The role of stimulus cross-splicing in an event-related potentials study. Misleading formant transitions hinder automatic phonological processing

Johanna Steinberg[a)]
Institute of Psychology, University of Leipzig, Seeburgstrasse 14-20, D-04103 Leipzig, Germany

Hubert Truckenbrodt
Centre for General Linguistics (ZAS) Berlin, Schützenstrasse 18, D-10117 Berlin, Germany

Thomas Jacobsen
Experimental Psychology Unit, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, P.O. Box 700822, D-22008 Hamburg, Germany

(Received 25 March 2011; revised 19 January 2012; accepted 3 February 2012)

The mental organization of linguistic knowledge and its involvement in speech processing can be investigated using the mismatch negativity (MMN) component of the auditory event-related potential. A contradiction arises, however, between the technical need for strict control of acoustic stimulus properties and the quest for naturalness and acoustic variability of the stimuli. Here, two methods of preparing speech stimulus material were compared. Focussing on the automatic processing of a phonotactic restriction in German, two corresponding sets of various vowel-fricative syllables were used as stimuli. The former syllables were naturally spoken while the latter ones were created by means of cross-splicing. Phonetically, natural and spliced syllables differed with respect to the appropriateness of coarticulatory information about the forthcoming fricative within the vowels. Spliced syllables containing clearly misleading phonetic information were found to elicit larger N2 responses compared to their natural counterparts. Furthermore, MMN results found for the natural syllables could not be replicated with these spliced stimuli. These findings indicate that the automatic processing of the stimuli was considerably affected by the stimulus preparation method. Thus, in spite of its unquestioned benefits for MMN experiments, the splicing technique may lead to interference effects on the linguistic factors under investigation.
 [http://dx.doi.org/10.1121/1.3688515]

PACS number(s): 43.71.Qr, 43.71.Es, 43.64.Sj, 43.70.Mn [TD] Pages: 3120–3140

I. INTRODUCTION

A. Mismatch negativity as tool for investigating speech processing

In cognitive neurosciences, auditory event-related brain potentials (ERP) are an eligible and commonly applied electrophysiological measure to investigate the mental organization of categorical linguistic knowledge and its involvement in different stages of speech processing (Phillips, 2001; Friederici, 2004; Martin *et al.*, 2008). In particular, the mismatch negativity (MMN) ERP component has become a well-established tool for investigating early, automatic spoken language processing (for reviews, see Näätänen, 2001; Pulvermüller and Shtyrov, 2006; Näätänen, Paavilainen, Rinne, and Alho, 2007; Shtyrov and Pulvermüller, 2007). The MMN reflects a preattentively operating memory-based mechanism that detects violations of an expectancy created by the regularity of the preceding acoustic stimulation (Näätänen, Gaillard, and Mantysalo, 1978; Näätänen, 1992, Chap. 4; Näätänen and Winkler, 1999; Winkler, 2007; Näätänen *et al.*, 2011). Usually, the MMN is elicited by infrequently occurring auditory events (deviants) whose mental representation mismatches the representation of the regularities extracted from frequently repeated standard stimuli. It is defined as the difference between the ERPs elicited by the standard and the deviant stimuli and it can be identified as a frontocentrally distributed negative deflection in the deviant-minus-standard difference wave occurring between 100 and 250 ms after deviation onset. If the deviation coincides with stimulus onset, the MMN is to be found within the time range of the N2 wave complex of the auditory ERP (e.g., Schröger, 1998; Picton *et al.*, 2000; Näätänen *et al.*, 2011). Since the MMN is sensitive to higher-order cognitive processes such as accessing contents of the long-term memory and because the underlying mechanism operates non-voluntarily, the MMN can be used to investigate what linguistic information is accessed when sounds are not in the focus of attention. Usually, the MMN is obtained using a passive oddball paradigm. However, because of the specific experimental requirements in obtaining MMN as an index of automatic speech processing (sufficient number of repetitions, adequate rate of deviant probability, control for different aspects and levels of deviation, control for exogenous contributions to the deviant and standard ERPs, withdrawal of participants' attention; see Schröger, 1998; Kujala *et al.*, 2007; Pulvermüller and Shtyrov, 2007), stimulation material for

[a)]Author to whom correspondence should be addressed. Present address: Experimental Psychology Unit, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, P.O. Box 700822, D-22008 Hamburg, Germany. Electronic mail: j.steinberg@uni-leipzig.de

passive oddball protocols is rather limited in duration and linguistic complexity compared to attentive protocols designed to investigate later language-dependent ERP components.

B. Methodological problems with speech stimuli in MMN experiments

One of the most crucial problems that one has to deal with when investigating speech material by means of the MMN is MMN being sensitive to several stimulus properties on different levels of abstraction at the same time (e.g., Schröger, 1998). Investigating abstract stimulus properties (such as grammatical categories) on the basis of acoustically complex stimulus material (such as speech sound sequences) requires careful control of the acoustic stimulus features (e.g., Jacobsen *et al.*, 2004; Jacobsen, Schröger, and Alter, 2004; Pettigrew *et al.*, 2004; Pulvermüller and Shtyrov, 2006). However, the speech signal is characterized by a high acoustic variability caused by many linguistic and paralinguistic factors. Apart from the pure acoustic sound properties such as sound amplitude, spectral energy distribution or the temporal dimension (so-called first-order features; e.g., Schröger, 1998) there commonly is considered at least one intermediate level of phonetic representation (e.g., Phillips, 2001). On this level, phonetic entities are encoded by organizing the speech sound signal into linguistically relevant perceptual categories. These phonetic correlates serve as cues in the phonological categorization process mapping the highly variable and continuous acoustic signal to invariant and discrete segmental and supra-segmental phonological categories such as phonological features, phonemes, syllables, pitch patterns, and so on. Distinctive phonological features such as [±back] are traditionally defined in terms of their articulatory consequences (Halle, 1992). The phonetic cues that give rise to their perception have also been described in terms of articulatory categories (Liberman and Mattingly, 1985) or in terms of specific combinations of acoustic signal properties (Jacobson, Fant, and Halle, 1952; for reviews of the field, see Raphael, 2005; Stevens, 2005; Clements and Hallé, 2010).

From a methodological point of view, an area of conflict is constituted between the technical requirement of strictly controlling all basic stimulus features on the one hand and preserving the naturalness and variability of the linguistic material under investigation on the other hand. Several attempts of dealing with the natural complexity inherent in every linguistic stimulus in MMN experiments have been developed as a result from this antagonism. One extreme possibility is to use synthesized stimulus material where as much acoustic properties of a stimulus are kept under strict control as possible by determining the specifications of every necessary linguistic parameter while excluding unnecessary (paralinguistic) features. Variants of this method have commonly been used to investigate isolated vowels (e.g., Näätänen *et al.*, 1997; Sharma and Dorman, 1998; Winkler *et al.*, 1999; Ikeda *et al.*, 2002; Hill *et al.*, 2004; Jacobsen, 2004; Jacobsen and Schröger, 2004; Jacobsen, Schröger, and Alter, 2004; Jacobsen, Schröger, and Sussman, 2004) or simple CV(C)-syllables (e.g., Dehaene-Lambertz, 1997; Sharma and Dorman, 1999; Phillips *et al.*, 2000; Jacobsen *et al.*, 2004; Pettigrew *et al.*, 2004; Pulvermüller and Assadollahi, 2007). However, it appears to be inappropriate for investigations of articulatorily more complex sound sequences. Moreover, evidence has been provided that natural and synthesized speech stimulus material might be processed in different manners (Benson *et al.*, 2001; Hertrich *et al.*, 2002). For more complex sequences, therefore, splicing techniques have become common. By cutting and replacing segments of sounds or sound sequences by other realizations one can keep the acoustic specification of a natural utterance constant and carefully control any influence of the sound context on the segments of interest (cf., e.g., Sharma and Dorman, 2000; Pulvermüller *et al.*, 2001; Mitterer and Blomert, 2003; Pulvermüller and Shtyrov, 2003; Menning *et al.*, 2005; Ylinen *et al.*, 2005, 2006; Flagg *et al.*, 2006; Hasting *et al.*, 2007; Kirmse *et al.*, 2008; Pulvermüller *et al.*, 2008; Garagnani *et al.*, 2009; Tavabi *et al.*, 2009). Finally, there is the possibility to neutralize randomly varying stimulus specifications by including controlled variability into the natural stimulation, which is mostly done by using several tokens of the same stimulus type (cf., e.g., Shestakova *et al.*, 2002; Eulitz and Lahiri, 2004; Jacobsen, Schröger, and Alter, 2004; Jacobsen, Schröger, and Sussman, 2004; Bonte *et al.*, 2005, 2007; Lipski and Mathiak, 2007, 2008; Steinberg *et al.*, 2010a,b, 2011). This method forces the mental system to abstract from varying acoustic stimulus specifications such as pitch, voice quality, or the speaker's sex, age, or voice effort that are not necessarily relevant for linguistic analysis. However, an MMN elicitation relies on a sufficient amount of invariant features within one stimulus category so that a stable representation of the preceding sound context can be created. In order to avoid overstraining the speech processing system, several stimulus properties such as duration, amplitude or pitch level and pitch contour are usually technically matched or varied in a controlled manner by means of signal processing techniques.

C. Intention of the present study

With the present study we aim to contribute deeper methodological insight in this area of research by testing comparatively two alternative methods of preparing linguistic stimulus material with respect to their applicability in MMN experiments. For this purpose, we build on previous research we conducted to investigate the mental organization of language-specific phonological knowledge and its involvement in automatic speech processing by means of the MMN (Steinberg *et al.*, 2010a,b, 2011).

1. Dorsal fricative assimilation

The topic of our research is the so-called dorsal fricative assimilation (DFA), which is an obligatory phonotactic restriction in German grammar. For phonological analysis of the phenomenon see Hall (1989, 1992), Macfarland and Pierrehumbert (1991), Merchant (1996), Noske (1997), and Féry (2001). Phonotactic restrictions are abstract principles regulating the co-occurrence of phonemes in sound sequences. By hypothesis, they distinguish categorically between possible sound sequences (words or pseudo-words) and impossible, i.e., ungrammatical structures (non-words) in a given language (Chomsky, 1988). According to phonological theory, this aspect of phonological grammar is represented independently of the set of possible phonemes (phoneme inventory) and is not included in the entries of the mental lexicon (Chomsky and Halle, 1968/1991; Kenstowicz, 1994). Within a prosodic word, DFA demands that a vowel and a following dorsal fricative agree in their phonological specifications for tongue backness [±back] (Féry, 2001). Accordingly, after front vowels, the palatal dorsal fricative [ç] occurs as in "Küche" [kʏçə] (kitchen); after back vowels, only the velar or uvular dorsal fricatives [x] or [χ] are appropriate, as in "Koch" [kɔx] (cook).

2. Previous experiments on DFA with naturally spoken stimuli

In our former MMN studies on the involvement of DFA in automatic speech processing, we decided to use carefully controlled naturally varying stimulus material. For stimulation, we chose phonotactically well-formed and ill-formed syllables consisting of a front or back vowel followed by one of the two dorsal fricatives or other fricatives. Stimuli were presented to participants that were native speakers of German using passive oddball paradigms. We consistently found the phonotactically ill-formed deviant syllables eliciting a specific negative-going deflection in the ERPs that was interpreted as a derivate of the MMN. This broadly distributed negative-going deflection within the deviant-minus-standard difference wave with its maximum over central scalp areas occurred about 100 ms after the onset of the inappropriate fricative, which is the approximate point in time when the violation of DFA became recognizable. This peak latency fits the expectable time range of MMN responses (e.g., Schröger, 1998). The stimuli were articulated by professional speakers, several tokens of each syllable type were included in the stimulation, and the stimuli were carefully controlled for duration, intensity, and pitch level. However, this method did not allow for precise control of the temporal development of the spectral characteristics during the syllable, especially the transitions of the second vowel formant (F2) towards the following fricative. Furthermore, the abstract status of ungrammaticality of an ill-formed stimulus syllable was confounded with its concrete and continuous phonetic realization: Rather than by the categorical ungrammaticality the MMN effect could have been driven by the unusual articulatory motions that had to be performed when the ill-formed syllables were articulated and that are variously encoded in the acoustic signal.

3. The present experiment

On the basis of these considerations we decided to test comparatively two methods of stimulus preparation as discussed above. Based on the phonotactic question of our recent research, we conducted an MMN experiment in which we ran a condition with natural stimulus material (hence referred as natural condition) and an alternative condition using a set of spliced stimulus syllables (spliced condition). Parts of the data obtained from the natural condition have already been published in Steinberg *et al.* (2011). In this article, we focused the discussion on the phonotactic question under investigation by resuming our previous studies on that matter. In the present paper, however, results from both parts of the experiment are presented and discussed mainly with respect to the effects of the stimulus preparation method. Crucial results from Steinberg *et al.* (2011) are reproduced here in order to render the present paper self-contained. By comparing the data from the natural condition with those from the spliced condition we reconsider the benefits and limits of the splicing technique for creating linguistic stimulus material for ERP experiments. We will present neurophysiological evidence for different processing of natural and spliced stimuli with respect to the phonological phenomenon under investigation. Our findings indicate that essential modifications of the phonetic stimulus properties resulting from the splicing technique were detected automatically during speech processing.

As depicted in Fig. 1(A,B), eight VC-syllables were created by exhaustively combining the rounded and lax vowels [ʏ œ ʊ ɔ] and the dorsal fricatives [ç x], resulting in sets of four well-formed and four ill-formed stimuli. None of these syllables have lexical meaning in German.

While for the natural condition each token was separately articulated and afterwards manipulated by means of signal processing techniques, for the spliced condition the vowel part of each syllable was replaced by a vowel originating from a neutral context. By the use of the splicing technique we intended to exclude any potential regressive coarticulatory influence of the following fricative on the phonetic specification of the vowel. Two benefits result from this: First we gained a better control over the time from when on the phonotactic violation could be recognized. Second, the abstract grammatical status of the syllable could be disentangled from its phonetic materializations within the vowel, most importantly the transitions of the F2. As a consequence, we aimed to estimate the relative impact of F2 transition and the spectral properties of the fricative on the detection of the phonotactic (un-)grammaticality of the syllables.

Concretely, we focus on two aspects here in an exploratory manner: First, are we able to replicate the results from the natural condition with the spliced stimulus material? If so, anticipatory F2 transitions would not be necessary to evaluate the grammatical status of the syllables. Second, does the method of stimulus preparation have a genuine impact on the processing of the stimuli irrespective of their grammatical correctness? Especially the latter question needs to be answered in order to interpret any result of the first aspect properly.

II. METHODS

A. Participants

Eighteen volunteers participated in the study, 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

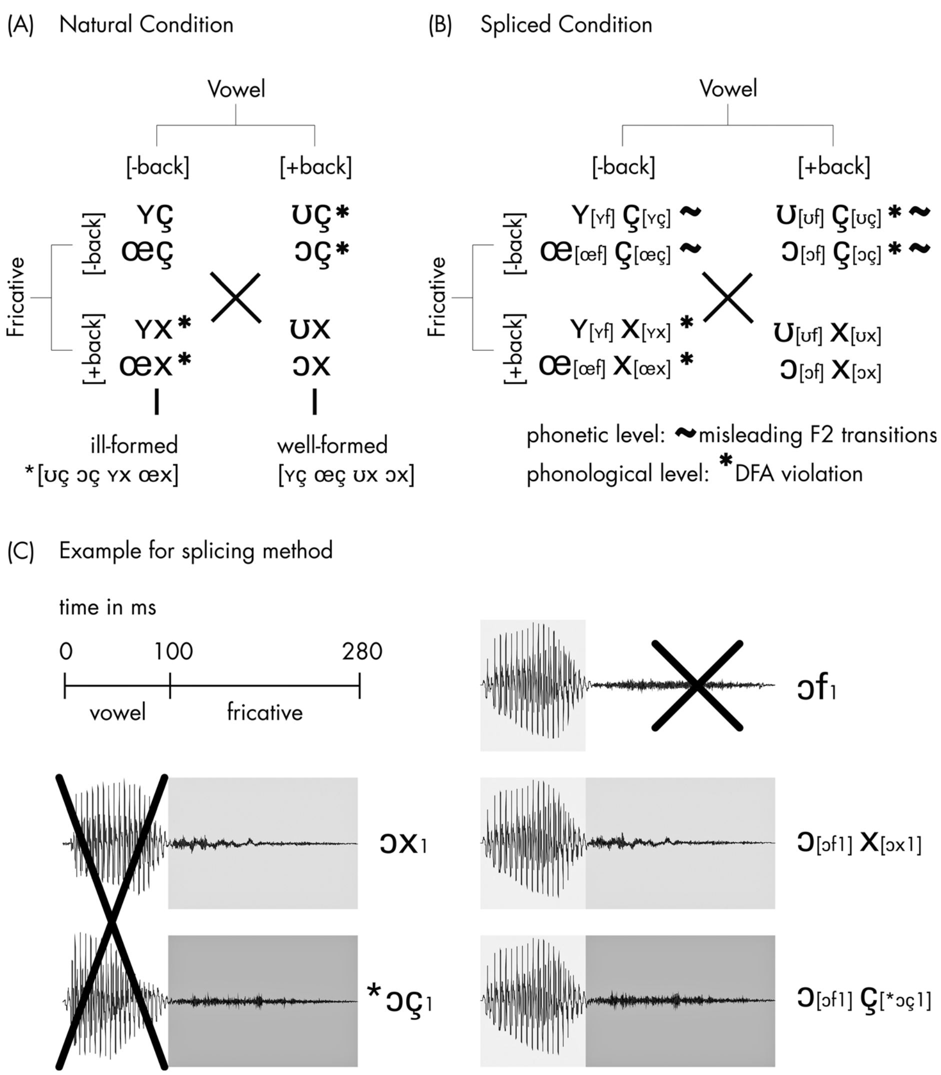

FIG. 1. Experimental design. Upper part: Stimulus syllables depicted in two by two tables, left for the natural condition (A) and right for the spliced condition (B). Syllables are differentiated by vowels (with respect to their horizontal tongue position: front [−back] versus back vowels [+back]) and by fricatives (the palatal [−back] and velar dorsal fricative [+back]). Each cell contains both, syllables with near close and open mid vowels. Phonotactically ill-formed syllables violating DFA are marked by an asterisk. Spliced syllables containing misleading phonetic information are indexed by a tilde. Subscripts in (B) indicate the original source of sound token. Lower part (C): example for the splicing method. Original vowels of the naturally spoken syllables [ɔx] and *[ɔç] were replaced by the same vowel [ɔ] stemming from a VC-syllable that contained the bilabial fricative [f]. The parts of the respective signals that are discharged in this process are crossed out for clarity. The signal parts that are acoustically identical are marked by the same grayscale shading: the vowel replacement in the splicing condition is highlighted in light gray, the velar fricatives are highlighted in middle gray, and the palatal ones are highlighted in dark gray. Segmental durations were normalized to 100 ms for the vowels and 180 ms for the fricatives, respectively.

or corrected-to-normal visual acuity, and neither neurological, nor psychiatric, nor other medical problems. They gave informed written consent and received either course credits or monetary compensation. Data of two participants had to be excluded from the analysis because of a poor signal-to-noise-ratio in the ERPs. The study conforms to the World Medical Association Declaration of Helsinki (2008). The mean age of the remaining sixteen participants (eight male, all right handed) was 26.25 years (range from 21 to 34). The same participants took part in both the spliced and the normal conditions (Steinberg *et al.*, 2011) since both conditions were recorded during the same experimental session.

B. Materials

1. Description of the stimulus material

Stimulus material is depicted in Fig. 1. Eight VC-syllables were used for stimulation, four phonotactically well-formed syllables [ʏç œç ʊx ɔx] and four syllables violating DFA *[ʏx œx ʊç ɔç]. The selection of the stimulus syllables was based on the need for a careful control of the phonetic properties and the phonological feature specifications of the sounds. Phonetically, the vowels differ with respect to the height of the tongue between nearly closed [ʊ ʏ] and open mid [ɔ œ]. With respect to the horizontal position of the tongue they differ between front [œ], near front [ʏ], near back [ʊ], and back [ɔ]. Phonologically, these properties fit into a fully crossed constellation of the feature specifications [±high] and [±back], the latter being crucial with respect to DFA. Furthermore, all vowels share the phonetic properties of lax articulation and lip rounding, phonologically characterized by [-tense] and [lab] (Hall, 1992). On the acoustic level, the tongue height directly corresponds with the height of the first vowel formant (F1), whereas the horizontal position of the tongue is indicated most prominently by the second vowel formant (F2). Since all vowels were characterized by lip rounding, we could abandon the position of the third formant (F3) which is sensitive to the lengthening of the front cavity of the vocal tract and the decrease of the cross-sectional area of its anterior end in rounded vowels (Stevens, 1998, pp. 290–294).

For the design of our study, the two dorsal fricative allophones available in the German sound inventory were used, both of them being characterized by the absence of vocal fold vibrations (phonologically specified as [-voiced], Hall, 1992). Phonetically, they differed with respect to the place of articulation, i.e., the horizontal position that the back of the tongue takes to perform a vocal tract constriction as it is necessary to elicit sufficient friction noise. The palatal fricative [ç] is characterized by a constriction of the tongue body at the middle of the hard palate (phonologically specified as [-back]), while the production of the velar fricative [x] demands a constriction between the back of the tongue and the soft palate (phonologically specified as [+back]) (Hall, 1992). We carefully paid attention not to select uvular realizations of the back dorsal fricative [χ] that are optional in German (Kohler, 1990). Acoustically, the shape and horizontal position of the constriction within the oral part of the vocal tract strongly affects the amplitude and spectral energy distribution of the friction noise (Stevens, 1998, p. 411; Johnson, 1997/2003, pp. 124–127). The further the constriction is located back in the vocal tract, the more the sound is influenced by the filtering specifications of the front cavity. As a consequence, both fricatives differ with respect to their typical spectral properties such as overall spectral shape, peak frequencies, amplitude, and others (cf. Strevens, 1960; Jassem, 1968; Gordon *et al.*, 2002). A common method to measure the properties of fricative spectra is to take the spectral moments, especially the center of gravity (COG) reflecting the spectral mean, the skewness indicating the spectral asymmetry, and the kurtosis as an indicator for the peakedness of the spectrum (cf. Forrest *et al.*, 1988; Jongman *et al.*, 2000; Tabain 2001; Gordon *et al.*, 2002; Maniwa *et al.*, 2009).

However, as we investigated vowel-fricative sequences rather than isolated sounds, progressive and regressive coarticulatory influences needed to be controlled additionally. On the one hand, the vowels were regressively affected by the place of articulation of the following fricative. Especially the final F2 transition depends on the required articulatory target position of the fricative as it reflects the respective horizontal movement of the tongue. The relative impact of the F2 transition for the identification of an adjacent fricative has been discussed conversely in the literature (cf. Wagner *et al.*, 2006, for an overview). Wagner and colleagues (2006) showed that its importance for fricative identification is language-specific as it depends on the fricative inventory of a given language. In German, F2 transitions seem to play a subordinated role for fricative identification. Moreover, fricative perception is also influenced by the segmental order of the neighboring sounds. The impact of F2 transitions has been shown to be considerably higher in CV-syllables (which have mostly been investigated) than in VC-syllables (Mann and Soli, 1991). By means of the splicing method we tried to minimize these F2 transitions when choosing the bilabial fricative [f] as consonantal context for the production of "neutral" vowels, since its articulation does not require any movement of the tongue and therefore this fricative should have the least effect on the F2 transitions of the preceding vowel at all. However, as the repertory of possible voiceless fricatives is quite limited and as the fricatives under investigation are rather different in articulation, this choice does not represent any optimal mean but seems the most applicable sound in our opinion. On the other hand, lip rounding and vowel height has been shown to influence the spectral properties of an adjacent fricative both in CV-syllables (Soli, 1981; Yeni-Komshian and Soli, 1981) and in VC-syllables (Whalen, 1983). Because the spectra of the dorsal fricatives appeared to be very sensitive to the articulatory specifications of the preceding vowels, especially lip rounding, it was not possible to create proper naturally appearing VC-syllables with fricatives from "neutral" articulatory environments. Therefore we decided to use the original fricatives and to control for any progressive effect of the preceding original vowels in the fricative spectra that could indicate the phonotactic violation.

2. Recording

Stimulus material was digitally recorded in an anechoic chamber in the PhonLab at the Centre for General Linguistics (ZAS, Berlin, Germany) at a 16 bit resolution with a sampling rate of 48 kHz using a DAT recorder (Tascam DA 20 M II). The isolated syllables were articulated numerous times by a professional female speaker of German who was instructed to keep her speaking rate, loudness, pitch, and intonation pattern as constant as possible. Five different tokens of each of the eight syllable categories were selected to be used as stimuli in the natural condition of the ERP experiment, resulting in a set of 40 stimulus syllables in total (5×8). Additionally, five tokens of each of the syllables [ɔf ʊf œf ʏf] were selected (5×4 in total) to be used as basis for replacing the original vowels in the spliced condition of the ERP experiment.

3. Editing

The digitized stimuli were edited with the software PRAAT (Boersma and Weenink, 2010). After low-pass filtering with a cut-off frequency of 12 kHz, duration manipulation of the stimulus syllables was performed using the time-domain pitch synchronous overlap addition (TD-PSOLA) algorithm available in PRAAT. The total duration of each of the original stimulus syllables was equated to 280 ms. In doing so, the vowel parts of the syllables were set to 100 ms and the fricatives to 180 ms duration. As a criterion for segmentation we chose the onset of a substantial friction noise in the wave form and in the spectrogram. Potential multiple glottal stops were removed from the onset of the vowel by cutting at zero crossing points. Measures of fricative onset are approximate, however, due to the unregulated acoustic variation in the natural spoken material. Finally, intensities were normalized using the root mean square of the whole sound file.

The set of spliced stimuli was created by replacing the 100 ms vowel parts of the natural stimulus syllables with tokens of the respective vowels extracted from the VC-syllables where the vowel was followed by the fricative [f]. For example, the vowels of the first token of *[ɔç] and the first token of [ɔx] both were replaced with the vowel extracted from the first token of [ɔf], as can be seen in Fig. 1(C). After replacing the vowels, duration of the replaced vowel part was set to 100 ms as described above, restoring the total stimulus duration of 280 ms. Importantly, the splicing point at 100 ms coincided with a zero crossing in the waveform. Finally, the intensities of the spliced stimuli were normalized as described for the natural tokens. Pitch contour and intonation pattern of the natural and the spliced syllables have not been manipulated at all.

4. Acoustic analysis of the stimuli

To get an insight in those acoustic specifications of the stimulus files that were likely to serve as phonetic cues for phonological categorization of the sounds, we conducted several acoustic measures of the stimuli. All acoustic analyses were performed with PRAAT. Prior to the spectral analyses, the signals were sampled down to 24 kHz. All collected measures are depicted in Fig. 2. Additionally to these analyses, intensity contours as well as pitch were plotted and inspected visually to ensure sufficient homogeneity within the stimulus material.

Formant measures were taken from each single stimulus file as mean values within 20 ms analysis windows by using the Burg method with a pre-emphasis frequency of 50 Hz. For the steady state F1 and F2 measures, the analysis window was centered at 50 ms after stimulus onset, which is the midpoint of the vowel. To estimate the F2 transition, F2 measures were taken subsequently two more times, at 70 ± 10 ms and at 90 ± 10 ms.

To analyze the spectral qualities of the fricatives, FFT power spectra were calculated using a 60 ms Hann window covering the mid third part of the respective fricative. From these spectra, the COG, skewness and kurtosis were obtained. During the splicing procedure, the fricative parts of the spliced syllable tokens, which should be identical with their counterparts from the natural spoken material, underwent an additional duration manipulation process that was performed to adjust the new vowel at 100 ms duration. Because of this, the spectral measures of the fricatives differed minimally between the spliced and the natural spoken syllable set. For statistical analysis we therefore used the mean of each spectral measure from the natural and the corresponding spliced fricative part.

For visualization, the FFT spectra were smoothed by means of linear predictive coding (LPC) with a pre-emphasis frequency of 50 Hz and 24 coefficients. Afterwards, the smoothed spectra were averaged separately for fricative type and tongue backness of the preceding vowel.

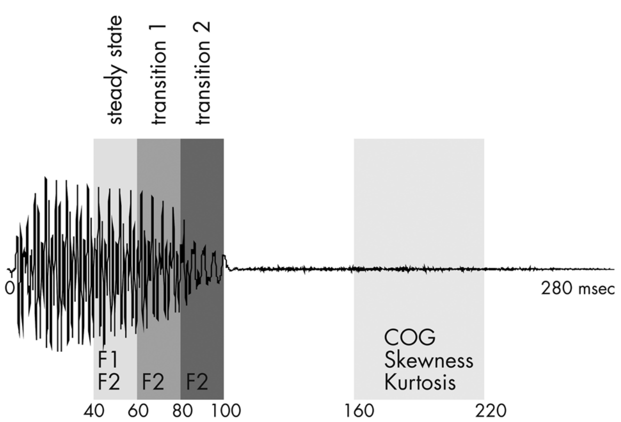

FIG. 2. Acoustic measures taken from the stimulus syllables. Vowels were analyzed by measuring the steady state F1 and F2 from a 20 ms window placed over the mid part of the vowel. Transitions of the F2 were measured by taking the F2 subsequently two times during the vowel. Fricatives were analyzed by measuring the FFT spectra within a 60 ms window centered over the mid part of the fricative. Spectral moments COG, skewness, and kurtosis were calculated from the spectra.

5. Statistic analyses of the stimuli

The transition of the F2 from its steady state position to the end of the vowel was analyzed by means of a three-way mixed-design analysis of variance (ANOVA) with the between-subject factors tongue backness (front/back), following fricative (palatal/velar/labiodental), and the repeated-measures factor transition (steady state/transition 1/transition 2). For further analysis of the factor transition, the design was broken down by the factors tongue backness and following fricative. Bonferroni-adjusted *post hoc* comparisons were performed in case of a significant main effect.

To investigate the spectral properties of the fricatives as well as their sensitivity to coarticulatory influences of the preceding vowel the spectral measures COG, skewness, and kurtosis were analyzed by means of a multivariate analysis of variance (MANOVA) with the factors fricative (palatal/velar) and tongue backness of the preceding vowel (front/back). Afterwards, univariate analyses of each depending measure were run.

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 (η_p^2) effect sizes are given for all reported effects.

6. Results of the acoustic analysis

The main results of the acoustic analyses are depicted in Fig. 3. Additionally, mean values of all taken measures are given in Table I (formants) and Table II (spectral moments) along with standard deviations.

The mixed-design ANOVA of the F2 revealed significant main effects for tongue backness ($F_{1,51} = 393.03$; $p < 0.001$; $\eta_p^2 = 0.885$), following fricative ($F_{2,51} = 98{,}95$; $p < 0.001$; $\eta_p^2 = 0.795$) and transition ($F_{2,102} = 71.85$; $\varepsilon = 0.628$; $p < 0.001$; $\eta_p^2 = 0.585$). As expected, front vowels showed higher F2 values than back vowels. Bonferroni-corrected *post hoc* tests revealed that the vowels followed by [ç] had significantly higher averaged F2 values compared with both the vowels followed by [f] and [x] while vowels followed by [x] showed the numerical lowest averaged F2 values, although not differing significantly from the vowels followed by [f]. Averaged F2 values rose significantly between the three points of measure during the vowel. Furthermore, interactions transition $\times$ tongue backness ($F_{2,102} = 33.25$; $p < 0.001$; $\eta_p^2 = 0.395$), and transition $\times$ following fricative ($F_{4,102} = 60.05$; $p < 0.001$; $\eta_p^2 = 0.702$) became significant. The interaction transition $\times$ tongue backness reflects that the effect of transition is stronger for the back vowels than for the front ones. This is because the steady state value of the F2 is higher in front vowels. The interaction transition $\times$ following fricative shows differences in F2 transition depending on the following fricative. Since the interaction tongue backness $\times$ following fricative has not become

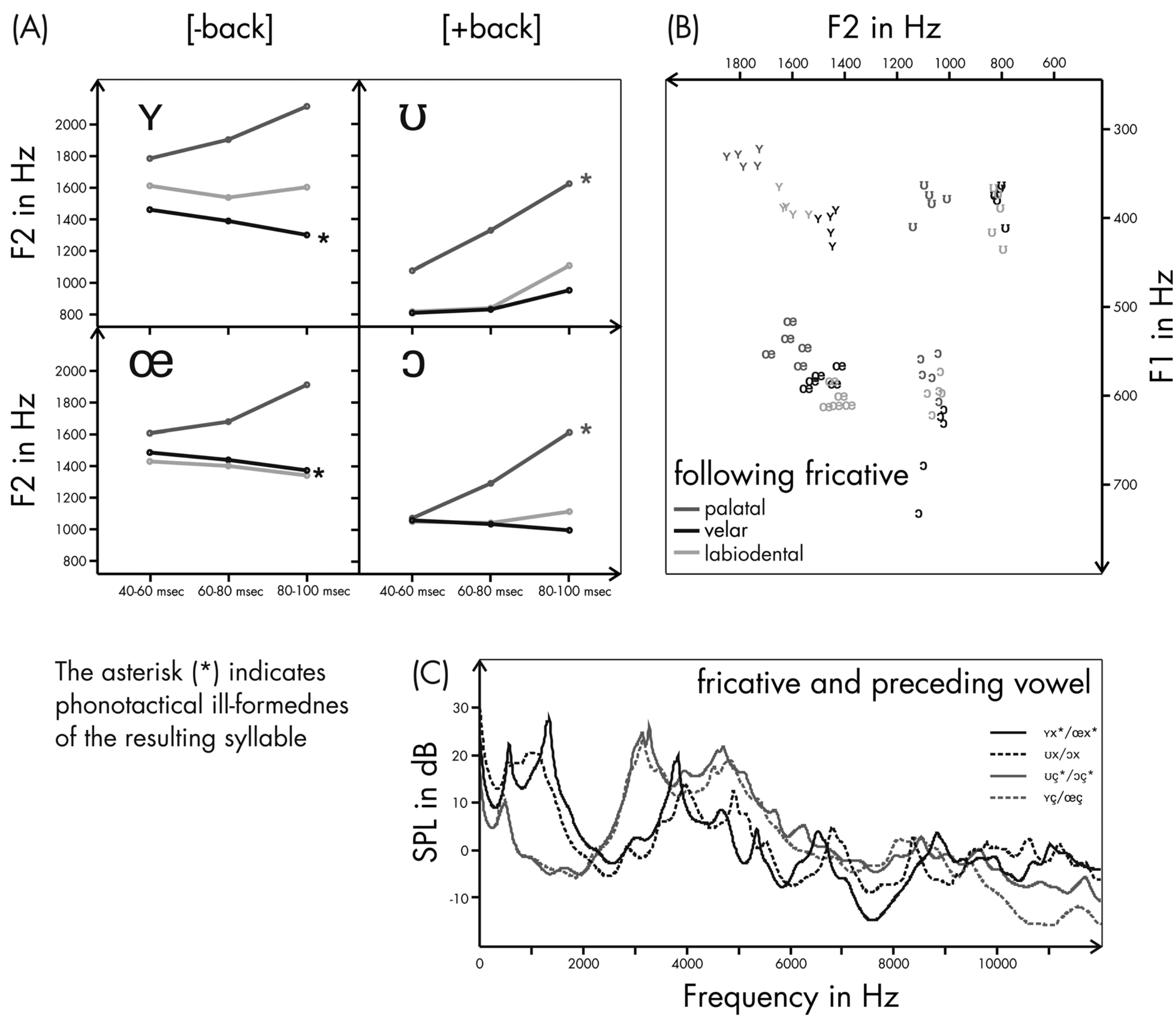

FIG. 3. Results from the acoustic analysis of the stimulus material. (A) Mean values of F2 from three subsequent measuring windows depicted for all vowels broken down by the following fricative. The asterisks represent the vowels stemming from phonotactically ill-formed syllables. (B) F1 by F2 scatter plot depicting the mean steady state formant values of all vowels broken down by the following fricative. The labiodental fricative context represents the vowels used in the spliced condition. (C) LPC spectra of the velar and palatal dorsal fricatives calculated from the mid third of the whole fricative and broken down by the backness of the preceding vowel. Dotted lines represent spectra from well-formed syllables; solid lines indicate spectra from ill-formed syllables (marked by an asterisk).

significant ($F_{2,51} = 0.25$; $p = 0.777$; $\eta_p^2 = 0.010$) we assume that the effect of the following fricative on F2 is independent from whether the vowel is a front or a back one.

Broken down one-way repeated-measures ANOVAs with the factor transition revealed the following results: Back vowels followed by the palatal fricative showed a significantly rising F2 transition ($F_{2,12} = 35.65$; $\varepsilon = 0.547$; $p < 0.01$; $\eta_p^2 = 0.856$; F2_50 versus F2_70: $p < 0.001$; F2_50 versus F2_90: $p < 0.01$; F2_70 versus F2_90: $p < 0.05$). By contrast, back vowels followed by the velar dorsal fricative did not show any significant F2 shift ($F_{2,18} = 1.26$; $\varepsilon = 0.570$; $p = 0.297$; $\eta_p^2 = 0.123$). The back vowels followed by the labiodental fricative that were used in the spliced condition showed a significant F2 rise ($F_{2,18} = 16.49$; $\varepsilon = 0.521$; $p < 0.01$; $\eta_p^2 = 0.647$; F2_50 versus F2_70: $p < 0.05$; F2_50 versus F2_90: $p < 0.05$; F2_70 versus F2_90: $p > 0.01$). Front vowels followed by the palatal dorsal fricative were characterized by a significant overall rising F2 transition ($F_{2,18} = 114.43$; $\varepsilon = 0.857$; $p < 0.001$; $\eta_p^2 = 0.927$; F2_50 versus F2_70: $p < 0.01$; F2_50 versus F2_90: $p < 0.001$; F2_70 versus F2_90: $p < 0.001$). By contrast, front vowels preceding the velar dorsal fricative showed a significant falling F2 transition ($F_{2,18} = 48.59$; $\varepsilon = 0.691$; $p < 0.001$; $\eta_p^2 = 0.844$; F_50 versus F2_70: $p < 0.001$; F2_50 versus F2_90: $p < 0.001$; F2_70 versus F2_90: $p < 0.01$). The F2 transition of the vowels from the spliced condition that were followed by the labiodental fricative appeared not to be significant ($F_{2,18} = 2.76$; $\varepsilon = 0.703$; $p = 0.111$; $\eta_p^2 = 0.237$).

The MANOVA of COG, skewness and kurtosis revealed a significant main effect of the factor fricative (Pillai's trace = 0.946; $F_{3,34} = 198.17$; $p < 0.001$; $\eta_p^2 = 0.946$) as expected. Furthermore, the interaction fricative × tongue backness of the preceding vowel became significant (Pillai's trace = 0.287; $F_{3,34} = 4.56$; $p < 0.01$; $\eta_p^2 = 0.287$). The univariate analyses revealed the following results: COG of the palatal fricatives was significantly higher compared with the velar fricatives in general, indicated by a significant main

TABLE I. Mean steady state F1 and F2 values and mean F2 transition values (in Hz) of all vowels are given along with standard deviations separately for the following fricative. Vowels followed by [x] or [ç] were used in the natural condition, vowels followed by [f] were used in the spliced condition.

Vowel	Following fricative	Steady state F1 mean/SD	Steady state F2 mean/SD	Transition 1 F2 mean/SD	Transition 2 F2 mean/SD	No.
ʊ	x	380/19	809/16	831/47	952/59	5
	ç	382/17	1079/47	1384/145	1623/362 ($n = 2$)	5
	f	396/29	818/17	839/30	1107/57	5
mean		386/22	902/133	1018/280	1128/271	15
ʏ	x	409/16	1459/26	1388/53	1300/34	5
	ç	334/9	1782/52	1901/69	2111/91	5
	f	388/13	1610/44	1536/50	1601/56	5
mean		377/35	1617/142	1608/230	1670/352	15
ɔ	x	657/49	1060/46	1036/16	998/31	5
	ç	575/21	1074/33	1293/81	1613/35	5
	f	597/17	1052/24	1044/34	1115/98	5
mean		610/47	1062/34	1125/132	1242/282	15
œ	x	582/10	1488/54	1441/58	1375/74	5
	ç	544/19	1609/54	1681/92	1914/140	5
	f	604/12	1432/33	1403/50	1343/87	5
mean		577/29	1510/89	1509/142	1544/287	15

effect of fricative ($F_{1,36} = 273.69$; $p < 0.001$; $\eta_p^2 = 0.884$). Furthermore, the interaction fricative × tongue backness of the preceding vowel became significant ($F_{1,36} = 4.34$; $p = 0.044$; $\eta_p^2 = 0.108$). Therefore, broken down ANOVAs were run for the palatal and velar fricatives separately. The backness of the preceding vowel did not have any significant influence on the COG of the palatal fricative. As for the velar fricatives, however, the COG of the fricatives following a front vowel was significantly higher compared to the fricatives following a back vowel (mean COG of [x] after front vowels: 2030 Hz, SD 545 Hz; mean COG of [x] after back vowels: 1533 Hz, SD 459 Hz; $F_{1,18} = 4.86$; $p < 0.05$; $\eta_p^2 = 0.213$). The analyses of the skewness and kurtosis of the FFT spectra revealed a significant main effect for fricative only (skewness: $F_{1,36} = 68.85$; $p < 0.001$; $\eta_p^2 = 0.657$; kurtosis: $F_{1,36} = 5.72$; $p < 0.05$; $\eta_p^2 = 0.137$), both without any effect of the backness of the preceding vowel and without any interaction with this factor.

7. Implications for the splicing method

The analysis of the crucial acoustic parameters (F2 steady state, spectral moments) revealed that both vowels and both fricatives differ significantly in these measures with respect to their horizontal place of articulation. The spectral properties of the fricatives turned out to be only little sensitive with regard to progressive coarticulatory effects by the vowels' backness as there was only an effect on the COG of the velar fricatives. In spite of this effect we considered it best to use the original fricative parts as basis for the splicing procedure. Pilot testing with isolated fricative realizations or realizations after a central vowel revealed unusable, unnatural sounding results.

As for the vowels, however, regressive within category influences of the following fricative were even present on the steady state formant level [as is depicted in the F1/F2-scatter plot in Fig. 3(B)] but were not analyzed statistically on the basis of five samples per category. Averaged F2 transitions of the back and the front vowels, however, were significantly affected by the type of the following fricative, the palatal fricative causing a much higher F2 rise than the velar fricatives as well as the labiodental ones that were used for the spliced condition. As a result, the acoustic indicator for the phonological place specification [-back] of the palatal fricative was already available during the vowel, that is

TABLE II. Mean values of the spectral moments COG, skewness, and kurtosis (in Hz) are given along with standard deviations for the palatal and the velar dorsal fricatives separately for the preceding vowel. The fricative parts were used in both, the natural and the spliced condition.

Fricative	Preceding vowel	COG in Hz mean/SD	Skewness mean/SD	Kurtosis mean/SD	No.
ç [−back]	ʊ [+back; +high]	3987/197	1.102/0.688	4.229/1.579	5
	ɔ [+back; −high]	3726/198	0.634/1.044	6.500/4.047	5
	ʏ [−back; +high]	3777/357	0.508/0.620	7.195/1.879	5
	œ [−back; −high]	3889/49	0.550/0.475	3.660/1.592	5
mean		3845/234	0.699/0.719	5.396/2.752	20
x [+back]	ʊ [+back; +high]	1704/551	2.632/1.008	8.088/5.839	5
	ɔ [+back; −high]	1362/312	2.904/0.544	9.151/3.558	5
	ʏ [−back; +high]	1700/266	2.994/0.858	11.937/7.261	5
	œ [−back; −high]	2360/570	2.372/0.850	5.701/4.882	5
mean		1781/552	2.726/1.273	8.719/5.589	20

considerably earlier than at the onset of the friction noise about 100 ms after stimulus onset in the natural stimuli. By means of the splicing procedure we consequently excluded any phonetic indicator of the following fricative being a palatal one from the vowel. So far, we used the term "neutral" to refer to the absence of any specific acoustic cues that could be used to make predictions on following sounds. However, we included an inconsistency on the phonetic level at the same time. When combining a vowel with "neutral" F2 transitions with a palatal fricative requiring specific rising F2 transitions we created a violation of phonetic expectations by means of a missing specific F2 transition pattern. Therefore, it needs to be stated that the removal of an acoustic cue can constitute a "mismatch" and subsequently cause a violation of predictions if this cue is enforced by the following sound. Henceforth we refer the missing F2 rise before palatal fricatives in the spliced condition as "mismatch" on the phonetic level. In our stimuli, such phonetic mismatch occurs only in the syllables containing the palatal fricative irrespective of its phonotactic well-formedness. Figure 1(B) indicates the several levels on which manipulations on the material have been performed due to the splicing procedure. Crucially for the purpose of the present study, phonetic (∼) and phonological (*) violations did not confound but varied across conditions within the quadruplet. As a consequence, possibly interfering effects of the phonetic violation (due to the splicing technique) on the primary research issue (the effect of the phonotactic violation on MMN) can be discovered at least by separately analyzing the stimuli with regard to the fricative type involved. Since phonetic issues are bound to the anatomical and physiological constitution of the human articulatory apparatus they cannot be manipulated in such a clear-cut manner. In our opinion, the best attempt is to keep factors as constant as possible and to control carefully for all phonetic specifications. Finally, the analyses revealed no reason to assume technical artifacts due to the splicing procedure.

C. Design and procedure of the EEG experiment

A complete experimental session included four conditions altogether: The linguistic contrast between well-formed and ill-formed syllable groups was presented twice in oddball blocks, each with reversed stimulus probabilities (see Fig. 4 for an exemplary description). This linguistic contrast, in turn, was run both with the natural stimulus set (natural condition) and with the spliced stimuli (spliced condition). Oddball stimulus sequences of 1600 trials in total were presented per condition. In every sequence, standard (85% of the trials) and deviants were delivered in a pseudo randomized order forcing at least two standards to be presented between successive deviants. Every syllable type occurred 340 times when serving as standard (this is 68 per every single token) and 60 times when serving as deviant (12 per token). Stimulus sequences were presented with a variable stimulus onset asynchrony randomly varying from 550 to 900 ms in units of 10 ms. Oddball blocks were split into two parts of 800 trials each, lasting approximately eight minutes. Altogether eight stimulus blocks were administered to the participants. The order of the blocks was counterbalanced across participants.

Participants were seated comfortably in a sound attenuated and electrically shielded experimental chamber and were instructed to ignore the auditory stimulation while

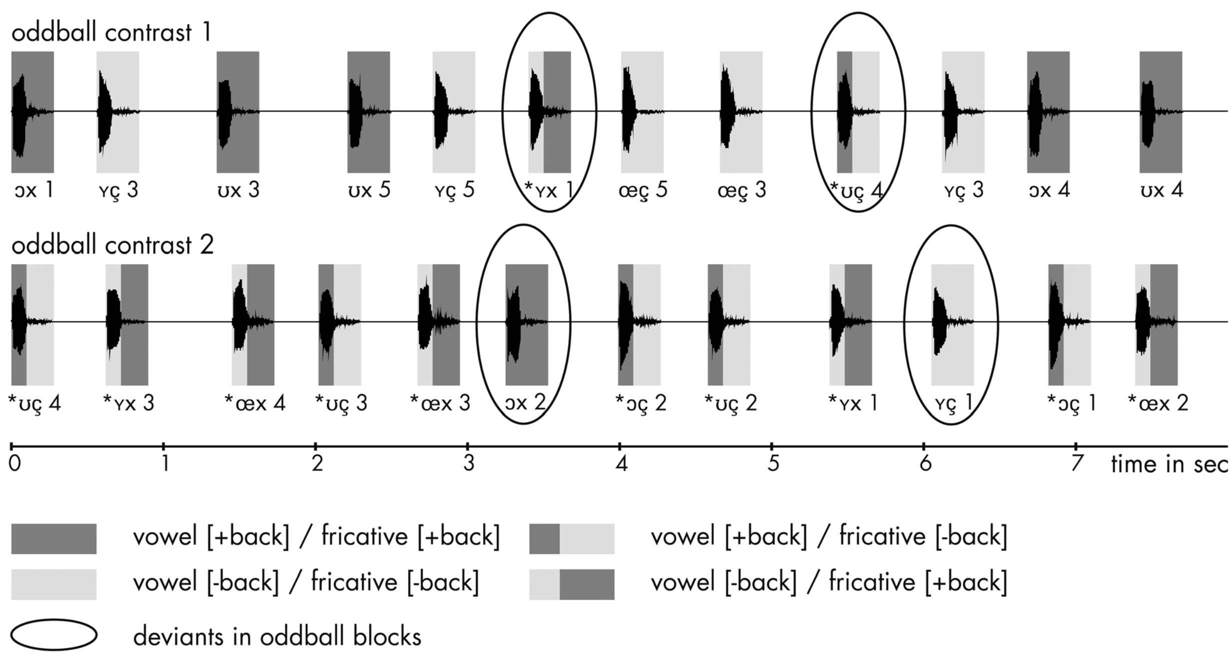

FIG. 4. Example oddball sequences of both oddball conditions presented as waveforms. Upper row: the ill-formed syllable tokens randomly presented as deviants among the well-formed tokens serving as standards. Lower row: the same stimulus syllables presented with reversed probabilities. Deviants (15%) are marked by an oval. Different tokens of a given syllable type are indicated by the number of tokens. Duration of each token was set to 280 ms. Stimulus onset asynchrony varied from 550 to 900 ms randomly in units of 10 ms. Segments with phonological feature specification of [+back] are highlighted in dark gray, phonological feature specification of [-back] is marked in light gray. Phonotactic ill-formedness is labeled by an asterisk.

watching a self-selected silent subtitled movie on a computer screen placed outside the chamber. Stimuli were presented binaurally at approximately 65 dB SPL (artificial head HMS III.0; HEAD acoustics, Herzogenrath, Germany) via headphones (HD 25-1 II, Sennheiser, Wedemark, 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. Every experimental session lasted approximately 1.5 h (plus additional time for electrode application and removal) including breaks of about 15 min in total.

D. Electrophysiological recording

The electroencephalogram (EEG) was recorded continuously with active Ag/AgCl electrodes from 32 standard scalp locations (Fp1, Fp2, AF3, AF4, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO3, PO4, O1, Oz, O2) according to the 10% extension of the International 10-20 system (American Electroencephalographic Society, 1994; Jasper, 1958). Electrodes were mounted in a nylon cap. Two electrodes placed left and right posterior to Cz were used as online-reference and as ground during the recording. External electrodes were placed at the tip of the nose, which served as off-line reference, and at the left and right mastoid sites (LM, RM). 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. EEG and EOG signals were amplified by BioSemi (Amsterdam, Netherlands) Active-Two amplifiers, and digitized with a sampling rate of 512 Hz.

E. Data analysis

Off-line signal processing was carried out using EEP 3.0. The raw 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 the experimental conditions in question and for each participant. With regard to the respective statistical analysis, averaging was performed in two different ways. First, the ERPs were averaged for the factor levels standard/deviant and phonotactically well-formed/ill-formed. Thus, the signals elicited by different syllable types were collapsed according to their classification of phonotactic correctness: [ʏç œç ʊx ɔx] versus *[ʏx œx ʊç ɔç]. Second, the data were averaged for standard/deviant and according to the respective fricative (palatal/velar), but were, to keep the signal-to-noise-ratio sufficient, collapsed across their respective phonotactic status: [*ʊç *ɔç ʏç œç] versus [ʊx ɔx *ʏx *œx]. The ERP response to the first five stimuli of each block as well as to standard stimuli immediately following deviants were not included in the analysis. Epochs showing an amplitude change exceeding 100 μV at any of the recording channels were rejected from averaging. Averaged over participants and conditions, 936 standards (standard deviation/SD 151) and 202 (SD 32) deviants were kept in the analysis after artifact rejection. Separately for experimental conditions (first averaging): well-formed syllables from natural condition 939 (SD 151) standards and 204 (SD 28) deviants; well-formed syllables from spliced condition 939 (SD 155) standards and 200 (SD 37) deviants; ill-formed syllables from natural condition 946 (SD 145) standards and 203 (SD 30) deviants; ill-formed syllables from spliced condition 920 (SD 167) standards and 202 (SD 37) deviants; (second averaging): standards with the velar fricative from natural condition 944 (SD 143), standards with the velar fricative from spliced condition 930 (SD 161), standards with the palatal fricative from natural condition 952 (SD 156), and standards with the palatal fricative from spliced condition 929 (SD 160). Grand-averages were subsequently computed from the individual-subject averages.

(1) Deviance-related effects (MMN) were examined on the basis of deviant-minus-standard difference waveforms that were calculated for the phonotactically well-formed and ill-formed condition separately (across oddball blocks) by subtracting the averaged standard ERPs from the respective deviant ERPs, that is *[ʊç ɔç ʏx œx] as deviant minus *[ʊç ɔç ʏx œx] as standard and [ʊx ɔx ʏç œç] as deviant minus [ʊx ɔx ʏç œç] as standard. For MMN quantification, ERP amplitudes were measured as the mean voltage in a fixed 40 ms time window. However, unlike the data from the natural condition, the data from the spliced condition did not show any distinct deviant-minus-standard differences in the grand-averaged difference waves present in the respective time range. Following the analysis routine that we applied for the data from the natural condition (cf. Steinberg *et al.*, 2011) we placed the analysis window *a posteriori* on the peak latency of the first observable negative-going deflection occurring after 200 ms in the deviant-minus-standard difference wave calculated from the ill-formed stimulus category only (peak latency was averaged over F7-, F3-, Fz-, F4-, F8-, T7-, C3-, Cz-, C4-, and T8 electrode sites). In order 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).

(2) Effects related to the stimulus manipulation method were examined on the basis of both, the data from the natural and those from the spliced condition. Analyses were performed on the re-referenced averaged standard ERPs only in order to minimize the potential deviance-related influence of the phonotactic violation. Two kinds of difference waveforms were calculated by subtracting the averaged standard ERPs elicited by the spliced syllables from the respective standard ERPs elicited by the natural syllables: first, separately for the well-formed and ill-formed condition irrespective of the fricative type, and second, separately for the fricative type irrespective of the phonotactic status. ERP amplitudes were measured as the mean voltages in a fixed 40 ms analysis

window reflecting the time range around the N2 waves of the ERPs. We centered this window *a posteriori* on the grand-averaged peak latency of the second negative-going deflection in the grand-average spliced-minus-natural difference waves of the well-formed and the ill-formed standard ERPs. Additionally, peak latencies were averaged over F7-, F3-, Fz-, F4-, F8-, T7-, C3-, Cz-, C4-, and T8 electrode sites.

F. Statistical analysis

1. Effects of phonotactic status on MMN

Possible deviance-related effects, that is the presence and magnitude of MMN responses, as well as their topographical distribution were analyzed by means of a four-way repeated-measures ANOVA with an electrode grid of 3 by 5 scalp electrodes (F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8). Factors were stimulus probability (standard/deviant), phonotactic status (ill-formed/well-formed), position (F-/C-/P-line), and lateral scalp location (7 -/3-/z-/4 -/8-line). All significant main effects and interactions were reported, but only results relevant to our hypotheses, i.e., main effect of stimulus probability and interactions with this factor involved were interpreted. 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 (η_p^2) effect sizes are given for all reported effects.

2. Effects of stimulus preparation method on N2

Possible effects related to the stimulus manipulation method were tested along with their topographical distribution by means of four-way repeated-measures ANOVAs with the same 3×5 electrode grid as described above. All significant main effects and interactions were reported but only results relevant to our hypotheses, i.e., main effect of the factor stimulus manipulation method and interactions with this factor involved were interpreted. 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 (η_p^2) effect sizes are given for all reported effects.

The first ANOVA was run with the factors phonotactic status (ill-formed/well-formed), stimulus manipulation method (spliced/natural), position (F-/C-/P-line), and lateral scalp location (7 -/3-/z-/4 -/8-line) on the basis of the standard ERPs from the spliced and natural conditions averaged according to the phonotactic status of the syllables.

Subsequently, a second ANOVA was calculated on the basis of the standard ERPs from the spliced and natural conditions averaged according to the following fricative, but collapsed over the phonotactic status of the syllables. Hence, factors were fricative (palatal/velar), stimulus manipulation method (spliced/natural), position (F-/C-/P-line), and lateral scalp location (7 -/3-/z-/4 -/8-line).

III. RESULTS

A. Deviance-related effects of phonotactic status

Descriptive statistics of the data from the spliced condition are given in Table III. The analysis window for the MMN was set to 200 ± 20 ms after stimulus onset for the natural condition (cf. Steinberg *et al.*, 2011) and to 231 ± 20 ms for the spliced condition. Figure 5 shows the re-referenced ERPs and difference waves elicited by the ill-formed (top) and well-formed (bottom) standards and deviants from both the natural condition (A) and the spliced condition (B) for a subset of electrodes. Note that the data from the natural condition have originally been published in Steinberg *et al.* (2011). Please consult this article for detailed

TABLE III. Mean amplitudes and standard deviations (in μV) of the re-referenced ERPs from the spliced condition measured at the analyzed three by five electrode grid within the time window of 211–251 ms, presented separately for stimulus type and phonotactic status.

	F7		F3		Fz		F4		F8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Deviant well-formed	0.055	0.704	−0.041	1.040	−0.180	1.083	0.052	1.115	0.303	0.865
Standard well-formed	0.160	0.499	0.139	0.731	0.114	0.759	0.189	0.710	0.372	0.611
Deviant ill-formed	0.039	0.715	−0.073	1.023	−0.100	1.088	−0.093	1.048	0.096	0.757
Standard ill-formed	0.109	0.536	0.115	0.795	0.098	0.834	0.249	0.795	0.282	0.708
	T7		C3		Cz		C4		T8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Deviant well-formed	−0.147	0.629	−0.203	0.939	−0.273	1.034	−0.105	1.010	0.270	0.750
Standard well-formed	0.010	0.365	−0.038	0.639	−0.058	0.722	0.078	0.674	0.323	0.654
Deviant ill-formed	−0.057	0.526	−0.199	0.821	−0.198	0.885	−0.089	0.886	0.007	0.755
Standard ill-formed	−0.042	0.419	−0.063	0.663	0.044	0.690	0.128	0.752	0.254	0.822
	P7		P3		Pz		P4		P8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Deviant well-formed	−0.316	0.518	−0.336	0.651	−0.276	0.765	−0.172	0.713	0.000	0.650
Standard well-formed	−0.135	0.334	−0.152	0.376	−0.104	0.443	−0.048	0.406	0.069	0.382
Deviant ill-formed	−0.003	0.406	−0.075	0.706	−0.158	0.637	−0.091	0.627	0.177	0.526
Standard ill-formed	−0.061	0.322	−0.078	0.362	−0.013	0.389	0.036	0.452	0.100	0.377

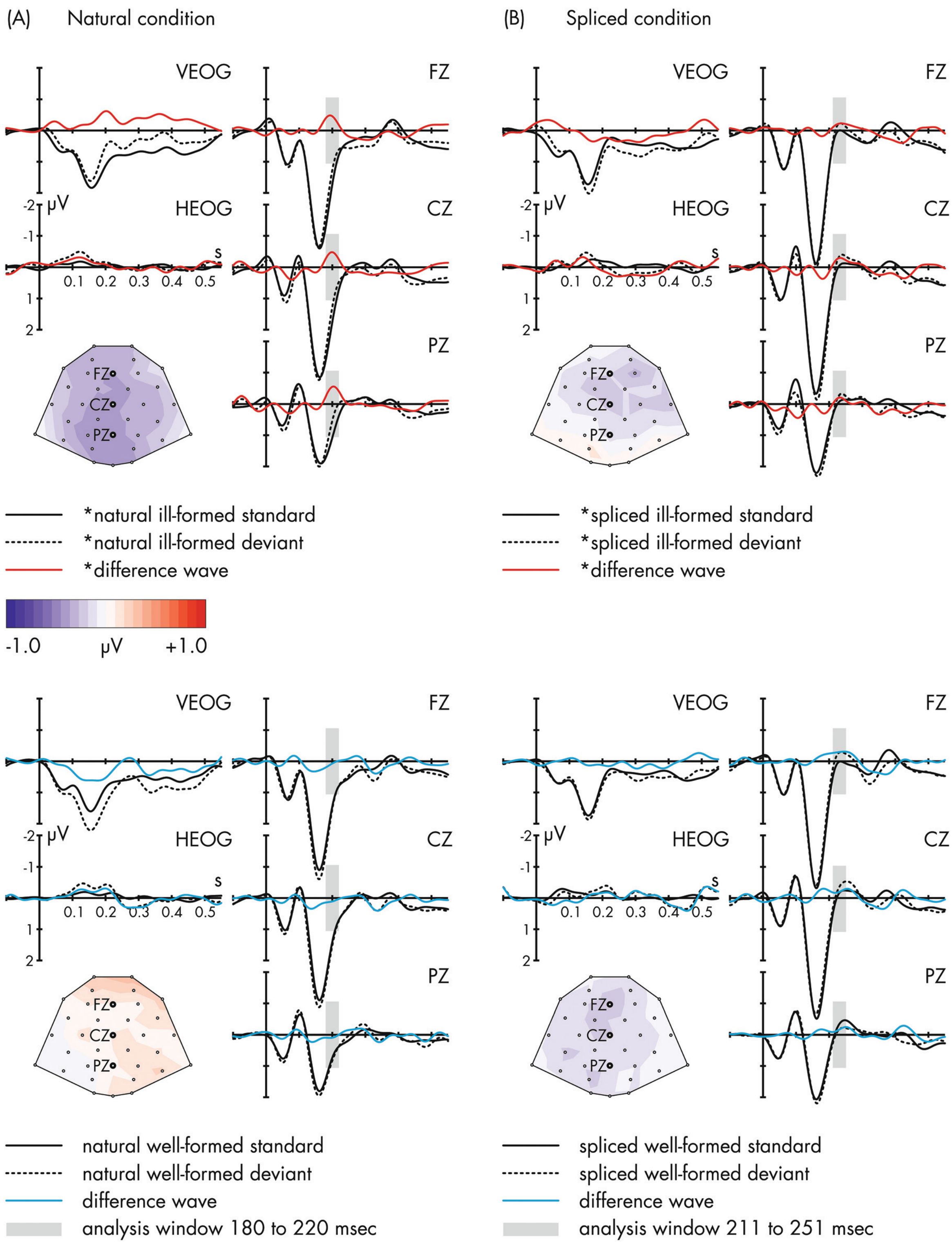

FIG. 5. Grand-averaged, re-referenced ERPs averaged separately for the ill-formed (red) and the well-formed syllables (blue) are depicted left for the natural condition (A) and right for the spliced condition (B). Shown are ERPs to the deviants (black dotted lines), ERPs to the standards (black solid lines), and deviant-minus-standard difference waves (colored solid lines). The gray bars mark the statistically analyzed time windows of 180–220 ms for the natural condition and 211–251 ms for the spliced condition. Scales are in ms and μV. Note that the results of the natural condition (A) are originally reported in Steinberg *et al.* (2011).

analysis and discussion of this part of the data. Nevertheless, the main statistic result with respect to the hypothesis is restated here to facilitate the comparison between both conditions.

The four-way repeated-measures ANOVA of the spliced data revealed a significant main effect of the factor lateral scalp location ($F_{4,60} = 5.49$; $\varepsilon = 0.496$; $p < 0.05$; $\eta_p^2 = 0.268$). Additionally, the interactions phonotactic status $\times$ position ($F_{2,30} = 4.78$; $\varepsilon = 0.575$; $p < 0.05$; $\eta_p^2 = 0.242$) and phonotactic status $\times$ position $\times$ lateral scalp location ($F_{8,120} = 2.73$; $\varepsilon = 0.67$; $p < 0.05$; $\eta_p^2 = 0.154$) became significant. Crucially, there was no interaction between the factors stimulus probability and phonotactic status at all ($F_{1,15} = 0.02$; $p = 0.889$; $\eta_p^2 = 0.001$). Crucial result from the natural condition: stimulus probability × phonotactic status $F_{1,15} = 5.55$, $p < 0.05$, $\eta_p^2 = 0.27$, indicating a significant effect of stimulus probability, i.e., MMN, for the ill-formed stimuli $F_{1,15} = 9.15$, $p < 0.01$, $\eta_p^2 = 0.38$, but not for the well-formed stimuli $F_{1,15} = 0.37$, $p = 0.55$, $\eta_p^2 = 0.02$.

B. Effects of stimulus preparation method on N2

The analysis window was set to 225 ± 20 ms according to the criteria stated in Sec. II E. Analyses of preceding time windows regarding effects of the stimulus preparation method on the N1 and P2 ERP components have been performed in addition. As these effects occurred too early to be attributed to any phonetic or phonological violation at the splicing point the results are not included in this article but are available as supplementary material.[1]

(1) Re-referenced grand-averaged ERPs of the spliced and natural standards are given in Fig. 6 separately for the ill-formed (A) and the well-formed (B) conditions along with topographical maps showing the distribution of the difference between spliced and natural ERP amplitudes within the analyzed time window.

Descriptive data of all tested electrode positions are given in Table IV. The four-way repeated-measures ANOVA of the ERPs elicited by the well-formed and ill-formed standard stimuli from both the natural and the spliced condition yielded significant main effects of the factors position ($F_{2,30} = 7.06$; $\varepsilon = 0.597$; $p < 0.05$; $\eta_p^2 = 0.320$) and stimulus manipulation method ($F_{1,15} = 24.16$; $p < 0.001$; $\eta_p^2 = 0.617$), the latter indicating significant differences between the ERP amplitudes of natural and spliced condition. Furthermore, the interactions stimulus manipulation method × position ($F_{2,30} = 14.46$; $\varepsilon = 0.577$; $p < 0.01$; $\eta_p^2 = 0.491$) and stimulus manipulation method × lateral scalp location ($F_{4,60} = 23.00$; $\varepsilon = 0.514$; $p < 0.001$; $\eta_p^2 = 0.605$) became significant. Note that there was neither a significant main effect of phonotactic status nor significant interactions with this factor. Further analyses were run for each position (F-line, C-line, P-line) separately, yielding significant main effects for stimulus manipulation method (F-line: $F_{1,15} = 38.03$; $p < 0.001$; $\eta_p^2 = 0.717$; C-line: $F_{1,15} = 22.21$; $p < 0.001$; $\eta_p^2 = 0.597$; P-line: $F_{1,15} = 7.27$; $p < 0.05$; $\eta_p^2 = 0.326$) and significant interactions stimulus manipulation method × lateral scalp location (F-line: $F_{4,60} = 12.41$; $\varepsilon = 0.589$; $p < 0.001$; $\eta_p^2 = 0.453$; C-line: $F_{4,60} = 18.70$; $\varepsilon = 0.533$; $p < 0.001$; $\eta_p^2 = 0.555$; P-line: $F_{4,60} = 13.47$; $\varepsilon = 0.617$; $p < 0.001$; $\eta_p^2 = 0.473$). Since the main effects were strongest for the F-line, further analysis was broken down to these data. Results of two-way repeated-measures ANOVAs with stimulus manipulation method and phonotactic status calculated separately for every lateral scalp location (7-, 3-, z-, 4-, 8-line) yielded significant main effects for stimulus manipulation method (F7: $F_{1,15} = 20.62$; $p < 0.001$; $\eta_p^2 = 0.579$; F3: $F_{1,15} = 32.95$; $p < 0.001$; $\eta_p^2 = 0.687$; Fz: $F_{1,15} = 44.19$; $p < 0.001$; $\eta_p^2 = 0.747$; F4: $F_{1,15} = 47.83$; $p < 0.001$; $\eta_p^2 = 0.761$; F8: $F_{1,15} = 21.33$; $p < 0.001$; $\eta_p^2 = 0.587$) but neither the factor phonotactic

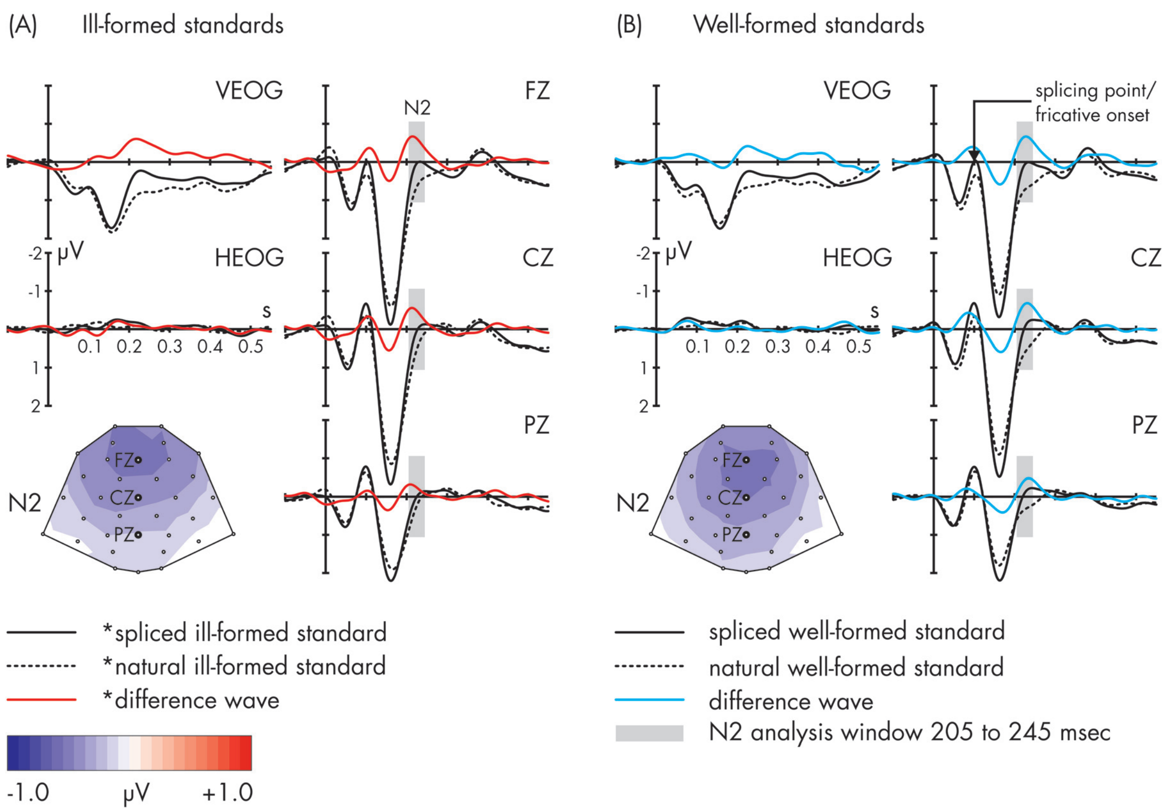

FIG. 6. Grand-averaged, re-referenced ERPs elicited by standard stimuli only are shown for a subset of electrodes. ERPs were averaged separately for the ill-formed (A) and the well-formed syllables (B) and for the spliced (solid lines) and natural (dotted lines) condition. The colored solid lines represent the spliced-minus-natural difference waves. The gray bar marks the statistically analyzed time window of 205 to 245 ms. Scales are in ms and μV. Topographical maps show the spliced-minus-natural-differences within the analyzed time window separately for the ill-formed and the well-formed stimuli.

TABLE IV. Mean amplitudes and standard deviations (in μV) of the re-referenced standard ERPs from the spliced and the natural condition measured at the analyzed three by five electrode grid within the time window of 205–245 ms, presented separately for stimulus manipulation method and phonotactic status.

	F7		F3		Fz		F4		F8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced well-formed	0.208	0.495	0.213	0.726	0.213	0.779	0.279	0.718	0.458	0.622
Natural well-formed	0.511	0.508	0.741	0.781	0.793	0.863	0.849	0.813	0.721	0.698
Spliced ill-formed	0.142	0.543	0.178	0.801	0.185	0.853	0.335	0.809	0.351	0.696
Natural ill-formed	0.527	0.535	0.728	0.810	0.755	0.851	0.876	0.825	0.749	0.777
	T7		C3		Cz		C4		T8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced well-formed	0.111	0.370	0.091	0.647	0.096	0.710	0.221	0.691	0.487	0.697
Natural well-formed	0.322	0.338	0.573	0.688	0.666	0.790	0.711	0.740	0.566	0.663
Spliced ill-formed	0.055	0.420	0.069	0.675	0.186	0.695	0.264	0.773	0.392	0.843
Natural ill-formed	0.240	0.434	0.552	0.742	0.633	0.763	0.645	0.752	0.553	0.704
	P7		P3		Pz		P4		P8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced well-formed	−0.060	0.316	−0.056	0.338	−0.006	0.393	0.035	0.366	0.145	0.349
Natural well-formed	0.126	0.304	0.271	0.417	0.363	0.524	0.324	0.469	0.187	0.428
Spliced ill-formed	0.012	0.333	0.014	0.378	0.085	0.407	0.119	0.456	0.162	0.385
Natural ill-formed	0.088	0.309	0.259	0.466	0.313	0.506	0.341	0.440	0.189	0.297

status nor the interaction became significant. The strongest effect occurred at the F4 electrode position.

(2) Figure 7 depicts re-referenced grand-averaged ERPs of the spliced and natural standards separately for the syllables containing the palatal (left) and the velar fricative (right) along with topographical maps showing the distribution of the difference between spliced and natural ERP amplitudes within the analyzed time window. Descriptive data of all tested electrode positions are given in Table V.

The four-way repeated-measures ANOVA revealed significant main effects for the factors fricative ($F_{1,15} = 38.46$;

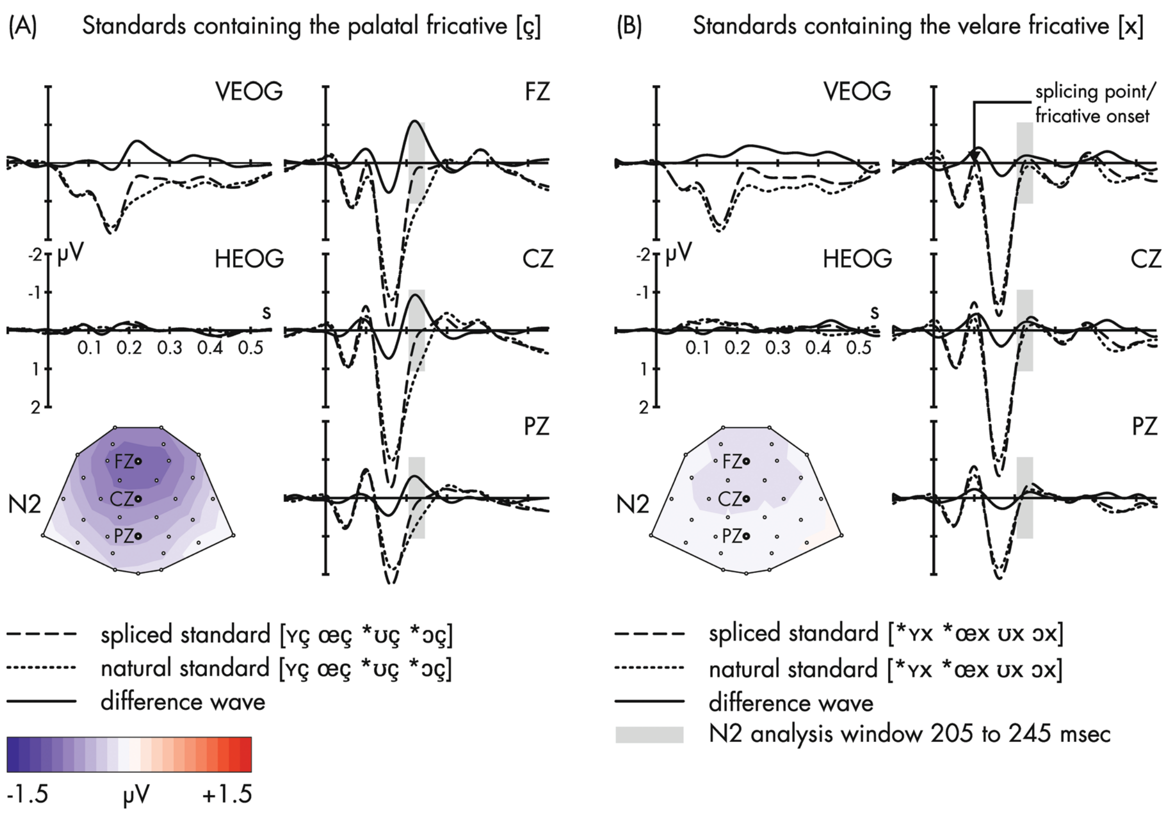

FIG. 7. Grand-averaged, re-referenced ERPs elicited by standard stimuli only are shown for syllables containing the palatal fricative (A) and syllables containing the velar fricative (B). ERPs were collapsed across ill-formed and well-formed syllables but separated depending on the involved fricative type and for the spliced (dashed lines) and natural (dotted lines) condition. The spliced-minus-natural difference wave is represented by the solid line. The gray bar marks the statistically analyzed time window of 205–245 ms. Scales are in ms and μV. Topographical maps show the spliced-minus-natural differences within the analyzed time window separately for the syllables containing the palatal fricative and those containing the velar fricative.

TABLE V. Mean amplitudes and standard deviations (in μV) of the re-referenced standard ERPs from the spliced and the natural condition measured at the analyzed three by five electrode grid within the time window of 205–245 ms, presented separately for stimulus manipulation method and the involved fricative (palatal/velar).

	F7		F3		Fz		F4		F8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced palatal	0.237	0.438	0.260	0.673	0.290	0.718	0.409	0.696	0.484	0.630
Natural palatal	0.828	0.489	1.180	0.737	1.273	0.841	1.336	0.828	1.060	0.733
Spliced velar	0.113	0.632	0.133	0.916	0.112	0.968	0.206	0.886	0.329	0.739
Natural velar	0.210	0.585	0.289	0.915	0.275	0.940	0.389	0.901	0.411	0.749
	T7		C3		Cz		C4		T8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced palatal	0.219	0.354	0.233	0.628	0.300	0.667	0.432	0.751	0.567	0.802
Natural palatal	0.535	0.370	0.997	0.699	1.129	0.793	1.125	0.738	0.793	0.612
Spliced velar	−0.055	0.452	−0.071	0.745	−0.017	0.784	0.055	0.767	0.316	0.767
Natural velar	0.026	0.440	0.128	0.762	0.168	0.808	0.230	0.808	0.328	0.775
	P7		P3		Pz		P4		P8	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Spliced palatal	0.094	0.286	0.121	0.348	0.205	0.397	0.253	0.463	0.240	0.452
Natural palatal	0.290	0.321	0.611	0.434	0.709	0.499	0.650	0.472	0.321	0.364
Spliced velar	−0.146	0.335	−0.165	0.380	−0.126	0.426	−0.098	0.394	0.068	0.313
Natural velar	−0.076	0.310	−0.082	0.491	−0.034	0.585	0.015	0.516	0.056	0.400

$p < 0.001$; $\eta_p^2 = 0.719$), stimulus preparation method ($F_{1,15} = 24.07$; $p < 0.001$; $\eta_p^2 = 0.616$), and position ($F_{2,30} = 7.09$; $\varepsilon = 0.597$; $p < 0.05$; $\eta_p^2 = 0.321$). Furthermore, the interactions stimulus preparation method × fricative ($F_{1,15} = 26.77$; $p < 0.001$; $\eta_p^2 = 0.641$), stimulus preparation method × position ($F_{2,30} = 14.28$; $\varepsilon = 0.577$; $p > 0.01$; $\eta_p^2 = 0.488$), stimulus preparation method × fricative × position ($F_{2,30} = 6.24$; $\varepsilon = 0.525$; $p < 0.05$; $\eta_p^2 = 0.294$), stimulus preparation method × lateral scalp location ($F_{4,60} = 22.90$; $\varepsilon = 0.515$; $p < 0.001$; $\eta_p^2 = 0.604$), fricative × lateral scalp location ($F_{4,60} = 13.24$; $\varepsilon = 0.542$; $p < 0.001$; $\eta_p^2 = 0.469$), and stimulus preparation method × fricative × lateral scalp location ($F_{4,60} = 7.21$; $\varepsilon = 0.522$; $p < 0.01$; $\eta_p^2 = 0.325$) became significant. Further analyses were carried out separately for the factor levels of fricative by breaking down the crucial interaction stimulus preparation method × fricative.

A three-way repeated-measures ANOVA with the factors stimulus preparation method, position, and lateral scalp location did not reveal any significant effects for the ERPs elicited by syllables containing the velar fricative, i.e., [ʊx ɔx *ʏx *œx].

Contrarily, for the syllables with the palatal fricative [*ʊç *ɔç ʏç œç], significant main effects for stimulus preparation method ($F_{1,15} = 50.84$; $p < 0.001$; $\eta_p^2 = 0.772$), position ($F_{2,30} = 10.07$; $\varepsilon = 0.660$; $p < 0.01$; $\eta_p^2 = 0.402$), and lateral scalp location ($F_{4,60} = 4.91$; $\varepsilon = 0.467$; $p < 0.05$; $\eta_p^2 = 0.247$) were found as well as significant interactions stimulus preparation method × position ($F_{2,30} = 14.52$; $\varepsilon = 0.523$; $p < 0.01$; $\eta_p^2 = 0.492$), and stimulus preparation method × lateral scalp location ($F_{4,60} = 24.71$; $\varepsilon = 0.561$; $p < 0.001$; $\eta_p^2 = 0.622$). Further analyses were run separately for the F-, C-, and P-line. Two-way repeated-measures ANOVAs revealed significant main effects for stimulus preparation method (F-line: $F_{1,15} = 45.80$; $p < 0.001$; $\eta_p^2 = 0.753$; C-line: $F_{1,15} = 44.36$; $p < 0.001$; $\eta_p^2 = 0.747$; P-line: $F_{1,15} = 21.28$; $p < 0.001$; $\eta_p^2 = 0.587$), and lateral scalp location (F-line: $F_{4,60} = 4.50$; $\varepsilon = 0.442$; $p < 0.05$; $\eta_p^2 = 0.231$; C-line: $F_{4,60} = 3.95$; $\varepsilon = 0.517$; $p < 0.05$; $\eta_p^2 = 0.208$; P-line: $F_{4,60} = 4.74$; $\varepsilon = 0.467$; $p < 0.05$; $\eta_p^2 = 0.240$), as well as a significant interaction of both factors (F-line: $F_{4,60} = 15.46$; $\varepsilon = 0.568$; $p < 0.001$; $\eta_p^2 = 0.508$; C-line: $F_{4,60} = 20.22$; $\varepsilon = 0.576$; $p < 0.001$; $\eta_p^2 = 0.574$; P-line: $F_{4,60} = 12.83$; $\varepsilon = 0.561$; $p < 0.001$; $\eta_p^2 = 0.461$). As the main effect stimulus preparation method was strongest for the F-line, further analyses were limited to these data. Separately for each electrode position, dependent two-tailed t-tests were performed between spliced and natural conditions for both the syllables with the palatal and those with the velar fricative. The spliced and natural syllables containing the palatal fricative [*ʊç *ɔç ʏç œç] differed significantly within the N2 window: F7 $t = 5.49$, $p < 0.001$; F3 $t = 6.52$, $p < 0.001$; Fz $t = 7.22$, $p < 0.001$; F4 $t = 7.01$, $p < 0.001$; F8 $t = 5.17$, $p < 0.001$. The strongest effect occurred at Fz electrode position. On the contrary, the very comparisons did not became significant for the syllables containing the velar fricative [ʊx ɔx *ʏx *œx] at all.

IV. DISCUSSION

A. Does splicing hinder the replication of the phonotactic MMN effect in this experiment?

Unlike the results from the natural condition reported by Steinberg *et al.* (2011), no deviance-related ERP effect (MMN) attributable to the violation of the phonotactic restriction of DFA in German grammar was found when using spliced syllables for stimulation [see Fig. 5(B)]. The phonotactically ill-formed spliced deviants (ERPs collapsed over all ill-formed syllables) did not elicit any specific negative-going response in the ERPs compared to the ERPs to the respective spliced standards.

The deviance-related negative-going ERP effect (MMN) elicited by the ill-formed deviants in the natural condition was interpreted as the outcome of an independent abstract phonotactic evaluation process resulting in the automatic detection of DFA violation. This process was assumed to operate on the basis of the outcome of any sub-categorical phonetic analysis processes concerning the involved sound segments by activating and applying implicit language-specific phonological knowledge such as DFA from the long-term memory.

The absence of any corresponding ERP effect in the data from the spliced condition indicated that stimulus preparation method interfered significantly with the results of the linguistic investigation. The phonotactic MMN effect might have been ruled out, inhibited or simply masked by any processing effect attributable to the splicing technique.

The first question to be discussed is whether the spliced syllables have not been phonotactically evaluated at all. However, this hypothesis seems to us highly unlikely for the following reasons: The acoustic analyses of the spliced stimuli revealed that sufficient phonetic information was available to easily identify the phonemes of the syllable even without the original F2 transitions indicating the frontness/backness of the following fricative. Subsequently, the satisfaction of DFA could have been evaluated in the spliced condition in any case. Furthermore, all participants reported informally that they perceived the stimulation as speech and that they did not detect any difference between spliced and natural conditions (presented block-wise in counterbalanced order). We therefore rather assume that an analogous MMN effect due to the phonotactic violation has been elicited by the ill-formed deviant syllables in the spliced condition, even if this effect was not measurable in our data.

For a further understanding of the differing results from the spliced and natural condition with respect to the linguistic topic under investigation, possible genuine effects caused by the splicing procedure have to be taken into account. The juncture between vowel and fricative in the spliced condition matches temporally with the fricative onset, i.e., the point in time where DFA could be evaluated. Artificial splicing effects would occur across all syllables irrespective of their phonotactic well-formedness and irrespective of whether they served as standard or deviant. In fact, a comparison of the ERPs from the spliced and the natural standard syllables revealed significant differences between both conditions within the time range of the N2 component. Spliced standard syllables elicited higher N2 amplitudes than natural standards irrespective of whether the syllables were phonotactically well-formed or ill-formed (see Fig. 6). This effect occurred to be maximal at approximately 225 ms after stimulus onset, i.e., 125 ms after the splicing point, and it was characterized by a fronto-central scalp distribution. Crucially, the N2 time range coincides with the time window in which the phonotactic MMN was to be expected in our data. Therefore, the robust splicing effect we found on all standard ERPs might account for the absence of any comparatively fragile MMN attributable to the phonotactic ill-formedness of deviants in the spliced data set.

B. Is the splicing effect attributable to misleading F2 transitions?

What acoustic consequences from the splicing procedure have accounted the robust N2 differences between the spliced and natural condition? During stimulus preparation, we carefully controlled every single stimulus for possible artifacts due to the splicing procedure such as ensuring that every juncture is performed at a zero crossing of the sound signal. Also, we took care for several aspects on the acoustic level such as loudness, pitch level and pitch contour, segmental duration, and articulatory elaboration. Finally, the acoustic analysis of the stimulation material did not reveal any artificial properties within the signals. Given this we assume that technical artifacts can be ruled out as cause of the N2 effect.

Instead of that the acoustic analysis of the stimulus syllables revealed that the crucial difference between spliced and natural syllables pertains to the F2 transitions mainly of those vowels that preceded a palatal fricative. While the spliced syllables containing the palatal fricative differ significantly from their naturally articulated counterparts with respect to the F2 transitions, the syllables containing the velar fricative lack this prominent phonetic indicator. If we consider the appropriateness of the F2 transition as the prominent measure for the phonetic violation, we will expect differences in the ERP data between these two syllable categories in such a way that the spliced syllables with the palatal fricative would elicit larger N2 amplitudes around the splicing point than the other ones. Accordingly, the split-up analysis of the splicing effect on the N2 with respect to the fricative type (collapsed across well-formed and ill-formed stimuli) resulted in the expected clear-cut pattern: Differences in N2 amplitude were found exclusively between ERPs from spliced and natural standard syllables with the palatal fricative, i.e., those syllables that differ with respect to the appropriateness of the F2 transition. On the contrary, spliced standard syllables with the velar fricative did not elicit any larger N2 responses than their natural counterparts (see Fig. 7).

We interpret this finding in terms of violated phonetic expectancies that were built on any regressive information within the vowel about the forthcoming fricative's place of articulation. Coarticulatory information about the forthcoming sound is inherent in the acoustic manifestation of the vowel. These phonetic correlates are assumed to be used by the mental system to create an expectation about what sounds are articulatorily likely to follow (Weber, 2001). The original vowels preceding the palatal fricative, on the one hand, *ex ante* highly restrict the set of possible following fricatives by indicating a frontal (and upwards) shift of the tongue body due to a rising F2 transition. On the other hand, the vowels followed by the velar fricative as well as all vowels used in the spliced condition (that originally preceded the labiodental fricative [f]) did not impose such strong limitations about the set of possibly following sounds, but excluded a certain continuant by indicating the absence of any frontal tongue motion. Given this, a vowel with rather steady-state F2 transitions (as used in the spliced condition) is likely to be followed by various fricatives except the palatal one. At least all spliced vowels preceding the palatal fricative, therefore, violated the expectations about articulatorily possible continuants due to missing proper phonetic indicators such as F2 transition [see Fig. 1(B)].

On the basis of the acoustic analysis of the stimulus material we consider the mismatch of the regressive phonetic indicator within the vowel (the improper steady-state F2 transition) and the actually realized palatal place of articulation in the spliced syllables [*ʊç *ɔç ʏç œç] as cause of the N2 effect.

C. Do the misleading F2 transitions hinder the replication of the phonotactic MMN in this experiment?

How is the processing on the phonetic and the phonological level related in the present experiment? Do the violations of the phonetic and the phonological level add up or does the violation on the phonetic processing level inhibit or even interrupt any further linguistic analysis processes? To test these hypotheses, we should rerun the original MMN analysis separately for the syllables containing the velar and the palatal fricative. Unfortunately, this procedure was not possible on the basis of the present data because of an insufficient signal-to-noise ratio. We primarily designed the present study to investigate the capacity of the mental system to abstract from specific syllables when conducting a phonotactic evaluation with respect to DFA. Therefore we decided to include as much syllable types pseudo-randomly into the stimulation as possible and to analyze the ERPs by averaging all syllables with respect to their grammaticality, their function as standard or deviant as well as their preparation mode. This experimental setup was not designed for running such differentiated analyses.

However, the data appeared to be sufficient enough to outline a recognizable pattern at least over fronto-central electrodes: As depicted in Fig. 8 for F4 (where the effect was numerically maximal), the MMN due to the DFA violation becomes visible for the phonotactically ill-formed deviants [*ʏx *œx] [Fig. 8(A)] whereas no MMN was elicited by the corresponding well-formed syllables [ʊx ɔx] [Fig. 8(B)]. This pattern is consistent with our general results obtained with the natural stimuli [Fig. 5(A)]. On the contrary, no MMN has been elicited by the ill-formed deviant syllables that additionally bear a violation on the phonetic level, namely, [*ʊç *ɔç] [Fig. 8(C)]. Dependent t-tests run between mean deviant and standard ERP amplitudes (quantified by means of a 40 ms analysis window adjusted following the criteria stated in Sec. II) on F4 revealed the following results: (A) spliced ill-formed [*ʏx *œx] $t = -3.224$, $p < 0.01$; (B) spliced well-formed [ʊx ɔx] $t = -0.197$; $p = 0.847$; (C) spliced ill-formed [*ʊç *ɔç] $t = -0.270$, $p = 0.791$; (D) spliced well-formed [ʏç œç] $t = 0.963$, $p = 0.351$.

We take this finding (even if statistically weak due to data quality) as strengthening of our hypothesis that the misleading F2 transitions within the vowels preceding the palatal fricative hindered any subsequent phonotactic analysis. For the syllables with the velar fricative, the difference in the F2 transition pattern between the original and the spliced vowels are considered to be perceptually marginal and consequently the fricative caused no mismatch with respect to any phonetic anticipation.

The fact that the MMN due to the phonotactic violation occurred about 27 ms delayed within the spliced condition (syllables with velar fricatives only) compared to the very effect in the natural condition (all ill-formed syllables) may be explained by the missing anticipatory information within the spliced vowel.

D. Does the detection of a phonetic violation interfere with higher-order speech processing?

The fact that mismatching phonetic cues (or even missing ones) have effects on speech comprehension is widely used in studies that behaviorally investigate the perceptual role of specific phonetic properties as perceptual cues for phonological categorization. Stimulus material is usually created by means of cross-splicing by either combining segments with substantial conflicting articulatory information (conflicting-cue stimuli) or by combining segments with a certain cue missing (deleted-cue stimuli) (for a review, see Smits *et al.*, 1996). Focussing on coarticulatory effects between vowels and obstruents, several studies investigated the impact of F2 transitions and acoustic parameters of the obstruents such as, for example, the spectrum of the friction noise (among many others Harris 1958; Whalen, 1983, 1984; Wagner *et al.*, 2006; Nowak, 2006), or the burst (for example, Cho and McQueen, 2006; Smits *et al.*, 1996). Furthermore, coarticulatory information has been manipulated by means of cross-splicing to behaviorally investigate the impact of phonological processes such as assimilations on speech processing (e.g., Fowler and Brown, 2000; Gow, 2001; Mitterer and Blomert, 2003; Gaskell and Snoeren, 2008; Hwang *et al.*, 2010), or to investigate the impact of phonetic analysis on higher-order processes such as lexical access (Whalen, 1991).

However, behavioral investigations on speech comprehension usually involve active processing aspects as, for instance, compensatory strategies to any kind of mismatches within the stimulation. On principle, they are bound to attentive speech processing because they require participants to perform specific tasks concerning the presented stimuli. By contrast, electrophysiological measures, such as the MMN, are often used for different purposes as they allow for investigating rather automatic and task-irrelevant speech processing. Depending on the specific linguistic phenomenon under investigation, the stimulus material must not contain interfering violations on every lower processing level.

Compared to behavioral research, electrophysiological evidence for effects of mismatches on a sub-categorical, phonetic processing level is rather poor. Flagg *et al.* (2006) found delayed evoked neuromagnetic activity (M50) to consonants within anomalous VC sequences that were constructed by means of cross-splicing compared to corresponding congruent ones. The anomaly of their stimuli was based on misleading phonetic expectations in terms of unmotivated vowel nasalization before oral consonants. Mismatching phonetic cues concerning nasality have also been investigated by Mitterer and Blomert (2003), and more recently by Tavabi *et al.* (2009), who *inter alia* addressed the effects of contextual appropriateness of facultative nasal place assimilation on

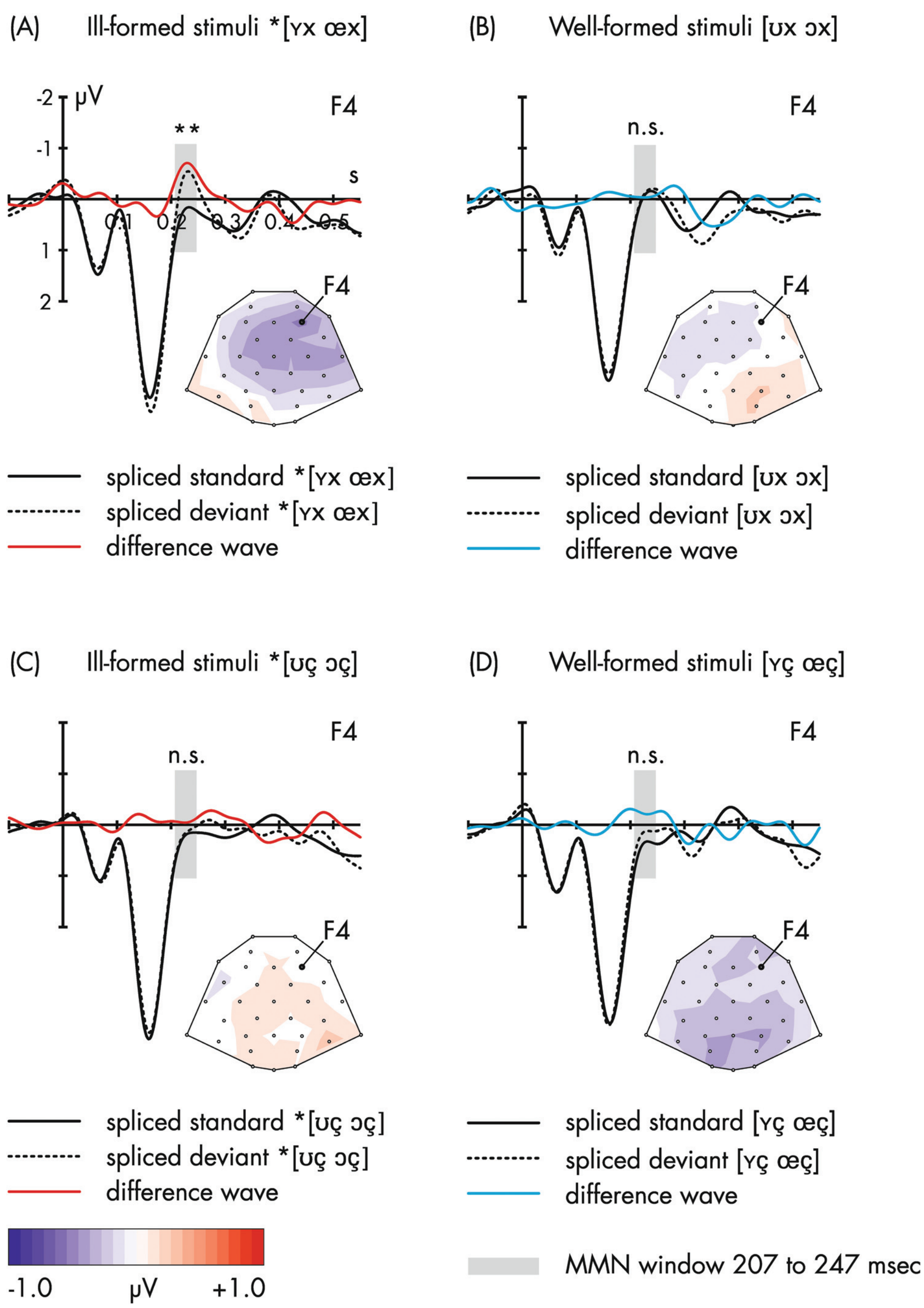

FIG. 8. Grand-averaged, re-referenced ERPs from the spliced condition separately averaged for the ill-formed (A, C) and the well-formed syllables (B, D) and for the involved fricative type (velar fricative: A, B; palatal fricative: C, D) are depicted for F4 electrode. Shown are ERPs to the deviants (black dotted lines), ERPs to the standards (black solid lines), and deviant-minus-standard difference waves (colored solid lines). The gray bar marks the statistically analyzed time window of 207–247 ms. Given are the results of dependent two-tailed t-tests between deviants and standards of each panel. Scales are in ms and μV.

automatic speech processing by means of MMN experiments. Both studies presented stimuli containing a nasal followed either by the bilabial plosive [b] or by a consonant with an alveolar place of articulation. Mitterer and Blomert used dutch compounds (tuimbank, tuinbank, tuinstoel, *tuimstoel), whereas Tavabi *et al.* used disyllabic pseudowords (onbo, ombo, ondo, *omdo), all stimuli were created by means of cross-splicing. Before the bilabial plosive [b], the nasal [n] is allowed to assimilate with respect to its place of articulation, i.e., turning to [m]. By contrast, before alveolar consonants such assimilation is not appropriate. Therefore, the nasal [m] restricts the possibly following consonants to the bilabial place of articulation, whereas the labiodental nasal [n] allows a large set of possible continuants. Both studies found MMN responses attributable to contextual inappropriate nasal-consonant sequences, indicated by misleading phonetic cues within the nasal that indicate bilabial place of articulation and consequently limiting the expectation of possible continuants. Contrarily, in cases where nasal place assimilation is appropriate, MMN responses were found to be attenuated or even missing. These results as well as the outcome of our study indicate that misleading phonetic transitions (that usually result from splicing techniques) have a considerable impact on automatic speech processing as can be measured by means of the MMN.

Despite this fact, splicing is widely used to create "natural sounding" speech sequences in MMN experiments focusing on subsequent linguistic processes on the higher processing levels, namely, the phonological, lexical, (morpho)syntactic or semantic level. According to our impression, however, the impact of the phonetic processing has not been accounted for sufficiently in every case. Given this, the evidence of findings from MMN studies that focus on higher-order speech processing using spliced stimuli that have not been analyzed phonetically in order to control for any conflicting phonetic cues should be challenged. Before setting up an MMN experiment to investigate higher-order linguistic processes, it needs to be clarified whether and how phonetic violations possibly interfere with these subsequent linguistic processing stages of interest.

V. SUMMARY AND CONCLUSION

The present ERP study addressed the primary phonological question whether and to what extent phonotactic constraints as part of the abstract and implicit phonological knowledge are involved in automatic speech processing. To this end we recorded human ERPs in a passive oddball paradigm and investigated the MMN ERP component. For stimulation we used two corresponding sets of various well-formed and ill-formed vowel-fricative syllables, one consisting of naturally spoken syllables and the other set containing spliced ones. Natural and spliced stimuli differed with respect to the regressive coarticulatory information within the vowels, especially regarding the appropriateness of the second vowel formant (F2) transition.

Second research interest of the present study was a methodological one. We aimed to examine the influence of both stimulus preparation techniques on the ERPs during automatic processing of our linguistic stimulation. For this purpose we compared the ERP results from the condition with naturally articulated stimuli with the ERPs stemming from the condition were spliced syllables were used.

A deviance-related ERP effect due to phonotactically ill-formed syllables was found only in the natural condition (Steinberg *et al.*, 2011), but not when using spliced syllables for stimulation. Spliced stimuli, however, elicited significant larger N2 amplitudes compared to the natural stimuli regardless of their phonotactic well-formedness and probability of occurrence. These findings indicate that the automatic processing of the stimulation material was considerably affected by the stimulus preparation method.

Crucially, the splicing technique resulted in a mismatch on the phonetic processing level by means of misleading F2 transitions for a certain subset of syllables. We found both of our results, i.e., the missing phonotactic effect in the spliced condition as well as the stronger N2 responses to standard stimuli, to be entirely triggered by those spliced syllables that contained misleading F2 transitions. The inherent phonetic violation in these syllables is assumed to provoke additional neurophysiological processing effort, even if the spliced stimulus material being accepted as proper speech behaviorally.

Furthermore, higher-order linguistic processing stages such as phonological evaluation are shown to be interfered by phonetic mismatch detection. Considering that splicing is commonly used to prepare acoustically controlled speech stimuli for electrophysiological experiments, our findings provide relevant insights into the risks and limits of this method. As to our knowledge, there exists no research so far expounding the problems of the splicing technique for ERP experiments (except for a promising conference proceeding presented by Hutch *et al.*, 2009). In spite of its unquestioned benefits for ERP experiments, the splicing technique implicates an artificial intervention that might have barely controllable effects on the investigated variables. Therefore, we would like to suggest taking particular caution when using spliced stimuli for neuro-linguistic research.

ACKNOWLEDGMENTS

We are grateful to Jörg Dreyer, Mariam Hartinger, Mira Müller, Anja Roye, Annekathrin Weise, and Andreas Widmann for technical help in various matters. Furthermore we would like to thank three anonymous reviewers for valuable comments and suggestions. This work was supported by the DFG SPP 1234 Grant No. JA1009/10-1 to T.J. and H.T.

[1]See supplemental material at http://dx.doi.org/10.1121/1.3688515 for effects of stimulus manipulation method on the N1 and N2 components of ERPs elicited by the spliced and natural standard stimuli.

American Encephalographic Society. (**1994**). "Guideline thirteen: Guidelines for standard electrode position nomenclature," J. Clin. Neurophysiol. **11**, 111–113.

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., and Liberman, A. M. (**2001**). "Parametrically dissociating speech and nonspeech perception in the brain using fMRI," Brain Lang. **78**, 364–396.

Boersma, P., and Weenink, D. (**2010**). "PRAAT: doing phonetics by computer (version 5.1.30) [computer program]," www.praat.org (last viewed March 24, 2010).

Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., and Blomert, L. (**2005**). "Auditory cortical tuning to statistical regularities in phonology," Clin. Neurophysiol. **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.

Cho, T., and McQueen, J. M. (**2006**). "Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants," J. Acoust. Soc. Am. **119**, 3085–3096.

Chomsky, N. (**1988**). *Language and Problems of Knowledge. The Managua Lectures* (MIT Press, Cambridge, MA), p. 25.

Chomsky, N., and Halle, M. (**1991**). *The Sound Pattern of English*, 2nd ed. (MIT Press, Cambridge, MA), Chaps. 1 and 2, pp. 3–55.

Clements, G. N., and Hallé, P. A. (**2010**). "'Phonetic bases of distinctive features': Introduction," J. Phonetics **38**, 3–9.

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

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

Féry, C. (**2001**). *Phonologie des Deutschen: eine optimalitätstheoretische Einführung. Teil 2 (Linguistics in Potsdam 11) (Phonology of German. An Optimally-Theoretic Introduction, Volume 2)* (Universitätsverlag Potsdam, Potsdam), Chap. 8.5, pp. 62–70.

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

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (**1988**). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," J. Acoust. Soc. Am. **84**, 115–123.

Fowler, C. A., and Brown, J. M. (**2000**). "Perceptual parsing of acoustic consequences of velum lowering from information for vowels," Percept. Psychophys. **62**, 21–32.

Friederici, A. (**2004**). "Event-related brain potential studies in language," Current Neurol. Neurosci. Rep. **4**, 466–470.

Garagnani, M., Shtyrov, Y., and Pulvermüller, F. (**2009**). "Effects of attention on what is known and what is not: MEG evidence for functionally discrete memory circuits," Front. Hum. Neurosci. **3**, Article 10, 1–12.

Gaskell, M. G., and Snoeren, N. D. (**2008**). "The impact of strong assimilation on the perception of connected speech," J. Exp. Psychol. Human. **34**, 1632–1647.

Gordon, M., Barthmaier, P., and Sands, K. (**2002**). "A cross-linguistic acoustic study of voiceless fricatives," J. Int. Phon. Ass. **32**, 141–174.

Gow, D. W. (**2001**). "Assimilation and anticipation in continuous spoken word recognition," J. Mem. Lang. **45**, 133–159.

Hall, T. A. (**1989**). "Lexical phonology and the distribution of German [ç] and [x]," Phonology **6**, 1–17.

Hall, T. A. (**1992**). *Syllable Structure and Syllable-related Processes in German* (Niemeyer, Tübingen), Chap. 5, pp. 220–235.

Halle, M. (**1992**). "Phonological features," in *International Encyclopedia of Linguistics*, edited by W. Bright (Oxford University Press, New York), Vol. 3, pp. 207–212.

Harris, K. S. (**1958**). "Cues for the discrimination of American English fricatives in spoken syllables," Lang. Speech **1**, 1–7.

Hasting, A. S., Kotz, S. A., and Friederici, A. D. (**2007**). "Setting the stage for automatic syntax processing: The mismatch negativity as an indicator of syntactic priming," J. Cognitive Neurosci. **19**, 386–400.

Hertrich, I., Mathiak, K., Lutzenberger, W., and Ackermann, H. (**2002**). "Hemispheric lateralization of the processing of consonant-vowel syllables (formant transitions): Effects of stimulus characteristics and attentional demands on evoked magnetic fields," Neuropsychologia **40**, 1902–1917.

Hill, P. R., McArthur, G. M., and Bishop, D. V. M. (**2004**). "Phonological categorization of vowels: A mismatch negativity study," Neuroreport **15**, 2195–2199.

Hutch, W., Fourie, R., Alku, P., and Pettigrew, C. (**2009**). "The quest for natural, finely controlled speech stimuli for MMN experiments: Is it possible?," Frontiers in Human Neuroscience, www.frontiersin.org/Community/AbstractDetails.aspx?ABS_DOI=10.3389/conf.neuro.09.2009.05.120 (last viewed March 25, 2011).

Hwang, S. K., Monahan, P. J., and Idsardi, W. J. (**2010**). "Underspecification and asymmetries in voicing perception," Phonology **27**, 205–224.

Ikeda, K., Hayashi, A., Hashimoto, S., Otomo, K., and Kanno, A. (**2002**). "Asymmetrical mismatch negativity in humans as determined by phonetic but not physical difference," Neurosci. Lett. **321**, 133–136.

Jacobsen, T. (**2004**). "Mismatch negativity to frequency changes: no evidence from human event-related brain potentials for categorical speech processing of complex tones resembling vowel formant structure," Neurosci. Lett. **362**, 204–208.

Jacobsen, T., Horváth, J., Schröger, E., Lattner, S., Widmann, A., and Winkler, I. (**2004**). "Pre-attentive auditory processing of lexicality," Brain Lang. **88**, 54–67.

Jacobsen, T., and Schröger, E. (**2004**). "Input to verbal working memory: preattentive construction of the central speech representation," Exp. Psychol. **51**, 231–239.

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

Jacobsen, T., Schröger, E., and Sussman, E. (**2004**). "Pre-attentive categorization of vowel formant structure in complex tones," Cognitive Brain Res. **20**, 473–479.

Jacobson, R., Fant, C. G. M., and Halle, M. (**1952**). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates.* (MIT Press, Cambridge, MA), pp. 1–64.

Jasper, H. H. (**1958**). "The ten-twenty electrode system of the International Federation," Electroencephalogr. Clin. Neurophysiol. **10**, 371–375.

Jassem, W. (**1968**). "Acoustical description of voiceless fricatives in terms of spectral parameters," in *Speech Analysis and Synthesis*, edited by W. Jassem (Panstwowe Wydawnictwo Naukowe, Warszawa), Vol. 1, pp. 189–206.

Johnson, K. (**2003**). *Acoustic and Auditory Phonetics*, 2nd ed. (Blackwell, Cambridge, MA), Chap. 7, pp. 120–134.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Kenstowicz, M. (**1994**). *Phonology in Generative Grammar* (Blackwell, Cambridge), Chap. 2, pp. 57–88.

Kirmse, U., Ylinen, S., Tervanierni, 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," Int. J. Psychophysiol. **67**, 131–143.

Kohler, K. (**1990**). "German," J. Int. Phon. Ass. **20**, 48–50.

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

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

Lipski, S. C., and Mathiak, K. (**2007**). "A magnetoencephalographic study on auditory processing of native and nonnative fricative contrasts in Polish and German listeners," Neurosci. Lett. **415**, 90–95.

Lipski, S. C., and Mathiak, K. (**2008**). "Auditory mismatch negativity for speech sound contrasts is modulated by language context," Neuroreport **19**, 1081–1084.

MacFarland, T., and Pierrehumbert, J. (**1991**). "On ich-Laut, ach-Laut and structure preservation," Phonology **8**, 171–180.

Maniwa, K., Jongman, A., and Wade, T. (**2009**). "Acoustic characteristics of clearly spoken English fricatives," J. Acoust. Soc. Am. **125**, 3962–3973.

Mann, V., and Soli, S. D. (**1991**). "Perceptual order and the effect of vocalic context on fricative perception," Percept. Psychophys. **49**, 399–411.

Martin, B. A., Tremblay, K. L., and Korczak, P. (**2008**). "Speech evoked potentials: From laboratory to the clinic," Ear Hearing **29**, 285–313.

Menning, H., Zwitserlood, P., Schoning, S., Hihn, H., Bolte, J., Dobel, C., Mathiak, K., and Lutkenhoner, B. (**2005**). "Pre-attentive detection of syntactic and semantic errors," Neuroreport **16**, 77–80.

Merchant, J. (**1996**). "Alignment and fricative assimilation in German," Linguist. Inq. **27**, 709–719.

Mitterer, H., and Blomert, L. (**2003**). "Coping with phonological assimilation in speech perception: Evidence for early compensation," Percept. Psychophys. **65**, 956–969.

Näätänen, R. (**1992**). *Attention and Brain Function* (Lawrence Erlbaum Associates, Hillsdale, NJ), Chap. 4, pp. 102–210.

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

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

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.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour-Luhtanen, 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., Paavilainen, P., Rinne, T., and Alho, K. (**2007**)."The mismatch negativity (MMN) in basic research of central auditory processing: A review," Clin. Neurophysiol. **118**, 2544–2590.

Näätänen, R., and Winkler, I. (**1999**). "The concept of auditory stimulus representation in cognitive neuroscience," Psychol. Bull. **125**, 826–859.

Noske, M. (**1997**). "Feature spreading as dealignment: The distribution of [ç] and [x] in German," Phonology **14**, 221–234.

Novak, P. M. (**2006**). "The role of vowel transitions and frication noise in the perception of Polish sibilants," J. Phonetics **34**, 139–152.

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

Pettigrew, C., Murdoch, B. M., Kei, J., Chenery, H. J., Sockalingam, R., Ponton, C. W., Finnigan, S., and Alku, P. (**2004**). "Processing of English words with fine acoustic contrasts and simple tones: A mismatch negativity study," J. Am. Acad. Audiol. **15**, 47–66.

Phillips, C. (**2001**). "Levels of representation in the electrophysiology of speech perception," Cognitive Sci. **25**, 711–731.

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," J. Cognitive Neurosci. **6**, 1038–1055.

Picton, T. W., Alain, C., Otten, L., Ritter, W., and Achim, A. (**2000**). "Mismatch Negativity: Different water in the same river," Audiol. Neurootol. **5**, 111–139.

Pulvermüller, F., and Assadollahi, R. (**2007**). "Grammar or serial order? Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity," J. Cognitive Neurosci. **19**, 971–980.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R. J., and Näätänen, R. (**2001**).

"Memory traces for words as revealed by the mismatch negativity," Neuroimage **14**, 607–616.

Pulvermüller, F., and Shtyrov, Y. (**2003**). "Automatic processing of grammar in the human brain as revealed by the mismatch negativity," Neuroimage **20**, 159–172.

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

Pulvermüller, F., Shtyrov, Y., Hasting, A. S., and Carlyon, R. P. (**2008**). "Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing," Brain Lang. **104**, 244–253.

Raphael, L. J. (**2005**). "Acoustic cues to the perception of segmental phonemes," in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell Publishers, Cambridge, MA), Chap. 8, pp. 182–206.

Schröger, E. (**1998**). "Measurement and interpretation of the mismatch negativity," Behav. Res. Meth. Instrum. Comput. **30**, 131–145.

Sharma, A., and Dorman, M. F. (**1998**). "Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential," J. Acoust. Soc. Am. **104**, 511–517.

Sharma, A., and Dorman, M. F. (**1999**). "Cortical auditory evoked potential correlates of categorical perception of voice-onset time," J. Acoust. Soc. Am. **106**, 1078–1083.

Sharma, A., and Dorman, M. F. (**2000**). "Neurophysiologic correlates of cross-language phonetic perception," J. Acoust. Soc. Am. **107**, 2697–2703.

Shestakova, A., Brattico, E., Houtilainen, M., Galunov, V., Soloviev, A., Sams, M., Ilmoniemi, R. J., and Näätänen, R. (**2002**). "Abstract phoneme representations in the left temporal cortex: magnetic mismatch negativity study," Neuroreport **13**, 1813–1816.

Shtyrov, Y., and Pulvermüller, F. (**2007**). "Language in the mismatch negativity design. Motivations, benefits, and prospects," J. Psychophysiol. **21**, 176–187.

Smits, R., ten Bosch, L., and Collier, R. (**1996**). "Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. 1. Perception experiment," J. Acoust. Soc. Am. **100**, 3852–3864.

Soli, S. D. (**1981**). "Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation," J. Acoust. Soc. Am. **70**, 967–984.

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

Steinberg, J., Truckenbrodt, H., and Jacobsen, T. (**2010b**). "Preattentive phonotactic processing as indexed by the mismatch negativity," J. Cognitive Neurosci. **22**, 2174–2185.

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.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA), Chaps. 6 and 8, pp. 257–322, 380–412.

Stevens, K. N. (**2005**). "Features in speech perception and lexical access," in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell Publishers, Cambridge, MA), Chap. 6, pp. 125–155.

Strevens, P. (**1960**). "Spectra of fricative noise in human speech," Lang. Speech **3**, 32–49.

Tabain, M. (**2001**). "Variability in fricative production and spectra: Implications for the hyper- and hypo- and quantal theories of speech production," Lang. Speech **44**, 57–94.

Tavabi, K., Elling, L., Dobel, C., Pantev, C., and Zwitzerlood, P. (**2009**). "Effects of place of articulation changes on auditory neural activity: A magnetoencephalography study," PLoS ONE **4**, e4452.

Wagner, A., Ernestus, M., and Cutler, A. (**2006**). "Formant transitions in fricative identification: The role of native fricative inventory" J. Acoust. Soc. Am. **120**, 2267–2277.

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

Whalen, D. H. (**1983**). "Vowel information in postvocalic fricative noises," Lang. Speech **26**, 91–100.

Whalen, D. H. (**1991**). "Subcategocrical phonetic mismatches and lexical access," Percept. Psychophys. **50**, 351–360.

Winkler, I. (**2007**). "Interpreting the mismatch negativity," J. Psychophysiol. **21**, 147–163.

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 Res. **7**, 357–369.

World Medical Association (**2008**). "World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects," www.wma.net/en/30publications/10policies/b3/17c.pdf (last viewed May 6, 2010).

Yeni-Komshian, G. H., and Soli, S. D. (**1981**). "Recognition of vowels from information in fricatives: Perceptual evidence of fricative-vowel coarticulation," J. Acoust. Soc. Am. **70**, 966–975.

Ylinen, S., Shestakova, A., Alku, P., and Huotilainen, M. (**2005**). "The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish," Lang. Speech **48**, 313–338.

Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P., and Näätänen, R. (**2006**). "Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study," Brain Res. **1072**, 175–185.