it seems, put simply, that the expectation for another standard is phonologically represented, or represented in more abstract terms (Näätänen 2001). The deviant, on the other hand, is just coming into the system and its initial processing is ongoing at the time when the mismatch against the standard arises.

Eulitz and Lahiri (2004), Scharinger et al. (2012), and Cornell, Lahiri and Eulitz (2011, 2013) have argued that the standard can in certain ways be seen as similar to a mental lexical entry, likewise abstractly represented, and that the deviant can be seen as similar to the incoming acoustic information that the system seeks to match to an abstract lexical entry. They have argued that lexical underspecification of features matters for MMN in a way that can be understood in these terms. A crucial aspect of the asymmetry for our experiment is that it provides a direction of application of final devoicing: If it applies in an oddball protocol in early pre-attentive processing, it should apply in pairs in which the standard corresponds to a possible mental lexical entry to which final devoicing could apply, and in which the deviant can be seen as similar to a spoken word to which final devoicing has applied. For ease of exposition we therefore adopt some notation of Cornell, Eulitz and Lahiri (2013). The standards are provided with slashes /./ and the deviants with squared brackets [.]. This is parallel to the phonological notation where /./ is used for mental lexical entries and [.] for what is heard.

The experiment reported here addressed German final devoicing in pre-attentive phonological processing. The stimuli employed were [vus], *[vuz], [vusə] and [vuzə] as depicted in Figure 4. We concentrated on four pair-wise contrasts each of which was employed twice with reversed roles of standard and deviant in the stimulations, resulting in a total of eight experimental conditions. As explained above, we marked the standard stimuli of the experimental conditions with /./ and the deviants with [.]. Our expectations were based on the similarities of standard stimuli to abstract phonological lexical representations on the one hand, and deviants to phonetic surface representations that are close to the acoustic input on the other hand (Näätänen 2001, Eulitz and Lahiri 2004).

Contrast 1 is what would be an alternation in German for an underlying voiceless /s/. There is no change in voicing of the fricative. Hence, we expected MMN elicitation both for the additional vowel in 1a and for the missing vowel in 1b.

In contrast 2, stimuli differ with respect to the voicing of the second fricative. In condition 2a, this change is phonologically unmotivated. Furthermore, the deviant differs due to its additional vowel. Here, we expected an MMN to be elicited by each of these changes. Run the other way around, as in condition 2b, the whole contrast between standard and deviant could be interpreted as an alternation in German for an underlying voiced /z/ in /vuzə/, with final devoicing in [vus]. While the phonetic differences were the same in 2a and 2b, we expected the respective MMN patterns to reflect that standard and deviant were phonologically related due to final devoicing in 2b but not in 2a.

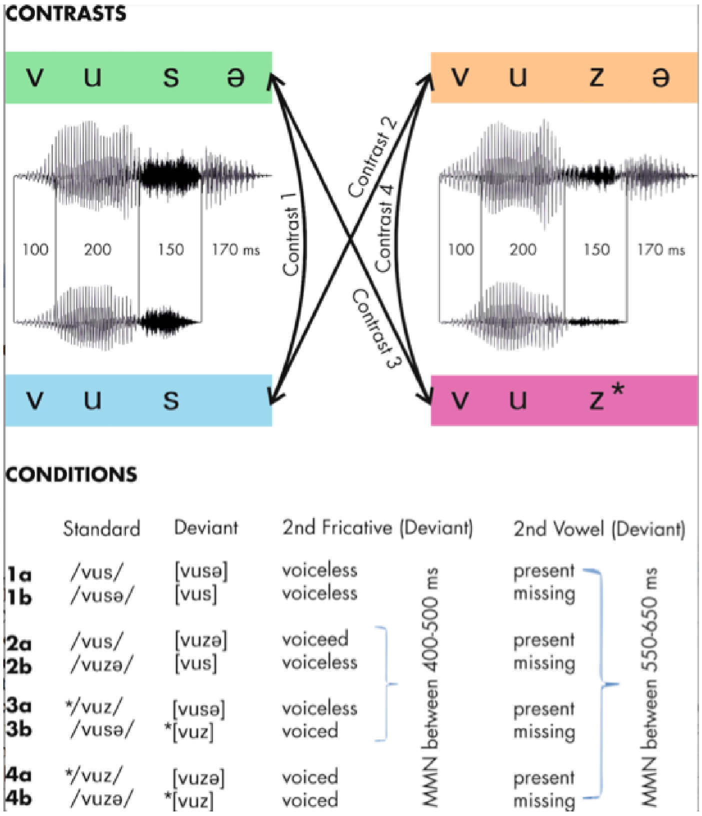

Figure 4. Stimuli and Design. Representative waveforms depict the segmental and temporal structure of each of the four pseudowords. Arrays indicate the employed pair-wise oddball contrasts. The experimental conditions are listed below stating the segmental deviation criteria between the respective standard and deviant along with the time-ranges in which MMN responses were expected in the difference-waves.

Contrasts 3 and 4 employ the ill-formed stimulus *[vuz]. It is ill-formed because final devoicing would obligatorily turn it into [vus] in German. In contrast 3, the deviant enters into an unmotivated voicing alternation of the second fricative in both conditions. In addition, the stimuli differed with respect to the presence or absence of the final vowel. Consequently, we expected to find MMN for each of these segmental differences in both conditions. Furthermore, we were interested in whether we would find an additional effect attributable to the ill-formedness of *[vuz] when being presented as deviant. As the phonological violation *[vuz] coincided with the absence of the second vowel, we expected these mismatch responses to overlay, as reflected by a larger MMN amplitude in condition 3b compared to 1b.

In contrast 4, there is no change in voicing with respect to the second fricative, so no effects were expected in the first time window. With respect to the difference in the final vowel, we were interested in whether condition 4a would show reduced MMN compared to the remaining a-conditions; this may be expected as it would reflect a remedy of the violation of final devoicing in the standard *[vuz]. In condition 4b, we again expect superimposed mismatch effects due to the ill-formedness of the deviant *[vuz] and due to the missing second vowel as in condition 3b.

3.3.1 Materials and methods

Participants

Sixteen volunteers participated in the study (four male; median age was 26 years, range from 22 to 33), all of them right-handed and monolingual native speakers of German. Handedness was assessed using an inventory adopted from Oldfield (1971). All participants reported normal auditory and normal or corrected-to-normal visual acuity and no neurological, psychiatric, or other medical problems. They gave informed written consent. The study conformed to The Code of Ethics of the World Medical Association (2013, Declaration of Helsinki).

Materials

As described, four pseudowords were used as stimuli: [vus], [vusə], *[vuz], and [vuzə]. The stimuli are phonotactically well-formed in German, except for the non-word *[vuz], which fails to have undergone final devoicing. The stimuli [vus], [vusə], *[vuz], and [vuzə] were articulated numerous times by a professional male speaker with a fundamental frequency (F0) of about 100 Hz, and digitally recorded with a 48 kHz sampling rate and a 16 bit resolution using a RME Fireface 800 recording device (Audio AG, Haimhausen, Germany) and a Neumann U87 Microphone (Georg Neumann GmbH, Berlin, Germany).

Stimulus preparation for ERP-experiments on speech processing is always a compromise. The point is to control for lower-level acoustic stimulus characteristics in order to avoid confounds with higher-level linguistic factors while on the same time keeping the stimuli as natural as possible and avoiding artifacts caused by manipulation. To assure some acoustic variability of the stimulus material, 5 different utterances of each pseudoword were selected, resulting in a set of 20 pseudoword stimuli in total (see Eulitz and Lahiri, 2004; Jacobsen, Schröger, and Alter, 2004; Steinberg et al. 2012). However, the conflicting methodological requirements mentioned above concern our study in a special way. The phonological issue under investigation (i.e. the voicing distinction between [s] and [z]) is also coded by the inherent durational differences both between the voided and voiceless fricatives and between the preceding vowels. Other sufficient voicing cues for fricative perception are the presence or absence of low frequency energy during the fricative in the acoustic signal and distinct F1 transitions on vowel-fricative and fricative-vowel boundaries. These cues are also highly reliable at least in intervocalic and final fricative position (as in our stimuli).

Based on these considerations the segmental durations of the stimuli were normalized across contrasts and the voicing distinction were based only on spectral phonetic parameters. Durational normalization was performed using the time-domain pitch synchronous overlap add (TD-PSOLA) algorithm provided by Praat software (Boersma and Weenink, 2010). Segmental durations were equated by setting the initial fricative to 100ms (mean original durations of [v] in ms: [vus] 129, [vusə] 119, [vuz] 102, [vuzə] 109), the full vowel to 200ms (mean original durations of [u] in ms: [vus] 200, [vusə] 188, [vuz] 285, [vuzə] 203), the second fricative to 150 ms (mean original durations of [s] in ms: [vus] 329, [vusə] 177; of [z] in ms: [vuz] 259, [vuzə] 119) and the final vowel to 170 ms (mean original durations of schwa in ms: [vusə] 189, [vuzə] 167). Afterwards, intensities were normalized using the root mean square (RMS) of the whole sound file.

Theoretically, the duration normalization bore two risks: First, originally voiced fricatives might be perceived as voiceless after the relative lengthening of the fricative and the shortening the preceding vowel. Second, the contrary effect might have occurred to the originally voiceless fricatives. However, our ERP data clearly indicate that a distinction in the fricative has been detected in both directions in contrasts 2 and 3 (see results section). Nevertheless, acoustic analyses were performed after the manipulation procedures to ensure that sufficient phonetic information was left in the stimulus material coding the voicing distinction between the fricatives [v] and [s] and to test potential interactions with the syllabic position of the fricative. Both offset F1 transitions of the first vowel, and the first two spectral moments of the fricative were tested.

Formant measures were taken from each single stimulus file as mean values within 20 ms analysis windows by using the linear prediction-based burg method (as implemented in Praat) with a pre-emphasis frequency of 50 Hz. F1 measures were taken from the mid part (190-210ms) and from the final part of the vowel (280-300ms) by automatically determining maximally two formants below 2000 Hz. F1-transitions were analyzed by means of a univariate mixed design analysis of variance (ANOVA) with the within-items factor TRANSITION (mid vowel /vowel offset) and the between-items factors FRICATIVE (voiceless/voiced) and SYLLABLE (mono-/bisyllabic). A main effect of the factor TRANSITION ($F_{1,16} = 6.2$; $p = .024$; $\eta p2 = .279$) and a significant interaction TRANSITION*FRICATIVE ($F_{1,16} = 5.9$; $p = .027$; $\eta p2 = .269$) were found. As expected from the literature, the first vowel formant showed a significantly falling pattern when preceding the voiced fricative (F1 mid vowel: 357 Hz/F1 vowel offset: 316 Hz; main effect TRANSITION in a broken down two-way ANOVA with TRANSITION and SYLLABLE: $F_{1,8} = 13.8$; $p = .006$; $\eta p2 = .633$) while the F1 transition remained steady-state when being followed by the voiceless fricative (F1 mid vowel: 356 Hz/F1 vowel offset: 355 Hz; no significant effects). Note that there was no significant effect by the factor SYLLABLE.

To analyze the spectral qualities of the fricatives, FFT power spectra were calculated using a 50 ms Hann window that was centered over the mid part of the fricative (350-400ms). From these spectra, center of gravity (COG) and standard deviation (SD) were obtained. The spectral measures of the fricatives were analyzed by means of a multivariate analysis of variance (MANOVA) with the between-items factors FRICATIVE and SYLLABLE as described before. A significant main effect of FRICATIVE indicates spectral differences between [s] and [z] (Pillai's trace = .532; $F_{2,15} = 8.5$; $p = .003$; $\eta_p^2 = .532$). The factor SYLLABLE did not show any significant effects. The univariate analyses revealed that voiceless fricatives were characterized by significantly higher COG frequencies ([s] 7712 Hz/[z] 5908 Hz: $F_{1,16} = 10.9$; $p = .004$; $\eta_p^2 = .406$), and lower Standard Deviation (([s] 2094Hz/[z] 2730 Hz: $F_{1,16} = 14.0$; $p = .002$; $\eta_p^2 = .466$) compared to the voiced fricatives. Based on this it was assumed that the voicing distinction in the stimulus material was sufficiently coded phonetically even though durational voicing cues had been neutralized by manipulation.

Experimental design and procedure

As described above, four different experimental contrasts were employed: [vus] vs. [vusə] (contrast 1), [vus] vs. [vuzə] (contrast 2), *[vuz] vs. [vusə] (contrast 3) and *[vuz] vs. [vuzə] (contrast 4). Each pair-wise contrast was presented twice in oddball sequences, both using one pseudoword as standard (85 % of the trials = 1360 items) and the other as deviant and the other way around (reversed oddball-design), resulting in 8 experimental conditions. Oddball sequences of 1600 trials in total were presented per condition, using all tokens of each pseudoword equally. Standard and deviant stimuli were delivered in pseudo-randomized order forcing at least two standards to be presented between successive deviants. Oddball conditions were then divided into two technical blocks each, resulting in a total of 16 stimulation blocks per participant. Sessions were split into two parts, so the second half of each condition was presented on a second day. Stimulus sequences were presented with a stimulus onset asynchrony randomly varying from 550 to 900 ms in units of 10 ms. The order of the experimental blocks was counterbalanced between participants. Participants were seated comfortably in a sound-attenuated and electrically shielded experimental chamber, and they were instructed to ignore the auditory stimulation while watching a self-selected silent subtitled movie. Stimuli were presented binaurally at 53 dB SPL through headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). Loudness was measured by means of an artificial head (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany). All participants reported that they were able to ignore the auditory stimulation. Informal questioning of the participants revealed that they had perceived all stimulus types as speech sounds. A whole experimental session lasted approximately 180 minutes (plus additional time for electrode application and removal) including ten short breaks of about two minutes each.

Electrophysiological Recordings

The EEG (Ag/AgCl electrodes, Falk Minow Services, V-Amp EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from 26 standard scalp locations according to the extended 10–20 system (American Electroencephalographic Society, 1994; FP1, FP2, F7, F3, FZ, F4, F8, FC5, FC1, FCZ, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2) and from the left and right mastoids. The reference electrode was placed on the tip of the nose, and an additional electrode placed at AFZ was used as ground during recording. Electroocular activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) was obtained from the right eye by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 10 kΩ. On-line band-pass filtering of the EEG and EOG signals was carried out using a

0.011 Hz high-pass and a 100 Hz low-pass filter. The signal was digitized with a 16 bit resolution at a sampling rate of 500 Hz.

Data Analysis

Off-line signal processing was carried out using EEP 3.0 (ANT Neuro, Enschede, Netherlands). EEG-data were band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass) (cf. Schröger 2005). EEG epochs with a total length of 1050 ms, time-locked to the onset of the stimuli and including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each stimulus probability (standard, deviant), for each pseudoword, and for each participant.

The ERP responses to the first five stimuli per block as well as to each standard stimulus immediately following a deviant were not included in the analysis. Epochs showing an amplitude change exceeding 100 μV at any of the recording channels were rejected. In the present study, an average of 15.1 % (standard deviation 6.2 %) of the trials per participant was rejected prior to ERP computation. Grand-averages were subsequently computed from the individual-subject averages.

To quantify the full MMN amplitude, the scalp ERPs were re-referenced to the averaged signal recorded from the electrodes positioned over the left and right mastoids. This computation results in an integrated measure of the total neural activity underlying the auditory MMN (e.g., Schröger, 2005).

Deviant-minus-standard difference waveforms were calculated for each pseudoword per oddball condition by subtracting the ERPs elicited by the standard point by point from the ERPs elicited by the original deviant obtained from the same oddball condition, i.e. the MMN elicited by [vusə] in condition 1a was quantified as difference between the deviant ERP from [vusə] and the standard ERP from [vus]. Original contrasts from the same block were opted in order to prevent superimposing effects from the block context to affect our comparisions.

Deviance-related effects (as the MMN) were quantified by measuring the ERP amplitudes as mean voltages in a fixed analysis window of 40 ms (for the width of the analysis window, cf. Luck, 2005, pp. 234). These windows were adjusted a posteriori on the basis of the grand-averaged deviance-minus-standard difference waves (cf. Picton et al., 2000). Separate windows for each condition and for each deviation were adjusted by identifying the peak latencies of any distinguishable negative-going deflection (averaged across F3, Fz, F4, C3, Cz, and C4 electrode positions) within a priori determined time ranges. First, any effect due to the voicing alternation in the second fricative was expected to occur between 400 and 500 ms post stimulus onset (note that this latency equals 100 to 200 ms after the onset of the differing fricatives). This voicing alternation only occurs in contrasts 2 and 3. Second, deviations due to the presence or absence of the second vowel were expected to affect processing within the time range of 550 to 650 ms post stimulus onset (i.e. 100 to 200 ms after the offset of the fricative/onset of the final vowel). In singular cases, additional earlier or later time windows were analyzed in an exploratory approach.

Statistical analyses were performed with SPSS (IBM SPSS Statistics 21). As the MMN is known to be maximal over frontal scalp areas (cf. Kujala, Tervaniemi, and Schröger 2007), the analyses were based only on the F-line positions by collapsing the ERPs obtained at F3, Fz, and F4 into one single measure. Separately for each analysis window, an overall univariate repeated-measures analysis of variance, henceforth ANOVA, was run including the within-subjects factors STIMULUS PROBABILITY (standard/deviant), CONTRAST (depends on the window), and VOWEL (additional vowel in the deviant is present/missing). Afterwards, analyses were broken down if appropriate. Finally, comparisons between conditions relating to the hypotheses were performed using repeated-measures ANOVAs with the factors introduced above. Only significant main effects of the factor STIMULUS PROBABILITY and interactions with this factor were reported. The level of type 1 error was set to $p < .05$ and, in case of multiple post-hoc comparisons, Bonferroni correction was applied. If the sphericity assumption was violated (indicated by the Mauchly test), the original degrees of freedom were provided along with the Greenhouse-Geisser-epsilon. Finally, partial eta-squared ($\eta p2$) effect sizes were given for all significant effects.

3.3.2 Results

The ERP results for all conditions are depicted in Figure 5. Also, this figure shows the respective analysis windows for each effect. The outcomes of the statistical analyses based on these windows are presented below separately for each analysis window. The outcomes of the statistical analyses based on these windows are presented below separately for each analysis window. In Figure 6, topographical maps of the analyzed MMN-effects are provided separately for each condition and time window.

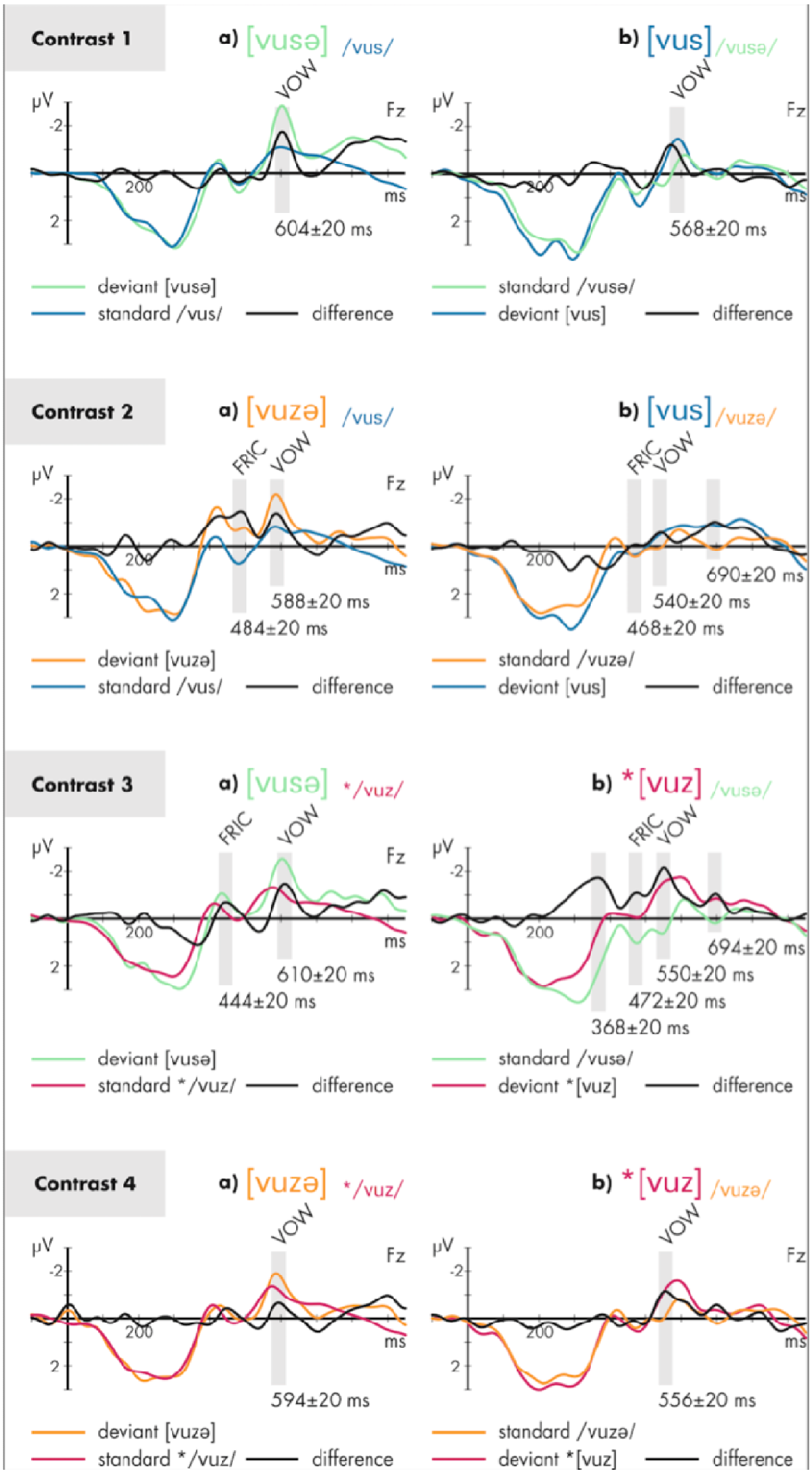

Figure 5. Grand-averaged and re-referenced ERPs elicited by the stimuli of Contrasts 1 to 4 depicted separately for conditions a (left) and b (right) at FZ electrode site. The color of the ERPs codes the stimulus that elicited that ERP. The black line represents the difference wave calculated for each condition. The grey bars indicate the time windows of statistical analyses. FRIC indicates that the marked effect is attributed to the voicing change in the second fricative, VOW indicates that the marked effect is attributed to the final vowel.

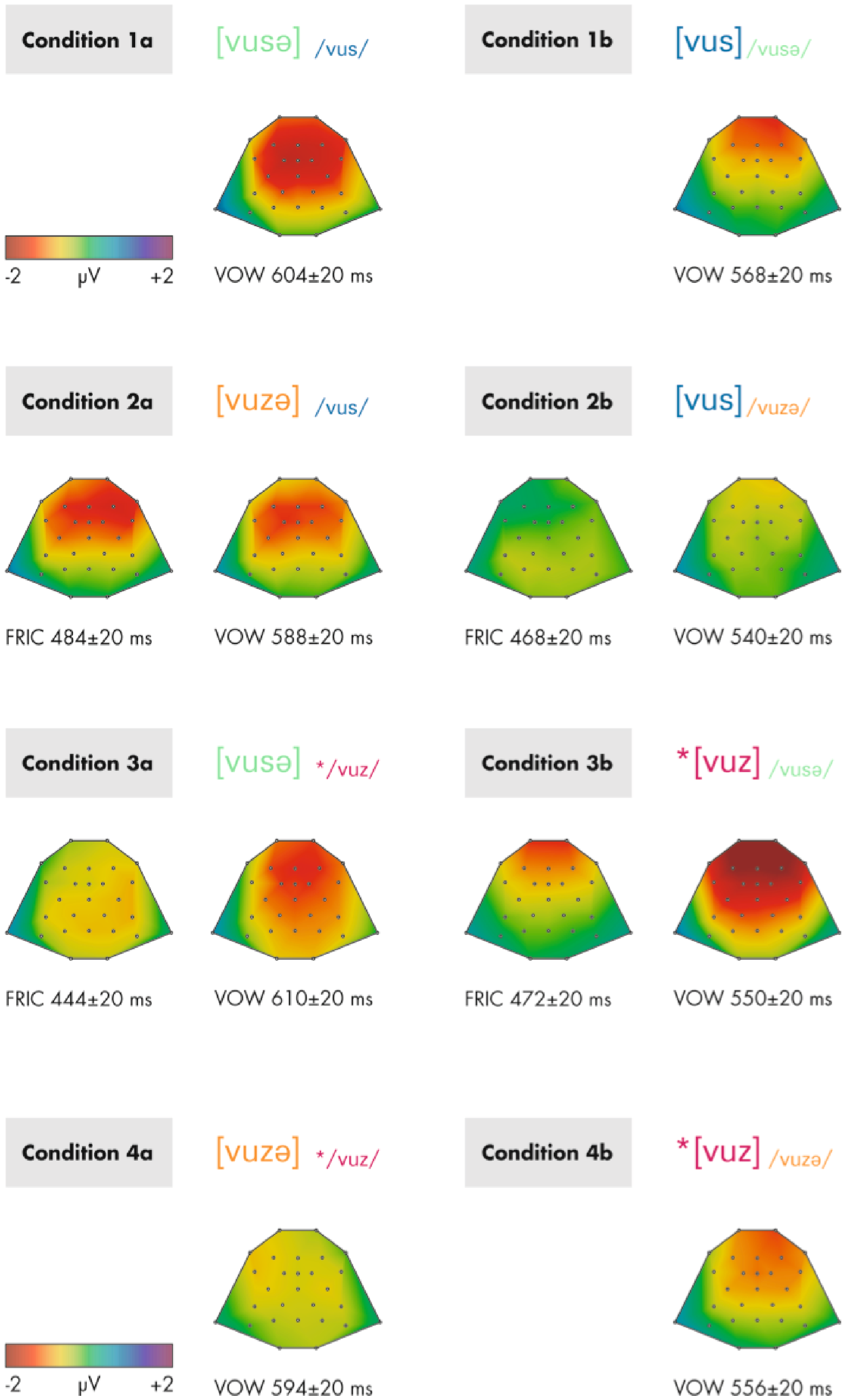

Figure 6. Topographical maps of the analyzed deviance-minus-standard differences obtained from grand-averaged re-referenced data within the time windows stated below each map. Depicted are contrasts 1 to 4 separately for conditions a (left) and b (right). FRIC indicates that the marked effect is attributed to the voicing change in the second fricative, VOW indicates that the marked effect is attributed to the final vowel.

- Analysis of the voicing change in the fricatives (contrasts 2 & 3)

For the MMN responses to the fricatives (FRIC in Figure 5) the overall ANOVA with the factors STIMULUS PROBABILTIY (standard/deviant), VOWEL (additional/missing), and CONTRAST (2/3) revealed a significant main effect of the factor STIMULUS PROBABILITY ($F_{1,15} = 17.9$; $p = .001$; $\eta p2 = .544$), indicating the presence of an MMN across all conditions, and a significant interaction STIMULUS PROBABILITY*VOWEL* CONTRAST ($F_{1,15} = 5.9$; $p = .028$; $\eta p2 = .284$), indicating different amplitudes of the MMN responses across conditions. Broken-down analyses were calculated separately for each contrast: In contrast 2, the main effect for STIMULUS PROBABILITY ($F_{1,15} = 9.1$; $p = .009$; $\eta p2 = .387$), and also the interaction STIMULUS PROBABILITY*VOWEL ($F_{1,15} = 5.9$; $p = .028$; $\eta p2 = .282$) were significant, the latter indicating a stronger MMN response in condition 2a compared to condition 2b. In contrast 3, only a significant main effect for STIMULUS PROBABILITY was obtained ($F_{1,15} = 9.3$; $p = .008$; $\eta p2 = .384$).

- Analysis of the effect due to the change in the final vowel (all contrasts)

For the MMN responses to the additional or missing vowel (VOW in Figure 5) the overall ANOVA with the factors STIMULUS PROBABILITY (standard/deviant), VOWEL (additional/missing), and CONTRAST (1/2/3/4) revealed a significant main effect for STIMULUS PROBABILITY ($F_{1,15} = 25.0$; $p < .001$; $\eta p2 = .625$), as well as significant interactions STIMULUS PROBABILITY*CONTRAST ($F_{3,45} = 4.0$; $p = .026$; $\varepsilon = .725$; $\eta p2 = .209$) and STIMULUS PROBABILITY*CONTRAST*VOWEL ($F_{3,45} = 3.1$; $p = .039$; $\varepsilon = .928$; $\eta p2 = .172$). Next, analyses were broken down by the factor VOWEL. Comparing the MMN amplitudes for the a-conditions, only a significant main effect for STIMULUS PROBABILTY ($F_{1,15} = 9.3$; $p = .008$; $\eta p2 = .382$) was found, but no interaction with this factor. For the b-conditions, the main effect STIMULUS PROBABILITY ($F_{1,15} = 21.1$; $p < .001$; $\eta p2 = .585$) and the interaction

STIMULUS PROBABILITY*CONTRAST ($F_{3,45} = 6.2$; $p = .002$; $\varepsilon =$; $\eta_p^2 = .292$) were significant. This interaction indicates differences in MMN amplitudes due to the missing final vowel across the contrasts. There was a priori only interest in potential differences between conditions 1b and 3b, both sharing the same legal standard [vusə]. A broken-down ANOVA with STIMULUS PROBABILITY and CONTRAST (1/3) revealed a significant main effect for STIMULUS PROBABILITY ($F_{1,15} = 35.4$; $p < .001$; $\eta_p^2 = .703$), and a significant interaction between both factors ($F_{1,15} = 5.1$; $p = .039$; $\eta_p^2 = .252$), indicating stronger MMN amplitudes for 3b compared to 1b.

- Explorative analyses of earlier and later effects (contrasts 2 & 3)

In conditions 2b and 3b, unexpected deviance-related effects were found in a time range later than 650ms post stimulus onset. These effects were analyzed as described above: a significant main effect of STIMULUS PROBABILTIY ($F_{1,15} = 19.5$; $p = .001$; $\eta_p^2 = .565$) was found but no interactions with this factor. Because of its latency, it seems possible that this effect reflects morphological processing (see Royle et al. 2010). This is conceivable if the additional vowel is processed as a morphological suffix.

Furthermore, a strong deviance-related effect was observed in condition 3b that appeared in an unexpected early time range before 400 ms, i.e. before the onset of the deviating fricative. Because of its latency, this effect seemed to be temporally related to the later part of the first vowel [u]. This effect was compared with a corresponding time window in condition 1b (360±20 ms) that shared the legal standard stimulus [vusə]. A significant main effect STIMULUS PROBABILITY ($F_{1,15} = 39.1$; $p < .001$; $\eta_p^2 = .722$) was found as well as a significant interaction STIMULUS PROBABILITY*CONRAST ($F_{1,15} = 9.0$; $p = .009$; $\eta_p^2 = .374$), indicating a stronger deviance-related response in 3b compared to 1b.

3.3.3 Discussion

- Remarks on contrast 1

Our statistical assessment employed condition 1b as a comparison condition for condition 3b (see section 4.4). However, contrast 1 (pairing the legal stimuli [vusə] and [vus] with no voicing change) is here also briefly considered on its own. Visual inspection of the difference waves for contrast 1 in Figure 5 shows distinct MMN responses that are attributable to the presence vs. absence of the final vowel, but no further effects, in particular no effects between 450–550 ms where differences attributable to the fricative would occur. This provides some assurance that effects attributed to fricatives in other conditions were not general consequences of our stimulus contrasts in which stimuli with and without a final vowel are compared. There is, for example, a distinction in syllabification. The a-conditions are syllabified like 1a: [vu.se] while the b-conditions are single syllables like 1b: [vus]. This distinction could in principle have phonetic correlates in regard to the extent of coarticulation of the [s/z] with the preceding vowel. Recall that the phonetic analysis of the stimuli did not detect any such differences. Condition 1 suggests that such differences, if they exist after all, also did not lead to observable effects in the difference wave.

- Evidence for the relevance of final devoicing in contrast 2

The following sketch shows condition 2b next to condition 2a. We included a dot to mark the syllable boundary in [vu.zə].

	Standard	Deviant	Comment
2a:	/vus/	[vu.zə]	significantly stronger effect for voicing change than 2b
2b:	/vu.zə/	[vus]	related by final devoicing

There is a significant difference between conditions 2a and 2b in the processing correlates of the voicing change in the fricative. The MMN effect due to the voicing mismatch in condition 2a was absent in condition 2b, where the voicing change was motivated by final devoicing. This significant difference between conditions 2a and 2b is here interpreted as evidence for the relevance of final devoicing in pre-attentive processing.

- Remarks about reactions to the final vowel in condition 2b

We turn to some remarks about the MMN response due to the additional/missing vowel in conditions 2a and 2b. The plots in Figure 5 suggest that the response attributable to the missing final vowel in the deviant of condition 2b was also reduced. We here want to comment this impression for the benefit of possible future experiments that might investigate such an effect more specifically. The observation suggests that the expectation of any upcoming auditory event, which is violated in the deviant and shown by the MMN, is not limited to the expectation of just another standard stimulus. It seems, instead, that this expectation can be modulated by what is found earlier in the deviant. The system seems to have related /vuzə/ and [vus] by final devoicing. If the system possesses knowledge of the environment of final devoicing, it will then expect the absence of a vowel following [vus], since final devoicing would not have applied in the presence of a following vowel. (Similar expectations could also be modulated by phonetic factors that might allow the anticipation of the absence of a final vowel. However, the reduced MMN response to the missing final vowel seems to be specific to condition

2b, where final devoicing has applied.) It is also possible, then, that the standard /vuz+ə/ and the deviant [vus] were processed as morphologically related by the omission of an inflectional element [ə] in the deviant, with phonological adjustment due to final devoicing. It seems conceivable that this was related to the late deviance-related effect that was observed about 250 ms after the missing vowel had become detectable.

We note that we have argued (see experiment three in this work; Jacobsen et al., 2013) against the assumption of successive MMN responses in case of mismatching monosyllabic vowel-consonant sequences, where both the vowel and the consonant differed. However, the case at hand is different in an important aspect: The second deviation in the present contrast pairs, namely the missing or additional final vowel in contrasts 2 and 3, did not just involve a distinct sound, but established a distinction in syllable structure between standard and deviant. By this, the present stimulus contrasts were clearly different not just at the segmental but also at suprasegmental representation levels.

- Evidence for the relevance of final devoicing in condition 3

It was seen in the presentation of the results that condition 3b and condition 1b both have MMN responses attributable to the missing vowel, and that both effects furthermore differ significantly in strength. This is illustrated in the following sketch.

	Standard	Deviant	Comment
1b:	/vusə/	[vus]	significant effect for missing final vowel
3b:	/vusə/	*[vuz]	significantly stronger effect for missing final vowel

It was suggested that this is evidence for a superposed effect of the ill-formedness of the deviant *[vuz] in condition 3b, which becomes manifest in the signal simultaneously with the absence of the final vowel. This distinction provides further evidence for the relevance of final devoicing in pre-attentive processing.

- Remarks on reactions to the first vowel in condition 3

The comparison between conditions 1b and 3b is repeated in the following, this time highlighting a significant distinction that was found post hoc: Condition 3b showed an effect at the time at which the second part of the vowel [u] is expected to be processed. The distinction to 1b was seen to be significant.

	Standard	Deviant	Comment
1b:	/vusə/	[vus]	no processing effect attributable to [u]
3b:	/vusə/	*[vuz]	significant effect attributable to [u]

This effect in 3b may be related to the anticipation of [z] during the vowel [u] due to coarticulatory cues. It is furthermore possible that phonetic factors allowed an early prediction of the syllable structure. The system might have noticed in *[vuz] during the vowel that there would be an upcoming voiced fricative within the same syllable, in violation of final devoicing. If so, the early strong MMN effect before 400 ms in condition 3b might already be a first electrophysiological response to the ill-formedness of the deviant.

3.4 Successive deviants

This study investigated whether two successive deviants occurring sequentially in two segments of a single hierarchically organized auditory object would elicit two Mismatch Negativities. To this end, natural samples of two vowel-consonant (VC-) syllables were used as stimuli.

With respect to streams of acoustic stimuli potentially eliciting MMN, two terms shall be distinguished in the context of the present paper: oddball sequences entailing double deviants, and those entailing successive deviants. The term double deviant shall refer to two features or feature dimensions within a given stimulus that are synchronously deviating from the standard, e.g. frequency, intensity or location. For several different feature dimensions, the MMN to double deviants has been reported to be as large as the sum of the MMNs to each of the individual deviants differing in only one feature (e.g., Paavilainen et al., 2001, Takegata et al., 1999, Wolff and Schröger, 2001). Therefore, the term double (or multiple) deviants is used referring to two (or more) deviations occurring synchronously within one auditory object.

Successive deviants, on the other hand, follow one another such that the two features do not differ from the standard at the same time, but successively. If the same stimulus was presented twice as a deviant in an oddball protocol, the second MMN has been reported to show a reduction in amplitude of about 50 %. This is called “deviance-repetition effect” (Sams et al., 1984; Müller et al., 2005a). If in such a protocol the second deviant differed from the first, the reduction of the second MMN was much smaller (Müller et al., 2005b; Näätänen et al., 2007; Nousak et al., 1996).

One interesting case of successive deviants pertains to instances of two deviations occurring within one auditory object. Müller and Schröger (2007) used triangle waves as stimuli and presented them in pairs. In one condition, the successive deviants were either in one pair or in different pairs. The first deviant at the second position of a pair elicited a larger MMN than the first deviant at the first position of a pair. Therefore, the authors concluded that deviants "occur less likely within an object than at the beginning of a new object". In their study, no clear MMN for the second deviant in one pair was observed.

Two important aspects governing the integration of stimuli into a single auditory object or into separate ones are the SOA (Stimulus Onset-Asynchrony, the temporal distance between the onsets of the stimuli), and the auditory grouping. Müller and Schröger (2007) found a decrease of the second MMN for 250 ms SOA intervals compared to 500 ms SOA intervals for the paired and non-paired condition which could result from integrated processing of successive sounds. Yabe et al. (1997) only found a clear MMN for stimulus omissions, when the SOA was shorter than 150 ms, which gave them a hint for indicating the temporal window of integration (TWI). Horvath et al. (2007) and Sussman et al. (1999) showed that the duration of the TWI is approximately between 200 and 250 ms. The TWI sets the upper time limit for the integration of separate elements into a single object (see Näätänen and Winkler [1999] for more information about stimulus representation, SOA and the temporal window of integration). In auditory grouping, the auditory system defines successively occurring sounds to specific auditory streams, which is based on similarities of pitch or location (Bregman, 1990).

It is an open question, however, which effects successive deviants being implemented in two separate elements of a hierarchically organized auditory object have on auditory deviance detection. This issue was addressed in the present study.

Speech is a special case of hierarchically organized auditory objects. A speech syllable can be analyzed as a phonological complex object with a nonlinear hierarchical structure (for reviews, see Blevins 1995, Zec, 2007). In nonlinear phonology, a syllabic sound structure is modeled with at least two different representation tiers that are related by means of association lines. On the one hand, the phonemes of a syllable are subsequently represented as nodes on the segmental tier. On the other hand, they are integrated under syllable nodes being represented on the syllable tier. Most models additionally assume several sub-syllabic constituents like syllable onset and rhyme, consisting of nucleus and coda (Fudge, 1969; Selkirk, 1982; Halle and Vergnaud, 1980, among many others). Simple VC sequences like the stimuli used in the present study can therefore be considered as phonologically complex auditory objects: They consist of clearly distinguishable phonological units (the phonemes) that are integrated into a phonological structure (the syllable) by means of abstract phonological principles. Being contrasted in oddball blocks, the deviant syllables entail deviations on both of the representation levels, the syllable tier and the segmental phonemic tier. Consequently, mismatch responses could either be elicited due to the change of the whole syllable or due to the onset of differing phonemes.

This study sets out to investigate whether phonemic changes on the segmental tier of a VC syllable would function as successive deviants despite being integrated into a single auditory object on the syllable tier at the same time. To this end, stimuli from two previous studies were re-used, [ɔx] and [ɛʃ] (Steinberg et al., 2010a; 2010b). In those studies, standard and deviant syllables were contrasted in a way that they differed only with respect to one phoneme. Deviants implemented as a vowel change at syllable onset ([ɛʃ] versus [ɔʃ]) elicited a Mismatch Negativity at about 160 ms successive to an N1 modulation at about 108ms (Steinberg et al., 2010a, Exp. 2; Steinberg et al., 2010b). Deviants occurring as a change in the second segment of the syllable ([ɔx] versus [ɔʃ]) elicited an MMN at about 216 ms after stimulus onset (Steinberg et al., 2010a, Exp. 1). These findings provided an estimate of the time courses of potential effects obtained in the present study. They also showed that it is possible, in principle, to obtain deviance effects at these positions with a single hierarchically structured auditory object (see also Kirmse et al. [2008] or Sussman et al. [2004] for comparable stimuli).

The goal of this study was to test, whether changes in the initial vowel and in the following consonant of a VC-syllable would elicit two separate MMNs.

Two different outcomes were possible. First, the deviants could elicit two MMNs, one triggered by the initial vowel change and another one due to the subsequent consonant change. Because phonemes are constructed automatically and function as a contrastive unit, they should also be processed reliably outside the focus of attention. The segmental tier and the syllable tier are represented separately and the segmental analysis should represent both phoneme changes. Moreover, the discussed previous studies showed that the difference between the vowels and the fricatives are detectable.

Alternatively, only one MMN might be elicited by the deviant syllable for two different reasons. On the segmental representation level, the change of the second phoneme, i.e. the fricative, was fully predictable in the oddball sequence because only one type of deviant syllable was used entailing two successive phonemic deviations. Furthermore, the syllables could be processed as phonologically integrated single objects. Given this, in turn, the fricative change might function as an anticipated double deviant or it could be of no consequence at all.

All in all, the following hypotheses were derived. (1) The change of the vowel should elicit an MMN at about 160 ms, possibly along with a previous deviance-related N1 modulation (as reported by Steinberg et al. 2010a, b). (2) If a second mismatch response would be elicited due to the fricative change, it should occur not before 200 ms (Steinberg et al, 2010a), considering the genuine MMN latency of about 100 to 250 ms after deviation onset (e.g., Schröger 1998) and the time that it takes to process the prior change of the vowel change. (3) If there wouldn't be any second MMN present in the signal, the processing of the fricative change could either contribute to the first MMN due to anticipation, or the fricative change could have had no genuine effect at all. These latter alternatives will be tested by comparing the results of the present study by those obtained by Steinberg et al (2010a, b) where the vowel-induced MMN was free from any anticipatory effects, as in those studies there was no subsequent phonemic change in the deviant syllables in the oddball design.

3.4.1 Materials and methods

Participants

Sixteen volunteers participated in the study (15 male; median age 25 years; range 22 to 27; all right-handed), all of them monolingual native speakers of German. Handedness was assessed using an inventory adopted from Oldfield (1971). All participants reported normal auditory and normal or corrected-to-normal visual acuity and no neurological, psychiatric, or other medical problems. They gave informed written consent. The study conformed to The Code of Ethics of the World Medical Association (2008, Declaration of Helsinki). Three additional participants had to be excluded from further analyses because more than 20 % of the trials in the EEG-signal had to be rejected due to artifacts (eye movements, blinks, etc.)

Materials

Two VC-syllables were used as stimuli: [ɔx] and [ɛʃ]. None of these syllables have lexical meaning in German. Ten different tokens of each syllable type were employed resulting in a total number of 20 stimuli per stimulation. The stimuli were matched in segmental duration and intensity. The vowel took 100 ms and the fricative lasted 180 ms. Note that stimuli designed for a previous study (see Steinberg et al [2010a] for a detailed description of stimulus preparation) were re-used.

Experimental design and procedure

The syllable types were contrasted in reversed oddball conditions. In each oddball sequence, one syllable type served as standard (85 % of the trials = 1360) and the other syllable type was used as deviant. Oddball stimulus sequences of 1600 trials in total were presented per condition, using all tokens of each syllable type equally. Standard and deviant stimuli were delivered in pseudo-randomized order forcing at least two standards to be presented between successive deviant syllables. Both oddball sequences were split into two technical blocks each, resulting in a total of four stimulation blocks per session. Stimulus sequences were presented with a stimulus onset asynchrony randomly varying from 550 to 900 ms in units of 10ms. The order of the experimental blocks was counterbalanced between participants. Participants were seated comfortably in a small separate experimental room, and they were instructed to ignore the auditory stimulation while watching a self-selected silent subtitled movie. Stimuli were presented binaurally at 56 dB (A) (58dB SPL) through headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). Loudness was measured with an artificial head (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany). All participants reported that they were able to ignore the auditory stimulation. Informal questioning of the participants revealed that they had perceived all stimulus types as speech sounds. An experimental session lasted approximately 50 minutes (plus additional time for electrode application and removal) including three breaks between the four blocks of about two minutes each.

Electrophysiological Recordings

The EEG (Ag/AgCl electrodes, Falk Minow Services, V-Amp EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from nine standard scalp locations according to the 10–20 system (American Electroencephalographic Society, 1994; F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) and from the left and right mastoids. The reference electrode was placed on the tip of the nose, and an additional electrode placed at FCz was used as ground during recording. Electroocular activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the right eye by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering of the EEG and EOG signals was carried out using a 0.011 Hz high-pass, a 100 Hz low-pass, and 50 Hz notch filter. The signal was digitized with a 16 bit resolution at a sampling rate of 500 Hz.

Data Analysis

Off-line signal processing was carried out using EEP 3.0. EEG-data were band-pass filtered with a finite impulse response filter: 2501 points, critical frequencies of 1.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 650 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard, deviant), for each of the syllable types, and for each participant. The ERP responses to the first five stimuli of each block as well as to each standard stimulus immediately following a deviant were not included in the analysis. Epochs showing an amplitude change exceeding 100 μV at any of the recording channels were rejected. Grand-averages were subsequently computed from the individual-subject averages.

To quantify the full MMN amplitude, the scalp ERPs were re-referenced to the averaged signal recorded from the electrodes positioned over the left and right mastoids. This computation results in an integrated measure of the total neural activity underlying the auditory MMN (e.g., Schröger, 1998).

Deviant-minus-standard difference waveforms were calculated for each syllable type by subtracting the ERPs elicited by the standard point by point from the ERPs elicited by the physically identical deviant obtained from the reversed oddball condition. Deviance-related effects were quantified by measuring the ERP amplitudes as mean voltage in a fixed analysis window of 40 ms. This window was centered a posteriori on the peak latency of the grand-average difference wave (averaged over C3, CZ, C4, F3, FZ and F4 electrode sites).

Statistical analyses of the present data were performed separately for each analysis window by means of univariate four-way repeated-measures analyses of variance (ANOVA) with the factors Condition (standard, deviant), Syllable ([ɔx], [ɛʃ]), Position (F-, C-, P-line) and Laterality (3-, z-, 4- line).

Additionally, the present data were compared with previous results both from Steinberg et al. (2010a) and Steinberg et al. (2010b). To this end, the respective deviance-related ERP responses were analyzed by means of univariate mixed-design ANOVAs with the repeated-measures factors Condition (standard, deviant), Syllable ([ɔx/ɔʃ], [ɛʃ]), Analysis Window (1st, 2nd), Position (F-, C-, P-line) and Laterality (3-, z-, 4- line), and the between-subjects factor Experiment (present study, previous study). With respect to the third set of hypotheses, crucial effects to test here were interactions in which the factors Condition, Analysis Window and Experiment would be involved.

The level of the type 1 error was set to $p < 0.05$. For effects with more than one degree of freedom, the original degrees of freedom are reported along with the corrected probability as well as the epsilon value (Greenhouse-Geisser). Finally, partial eta-squared (ηp2) effect sizes are given for all reported effects. Only significant results relevant to our hypotheses, i.e. effects involving the factor Condition (as for the comparisons with previous data sets: interactions involving the factors Condition, Analysis Window, and Experiment), were reported and discussed in the text of the article.

3.4.2 Results

An average of 2.9 % (standard deviation 4.8%) of the trials per participant was rejected prior to ERP computation. The deviant-minus-standard difference waveforms showed two negative-going deflections with maximal amplitudes at F4 for [ɛʃ] and at C4 for [ɔx]. The deviance-related effects of the syllable [ɛʃ] at F4 were found at 116 ms after stimulus onset with a peak amplitude of -0,855 μV (first peak), and at 174 ms with a maximal amplitude of -1,754 μV (second peak). For [ɔx], the deflections were maximal at C4 with a latency of 98 ms and a peak amplitude of -0,462 μV (first peak) and at 160 ms with a peak amplitude of -0,611 μV (second peak).

According to the criteria stated above, the analysis windows were set to 108 ± 20 ms (first peak) 168 ± 20 ms (second peak).

For the first window, the four-way repeated-measures ANOVA yielded a significant main effect for the factor Condition ($F_{1,15} = 47.04$, $p < .001$, $\eta p^2 = .76$), indicating significant differences between standards and deviant ERPs of both syllable types. Furthermore, the interactions Condition × Syllable × Position ($F_{2,30} = 6.87$, $\varepsilon = 0.597$, $p < .05$, $\eta p^2 = .31$) and Condition × Laterality ($F_{2,30} = 5.53$, $\varepsilon = 0.844$, $p < .05$, $\eta p^2 = .27$) were significant. Subsequent analyses were run separately for each level of the factor Position, yielding in significant main effects of the factor Condition for all subsets of electrodes (F-line: $F_{1,15} = 25.06$, $p < .001$, $\eta p^2 = .63$; C-line: $F_{1,15} = 48.74$, $p < .001$, $\eta p^2 = .77$; P-line: $F_{1,15} = 46.56$, $p < .001$, $\eta p^2 = .76$), whereas the interactions Condition × Syllable (F-line: $F_{1,15} = 6.04$, $p < .05$, $\eta p^2 = .29$) and Condition × Syllable × Laterality (F-line: $F_{2,30} = 3.98$, $\varepsilon = 0.962$, $p < .05$, $\eta p^2 = .21$) became significant only at frontal electrodes. At the C-line, the interaction Condition × Laterality was found to be significant ($F_{2,30} = 3.68$, $\varepsilon = 0.841$, $p < .05$, $\eta p^2 = .20$).

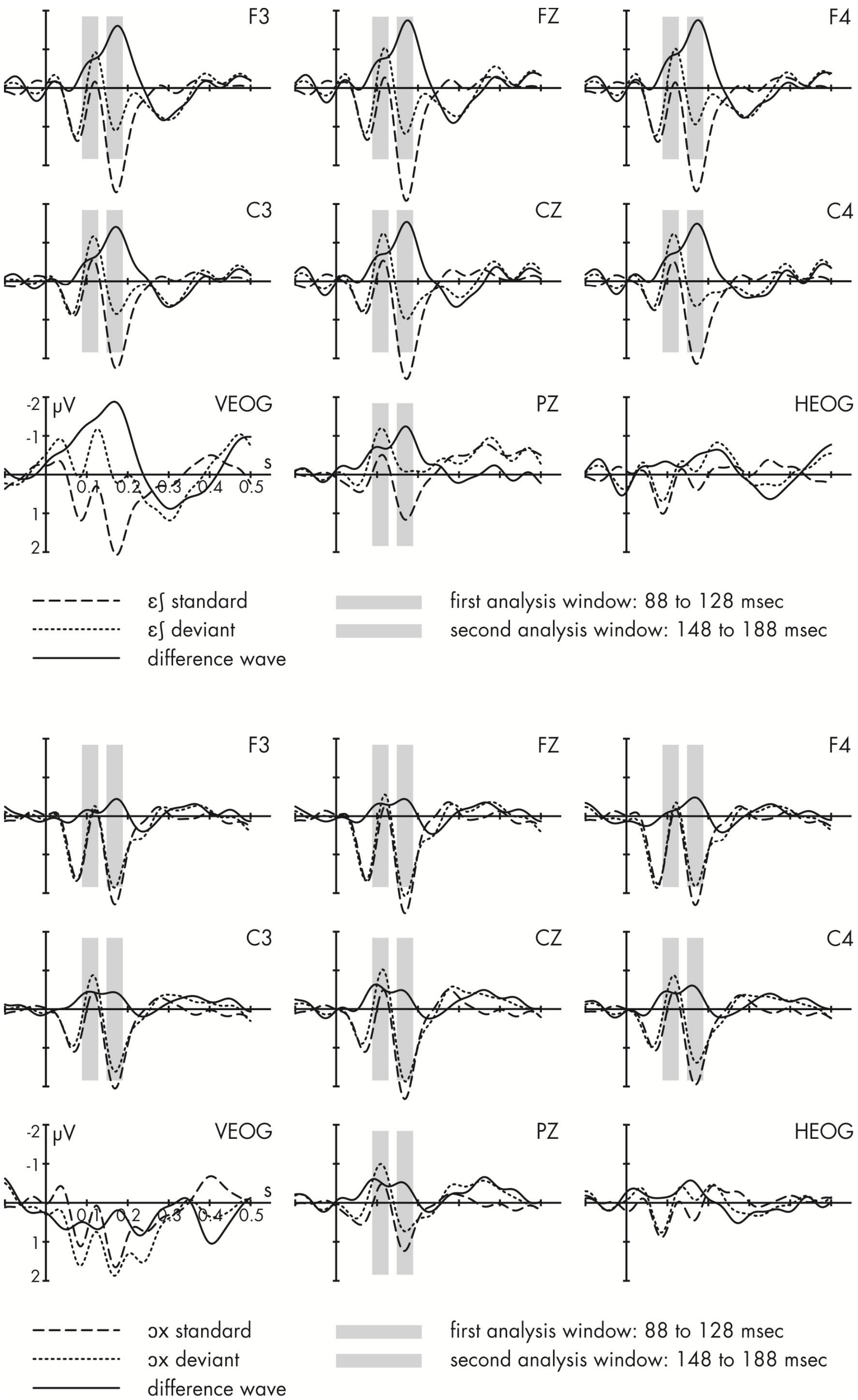

Figure 7. Grand-averaged, re-referenced ERP responses are represented separately for the stimulus syllable [εʃ] (above) and [ɔx] (below) for a subset of electrodes. Shown are ERPs elicited by the deviants (dotted lines), ERPs elicited by the standards (dashed lines), and Deviant-minus-Standard difference waves (solid lines). The bars mark the statistically analyzed time windows of 88 to 128ms (N1) and 148 to 188ms (vowel-induced MMN). Scales are in milliseconds and microvolt.

The ANOVA calculated for the second time window also resulted in a significant main effect for the factor Condition ($F_{1,15} = 65.77$, $p < .001$, $\eta p2 = .81$). Additionally, the interaction Condition × Syllable became significant ($F_{1,15} = 17.97$, $p < .01$, $\eta p2 = .55$), indicating a stronger deviance-related effect for [ɛʃ]. Furthermore, the factor Condition interacted significantly with the factors Position ($F_{2,30} = 7.52$, $\varepsilon = 0.611$, $p < .05$, $\eta p2 = .33$), and Syllable × Position ($F_{2,30} = 7.96$, $\varepsilon = 0.613$, $p < .01$, $\eta p2 = .35$). Subsequent analyses were run separately for each level of the factor Position. Significant effects of the factor Condition were found for all electrode sets (F-line: $F_{1,15} = 54.71$, $p < .001$, $\eta p2 = .79$; C-line: $F_{1,15} = 61.73$, $p < .001$, $\eta p2 = .81$; P-line: $F_{1,15} = 46.35$; $p < .001$, $\eta p2 = .76$). Furthermore, the interaction Condition × Syllable was found to be significant at all positions (F-line: $F_{1,15} = 18.11$, $p < .01$, $\eta p2 = .55$; C-line: $F_{1,15} = 19.74$, $p < .001$, $\eta p2 = .57$; P-line: $F_{1,15} = 8.48$, $p < .05$, $\eta p2 = .36$), and finally, the interaction Condition × Laterality was significant only at the F-line ($F_{2,30} = 4.32$, $\varepsilon = 0.787$, $p < .05$, $\eta p2 = .22$). Subsequent analyses were run separately for [ɔx] and [ɛʃ] by means of three-way repeated-measures ANOVAs broken down by the factor Syllable. For both of the syllables, a significant main effect of the factor Condition was found ([ɔx]: $F_{1,15} = 21.78$, $p < .001$, $\eta p2=.59$; [ɛʃ]: $F_{1,15} = 46.58$, $p < .001$, $\eta p2=.76$). For [ɛʃ], the interaction Condition × Position became significant as well ($F_{2,30} = 14.55$, $\varepsilon = 0.581$, $p < .01$, $\eta p2=.49$).

The mixed-design ANOVA comparing the present data with the findings of Steinberg et al. (2010a) revealed that the crucial interaction Condition × Analysis Window × Experiment did not reach significance ($F_{1,30} = 3.30$, $p = .08$, $\eta p2=.10$), but the interactions of these factors with Position ($F_{2,60} = 5.97$, $\varepsilon = 0.913$, $p < .01$, $\eta p2=.17$), and with Position × Syllable ($F_{2,60} = 15.80$, $\varepsilon = 0.750$, $p < .001$, $\eta p2=.35$) were found to be significant. Further analyses were carried out separately for each syllable by breaking down the interactions with this factor. Only for [ɛʃ], the crucial interaction Condition × Analysis Window × Experiment was found to be significant ($F_{1,30} = 5.53$, $p < .05$, $\eta p2=.16$). No further interactions of this factor combination with topographical factors were found to be significant for any of the syllables. Comparing the findings of Steinberg et al. (2010b) with the present data set, a respective mixed-design ANOVA yielded a significant interaction Condition × Analysis Window × Experiment × Syllable ($F_{1,30} = 4.81$, $p < .05$, $\eta p2=.14$), whereas the interaction Condition × Analysis Window × Experiment per se did not reach significance ($F_{1,30} = 3.12$, $p = .087$, $\eta p2=.09$). Broken down analyses revealed the interaction Condition × Analysis Window × Experiment to be significant for [εʃ] ($F_{1,30} = 6.74$, $p < .05$, $\eta p2=.18$), but not for [ɔx]. There were no further interactions of this factor combination with topographical factors significant for any of the syllables. The crucial results of the comparisons between the data sets are depicted in Figure 8.

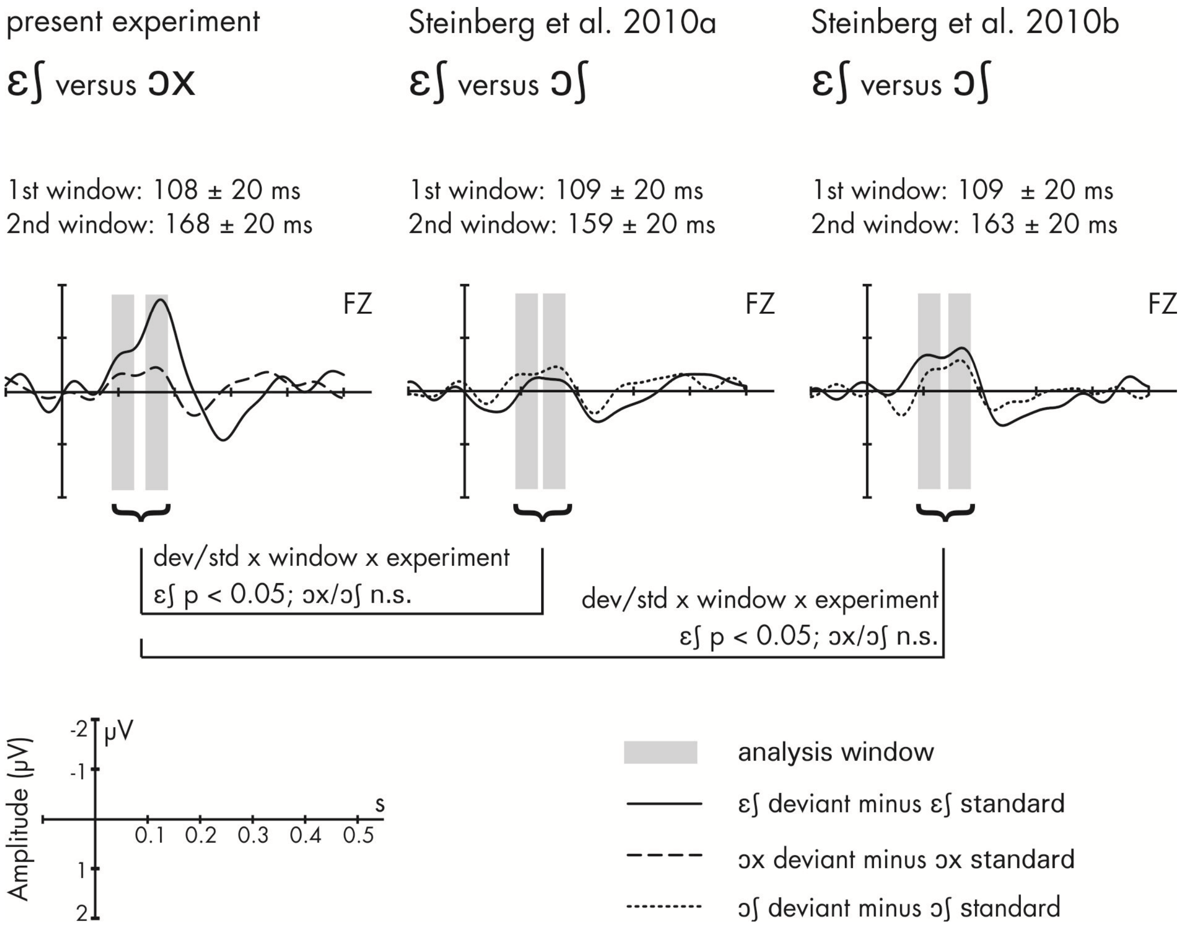

Figure 8. Standard-minus-deviant-difference waves calculated from grand-averaged, re-referenced ERPs are given for the present data set (left) and for data sets of previous experiments (middle, right) for comparison. Difference waves representing deviance-related effects elicited by the syllable [εʃ] are depicted as solid lines, the results for [ɔx] are indexed as dashed line, and the results for [ɔʃ] from the previous studies are given as dotted lines. The critical outcome of the respective statistic comparisons are indexed below the graphs. The bars mark the respective analysed time windows.

3.4.3 Discussion

(1) The present study focused on investigating the auditory processing of successive deviants occurring within one hierarchically organized auditory object. For stimulation, speech syllables consisting of two different phonemes, namely a vowel and a fricative, were used. The syllables functioned as successive deviants on the segmental phonemic level while they represent an integrated holistic object on the syllabic analysis level. In contrast to double deviants, where changes occur at the same time, in successive deviants, the deviations occur one after another. Using a passive oddball paradigm, two different VC-syllables were presented as standards and deviants such that both the vowel and the fricative changed successively in the deviant stimuli. After the occurrence of the first deviating event in the deviant syllable, the second sound change was fully predictable.

(2) Data revealed a replication of previous findings by Steinberg et al. (2010a, experiment 2; 2010b). In those studies, a deviating vowel was found to elicit a biphasic negative-going ERP deflection. In comparable time ranges, the present data revealed two negative maxima in the difference waves, which were interpreted consistently to the previous findings. The first peak found at 98 ms for [ɔx] and at 116 ms for [ɛʃ] was interpreted as deviance-related N1 effect, whereas the second effect occurring at 160/174 ms after stimulus onset was taken as the MMN attributable to the vowel change. However, any further discernible negative-going ERP deflections attributable to the subsequent fricative deviation were not observed. In a previous study (Steinberg et al. 2010a, experiment 1), deviant syllables differing only with respect to the fricative from the standard (but not in both sounds as in the present experiment) elicited an MMN not before 210 ms after stimulus onset (i.e. 110 ms after deviation onset). Taking this, any MMN due to the second sound deviation in the present data would not be expectable in earlier time ranges but rather later since the change of the vowel had to be processed first. So it

is concluded that there was no second MMN response elicited by the successive deviant in the present data.

(3) However, the predictable change of the fricative might have contributed to the MMN elicited at the deviating stimulus onset due to anticipation. Thus, the fricative change might have rather superimposed the initial MMN like a double deviant than having functioned as a successive deviant. If this would be the case, the vowel-induced MMN in the present data should contain some additive mismatch response due to the second sound change as compared to the data of Steinberg et al. (2010a, experiment 2; 2010b) that were obtained from oddball contrasts without any second sound deviation. To test this hypothesis, conjoined analyses of the present data and the data by Steinberg et al. (2010a, b) have been performed. Due to differences in data acquisition between the data sets, the above mentioned hypothesis would be confirmed by means of an ANOVA interaction between the two time windows of the biphasic negative going effect (N1 and MMN) and the experiments. In contrast to the MMN (second time window), which is known to be sensitive to categorical linguistic factors, the N1 effect (first time window) should not have been affected by any change of the upcoming fricative and, consequently, should not differ between the compared data sets. In fact, this critical interaction was found for the deviant syllable [ɛʃ] but not for [ɔx], as can also be seen in Figure 8. At least for the syllable [ɛʃ], there appeared to have been anticipatory effects of the fricative present in the initial MMN.

(4) Unlike the findings from Steinberg et al. the present data show a considerable difference in the amplitudes of the MMNs elicited by the stimulus syllables, which is also reflected by the significant interaction Stimulus x Syllable in the analysis of the present data. The deviant syllable [ɛʃ] presented among the standard syllable [ɔx] elicited a larger MMN response due to the differing vowel than the deviant [ɔx] from the reversed oddball condition. Since the very same

vowel contrast ([ɛ] versus [ɔ]) was used to elicit the initial MMN in both the present one and the previous experiments, the present asymmetry seemed to be caused by the subsequent change of the predictable fricatives. In fact, this asymmetrical pattern is consistent with the results obtained in Steinberg et al. (2010a, experiment 1). In this former study, the contrast between the fricatives [x] and [ʃ] resulted in much larger MMN amplitudes for the syllable [ɔʃ] compared to its counterpart [ɔx]. It has been assumed that genuine differences between the spectral and dynamical properties of the fricatives should have caused this effect. Applied to the present findings, the anticipation of [ʃ] in [ɛʃ] might have contributed to the higher MMN amplitude compared to [ɔx] due to the higher dynamical range and the spectral energy concentration at higher frequencies being typical for this fricative (cf. Gordon, Barthmaier, & Sands, 2002; Johnson, 2002; Jongman, Wayland, & Wong, 2000).

(5) To summarize, VC-syllables were used to investigate the processing of successive deviants within one hierarchically organized auditory object, because they contain successive deviating events (the phonemes on the segmental tier) that are integrated into one hierarchically organized object (the syllable on the syllable tier). The main result of the present study is that the deviant syllable did not elicit two separable MMN responses despite it differed from the standard syllable with respect to both phonemes, i.e. it entailed two differing sound events subsequently in time. This finding indicates that the syllables were not processed as linear chains of differing phonemes but rather as internally structured auditory objects. However, anticipatory effects on the stimulus-initial MMN were obtained, which seem to reflect some processing related to the subsequent fricative change within the deviant syllables. In consistence with Müller and Schröger (2007) who only found two MMNs, when the deviants were in two different auditory objects, the deviant syllables of the present experiment seem to have been analyzed as double deviants rather than as successive deviants. However, the reason for the missing successive MMN effect could have alternatively been caused by the predictability of the

second sound change. An interesting question would therefore be, whether multiple successive deviants (like VC-syllables with varying vowels and fricatives) with at least the second deviating event being unpredictable would elicit a second MMN response.

Part III

4 Summary Section

4.1 General Discussion

With the experiments presented, new data points were added to the literature that establishes the relevance of phonotactic violations in pre-attentive processing: Dehaene-Lambertz, Dupoux and Gout 2000 (who also investigated a phonotactic repair but did not use a clearly pre-attentive protocol), Flagg et al. 2006 (who differ in investigating phonetic cooccurrence effects, in contrast to the phonological ones in these investigations), Mitterer and Blomert 2003 (who used postlexical assimilation in contrast to this lexical phonological process). Steinberg et al. (2010a, 2010b, 2011) had similar results while analyzing the processing of the dorsal fricative assimilation rule. The ill-formed deviant *[ɛx] showed an additional MMN response attributable to the fricative when presented with the standard [ɔx]. The ill-formed deviant [*ɔç] showed an additional MMN response as well when presented with the standard [ɛç]. These responses are temporally separated from the MMN elicited by the distinct vowel or fricative change and attributed to the abstract phonotactic ill-formedness of the deviant. That means that there was only a violation detection and no repair.

For the processing of illegal stimuli, Bonte et al. (2005) argued that less likely/frequent consonant clusters in the deviant lead to less MMN (notkel vs. notsel). This would be the inverse of the effect found here, if it is construed as a frequency effect. On the other hand, Cornell, Eulitz and Lahiri (2011) found opposite results on frequency: low phonotactic probability in the deviant leads to higher MMN. They point out that the results of Bonte et al. might also be an asymmetry due to phonological underspecification. On this perspective, the ill-formedness could in principle be the end-point of a frequency effect.

However, the empirical evidence from these experiments is not yet sufficiently explanatory to fully understand the underlying processes. "The crucial questions under debate concern the time course of phonological decoding, the format in which the incoming information is delivered to further processing, and

[…] the impact of the immediate linguistic context on stimulus perception."[6] There are contrary opinions about the point in time in which the illegal percept is modified. Dehaene-Lambertz et al. (2000) argued that this happens early during perception according to the native phonotactic requirements, whereas Breen et al. (2013) stated that this is performed post-perceptional on the basis of a categorical feature-based representation. The experiments used here are unfortunately not informative enough to fully explain the underlying processes and to show which hypothesis is correct.

The general problem with these studies presented is mainly the production of the stimuli. Stimuli with separate sounds from different recordings, which are concatenated afterwards, do usually not sound natural. These spliced stimuli can cause own effects and sometimes do not have the MMNs which experiments with natural stimuli had (Steinberg et al, 2013). But it is also difficult to articulate ill-formed stimuli, because a lot of speakers would hyperarticulate that point where a phonological rule is violated. The transitions between the sounds are another difficulty, because it is possible that EEG-effects, which should occur because of a sound change, appear earlier, because the transitions of the sounds before that change could contain information about what will happen.

The first experiment showed that there are effects between speakers. The same stimulus set was recorded with a professional and a non-professional speaker. There was an overall main effect between the data sets, but it did not interact with the phonological factors tested. There were yet some interactions in the individual data set from the non-professional speaker which appear to lack statistical power, even if they point in the intended direction. But that shows that it is very difficult for untrained speakers to produce illegal stimuli, but the stronger effects in the data set of the professional speaker may also result from hyperarticulation.

---

[6] Derived from Paper 1

Another problem concerns studies, in which phonological rules are tested, where sounds are deleted or added. The stimuli have different durations. This is especially a problem when the effect you wanted to test and the detection of the different durations are nearly at the same time and therefore result into one MMN, maybe with at least a higher amplitude. This was the case in the second experiment, where the detection of the ill-formed pseudoword took place at the same time as the detection of the different durations.

There was a plan to further investigate the processing of DFA, but the last experiment concerning successive deviants showed only one MMN instead of our expected two. It would be interesting though to have a look at the processing difference between an experiment with [ɔx] and [ɛç] and an experiment with [ɔx] and [ɛʃ] for example. The change of the vowels would elicit an MMN, but the allophonic fricative change in the first experiment is affected by DFA and could therefore elicit no MMN. The last thought here was to test only the fricatives as for example an experiment with only [x] and [ç] and as a control [x] and [ʃ]. It is not clear if these stimuli would be processed as speech sounds and furthermore the DFA-rule actually requires a front or a back vowel. This made it complicated to formulate hypotheses.

The last experiment could also be repeated with a different design, so that the second change would not be fully predictable. Instead of one deviant with a vowel and a fricative change there could be another deviant, in which only one of these sounds change. The system would not know which of the deviants is presented when they are in a random order. Nevertheless the experiment with successive deviants shows a disadvantage of the experimental design. The effects are very small, so every stimulus has to be repeated very often. But this repetition makes both the standard and the deviant predictable, although varieties of one stimulus were recorded and presented.

It would also be interesting to test other phonological phenomenon in German like the nasal assimilation, in which a nasal is assimilated by the following plosive. Possible stimuli would be *[onp] and *[omp]. An MMN-effect would only be expected in the first case, because the change from the alveolar to the labial nasal would here be motivated by the assimilation to the labial plosive. But these stimuli are both illegal stimuli, because there is neither a /np/ cluster nor a /mp/ cluster at syllable coda in German, which shows another difficulty: there are a few phonological rules left, but it is difficult to find good stimuli to test them. The more complex they are, the more effects could appear, which make the results hard to evaluate. A possible change of stimuli would be the insertion of [f] at the end of these pseudowords (to keep a design with stimuli containing one syllable) or a vowel like [ə].

Dialectal varieties, like the different /r/-realizations or the g-spirantization, with participants from different regions in Germany could also be a field to investigate. Like in cross-language studies, it could be that some combinations of sounds elicit MMN for participants from southern Germany, but not for people from northern Germany.

One of the last steps then would be the comprehension of fast spoken language, where assimilation and consonant or vowel epenthesis occur. Is the auditory system used to this kind of language, so that pseudowords, which are similar to natural words, will be repaired, if for example [ə] is missing in a pair like [klɪsən] and [klɪsn]?

But in the end the question is whether this method is really appropriate to test these questions, because the effects as such were very small throughout the experiments. Most of the peaks were about 0.5 µV at most, which is very less compared to other EEG experiments. That is why lots of participants have to do these experiments, in which lots of stimuli have to be tested because of the high number of repetitions, which have to be averaged. But testing phonological rules passively is the only way to have a look inside the system and how it works. In active protocols, repairs could already been made so that the answers are not that reliable.

4.2 General Summary

In this dissertation three experiments were presented and discussed, in which the preattentive processing of phonological rules was examined. In all experiments, event-related brain potentials were used, which are time-locked signal changes in the EEG-signal. The MMN was used as a tool, because it is sensitive to higher order cognitive processes and can reflect the activation of linguistic knowledge. It also provides a temporally highly resolved on-line measure of speech processing and it can be used to investigate fairly automatic processes in passive oddball protocols.

Pair-wise contrasts of different stimuli were presented in oddball sequences, using one pseudoword as standard and the other as deviant and the other way around (reversed oddball-design). The participants were instructed to ignore the auditory stimulation while watching a self-selected silent subtitled movie.

In the first two experiments it was shown that phonological rules are processed automatically outside the focus of attention. Two different German rules were examined and showed similar results, which were also similar to those of Steinberg et al (2010a, 2010b, 2011), who investigated the dorsal fricative assimilation rule (DFA). The first experiment showed that the processing of the illegal pseudoword [*ɔŋo] depends on the context with which it was presented. It caused an enhanced MMN effect when it was presented with [ɔno], which was nearly missing while it was in a condition with [ɔŋgo], which was interpreted as an activation of the g-deletion constraint. The phonotactic modifications were observed only if the relevant phonotactic information was already pre-activated due to the previous context; otherwise, the input was represented veridically and evaluated with respect to its phonological well-formedness afterwards.

The second experiment showed the processing of final devoicing. Conditions with different fricatives showed effects because of the sound change, but the condition with standard [vuzə] and deviant [vus] showed no effect because of the final devoicing constraint. There was also an overlaid effect of the ill-formed [*vuz], but it was difficult to detect it because of the duration differences.

That means that ill-formed syllables or pseudowords will automatically be detected by the mental system, but could also be repaired in a context with the appropriate legal stimulus. These results have also parallels in MMN studies about syntactic violations, where ungrammatical deviants also have MMN effects (Pulvermüller & Assadollahi 2007, Pulvermüller & Shtyrov 2003, Shtyrov, Pulvermüller, Näätänen & Ilmoniemi 2003).

The last experiment tested whether it is possible to elicit two Mismatch Negativities when two sounds in the deviant differ from the standard successively. The syllables [ɛʃ] and [ɔx] were tested, so that the vowel fricative and the fricative changed in the deviant. The second change was fully predictable and the higher amplitude of the MMN only obtained anticipatory effects, it was therefore analyzed as a double deviant.

5 References

American Encephalographic Society (1994). Guideline thirteen: Guidelines for standard electrode position nomenclature. Journal of Clinical Neurophysiology, 11, 111–113.

Avery, P., and Rice, K. (1989). Segment structure and coronal underspecification. Phonology, 6, 179–200.

Beckmann, J., Jessen, M., and Ringen, C. (2009). German fricatives: coda devoicing or positional faithfulness? Phonology, 26, 231–268.

Blevis, J. (1995). The syllable in Phonological Theory. In: J. A. Goldsmith (Ed.), The Handbook of Phonological Theory, Cambridge MA: Blackwell, 206–244.

Boersma, P., and Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.1.44) [Computer program]. Retrieved October 20, 2010, from http://www.praat.org/.

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

Bonte, M.L., Poelmans, H., and Blomert, L. (2007). Deviant neurophysiological responses to phonological regularities in speech in dyslexic children. Neuropsychologia, 45, 1427–1437.

Breen, M., Kingston, J., & Sanders, L. D. (2013). Perceptual representations of phonotactically illegal syllables. Attention Perception & Psychophysics, 75(1), 101–120.

Bregman, A., (1990). Auditory Scene Analysis, 1st edn., Cambridge MA: MIT Press.

Cho, T., & McQueen, J. M. (2006). Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants. Journal of the Acoustical Society of America, 119, 3085–3096.

Coleman, J., and Pierrehumbert, J.B. (1997). Stochastic phonological grammars and acceptability. In: J. Coleman (Ed.), Computational phonology. Third meeting of the ACL special interest group in computational phonology (p. 49–56). Somerset, NJ: Association for computational linguistics.

Cornell, S.A., Lahiri, A., and Eulitz, C. (2011). What you encode is not necessarily what you store: Evidence for sparse feature representations from mismatch negativity. Brain Research, 1394, 79–89.

Cornell, S.A., Lahiri, A., and Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. Journal of Experimental Psychology: Human perception and performance, 39, 757–772.

Cuetos, F., Hallé, P. A., Dominguez, A., & Segui, J. (2011). Perception of prothetic /e/ in #sC utterances: Gating data. Paper presented at the The 17th International Congress of Phonetic Sciences (ICPhS XVII).

Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic constraints. Journal of experimental psychology: Human Perception and Perfomance, 37, 270–282.

Davidson, L., & Shaw, J. A. (2012). Sources of illusion in consonant cluster perception. Journal of Phonetics, 40, 234–248.

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

De Lacy, P. (2007). The Cambridge handbook of phonology. Cambridge, UK: Cambridge University Press.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? Journal of Experimental Psychology: Human Perception and Performance, 25, 1568–1578.

Dupoux, E., Parlato, E., Frota, S., Hirose, Y., & Peperkamp, S. (2011). Where do illusionary vowels come from? Journal of Memory and Language, 64, 199–210.

Durand, J. (1990). Generative and non-linear phonology. London & New York, NY: Longman.

Eulitz, C., and Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. Journal of Cognitive Neuroscience, 16, 577–583.

Féry, C., Hohmann, C., & Stähle, K. (2009). Gradient dorsal nasal in Northern German. In: F. Kügler, C. Féry & R. v. d. Vijver (Eds.), Variation and gradience in phonetics and phonology (pp. 185-214). Berlin: Mouton de Gruyter.

Flagg, E.J., Cardy, J.E.O., and Roberts, T.P.L. (2006). MEG detects neural consequences of anomalous nasalization in vowel-consonant pairs. Neuroscience Letters, 397, 263–268.

Frisch, S.A., Pierrehumbert, J.B., and Broe, M.B. (2004). Similarity avoidance and the OCP. Natural Language and Linguistic Theory, 22, 179–228.

Fudge, E. C., (1969). Syllables. Journal of Linguistics, 5, 253–286.

Gordon, M., Barthmaier, P., and Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. Journal of the International Phonetic Association, 32, 141–174.

Hall, T.A. (1992). Syllable Structure and Syllable-Related Processes in German. Tübingen: Max Niemeyer.

Hall, T.A. (2000). Phonologie. Berlin: de Gruyter.

Halle, M., and Vergnaud, J.-R. (1980). Three dimensional phonology. Journal of Linguistic Research, 1, 83–105.

Hallé, P. A., & Best, C. T. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. Journal of the Acoustical Society of America, 121, 2899–2914.

Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? Journal of Experimental Psychology: Human Perception and Performance, 24, 592–608.

Hayes, B., and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. Linguistic Inquiry, 39, 379–440.

Horvath, J., Czigler, I., Winkler, I., and Teder-Sälejärvi, W. (2007). The Temporal window of integration in elderly and young adults. Neurobiology of Aging, 28, 964–975.

Hwang, S.-O.K., Monahan, P., and Idsardi, W. (2010). Underspecification and asymmetries in voicing perception. Phonology, 22, 205–224.

Itô, J. (1986). Syllable theory in prosodic phonology. New York: Garland Publishing.

Jacobsen, T., Schröger, E., and Alter, K. (2004). Pre-attentive perception of vowel phonemes from variable speech stimuli. Psychophysiology, 41, 654–659.

Jacobsen, T.K., Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (2013). Mismatch Negativity (MMN) to successive deviants within one hierarchically structured auditory object. International Journal of Psychophysiology, 87, 1–7.

Jessen, M. (1998). Phonetics and Phonology of Tense and Lax Obstruents. Amsterdam: John Benjamins.

Jessen, M., and Ringen, C. (2002). Laryngeal features in German. Phonology, 19, 189–218.

Johnson, K., (2002). Acoustic and auditory phonetics, 2nd edn., Oxford: Blackwell.

Jongman, A., Wayland, R., and Wong, S. (2000). Acoustic characteristics of English fricatives. Journal of the Acoustical Society of America, 108, 1252–1263.

Kenstowicz, M. (1994). Phonology in generative grammar. Cambridge, MA: Blackwell.

Kharlamov, V., Campbell, K., and Kazanina, N. (2011). Behavioral and electrophysiological evidence for early and automatic detection of phonological equivalence in variable speech inputs. Journal of Cognitive Neuroscience, 23:11, 3331–3342.

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

Kohrt, M. (1980). Rule ordering, rule extension, and the history of present state of /ng/-clusters in German. Linguistics, 18, 777–802.

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

Lahiri, A., and Reetz, H. (2002). "Underspecified recognition," in Laboratory Phonology VII, eds. C. Gussenhoven & N. Werner. Berlin: Mouton, 637–675.

Lahiri, A., and Reetz, H. (2010). Distinctive features: Phonological under-specification in processing. Journal of Phonetics, 38, 44–59.

Lombardi, L. (1991). Laryngeal features and laryngeal neutralization. Doctoral dissertation, University of Massachusetts, Amherst.

Lombardi, L. (1999). Positional faithfulness and voicing assimilation in Optimality Theory. Natural Language and Linguistic Theory, 17, 267–302.

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

Maniwa, K. and Jongman, A. (2009). Acoustic characteristics of clearly spoken English fricatives. Journal of the Acoustical Society of America, 125, 3962–3973.

Mitterer, H., and Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. Perception & Psychophysics, 65, 956–969.

Müller, D., Widmann, A., and Schröger, E. (2005a). Deviance-repetition effects as a function of stimulus feature, feature value variation, and timing: a mismatch negativity study. Biological Psychology, 68, 1–14.

Müller, D., Widmann, A., and Schröger, E. (2005b). Auditory streaming affects the processing of successive deviant and standard sounds. Psychophysiology, 42, 668–676.

Müller, D., and Schröger, E. (2007). Temporal grouping affects the automatic processing of deviant sounds. Biological Psychology, 74, 358–364.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology, 38, 1–21.

Näätänen, R. (1992). Attention and brain fuction. Hillsdale: Lawrence Erlbaum.

Näätänen, R., Gaillard, A., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. Acta Psychologica, 42, 313–329.

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

Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. Psychophysiology, 42, 25–32.

Näätänen, R., and Winkler, I. (1999). The Concept of Auditory Stimulus Representation in Cognitive Neuroscience. Psychological Bulletin, 125 (6), 826–859.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. Clinical Neurophysiology, 118, 2544–2590.

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

Nousak, J., Deacon, D., Ritter, W., and Vaughan Jr., H. (1996). Storage of information in transient auditory memory. Cognitive Brain Research, 4, 305–317.

Obleser, J., Lahiri, A., and Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. Journal of Cognitive Neuroscience, 16, 31–39.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychololologia, 9, 97–113.

Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. Journal of the Acoustical Society of America, 100, 3831–3842.

Paavilainen, P., Valppu, S., and Näätänen, R. (2001). The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity: 1+1≈2 but 1+1+1<3. Neuroscience Letters, 301, 179–182.

Peltola, M.S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., and Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. Neuroscience Letters, 352, 25–28.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., Mcginnis, M., and Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. Journal of Cognitive Neuroscience, 12, 1038–1055.

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

Poeppel, D., and Monahan, P.J. (2011). Feedforward and feedback in speech perception: revisiting analysis by synthesis. Language and cognitive processes, 26, 935–951.

Port, R.F., and O'Dell, M.L. (1985). Neutralization of syllable-final voicing in German. Journal of Phonetics, 13, 455–471.

Pulvermüller, F., and Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. Progress in Neurobiology, 79, 49–71.

Ramers, K.-H., and Vater, H. (1995). Einführung in die Phonologie (4th ed.). Hürth: Gabel.

Roberts, T.P., Flagg, E.J., and Gage, N.M. (2004). Vowel categorization induces departure of M100 latency from acoustic predictions. Neuroreport, 15, 1679–1682.

Royle, P., Drury, J.E., Bourguignon, N., and Steinhauer, K. (2010). Morphology and word recognition: an ERP approach. Proceedings of the 2010 annual conference of the Canadian Linguistic Association.

Rubach, J. (1990). Final devoicing and cyclic syllabification in German. Linguistic Inquiry, 21, 79–94.

Sams, M., Alho, K., Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. Psychophysiology, 21, 434–441.

Scharinger, M., Bendixen, A., Trujillo-Baretto, N.J., and Obleser, J. (2012). A sparse neural code for some speech sounds but not for others. PLoS ONE, 7, e40953.

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

Schröger, E. (2005). The mismatch negativity as a tool to study auditory processing. Acta Acustica united with Acustica, 91, 490–501.

Schröger, E., Bendixen, A., Denham, S.L., Mill, R.W., Bohm, T.M., and Winkler, I. (2013). Predictive regularity representations in violation detection and auditory stream segregation: From conceptual to computational models. Brain Topography.

Segui, J., Frauenfelder, U., & Hallé, P. A. (2001). Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing. In: E. Dupoux (Ed.), Language, brain, and cognitive developement. Essays in honor of Jacques Mehler (pp. 195-208). Cambridge, MA: MIT Press.

Selkirk, E. O. (1982). Syllables. In: H. van der Hulst & N. Smith (Eds.), The Structure of Phonological Representations, vol. 2, Dordrecht: Foris, 337–383.

Sharma, A., and Dorman, M.F. (2000). Neurophysiologic correlates of cross-language phonetic perception. Journal of the Acoustical Society of America, 107, 2697–2703.

Shtyrov, Y., and Pulvermüller, F. (2007). Language in the mismatch negativity design: Motivations, benefits and prospects. Journal of Psychophysiology, 21, 176–187.

Steinberg, J. (2014). Phonotaktisches Wissen. Zur prä-attentiven Verarbeitung phonotaktischer Illegalität (Vol. 15). Berlin: Walter de Gruyter.

Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (2010a). Preattentive phonotactic processing as indexed by the mismatch negativity. Journal of Cognitive Neuroscience, 22, 2174–2185.

Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (2010b). Activation and application of an obligatory phonotactic constraint in German during automatic speech processing is revealed by human event-related potentials. International Journal of Psychophysiology, 77, 13–20.

Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (2011). Phonotactic constraint violations in German grammar are detected automatically in auditory speech processing: A human event-related potentials study. Psychophysiology, 48, 1208–1216.

Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (2012). The role of stimulus cross-splicing in an event-related potentials study. Misleading formant transitions hinder automatic phonological processing. Journal of the Acoustical Society of America, 131, 3120–3140.

Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. Manuscript, UCLA. http://www.linguistics.ucla.edu/people/steriade/papers/PhoneticsInPhonology.pdf

Stevens, K.N., Blumstein, S.E., Glicksman, L., Burton, M., Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. Journal of the Acoustical Society of America, 91, 2979–3000

Sussman, E., Winkler, I., Ritter, W., Alho, K., Näätänen, R. (1999). Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. Neuroscience Letters, 264, 161–164.

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

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

Vaux, B. (1998). The laryngeal specifications of fricatives. Linguistic Inquiry, 29, 497–511.

Wagner, M., Shafer, V.L., Martin, B., and Steinschneider, M. (2012). The phonotactic influence on the perception of a consonant cluster /pt/ by native Englisch and native Polish listeners: A behavioral and event related potential (ERP) study. Brain and Language, 123, 30–41.

Weber, A. (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. Language and Speech, 44, 95–118.

Wiese, R. (1996). The phonology of German. Oxford: Oxford University Press.

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

Winkler, I. (2007). Interpreting the Mismatch Negativity. Journal of Psychophysiology, 21, 147–163.

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

World Medical Association (2013). World Medical Association Declaration of Helsinki - Ethical principles for medical research involving human subjects. http://www.wma.net/en/30publications/10policies/b3/ (last viewed April 17, 2014).

Wurzel, W. U. (1981). Phonologie: Segmentale Struktur. In: K. E. Heidolph, W. Flämig & W. Motsch (Eds.), Grundzüge einer deutschen Grammatik (pp. 898-990). Berlin: Akademie-Verlag.

Yabe, H., Tervaniemi, M., Reinikainen, K., Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. Neuroreport, 8, 1971–1974.

Yip, M. (1991). Coronals, consonant clusters, and the coda condition. In: Phonetics and phonolgy: The special status of coronals, eds. C. Paradis & J.-F. Prunet. San Diego: Academic Press, 61–78.

Zeg, D. (2007). The syllable. In: P. de Lacy (Ed). The Cambridge Handbook of Phonology, Cambridge: Cambridge University Press, 161–194.

6 Appendix (Experimental protocols)

Vertraulich

Experiment:

Name der Versuchsperson: ______________________________ VP-Code:

Name des Versuchsleiters:

Erklärung

Hiermit erkläre ich, an den/m Experiment(en) am Institut für Allgemeine Psychologie der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg freiwillig teilgenommen zu haben. Ich wurde vor Beginn des Experiments vollständig über die Natur des Experiments aufgeklärt.

Mir wurde mitgeteilt, dass meine zur Identifizierung notwendigen persönlichen Daten vom Versuchsleiter vertraulich behandelt werden und dass er/sie keinen anderen Personen als den mit der Durchführung des Experiments Betrauten Zugang zu diesen Daten gewähren wird. Ich bin damit einverstanden, dass die erhobenen Daten anderen Forschern zur Verfügung gestellt werden, soweit diese nicht zur Identifizierung meiner Person ausreichen.

Ich bewahre mir das Recht, das Experiment jederzeit abzubrechen. In diesem Fall sollen alle erhobenen Daten gelöscht werden. Ich bin mir bewusst, dass ich in diesem Fall keinen vollen Anspruch auf Bezahlung/Gutschrift von Versuchspersonenstunden habe.

Ich erkläre, meinem Wissen nach nicht an neurologischen Störungen oder Erkrankungen zu leiden. Ich stehe nicht unter dem Einfluss von Beruhigungsmitteln oder Medikamenten die auf das zentrale Nervensystem wirken.

Ich weiß, dass die Daten in diesem Experiment ausschließlich für Forschungszwecke und nicht zur Diagnostik erhoben werden, und ich werde keine Auskunft oder Expertenmeinung dieser Art fordern.

Hamburg, ________________________ ______________________________

Datum Unterschrift

Experiment _______

Name:	**Datum:**
Vp-Code:	**Zeit:**

Geschlecht	**Alter**	**Beruf**	**Hörvermögen**	**Sichtigkeit**	**Händigkeit**	**Status**
❑ männlich			❑ normal	❑ normal	❑ rechts	❑ gesund
❑ weiblich			❑ korrigiert	❑ korrigiert	❑ links	
	....	....	❑ beeinträchtigt			

Oldfield	Rechte Hand	Teilweise linke Hand	Linke Hand
Händigkeit der Eltern/Geschwister	❑	❑	❑
Schreiben / Zeichnen	❑	❑	❑
Schneiden	❑	❑	❑
Zähne putzen	❑	❑	❑
Werfen	❑	❑	❑
Löffel	❑	❑	❑
Behälter öffnen	❑	❑	❑

Allgemeine Fragen

Zigaretten	❑ Ja		
	❑ Nein	wenn ja, wieviel tgl.	wann zuletzt ....
Alkohol	❑ Ja		
	❑ Nein	wenn ja, wieviel tgl.	wann zuletzt ....
Medikamente	❑ Ja		
	❑ Nein	wenn ja, wieviel tgl.	wann zuletzt ....
			Art ....
Kaffee	❑ Ja		
	❑ Nein	Wenn ja, wieviel tgl.	wann zuletzt ....
Schlaf der letzten Nacht	❑ gut		
	❑ mäßig		
	❑ schlecht		
Schon vorher EEG	❑ Ja	Wenn ja, welche Art von Experiment	❑ Sprache
	❑ Nein		❑ Visuell
			❑ Geräusche
Konzentration	❑ gut		
	❑ mäßig		
	❑ schlecht		

Impedanzen/Dauer:

Impedanzen		**Dauer**	
> 5 kΩ Beginn	....	Vorbereitung	....
> 5 kΩ Ende	....	Experiment	....
		Gesamt	....

Film:

Bemerkungen bitte auf die Rückseite!

Versuchsleiter: ....

Fragebogen nach dem Experiment (______) VP-Code:

Bitte beantworte die folgenden Fragen der Reihe nach!

1. Was hast Du gehört?

2. Glich das Gehörte… eher Geräuschen ☐ eher Sprache ☐

3. Beschreibe bitte die Stimuli, die Du gehört hast!

4. Wie störend hast Du die Stimulation empfunden?

kaum störend 1 2 3 4 5 stark störend

5. Wie gut konntest Du die Stimulation ignorieren? Ist es Dir schwergefallen, die Stimulation zu ignorieren? Wenn ja, wann?

6. Bitte nenne alle Fremdsprachen, die Du gelernt hast (z.B. in der Schule, durch Auslandsaufenthalte, Sprachkurse, im Studium etc.)

7. Wo bist Du aufgewachsen? (hier geht es um Deinen dialektalen Hintergrund. Es reichen regionale Angaben, z.B. Bundesland und nächste größere Stadt.)

8. Hast Du (z.B. im Studium) eine linguistische Ausbildung erhalten? Welche?

Vielen Dank, dass du an dem Experiment teilgenommen hast!

Lebenslauf / Curriculum vitae

Thomas Konstantin Jacobsen
Geburtsdaten: 03.05.1986 in Uelzen
Anschrift: Schellingstraße 14
22089 Hamburg
Telefon: 040 21116900
E-Mail: jacobstk@hsu-hh.de

Werdegang:

Seit 06/2014	Knowledge base designer bei der novomind AG
10/2010 – 05/2014	Wissenschaftlicher Mitarbeiter im DFG- Forschungsprojekt „Sprachlautliche Kompetenz: Zwischen Grammatik, Signalverarbeitung und neuronaler Aktivität“ an der Professur für Allgemeine Psychologie der Helmut-Schmidt-Universität
01/2008 – 07/2010	Studentische Hilfskraft am Institut für Phonetik und digitale Sprachverarbeitung Kiel
10/2005 – 07/2010	Magister Artium an der Christian-Albrechts Universität Kiel Hauptfach: Phonetik und digitale Sprachverarbeitung 1. Nebenfach: Allgemeine Sprachwissenschaft 2. Nebenfach: Informatik
2005	Abitur am Lessing-Gymnasium Uelzen

Publikationen:

Paper 1: Steinberg, J., Jacobsen, T.K., Truckenbrodt, H., and Jacobsen, T. (2015). Repair or violation detection? Pre-attentive processing strategies of phonotactic illegality demonstrated on the constraint of g-deletion in German. *Journal of Speech, Language, and Hearing Research*, accepted for publication.

Paper 2: Truckenbrodt, H., Steinberg, J., Jacobsen, T.K., and Jacobsen, T. (2014). Evidence for the role of German final devoicing in pre-attentive speech processing: A mismatch negativity study. *Frontiers in Psychology*, **5**, 1-11.

Paper 3: Jacobsen, T.K., Steinberg, J., Truckenbrodt,, H. and Jacobsen, T. (2013). Mismatch Negativity (MMN) to successive deviants within one hierarchically structured auditory object, *International Journal of Psychophysiology*, **87**, 1–7.

Vortrag:

(2011, September). *Pre-attentive phonotactic processing*. Paper presented at the 6th annual meeting of the SPP 1234, Marburg.

Versicherung an Eides Statt

Hiermit erkläre ich, Thomas Konstantin Jacobsen, geb. am 03.05.1986 in Uelzen, dass ich die vorliegende Dissertation selbstständig und ohne unzulässige Hilfe verfasst habe. Ich habe insbesondere nicht die Hilfe von Vermittlungs- und Beratungsdiensten in Anspruch genommen. Außerdem habe ich im Zusammenhang mit dem Promotionsverfahren und seiner Vorbereitung keine Entgelte gezahlt oder Dienste unentgeltlich in Anspruch genommen, die dem Sinn und Zweck eines Promotionsverfahrens widersprechen.

Ich habe wörtliche und sinngemäße Zitate als solche gekennzeichnet und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt.

Die vorliegende Arbeit ist in dieser oder anderer Form zuvor nicht als Prüfungsarbeit zur Begutachtung vorgelegt worden.

Hamburg, 04.11.2015