



# Dialect Variation Influences the Phonological and Lexical-Semantic Word Processing in Sentences. Electrophysiological Evidence from a Cross-Dialectal Comprehension Study

Manuela Lanwermeyer<sup>1\*</sup>, Karen Henrich<sup>1,2</sup>, Marie J. Rocholl<sup>1</sup>, Hanni T. Schnell<sup>1</sup>, Alexander Werth<sup>1</sup>, Joachim Herrgen<sup>1</sup> and Jürgen E. Schmidt<sup>1</sup>

<sup>1</sup> Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg, Marburg, Germany, <sup>2</sup> Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

## OPEN ACCESS

### Edited by:

Josef P. Rauschecker,  
Georgetown University School of  
Medicine, USA

### Reviewed by:

Mirko Grimaldi,  
University of Salento, Italy  
Dan Zhang,  
Tsinghua University, China

### \*Correspondence:

Manuela Lanwermeyer  
manuela.lanwermeyer@uni-marburg.de

### Specialty section:

This article was submitted to  
Auditory Cognitive Neuroscience,  
a section of the journal  
Frontiers in Psychology

Received: 23 December 2015

Accepted: 03 May 2016

Published: 27 May 2016

### Citation:

Lanwermeyer M, Henrich K,  
Rocholl MJ, Schnell HT, Werth A,  
Herrgen J and Schmidt JE (2016)  
Dialect Variation Influences the  
Phonological and Lexical-Semantic  
Word Processing in Sentences.  
Electrophysiological Evidence from a  
Cross-Dialectal Comprehension  
Study. *Front. Psychol.* 7:739.  
doi: 10.3389/fpsyg.2016.00739

This event-related potential (ERP) study examines the influence of dialectal competence differences (merged vs. unmerged dialect group) on cross-dialectal comprehension between Southern German dialects. It focuses on the question as to whether certain dialect phonemes (/ôâ/, /ôû/), which are attributed to different lexemes in two dialect areas (Central Bavarian, Bavarian-Alemannic transition zone) evoke increased neural costs during sentence processing. In this context, the phonological and semantic processing of lexemes is compared in three types of potentially problematic communication settings (misunderstanding, incomprehension, allophonic variation = potential comprehension). For this purpose, an oddball design including whole sentences was combined with a semantic rating task. Listeners from the unmerged Central Bavarian dialect area heard sentences including either native or non-native lexemes from the merged neighboring dialect. These had to be evaluated with regard to their context acceptability. The main difference between the lexemes can be attributed to the fact that they have different meanings in the respective dialect areas or are non-existent in the linguistic competence of the Central Bavarians. The results provide evidence for the fact that non-native lexemes containing the /ôâ/-diphthong lead to enhanced neural costs during sentence processing. The ERP results show a biphasic pattern (N2b/N400, LPC) for non-existent lexemes (incomprehension) as well as for semantically incongruous lexemes (misunderstanding), reflecting an early error detection mechanism and enhanced costs for semantic integration and evaluation. In contrast, allophonic /ôû/ deviations show reduced negativities and no LPC, indexing an unproblematic categorization and evaluation process. In the light of these results, an observed change of /ôâ/ to /ôû/ in the Bavarian-Alemannic transition zone can be interpreted as a facilitation strategy of cross-dialectal comprehension to reduce both misunderstandings as well as neural costs in processing, which might be interpreted as the initial trigger for this particular phoneme change.

**Keywords:** dialect perception, phonological variation, semantic processing, N200, N400, late positive component (LPC)

## INTRODUCTION

### Phoneme Change As a Result of Dialect Contact

Despite the intensive preoccupation with the phenomenon of linguistic change, the question as to why linguistic units change and which factors influence this process is still a matter of debate. Several studies show that lexically irregular changes are primarily the result of dialect contact (cf. for example Trudgill, 1986; Wang and Lien, 1993; Schmidt and Herrgen, 2011)<sup>1</sup>. These changes result from the interference between systems and the interaction of speakers with different phonological competences.

In this context, one interesting phenomenon is the merger of a phonemic contrast in one dialect, which is still maintained in a related dialect. The expansion of unconditioned mergers can be explained by the close contact between merged and unmerged speech communities. For instance, the actuation of the low back merger in Pennsylvania is explained by a massive influx of foreign-born immigrants who had difficulties in acquiring the distinction between long and short /o/ in words like *cot* and *caught* due to their own reduced vowel systems. Thus, the expansion of the merger is the result of repeated misunderstandings of productions of one-phoneme-speakers by two-phoneme-speakers in face-to-face communication (cf. Herold, 1997; Labov, 2010). Overall, misunderstandings between groups of regional speakers are often motivated by differences in their linguistic competences (cf. Labov, 2010; Schmidt, 2010; Schmidt and Herrgen, 2011).

In order to investigate the relationship between cross-dialectal comprehension and phoneme change, it is useful to study similar varieties, which differ only in few phoneme contrasts so that the general understanding between the speaker groups is ensured. This particular setting can be found in two Southern German dialects. Therefore, the current study investigates the dialect / $\text{ɔ̄a}$ /-/ $\text{ɔ̄u}$ / contrast, which has developed differently in these areas. In the Central Bavarian dialect (henceforth CB), it is a stable contrast, while in the neighboring Bavarian-Alemannic transition zone (henceforth BA) only / $\text{ɔ̄a}$ / occurs before obstruents, brought forth by a merger of Middle High German (MHG)  $\text{ô}$  and  $\text{ei}$  (see **Figure 1** for the geographic location of the dialect areas). Interestingly, a phoneme change to either / $\text{ɔ̄u}$ / or / $\text{o}$ :/ can be observed in certain lexemes in BA, which is possibly due to the dialect contact with CB (Schmidt and Herrgen, 2011). Even if this development has been documented by production data, no perception study has tested this assumption so far. To investigate this gap in research, a study employing event-related potentials (ERPs) was conducted focusing on cross-dialectal comprehension between both of these dialect groups.

Our study focuses on the question whether the usage of dialect phonemes (/ $\text{ɔ̄a}$ /, / $\text{ɔ̄u}$ /) which are attributed to different lexemes in the two contiguous dialect areas, leading to minimal pairs between these areas, evoke increased neural costs during sentence processing. If so, this would indicate difficulties in

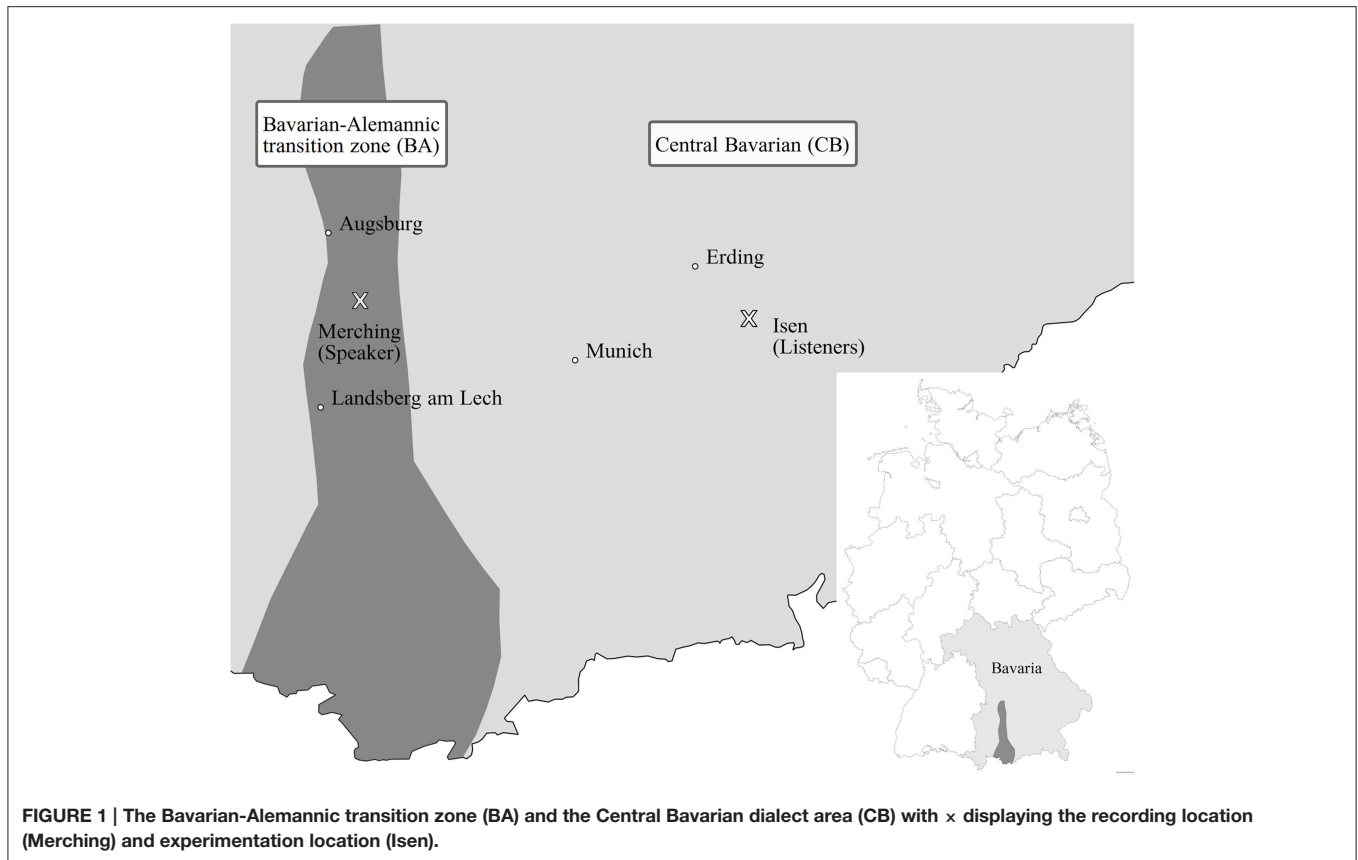
cross-dialectal comprehension. We are especially interested in semantic processing differences elicited by minimal phonological differences between different phoneme contact settings (misunderstanding, incomprehension, allophonic variation) in the form they can appear in everyday communication situations.

### The Impact of Regional Variation on Phoneme Perception in Online Observation

Cross-dialectal comprehension is highly related to the capacity of listeners to deal with acoustic variability in pronunciation resulting from dialectal variation. A listener's ability to perceive different speech sounds as phonemes mainly depends on the phoneme inventory of his own native language (cf. e.g., Buchwald et al., 1994 concerning the discrimination between /r/ and /l/ in Japanese). Studies using the electroencephalography (EEG) technique provide evidence that this assumption can be adapted to non-native regional phonemic contrasts within a language, as well. For instance, Brunellière et al. (2011) compared the / $\text{e}$ /-/ $\text{ɛ}$ / contrast in word-final open syllables (e.g., / $\text{epɛ}$ / 'sword' vs. / $\text{epɛ}$ / 'thick') in merged and unmerged French speaker groups. They found processing differences concerning the cortical topographies, indicating that in contrast to unmerged speakers, merged speakers associate the two forms with only one semantic representation (homophones). These clear differences between the groups support the assumption that the access to lexical meaning in spoken word recognition heavily depends on the listeners' native regional accent. The influence of the native phoneme inventory on phoneme perception was also investigated by Conrey et al. (2005), who focus on mechanisms of semantic integration and phonological decision processes using the example of the / $\text{i}$ / and / $\text{ɛ}$ / merger before nasal consonants in American English (the so-called *pin-pen* merger). The results show that in contrast to the unmerged group neither behavioral nor neural differences could be detected in the merged group. In contrast to previous studies, these results suggest that the different groups process the stimuli differently at a conscious, decisional level.

Previous studies dealing with dialect contrasts mainly used the Mismatch Negativity (MMN) component to examine vowel discrimination (cf. Brunellière et al., 2009, 2011; Scharinger et al., 2011). The MMN is a fronto-central negative component, usually peaking at 150–250 ms from change onset, when infrequent deviations (deviant stimuli) occur among frequently repeated sound patterns (standard stimuli) in a passive oddball design. The MMN is elicited regardless of the participant's direction of attention and thus reflects an automatic, pre-attentive response to any change in auditory stimulation. A basic prerequisite for MMN elicitation is the creation of a short-term memory trace in the auditory cortex, i.e., a representation of the repetitive standard stimulus. The MMN is the reflection of a discrimination process as the representation is violated by an infrequent deviant, indicating that the deviant is found to be incongruent with the memory representation of the preceding series of standard stimuli (cf. Näätänen et al., 2007). Using cross-linguistic oddball designs, several studies could establish language-specific memory

<sup>1</sup>Dialects can be defined as the least standard and most local (regionally restricted) varieties of a language and consist of independent prosodic, phonological and morpho-syntactic structures (cf. Schmidt, 2011).



traces for phonemes. The MMN deflection is increased when the deviant is a vowel category in the subject's native language in contrast to non-native vowel categories (Nääänen et al., 1997). Moreover, Kazanina et al. (2006) investigated the [t]-[d] contrast, which is mapped onto distinct phoneme categories in Russian, while it is an allophonic contrast in Korean. An MMNm was only elicited for the Russian listeners, indicating a rapid separation of these sounds into two categories, while the Korean listeners do not show any immediate sensitivity to the contrast. Furthermore, the MMN is also modulated by dialectal categories. Miglietta et al. (2013) compared the phonemic contrast [e]-[i] to the allophonic variation [ɛ]-[e] present in the Tricase dialect located in Southern Italy and found—in contrast to Kazanina et al. (2006)—an MMN response for both conditions. However, the latency of the phonemic condition was significantly earlier, pointing to a facilitated short-term memory trace formation in contrast to the allophonic condition.

In sum, these studies show that the electrophysiological investigation of allophonic and phonemic dialect contrasts is quite promising, since phonological and lexical processing stages are highly influenced by the listeners' regional accent.

## The Present ERP Study

In Southern Germany, the Bavarian (including CB) and the Alemannic dialect area adjoin each other. Between both areas, a

transition zone (BA) is located, in which phonological forms of both dialects interact with each other (cf. Wiesinger, 1983).

The investigated / $\delta\text{a}$ /-/ $\delta\text{u}$ / contrast distinguishes BA and CB since it is a stable contrast in CB, while in BA / $\delta\text{a}$ / occurs exclusively. Thus, the crucial point is how the respective phonemes have been assigned to lexemes in the dialect areas (see Supplementary Material for a precise diachronic description). For instance, / $r\delta\text{as}\eta$ / means 'roses' and 'journeys' in BA, while in CB it only corresponds to 'journeys'. Moreover, / $\text{ʃtr}\delta\text{a}$ / means 'straw' in BA, while in CB it is pronounced as / $\text{ʃtr}\delta\text{u}$ /.

Thus, the differences in the speakers' competences lead to different form-meaning-associations in the dialect areas. As a result, they might evoke different kinds of dialect-related communication difficulties when lexemes containing the / $\delta\text{a}$ -phoneme traced back to MHG  $\delta$  are used by speakers of BA. Three types of potential communication difficulties are reflected in our experimental conditions:

- (1) On the one hand, there might occur misunderstandings when a lexeme has different meanings in the two dialect areas. The usage of such lexemes in cross-dialectal communication leads to faulty decoding by the listeners. For example, the BA lexeme / $r\delta\text{as}\eta$ / for 'roses' means 'journeys' in CB.
- (2) On the other hand, there might arise incomprehension, since many BA words containing the / $\delta\text{a}$ -diphthong do not exist

in CB. In these cases, the listeners do not have a lexical entry for the lexeme and cannot decode it. For example, the BA lexeme /ʃtrôa/ 'straw' is not part of the CB speakers' competence.

- (3) A further differentiation between the dialects relevant for the present study affects the pronunciation of MHG *ô* before nasals (e.g., *Lohn* 'wage', *Bohne* 'bean'). Among other variants, one common pronunciation is /ôʊ/ in BA and /o:/ in CB. In cross-dialectal communication, it seems likely that articulatory very similar contrasts which are not integrated in the change of meaning between the dialects do not lead to erroneous decoding (cf., the BA variant /lôû/ versus the CB /lô:/ for 'wage').

An ERP study using an oddball design was conducted in order to study the effect of the different phoneme to lexeme assignments on cross-dialectal comprehension caused by the merger of MHG *ô* and MHG *ei* in BA (condition 1 and 2) in contrast to a probable pure allophonic contrast (condition 3).

In a typical oddball design, isolated phonemes, syllables or meaningful words are presented to participants while they watch a silent film. However, since in everyday communication listeners are faced with words embedded in complex sentences, we were interested in vowel perception during auditory sentence processing in order to investigate a more natural setting. This way, the influence of phonological differences on lexical-semantic processes should be made ascertainable.

It is however doubtful whether sentences can constitute an equally invariant acoustic context as isolated syllables or words against which deviants are normally compared. So far, only few studies have adapted the oddball design to questions of phonemic and semantic processes during sentence comprehension (e.g., Menning et al., 2005; Boulenger et al., 2011; Bendixen et al., 2014). Their results indeed demonstrate the sensitivity of the MMN to complex linguistic material and suggest that during natural speech processing, the brain rapidly extracts phonetic information from the continuous signal and forms memory traces in the auditory cortex. Thus, it seems that memory traces can also develop for complex sentences, which include large-scale details about phonetic features of the speech signal.

Following Bendixen et al. (2014), an experimental design combining a classic oddball paradigm and a semantic rating task was developed for the current study. In contrast to Bendixen et al. (2014) as well as Boulenger et al. (2011), the material used in this study also included semantic violations that might involve higher levels of processing such as semantic contextual integration. Effects of semantic integration are indicated by a rather late negativity peaking around 400 ms after stimulus onset in the EEG. The N400 is distributed primarily over centro-parietal sites and is typically elicited by sentence-final words that are semantically anomalous or of low cloze probability (cf. Kutas and Hillyard, 1980, 1984; Connolly and Phillips, 1994; Lau et al., 2008; Kutas and Federmeier, 2011). Thus, it displays the violation of predictions and expectations built up by the preceding sentence context since predictable words are easier to access from memory and it requires more resources to process an implausible or infrequent continuation (cf. Lau et al., 2008). In

the design of Boulenger et al. (2011), late ERP effects in the N400 time window were found indicating the involvement of semantic integration mechanisms. Thus, an oddball design including a task seems very promising to address the interaction between early acoustic processes and late mechanisms of semantic integration.

In the current study, a member of the N200 family (MMN, N2b) might be evoked as in the comparable study of Bendixen et al. (2014) for the conditions misunderstanding and incomprehension. However, due to the embedded semantic violations, which form an important difference to that study, we expect to also find an N400 in the misunderstanding condition. In contrast, we expect the fewest costs in semantic and lexical processing for the potential allophonic variation (condition 3) since the deviant might be categorized as a potential, allophonic form of the standard.

## MATERIALS AND METHODS

In the current study, listeners' perception of four different lexemes is investigated during sentence processing. The following three conditions reflect the special phoneme contact between the dialects of BA and CB (see Table 1). The experiment was conducted in CB, subsequently the terms 'standard' and 'deviant' are used with regard to the Central Bavarian dialect. The Central Bavarian lexemes served as the standard (2/3), while the Bavarian-Alemannic lexemes form the deviants (1/3).

### (1) Condition 1: Misunderstanding

Whereas in BA the (former) homophonic lexeme /rôasɲ/ signifies both 'roses' and 'journeys', in CB it only bears the meaning 'journeys'. When speakers from both areas communicate with one another, the usage of /rôasɲ/ 'roses' could lead to misunderstandings because it is always understood as 'journeys'.

- (a) All sentences presented prime for the meaning 'roses'. The Central Bavarian variant /rôûsɲ/ 'roses' is infrequently interrupted by the Bavarian-Alemannic variant /rôasɲ/, which means 'journeys' in Central Bavarian.

e.g., *Was im Garten viel Pflege braucht, sind Rosen.*  
'That which needs much care in the garden, are roses'

- (b) All sentences presented prime for the meaning 'journeys'. The Central Bavarian variant /rôasɲ/ 'journeys' is infrequently interrupted by the lexeme /rôûsɲ/ 'roses'.

e.g., *Wofür er seinen Koffer packt, sind Reisen.* 'That for which he is packing his suitcase, are journeys'

TABLE 1 | Experimental conditions.

Condition	Standard	Deviant
1a (misunderstanding)	/rôûsɲ/ 'roses'	/rôasɲ/ 'journeys'
1b (misunderstanding)	/rôasɲ/ 'journeys'	/rôûsɲ/ 'roses'
2 (incomprehension)	/lôûs/ 'sow'	/lôas/ pseudo-word
3 (potential comprehension)	/lô:/ 'wage'	/lôû/ potential allophonic form

## (2) Condition 2: Incomprehension

Whereas in BA the MHG lexeme *lôs* ‘sow’ is pronounced like /l<sup>h</sup>as/, in CB the dialectal lexeme is /l<sup>h</sup>ö:s/. In regional communication, the usage of /l<sup>h</sup>as/ may lead to incomprehension because this form is not part of the phonological competence of speakers from CB. Therefore, the lexical access fails.

Half of the sentences prime for the meaning ‘sow’, half are neutral sentences. The Central Bavarian variant /l<sup>h</sup>ö:s/ is infrequently interrupted by the Bavarian-Alemannic variant /l<sup>h</sup>as/.

priming: e.g., *Was die kleinen Ferkel säugt, ist die Lous.*  
‘That which is nursing the little piglets, is the sow’

neutral: e.g., *Was er ihr genau beschreibt, ist die Lous.*  
‘What he is describing to her exactly, is the sow’

## (3) Condition 3: Potential Comprehension

Whereas in BA ‘wage’ is pronounced like /l<sup>h</sup>ö:/, in CB the variant is /l<sup>h</sup>ö:/. Although the form is non-native to the listeners of CB, it is likely that it does not lead to erroneous decoding.

Half of the sentences prime for the meaning ‘wage’, half are neutral sentences. The Bavarian variant /l<sup>h</sup>ö:/ is sporadically interrupted by the Bavarian-Alemannic variant /l<sup>h</sup>ö:/.

priming: e.g., *Was man am Monatsende bekommt, ist der Lohn.* ‘That which you receive at the end of the month, is the wage’

neutral: e.g., *Worüber sie beim Treffen reden, ist der Lohn.*  
‘That which they are talking about at the meeting, is the wage’

## Pretests

A pool of German sentences ending in *Rosen* ‘roses’ (77), *Reisen* ‘journeys’ (104), *Lohn* ‘wage’ (85) or *Muttersau* ‘sow (female pig)’ (75) was developed using standard language vocabulary and syntax. All sentences followed the same global structure of topicalized relative clauses with the critical item in sentence final position (e.g., *Was wild in der Hecke blüht, sind Rosen.* ‘That which is blossoming wildly in the hedge, are roses’). The sentences were designed to create a semantic expectation of the critical items. In addition, neutral sentences were created, which were integrated into conditions 2 and 3, in which the lexical meaning of the sentence-final lexemes does not differ, in order to keep the participants’ attention. Using exclusively well-matched priming sentences would have been too invariant for the semantic rating task, so this was counterbalanced by adding neutral sentence contexts.

To guarantee that the sentences were classified as either priming or neutral context sentences, the high and low cloze probability of the sentences was surveyed within an online rating procedure. The full sentences were presented visually on a computer screen and participants had to evaluate on a scale from 1 to 7 whether the sentence-final word fit the context, with 1 indicating that the respective word did not fit at all and 7 indicating that the word fit very well (task 1). In a second task, participants had to decide whether the critical item fit

better, worse or equally well into the sentence in comparison to other possible options (task 2). If the participants stated that other words fit better into the context, they were asked to write them down (task 3). All sentences from all conditions were mixed up randomly into equal parts and separated into 8 groups, so that each participant had to judge approximately 60 sentences. Altogether, 78 speakers of the standard variety participated in the rating task (54 women, mean age 32.03 (SD 12.51)). In the subsequent steps, only sentences which were evaluated as very suitable regarding the sentence-final items were selected for priming conditions, reflected by the mean of >6 respectively >5.5 for *Muttersau*. In contrast, for neutral context conditions, moderately well-judged sentences were selected, reflected by a mean of 1.8–4.9 (*Muttersau*) respectively 2.9–5.5 (*Lohn*). This rating procedure resulted in 35 sentences for each condition, in total 210 sentences.

## Stimuli

All 210 sentences were recorded several times by a male native speaker (year of birth 1963) who was born and raised in the Bavarian-Alemannic transition zone (Merching). He adapted the sentences to the dialect lexically and phonologically and produced the critical items in two variants each—his own and the Central Bavarian one. During the recordings, a natural pronunciation and a normal speech rate was ensured, as well as a comparable realization of the native and non-native lexemes with regard to their intensity and pitch. All stimuli were digitally recorded with a sampling rate of 44.1 kHz and a 16 bit (mono) sample size, using an electret microphone (Sony ECM-MS957).

For each condition, the best 30 sentences were selected from the recorded auditory material, leading to 180 sentences in total. Furthermore, 10 critical tokens of each item were chosen due to their F1 and F2 values being as similar as possible (see **Table 2** for mean values). This acoustic variability was chosen in order to create a more natural speech perception and memory trace for the standard condition. This point is essential as it has been demonstrated that a higher and hence more natural variability in standard items allows for a more reliable abstraction or trace form from the different acoustic stimuli (cf. Phillips et al., 2000; Scharinger et al., 2011). Finally, four speakers from CB, who did not participate in the experiment, were asked to listen to the sentences to ensure that they are generally acceptable and comprehensible.

**TABLE 2 |** Phonetic values of the critical items (means).

Critical item	Stimulus length [ms]	Vowel length [ms]	F1 [Hz]		F2 [Hz]	
			20%	80%	20%	80%
/r <sup>h</sup> ö:sɪ/	656	150	451.2	329.2	1114.7	1024.8
/r <sup>h</sup> asɪ/	633	147	378.4	505.5	807.1	1278.8
/l <sup>h</sup> ö:s/	628	118	398.7	320.0	1023.0	1042.1
/l <sup>h</sup> as/	625	152	356.8	468.9	928.8	1238.5
/l <sup>h</sup> ö:/	527	186	348.4	335.7	736.1	704.2
/l <sup>h</sup> ö:/	527	209	467.5	321.7	1154.1	759.0

The selected sentences were cross-spliced in order to use the same carrier sentence for standard and deviant conditions. To avoid different context inferences, a defined pause of 100 ms was inserted in front of the verb preceding the critical stimulus. In total, the sentences have an average length of 2.4 s. The dynamic range was manipulated in order to create a consistent sound because of the acoustic variation between and within sentences as a result of splicing. The pitch, duration or formants were not manipulated. Finally, all of the chosen items were controlled for and normalized in intensity. All of the adjustments were carried out using the software Adobe Audition CS6 (version 5.0.2).

## Procedure

During the experiment, participants were seated comfortably in a dimly lit and quiet room. A computer screen was placed in front of them. Participants were instructed to listen to the auditorily presented sentences and to evaluate on a four-point-scale how well the sentence-final word fit the sentence context after the offset of each sentence. For each word pair of the misunderstanding condition, 180 sentences (30 prime sentences, presented four times = 120 primes; 30 deviant sentences presented twice = 60 deviants) were presented, distributed over 2 blocks containing 90 sentences each. For the word pairs of the incomprehension and potential comprehension conditions, 360 sentences (30 prime sentences, presented four times = 120 primes; 30 deviant sentences, presented twice = 60 deviants / 30 neutral sentences, presented four times = 120 neutral sentences; 30 deviant sentences, presented twice = 60 neutral deviants) in total, were distributed over 4 blocks containing 90 sentences each. All blocks of one word pair were presented directly following each other with only short breaks in between. The two word pairs of the misunderstanding condition did not directly follow each other in the presentation. In total, 1080 sentences were presented in 12 blocks consisting of 90 sentences each (60 standards / primes, 30 deviants), with each block of approximately 7 min duration. Between separate blocks, participants were offered a short break to rest their eyes. In order to avoid sequence effects, the block order was varied across participants.

Before the experiment started, the participants completed a short practice phase to ensure that the given task and further instructions regarding eye blink phases were understood. Thereafter, the first experimental block started with the request to the participant to click any key to begin the experiment to ensure the participants' full attention when each block started. Each trial began with the presentation of a fixation cross in the center of the computer screen for 500 ms. A stimulus embedded in a carrier sentence was played via two loudspeakers while the fixation cross remained displayed on the screen to minimize eye movements. After the offset of each sentence, the fixation cross was replaced by a question mark which gave the signal for the participants to rate how well the sentence-final word fitted the sentence context as accurately and as quickly as possible by pressing one of four buttons within maximally 2000 ms. The assignment of buttons to four possible answers (very well, rather well, rather badly, very badly) was counterbalanced across participants. During the question mark phase, participants were allowed to blink and rest

their eyes. 1000 ms after each response or time-out, the next trial started with an upcoming fixation cross. All procedures were performed in compliance with relevant laws and institutional guidelines.

## Participants

Twenty (13 women; mean age 44.5 (SD 4.87), age range 34–53) right-handed monolingual native speakers of German with normal or corrected-to-normal vision participated in the experiment. None of the participants had hearing deficits. All participants were born and raised in Isen (located in CB) and still live there. Their dialect competence was tested via a dialect pre-test. All participants gave their informed consent to this study and privacy rights were thoroughly obeyed. Each participant received monetary compensation for taking part in the study.

## ERP Recording and Data Processing

An electroencephalogram (EEG) was recorded from 26 Ag/AgCl electrodes, mounted on an elastic cap (EasyCap), according to the 10–20 system (F7, F3, Fz, F4, F8, FC5, FC1, FCz, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CPz, CP2, CP6, P7, P3, Pz, P4, P8, POz) with a *BrainVision* (Brain Products GmbH) amplifier. The C2 electrode served as ground electrode, the reference electrode was placed at the tip of the nose. Two further electrodes were placed at the left and right mastoid sites. To measure the electrooculogram (EOG), two electrodes were placed below and above the left eye and two further electrodes laterally to the outer canthi of both eyes in order to control for horizontal and vertical eye movements. All electrode impedances were kept below 5 k $\Omega$ . EEG and EOG were recorded continuously with a sampling rate of 500 Hz and filtered offline with a 0.3–20 Hz bandpass filter. This filter setting was chosen in order to remove slow drifts from the signal and the 20 Hz low pass filter was chosen following previous ERP studies investigating natural speech processing (cf. Menning et al., 2005; Partanen et al., 2011). Then, EEG recordings were re-referenced offline to the linked mastoids.

Prior to data analysis, all individual EEG recordings were scanned for artifacts from body movements or eye blinks. All artifacts exceeding an amplitude of 40 microvolt were automatically removed from the data set. A subsequent manual inspection of all single-trial waveforms scanned for further artifacts in all EEG channels. Data sets with more than 25% artifacts within one condition were excluded from further analyses. As a result of these inspections, the data set of 1 participant (1 male) had to be excluded from the misunderstanding condition and the potential comprehension condition; 2 data sets (1 male) had to be excluded from analysis for the incomprehension condition. These data sets were also excluded from the respective behavioral data analyses (see Supplementary Material for exact numbers of rejections). From the overall data set of the remaining participants, 5.6% of the misunderstanding condition stimuli, 6.0% of the incomprehension condition stimuli and 5.0% of the potential comprehension condition stimuli were excluded from analysis.

## Data Analyses

The arithmetical mean of all responses for each condition was calculated by allocating a numerical value to each of the four possible response levels: 1  $\hat{=}$  very well, 2  $\hat{=}$  rather well, 3  $\hat{=}$  rather badly, and 4  $\hat{=}$  very badly. The arithmetical means were analyzed with an ANOVA separately for each condition pair, the respective factors are therefore presented in the results section of each condition (see Sections Behavioral data Condition 1: Misunderstanding–Behavioral data Condition 3: Potential comprehension). Further analyses of comparisons of each word pair were conducted using the Wilcoxon signed-rank test with a Bonferroni correction for the  $p$ -values.

In order to prevent movement artifacts, the evaluation response was given with a delay after the offset of the sentence. Due to this temporal distance between the perception of each critical item and the response, the measured reaction times for each evaluation response are not meaningful and not reported here.

For the EEG data, a multifactorial repeated-measures ANOVA was calculated with the factors CONDITION (standard vs. deviant) and REGION [frontal (F3, FZ, F4), central (C3, CZ, C4), and parietal (P3, PZ, P4)]. Averages were calculated from onset of each sentence-final word up to 1000 ms thereafter for the misunderstanding condition and up to 900 ms thereafter for the two other conditions, with a baseline of 100 ms preceding the onset. The analysis was conducted with consecutive epochs of 50 ms from 0 ms up to 900 ms and 1000 ms respectively. Moreover, time windows for each paired comparison were also chosen based on hypotheses taken from the literature with similar experimental set-ups (Domahs et al., 2009; Boulenger et al., 2011; Bendixen et al., 2014) and were adjusted on the basis of visual inspection of the grand average curves. For effects with more than one degree of freedom, Huynh and Feldt (1976) corrections were applied to the  $p$ -values.

## RESULTS

### Condition 1: Misunderstanding

#### Behavioral Data

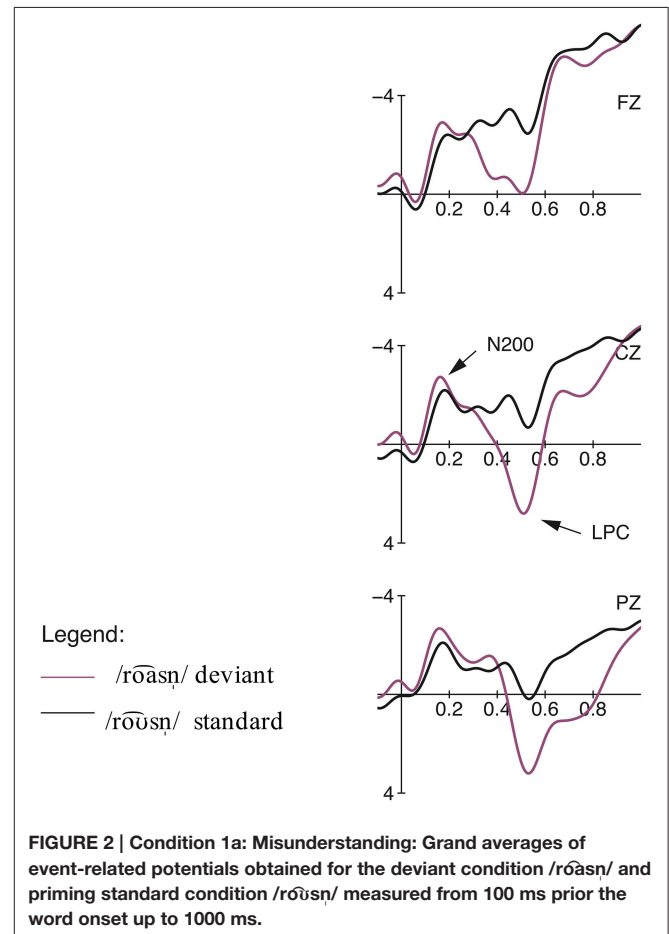
The ANOVA for the misunderstanding condition with the factors DIPHTHONG (/rōʊsn̩/ ‘roses’ vs. /rōʌsn̩/ ‘journeys’) and EXPECTANCY (priming fulfilled vs. deviant) revealed only a main effect for the factor EXPECTANCY [ $F_{(1, 18)} = 605.67, p = 0.000$ ], but not for the factor DIPHTHONG [ $F_{(1, 18)} = 0.56, p > 0.05$ ], i.e., the diphthong itself did not have a significant influence on the participants’ evaluation but only the compliance or non-compliance with the built-up expectancy due to the priming sentence context. Further analyses support this finding as they show that both correct priming conditions primed equally well regardless of the diphthong [prime /rōʊsn̩/ ‘roses’: mean 1.12 (SD 0.20) vs. /rōʌsn̩/ ‘journeys’: mean 1.15 (SD 0.17);  $Z_{(18)} = -1.01, p > 0.05$ ] while the non-compliant deviants were evaluated as significantly less acceptable than the correct sentence-final words in the priming condition sentences [on a scale from 1 = acceptable to 4 = unacceptable; prime /rōʊsn̩/ ‘roses’ mean 1.12 (SD 0.20) vs. deviant /rōʌsn̩/ ‘journeys’: mean 3.72 (SD 0.79);  $Z_{(18)} = -3.79, p = 0.000$  / prime /rōʌsn̩/ ‘journeys’: mean 1.15

(SD 0.17) vs. deviant /rōʊsn̩/ ‘roses’ mean 3.90 (SD 0.08);  $Z_{(18)} = -3.82, p = 0.000$ ].

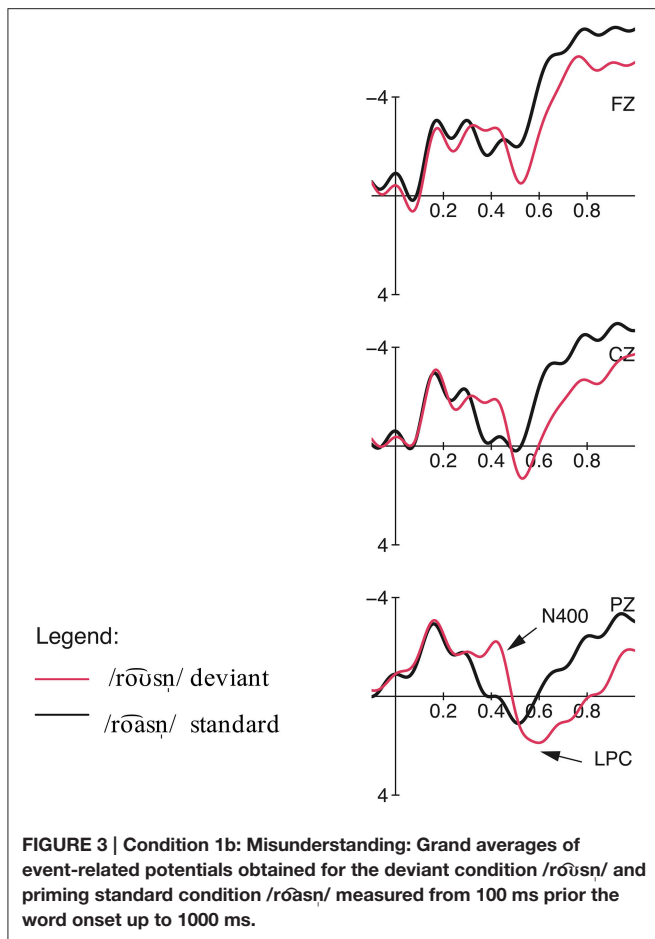
#### ERP Data

The comparison of the standard /rōʊsn̩/ ‘roses’ and the deviant /rōʌsn̩/ ‘journeys’ elicited an early negativity effect in the time window from 100 to 200 ms, followed by a late positivity effect from 400 to 900 ms (cf. **Figures 2, 4**). The calculation of a repeated measures ANOVA revealed a significant main effect for the factor EXPECTANCY (priming fulfilled vs. deviant) [ $F_{(1, 18)} = 11.52, p = 0.003, \eta^2 p = 0.05$ ] but no significant interaction between the two factors EXPECTANCY and REGION in the early time window [ $F_{(2, 36)} < 1, p > 0.05, \eta^2 p = 0.00$ ]. In the later time window, the statistical analysis showed a main effect for both conditions [EXPECTANCY:  $F_{(1, 18)} = 26.35, p = 0.000, \eta^2 p = 0.07$ ; REGION:  $F_{(2, 36)} = 28.08, p = 0.000, \eta^2 p = 0.23$ ] as well as a significant interaction between them [ $F_{(2, 36)} = 15.49, p = 0.000, \eta^2 p = 0.01$ ]. The *post-hoc* analysis of this interaction by REGION revealed a stronger occurrence of the late positive component in the centro-parietal regions [frontal:  $F_{(1, 18)} = 7.37, p < 0.05, \eta^2 p = 0.01$ ; central:  $F_{(1, 18)} = 26.55, p = 0.000, \eta^2 p = 0.12$ ; parietal:  $F_{(1, 18)} = 44.45, p = 0.000, \eta^2 p = 0.28$ ].

For the comparison of the standard /rōʌsn̩/ ‘journeys’ and the deviant /rōʊsn̩/ ‘roses’ (see **Figures 3, 4**), the ANOVA showed



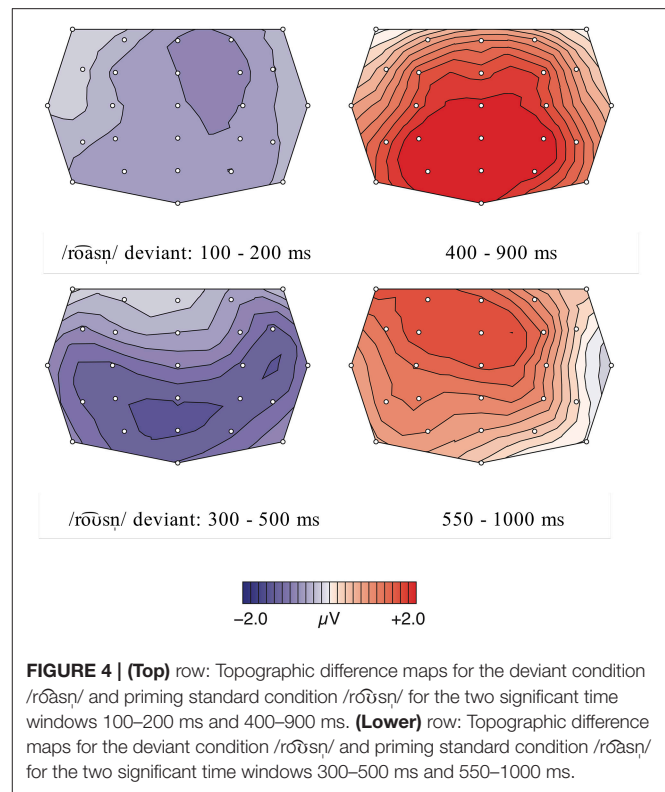
**FIGURE 2 | Condition 1a: Misunderstanding: Grand averages of event-related potentials obtained for the deviant condition /rōʌsn̩/ and priming standard condition /rōʊsn̩/ measured from 100 ms prior to the word onset up to 1000 ms.**



no significant effects in the early negativity time window (100–200 ms) but a significant negativity effect was elicited by the deviant condition in a later time window from 300 to 500 ms [EXPECTANCY:  $F_{(1, 18)} = 9.19, p = 0.007, \eta^2 p = 0.04$ ; REGION:  $F_{(2, 36)} = 7.47, p = 0.012, \eta^2 p = 0.07$ ]. The significant interaction between the factors EXPECTANCY and REGION [ $F_{(2, 36)} = 5.37, p = 0.028, \eta^2 p = 0.01$ ] resolved by REGION revealed a stronger occurrence of the early negativity in the centro-parietal region [frontal:  $F_{(1, 18)} = 1.33, p > 0.05, \eta^2 p = 0.00$ ; central:  $F_{(1, 18)} = 10.53, p < 0.01, \eta^2 p = 0.08$ ; parietal:  $F_{(1, 18)} = 14.96, p < 0.01, \eta^2 p = 0.13$ ]. In the time window from 550 to 1000 ms, a positivity effect was elicited by the deviant condition [EXPECTANCY:  $F_{(1, 18)} = 22.22, p = 0.000, \eta^2 p = 0.04$ ; REGION:  $F_{(2, 36)} = 20.19, p = 0.000, \eta^2 p = 0.21$ ] but there was no significant interaction between the two factors [ $F_{(2, 36)} = 2.66, p > 0.05, \eta^2 p = 0.00$ ].

## Discussion

In the misunderstanding condition, the deviant  $rɔʊsn̩/rɔʌsn̩$  elicited a negativity effect in the latency range between 300 and 500 ms. Due to its latency as well as its centro-parietal scalp distribution, which is typical for the context-dependent N400 effect (cf. Lau et al., 2008; Kutas and Federmeier, 2011),



we interpret this negativity as an N400 reflecting the semantic mismatch between the expected continuation of the sentence and the perceived input. The semantic priming of ‘journeys’ builds up an expectation for the correct word form /rɔʌsn̩/. The semantically incongruous word /rɔʊsn̩/ leads to a strong semantic mismatch between the context-based information held in working memory and the unfitting item. The N400 effect indexes an increased integration difficulty of the (incongruous) critical item with the prior sentence context (cf. Kutas and Hillyard, 1980; Brown and Hagoort, 1993; Kutas and Federmeier, 2000; Lau et al., 2009). Moreover, the N400 amplitude is modulated by the ease of accessing information from long-term memory (cf. Kutas and Federmeier, 2000; Lau et al., 2009). Due to the priming context, a congruous ending is pre-activated which is then disrupted by the incongruous sentence-final word. This disruption leads to higher processing costs, displayed by the pronounced N400 amplitude. The semantic mismatch between expectancy and perceived input is also reflected by the behavioral data, since sentences with unexpected final lexeme /rɔʊsn̩/ were evaluated significantly worse than the sentences ending with the predictable word /rɔʌsn̩/.

Interestingly, in the reversed condition our results show an earlier negativity effect between 100 and 200 ms for the deviant  $rɔʌsn̩/rɔʊsn̩$ . In previous studies, early negativities in similar time windows have been interpreted as detections of sudden changes in acoustic features of speech sounds embedded in sentences and have been classified as members of the N200 family (cf. Boulenger et al., 2011 for reversed speech; Bendixen et al., 2014 for omissions). The N200 component is distinguished into



three sub-components (N2a, N2b, N2c) and is typically evoked between 180 and 325 ms respectively 100 and 200 ms (N2a) (cf. Patel and Azzam, 2005). Beside differences in latency and topography, early negativity effects are primarily separated on the basis of their sensitivity to varying task conditions. While the MMN (N2a) reflects pre-attentive passive change detections, the N2b is elicited by task relevant physical mismatches and thus requires attention to the stimuli. Thus, both components index different stages of mismatch detection (cf. Ritter et al., 1984; Pritchard et al., 1991; Folstein and Van Petten, 2008). Bendixen et al. (2014) support the component's dependency on active listening and its reflection of conscious error and mismatch detection, as well. In line with this perspective, we interpret the early negativity effect found for  $r\text{öas}\eta/r\text{öüs}\eta$  as an instance of the N2b, a reflection of a general rule-governed error detection mechanism. Since participants' attention was directed explicitly to the critical items due to the rating task, the N200 effect reflects an active discrimination and classification process, elicited by the deviation from a mentally-stored expectation of the standard stimulus (cf. Patel and Azzam, 2005), which more generally reflects the recognition of deviations and violations in regular structures (cf. Bohn et al., 2013; Henrich et al., 2014 for rhythmic irregularities). Thus, the semantic priming of 'roses' builds up an expectation for the correct word form  $/r\text{öüs}\eta/$ . The deviation  $/r\text{öas}\eta/$  is perceived as being different from the activated standard stimulus  $/r\text{öüs}\eta/$  and thus not fitting the activated memory trace.

Furthermore, late positive components (LPC) were elicited for  $r\text{öüs}\eta/r\text{öas}\eta$  (550–1000 ms) and  $r\text{öas}\eta/r\text{öüs}\eta$  (400–900 ms). We interpret these positivity effects as members of the P300 family, reflecting the evaluation process related to the given task requirements (Bentin et al., 1999; Knaus et al., 2007; Roehm et al., 2007; Domahs et al., 2008, 2009; Bohn et al., 2013; Henrich et al., 2014). Recall that participants had to perform a semantic rating task, i.e., their attention was directed consciously toward the linguistic material. Thus, the elicited positivity reflects the match or mismatch between the expected and encountered word form, i.e., the comparison of the critical stimulus with the expectation built-up by the memory trace of the standard form and—in priming contexts—the semantic information from the sentence context. The P300 is a positive deflection evoked by meaningful, task-relevant stimuli only when the subjects' attention is required for the task (cf. Picton, 1992). The P300 can be situated in processes of categorization, decision making, and context updating (cf. Coulson et al., 1998). Furthermore, the LPC may reflect a reanalysis process with regard to the semantic correctness of the presented sentences (cf. Domahs et al., 2008; Henrich et al., 2014). A deviating word form in sentence-final position requires a reanalysis and reevaluation of the previously built-up structure because the repetitively presented standard and the priming sentence context build up high expectations for a certain form. In this respect, the amplitude of the LPC is additionally modulated by the degree of the required reanalysis process, i.e., by the degree of deviation between what was expected and what was encountered.

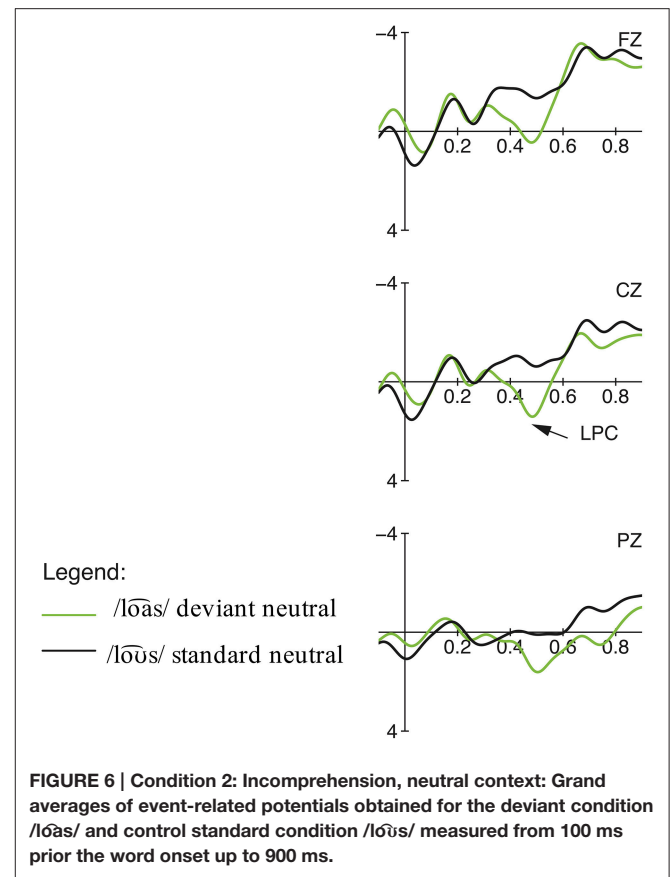
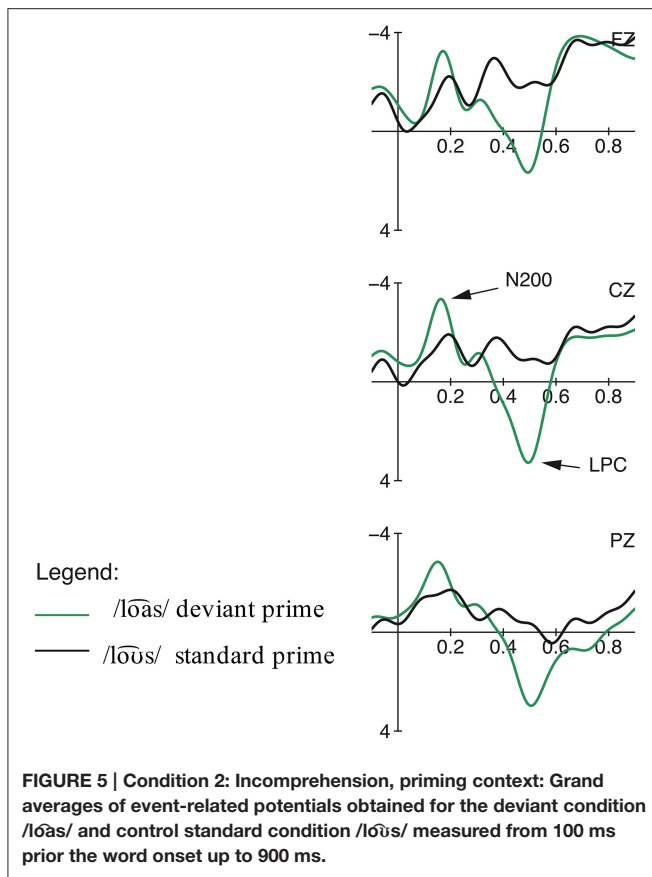
## Condition 2: Incomprehension

### Behavioral Data

The ANOVA for the incomprehension condition with the factors WORD TYPE (correct word  $/\text{löüs}/$  'sow' vs. pseudo-word  $/\text{löäs}/$ ) and SENTENCE CONTEXT (priming context vs. neutral context) showed main effects for both factors [WORD TYPE:  $F_{(1, 17)} = 6.06$ ,  $p = 0.025$ ; SENTENCE CONTEXT:  $F_{(1, 17)} = 15.14$ ,  $p = 0.001$ ] but no significant interaction between the two factors [ $F_{(1, 17)} < 1$ ,  $p > 0.05$ ]. These results confirm that structures with the diphthong  $/\text{öä}/$  were detected as non-existent pseudo-words and evaluated as significantly less acceptable than real words with the diphthong  $/\text{öü}/$ . A closer look at the arithmetical means shows that pseudo-words were apparently evaluated as less acceptable than real words when the priming sentence context built up expectancies which word should end the sentence. However, due to a high dispersion of the standard deviation, this difference is not statistically significant [correct word  $/\text{löüs}/$ , mean 1.74 (SD 0.95) vs. pseudo-word  $/\text{löäs}/$ , mean 2.72 (SD 1.31);  $Z_{(17)} = -1.68$ ,  $p > 0.05$ ]. In the neutral sentence contexts, pseudo-words were significantly rejected more often as unacceptable sentence-final words in comparison to the correct word form [correct word  $/\text{löüs}/$ , mean 2.41 (SD 0.74) vs. pseudo-word  $/\text{löäs}/$ , mean 3.07 (SD 0.91);  $Z_{(17)} = -2.33$ ,  $p = 0.020$ ]. The paired contrast of correct word forms depending on the embedded sentence context showed that the correct word form was evaluated as more acceptable in the priming sentence context [priming  $/\text{löüs}/$ , mean 1.74 (SD 0.95) vs. neutral  $/\text{löüs}/$ , mean 2.41 (SD 0.74);  $Z_{(17)} = -2.24$ ,  $p = 0.025$ ]. This shows that the sentence context had a significant influence on the lexical categorization process, i.e., real words were identified as such more easily when the sentence context gave specific cues and helped in building up expectations toward the incoming speech signal.

### ERP Data

The statistical analysis of the incomprehension condition calculating a repeated measures ANOVA with the factors WORD TYPE (real word vs. pseudo-word) and REGION revealed an early negativity effect between 100 and 200 ms, elicited by the deviant pseudo-words (see Figures 5–7). This early negativity effect is only significant in the priming condition, leading to a significant main effect for the factor WORD TYPE [ $F_{(1, 17)} = 28.88$ ,  $p = 0.000$ ,  $\eta^2 p = 0.08$ ]. In contrast, there were no significant main effects in the neutral sentence context condition [WORD TYPE:  $F_{(1, 17)} = 1.17$ ,  $p > 0.05$ ,  $\eta^2 p = 0.01$ ; REGION:  $F_{(2, 34)} = 1.90$ ,  $p > 0.05$ ,  $\eta^2 p = 0.02$ ]. In both contexts, no significant interaction between the two factors was observed [priming context:  $F_{(2, 34)} = 1.91$ ,  $p > 0.05$ ,  $\eta^2 p = 0.00$  / neutral context:  $F_{(2, 34)} < 1$ ,  $p > 0.05$ ,  $\eta^2 p = 0.00$ ]. In the priming context condition, this early negativity was followed by a late positive component (350–600 ms) [WORD TYPE:  $F_{(1, 17)} = 23.33$ ,  $p = 0.000$ ,  $\eta^2 p = 0.15$ ; REGION:  $F_{(2, 34)} = 19.66$ ,  $p = 0.000$ ,  $\eta^2 p = 0.09$ ]. This late positivity was also significant in the neutral context condition [WORD TYPE:  $F_{(1, 17)} = 22.96$ ,  $p = 0.000$ ,  $\eta^2 p = 0.08$ ; REGION:  $F_{(2, 34)} = 29.14$ ,  $p = 0.000$ ,  $\eta^2 p = 0.15$ ].

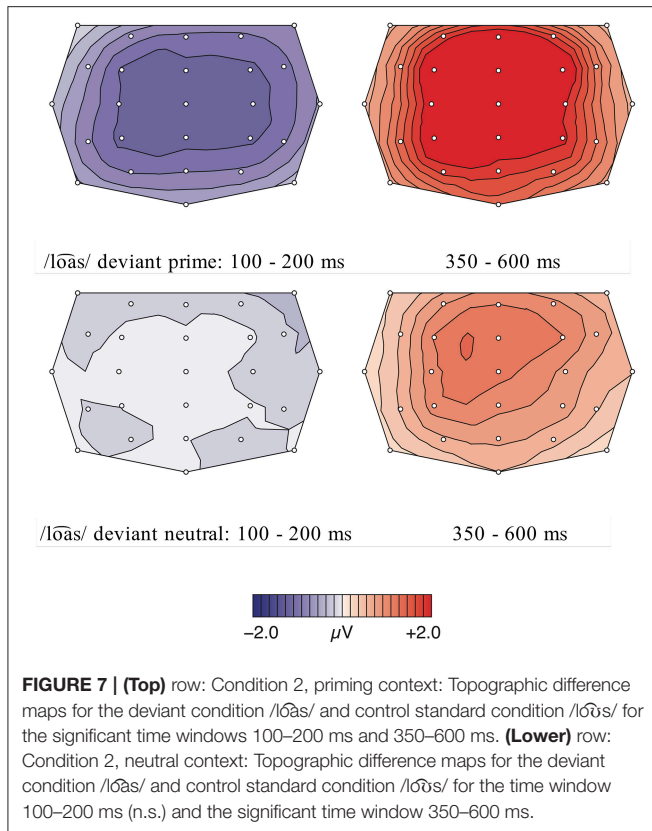


## Discussion

The ERP results for the deviant  $l\hat{o}as/l\hat{o}ʊs$  in the priming condition show an early negative effect between 100 and 200 ms, which is absent in the neutral sentence context. We interpret this early negativity analogous to the negativity found for  $r\hat{o}as\eta/r\hat{o}ʊs\eta$  as the reflection of an early error detection mechanism. Due to the priming sentence context, the listener builds up strong expectancies toward the following incoming speech signal, i.e., for  $l\hat{o}ʊs$ , which are violated when the deviating form  $l\hat{o}as$  is perceived instead. Since only two different alternating phonological forms were presented, participants were able to differentiate the two lexemes very early, as  $l\hat{o}as$  and  $l\hat{o}ʊs$  diverge from the onset of the diphthong in the F1 and F2 values (see **Table 2**), a fact reflected by the early latency of the negativity effect for  $l\hat{o}as/l\hat{o}ʊs$ . The absence of the N200 effect for neutral sentences emphasizes the influence of semantic priming on perception. Since the context does not provide any hints about the continuation of the sentence, there are no high expectancies built up and thus not violated.

Furthermore, an LPC (350–600 ms) was elicited for  $l\hat{o}as/l\hat{o}ʊs$  in the priming condition and in the neutral condition. We interpret the late positive deflection similarly as those elicited for condition 1 (see Section Discussion Condition 1: Misunderstanding). The amplitude of the LPC can also be modulated by the degree of complexity and difficulty, i.e.,

the resolvability of the given task: A more pronounced effect correlates with the simplicity of a stimulus according to the task requirement and thus with an easy evaluation (cf. Bentin et al., 1999; Domahs et al., 2009; Bohn et al., 2013). This relation is apparent in the ERP results for  $l\hat{o}as/l\hat{o}ʊs$ , which show an enhanced positivity effect for primed sentences in contrast to neutral sentences. This indicates that the categorization of a stimulus is easier when participants are able to build up an expectation through priming. In neutral sentences, the context does not give indications about the sentences' progression, so participants cannot rely on clear cues for their evaluation. Thus, the rejection of the non-word  $l\hat{o}as$  is more difficult for the participants, which is indicated by a reduced late positive component for  $l\hat{o}as/l\hat{o}ʊs$  in neutral sentences. Furthermore, the observed amplitude differences within this condition can also be explained by a reflected reanalysis process: in the priming context, the semantic cues enhance the expectations for a specific form; thus, the mismatch between the expected and encountered form is even bigger than in the neutral sentence context in which only the memory trace of the standard form built up an expectation. The stronger cues in the priming condition and thus the stronger degree of deviation are reflected in the more pronounced amplitude of the LPC in this sentence context condition.



### Condition 3: Potential Comprehension Behavioral Data

The analysis of the potential comprehension condition calculating an ANOVA with the factors PHONEME and SENTENCE CONTEXT revealed only a significant main effect for the factor SENTENCE CONTEXT but no effect for the factor PHONEME [SENTENCE CONTEXT:  $F_{(1, 18)} = 40.63$ ,  $p = 0.000$ ; PHONEME:  $F_{(1, 18)} < 1$ ,  $p > 0.05$ ] and no significant interaction between the two factors [ $F_{(1, 18)} < 1$ ,  $p > 0.05$ ]. Thus, the acceptability and categorization process of the sentence-final word did not depend on the included phoneme but only on the sentence context the words were embedded in. Further paired comparisons of the arithmetical means showed that the standard form /lɔ̃:/ and the deviating form /lɔ̃ʊ/ were evaluated as equally acceptable sentence-final words in both sentence contexts [priming context: /lɔ̃:/, mean 1.31 (SD 0.45) vs. /lɔ̃ʊ/, mean 1.32 (SD 0.41);  $Z_{(18)} = -0.28$ ,  $p > 0.05$  / neutral context: /lɔ̃:/, mean 2.31 (SD 0.69) vs. /lɔ̃ʊ/, mean 2.33 (SD 0.60);  $Z_{(18)} = -0.24$ ,  $p > 0.05$ ]. However, the categorization was significantly easier in the priming context condition for both phoneme types [priming context: /lɔ̃:/, mean 1.31 (SD 0.45) vs. neutral context: /lɔ̃:/, mean 2.31 (SD 0.69);  $Z_{(18)} = -3.82$ ,  $p = 0.000$  / priming context: /lɔ̃ʊ/, mean 1.32 (SD 0.41) vs. neutral context: /lɔ̃ʊ/, mean 2.33 (SD 0.60);  $Z_{(18)} = -3.74$ ,  $p = 0.000$ ].

### ERP Data

The comparison of the standard form and the deviant form in the potential comprehension condition revealed two

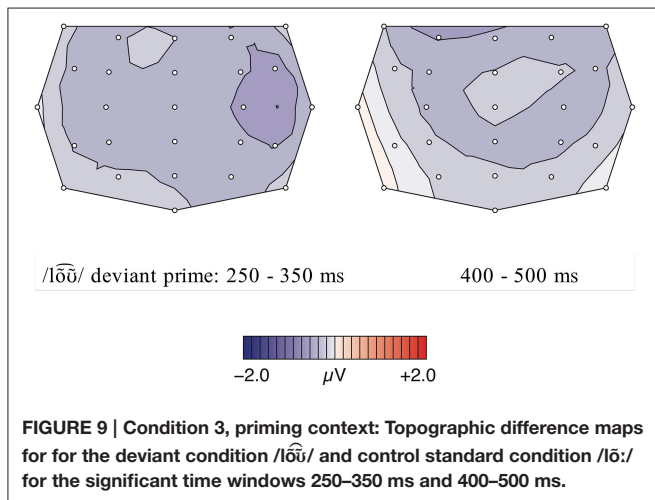
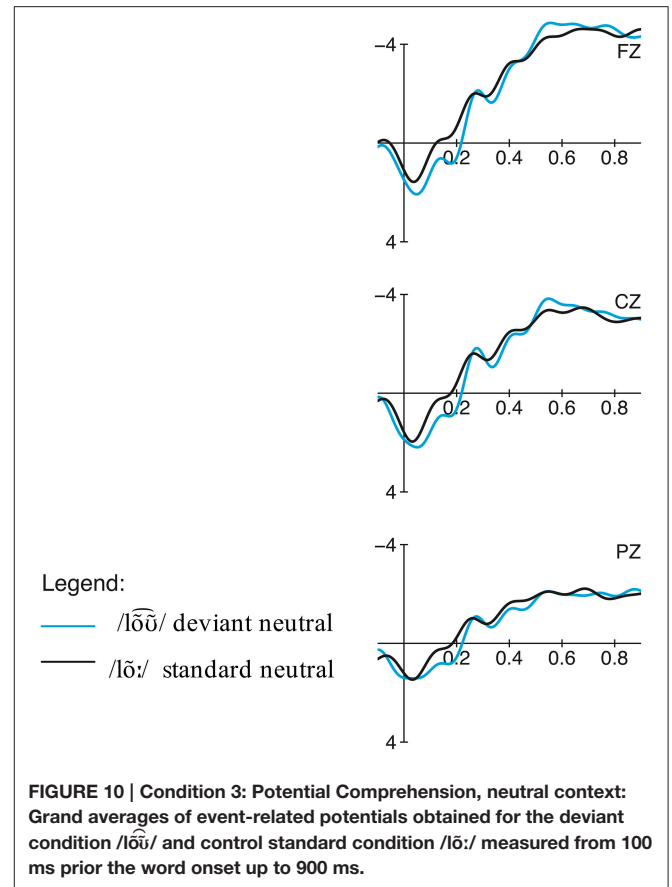
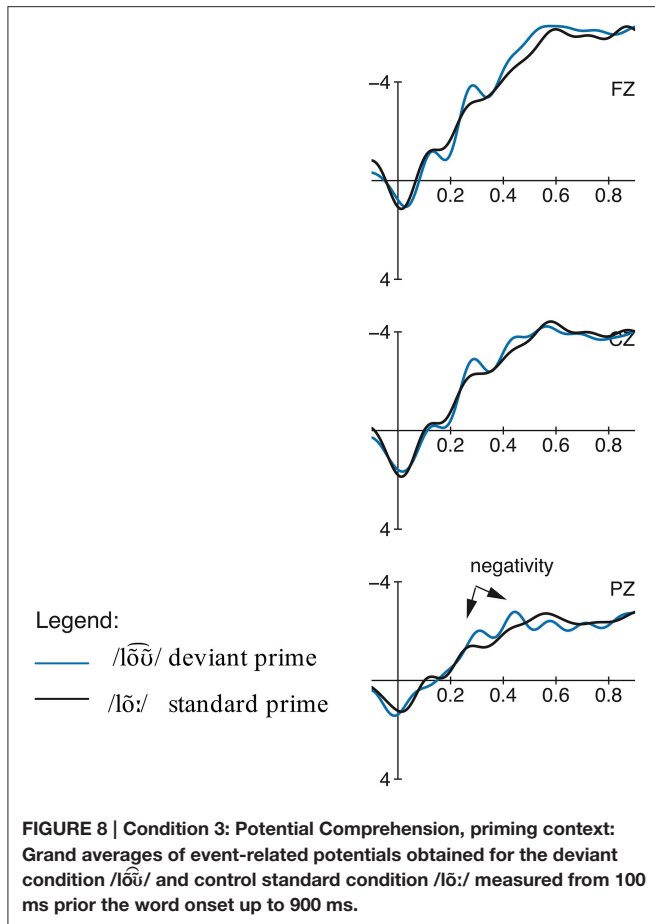
moderate consecutive negativity effects, elicited by the deviant condition in the time windows 250–350 ms and 400–500 ms (see **Figures 8–10**). The repeated measures ANOVA showed significant main effects for the factors PHONEME (standard /lɔ̃:/ vs. deviant /lɔ̃ʊ/) and REGION in the first time window [PHONEME:  $F_{(1, 18)} = 5.26$ ,  $p = 0.034$ ,  $\eta^2 p = 0.01$ ; REGION:  $F_{(2, 36)} = 14.67$ ,  $p = 0.001$ ,  $\eta^2 p = 0.11$ ] as well as in the second time window [PHONEME:  $F_{(1, 18)} = 7.46$ ,  $p = 0.014$ ,  $\eta^2 p = 0.01$ ; REGION:  $F_{(2, 36)} = 15.93$ ,  $p = 0.001$ ,  $\eta^2 p = 0.15$ ] but no significant interactions between these two factors. The same comparison embedded in neutral sentence context did not reveal any significant effects in any time window.

### Discussion

With regard to the ERP results, two moderate negativities for lɔ̃ʊ/lɔ̃: between 250–350 ms and 400–500 ms were obtained, which are absent in the neutral condition. In contrast to the other conditions, the amplitude of these negative deflections is much less pronounced. Although /lɔ̃as/ as well as /lɔ̃ʊ/ both do not match the native Central Bavarian lexemes, they differ in their phonetic distance to the native variants. In contrast to pseudo-words, the lexemes /lɔ̃:/ and /lɔ̃ʊ/ are related and thus do not elicit the same pronounced negativity as unrelated words do (cf. Kutas and Van Petten, 1994). Thus, the reduced negative deflections in contrast to the more pronounced one for lɔ̃as/lɔ̃ʊs can be interpreted as simply reflecting the perception of a phonetic deviation and thus a slightly hindered lexical access caused by the phonetically deviating form. This is in contrast to results from an MMN study on Italian vowel contrasts by Miglietta et al. (2013) who found similarly pronounced effects for both an allophonic ([ɛ]-[e]) and phonemic ([e]-[i]) condition. However, since this study concentrated on the processing of isolated vowels without lexical or semantic context, this does not refute our findings. This is in line with the fact that the sentence context plays a decisive role, since the negativities are only elicited when the meaning ‘wage’ has already been pre-activated through priming. In neutral contexts, however, no negative effects could be observed since no strong expectations toward a special form had been built up in these sentences.

Furthermore, the behavioral data show no difference in the rating between the match and mismatch condition, indicating that the deviant /lɔ̃ʊ/ is also interpreted in terms of the meaning ‘wage’. However, the evaluation process was easier in the priming context, i.e., the cues of semantic expectation facilitated the categorization process for both forms in these sentences.

**Table 3** gives an overview of the most important results of the behavioral data. **Table 4** displays an overview of relevant time windows and significant ERP results for all conducted comparisons of the three conditions described in the sections ERP data Condition 1: Misunderstanding–ERP data Condition 3: Potential comprehension. **Table 5** contains mean latency and amplitude values measured from the difference waves of all statistically significant comparisons.



## GENERAL DISCUSSION

The present paper explores the potential difficulties in cross-dialectal comprehension attributable to the close contact of a merged and an unmerged speech community. In contrast to previous studies, the investigated phonemes /ōā/ and /ōũ/ occur

in both dialect areas, but are assigned to different lexemes. In fact, we found different ERP effects for three different types of potential communication difficulties. These are discussed first with regard to their asymmetric latency distribution found within the misunderstanding condition and second in terms of the connection between cross-dialectal comprehension and phoneme change.

### Asymmetric ERP Effects in Misunderstanding Condition

In the misunderstanding condition, we expected to find an N400 effect for both deviant types due to the identical semantic priming context. However, while we indeed found a centro-parietal negativity between 300 and 500 ms for rōũsŋ/rōāsŋ, the negativity for rōāsŋ/rōũsŋ was elicited in an earlier time window between 100 and 200 ms. This is somewhat unexpected, since the same phonetic material for the critical items was used and behavioral data show a significant difference between both predictable and unpredictable sentence-final lexemes, suggesting that the forms were perceived differently and were easy to distinguish. One reason for the absence of an N400 effect might be the low variability of semantically alternating sentence-final word forms, as Bendixen et al. (2014) suggest. Therefore, it is possible that both words were semantically activated during the experiment. Furthermore, studies have shown that a high

**TABLE 3 | Behavioral data: mean evaluations of all responses for each standard-deviant pair of each condition (scale from 1 = acceptable to 4 = unacceptable).**

Condition	Sentence context	Evaluation (mean)	p-value	Critical items
Misunderstanding	Priming /rɔ̃asɲ/	standard: 1.15 deviant: 3.90	0.000***	/rɔ̃asɲ/ 'journeys' vs. /rɔ̃ʊsɲ/ 'roses'
	Priming /rɔ̃ʊsɲ/	standard: 1.12 deviant: 3.72	0.000***	/rɔ̃ʊsɲ/ 'roses' vs. /rɔ̃asɲ/ 'journeys'
Incomprehension	Priming /lɔ̃ʊs/	standard: 1.74 deviant: 2.72	>0.05 n.s.	/lɔ̃ʊs/ 'sow' vs. /lɔ̃as/ pseudo-word
	Neutral	standard: 2.41 deviant: 3.07	0.020*	
Potential comprehension	Priming /lɔ̃:/	standard: 1.31 deviant: 1.32	>0.05 n.s.	/lɔ̃:/ 'wage' vs. /lɔ̃ʊ/ potential allophonic form
	Neutral	standard: 2.31 deviant: 2.33	>0.05 n.s.	

Statistical significance is indicated by \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

**TABLE 4 | ERP effects in different time windows for three different conditions.**

Condition	Sentence context	Negativity	Positivity	Critical items
Misunderstanding	Priming /rɔ̃asɲ/	100–200**	400–900***	/rɔ̃ʊsɲ/ 'roses' vs. /rɔ̃asɲ/ 'journeys'
	Priming /rɔ̃ʊsɲ/	300–500**	550–1000***	
Incomprehension	Priming /lɔ̃ʊs/	100–200***	350–600***	/lɔ̃ʊs/ 'sow' vs. /lɔ̃as/ pseudo-word
	Neutral	100–200 n.s.	350–600***	
Potential comprehension	Priming /lɔ̃:/	250–350* 400–500*	–	/lɔ̃:/ 'wage' vs. /lɔ̃ʊ/ potential allophonic form
	Neutral	–	–	

Statistical significance is indicated by \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

**TABLE 5 | Peak latencies and amplitudes of significant ERP components measured over nine electrodes (F3, FZ, F4, C3, CZ, C4, P3, PZ, P4), standard deviations given in brackets.**

Condition	Comparison	Mean peak latency in ms (SD)	Mean peak amplitude in $\mu$ V (SD)	Critical items
Misunderstanding	rɔ̃asɲ/ rɔ̃ʊsɲ	Neg: 132 (6.82) Pos: 517 (91.21)	Neg: –0.88 (0.12) Pos: 2.94 (0.53)	/rɔ̃asɲ/ 'journeys' vs. /rɔ̃ʊsɲ/ 'roses'
	rɔ̃ʊsɲ/ rɔ̃asɲ	Neg: 400 (14.58) Pos: 664 (82.96)	Neg: –1.77 (0.44) Pos: 2.14 (0.40)	/rɔ̃ʊsɲ/ 'roses' vs. /rɔ̃asɲ/ 'journeys'
Incomprehension	Priming context lɔ̃as/ lɔ̃ʊs	Neg: 154 (3.71) Pos: 500 (3.71)	Neg: –1.43 (0.19) Pos: 3.41 (0.39)	/lɔ̃ʊs/ 'sow' vs. /lɔ̃as/ pseudo-word
	Neutral context lɔ̃as/ lɔ̃ʊs	Pos: 485 (18.92)	Pos: 1.68 (0.24)	
Potential comprehension	Priming context lɔ̃ʊ/ lɔ̃:/	Neg: 292 (15.46) Neg: 450 (34.23)	Neg: –0.74 (0.15) Neg: –0.61 (0.11)	/lɔ̃:/ 'wage' vs. /lɔ̃ʊ/ potential allophonic form
	Neutral context lɔ̃ʊ/ lɔ̃:/	–	–	

degree of repetition is responsible for abolishment or decrease of the N400 amplitude (cf. Conrey et al., 2005; Renoult et al., 2012). However, since the deviant  $r\hat{o}\hat{u}s\eta/r\hat{o}\hat{a}s\eta$  elicited a centro-parietal negativity effect in the typical latency range of the N400 component, these explanations can be regarded as rather unlikely.

Another possible reason for the absence of an N400 for this particular deviant form is its frequency. It is likely that the dialectal lexeme for 'journeys' has a lower frequency than the one for 'roses'. In fact, the modulation of the N400 is known to be sensitive to word frequency (cf. Kutas and Federmeier, 2000, 2011; Lau et al., 2008). Unfortunately, no data are available for word frequencies in dialects, which may differ strongly from the standard variety so that this interpretation can neither be supported nor weakened by word frequency data.

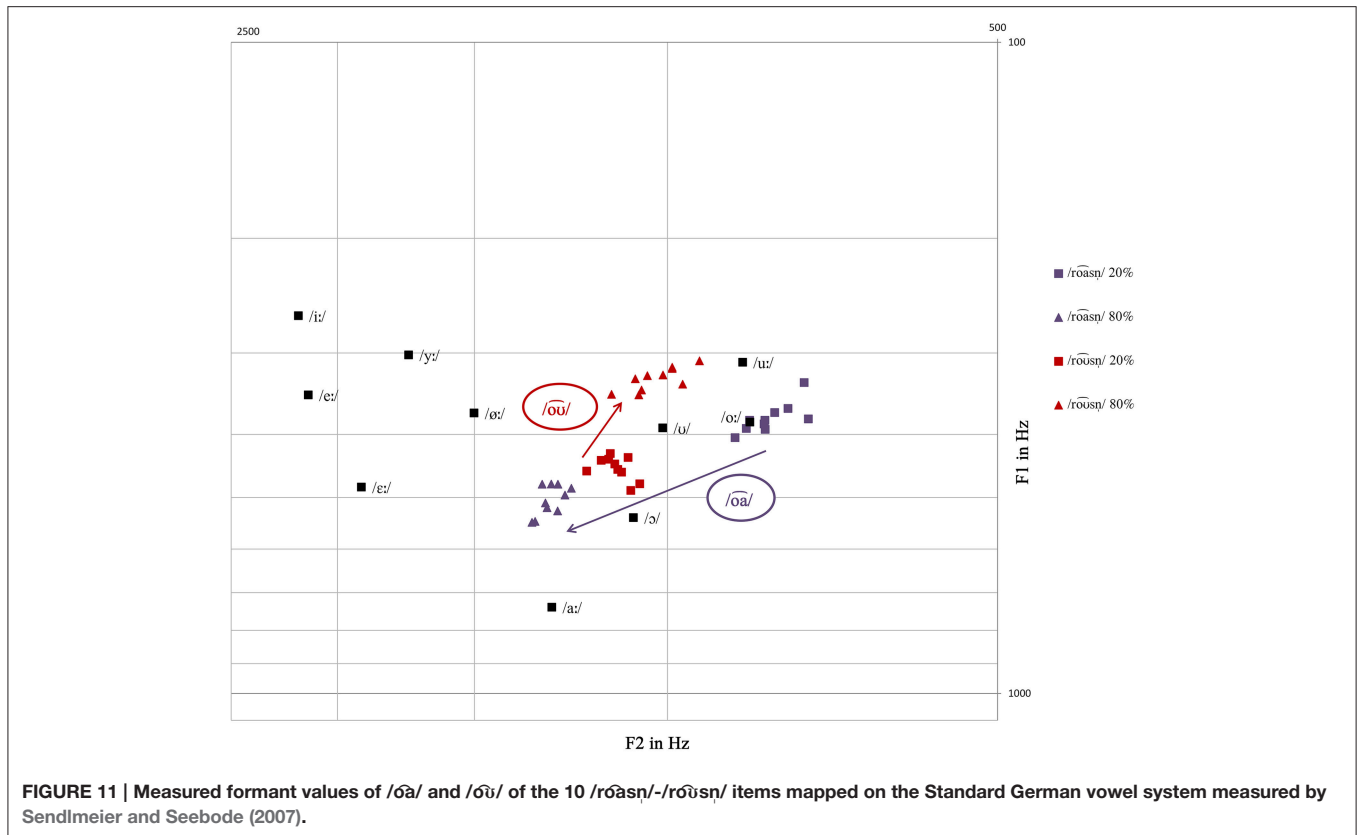
The latency differences might also be due to the diphthong itself, since studies show that early negative components are modulated by specific features of the respective phonemes. In the featurally underspecified lexicon model (FUL) by Lahiri and Reetz (2002, 2010), asymmetries in vowel perception are explained by underlying phonological features. Eulitz and Lahiri (2004) argue that for an underspecified deviant preceded by a specified standard, the MMN reflects an acoustic and feature mismatch, while no mismatch is given for the reverse direction (see also Scharinger et al., 2012). This model cannot be adapted to the  $/\hat{o}\hat{a}/-/\hat{o}\hat{u}/$  contrast one-to-one, but underspecification can be understood as an overall representation of a speech sound, as well. Thus, articulatory more specific speech sounds are reflected by specific tongue positions, which require more detailed motor planning (cf. Scharinger et al., 2012). Applied to the diphthongs in our study,  $/\hat{o}\hat{a}/$  and  $/\hat{o}\hat{u}/$  differ in the articulatory effort of their respective tongue movement, indicated by the F1 and F2 values. The difference between the formant values between 20% and 80% of the diphthong is  $-127.1$  (F1) and  $-471.7$  (F2) for  $/\hat{o}\hat{a}/$ , respectively  $122.1$  (F1) and  $89.9$  (F2) for  $/\hat{o}\hat{u}/$ , indicating that the tongue movement in the diphthong  $/\hat{o}\hat{u}/$  is less pronounced than in  $/\hat{o}\hat{a}/$ . In this regard, it can be assumed that  $/\hat{o}\hat{a}/$  requires more specific motor planning than  $/\hat{o}\hat{u}/$ , so that the asymmetric effects found for  $r\hat{o}\hat{a}s\eta/r\hat{o}\hat{u}s\eta$  and  $r\hat{o}\hat{u}s\eta/r\hat{o}\hat{a}s\eta$  may also reflect the different articulatory effort.

The fact that the lexical-semantic processing of the two lexemes is heavily influenced by the two different diphthongs is further supported by data on asymmetric vowel perception from Polka and Bohn (2003). For infants, the discrimination of a vowel change presented in the direction from a less peripheral to a more peripheral vowel, i.e., closer to the limits or corners of the vowel space, is significantly better than in the reversed direction (see also Polka and Werker, 1994; Polka and Bohn, 1996). The authors conclude that the relatively more peripheral vowel in a contrast serves as a reference point. This is related to the assumption that asymmetries in vowel perception result from the salience of vowels, which is innately more pronounced for vowels tied to extremes of the vowel space. The peripheral hypothesis can be confirmed on the lexical level as well (Sebastián-Gallés et al., 2005; Larsson et al., 2008), indicating

that such perceptual asymmetries reflect higher-level processing at phonetic and lexical levels rather than a low-level general auditory processing (Polka and Bohn, 2011). This assumption fits well with the results obtained in our study examining full lexical words. The formant map of  $/\hat{o}\hat{a}/$  and  $/\hat{o}\hat{u}/$  shows that at the 20% measurement point,  $/\hat{o}\hat{a}/$  is the more peripheral vowel in this contrast (see **Figure 11**). Assuming the peripheral hypothesis, the deviant  $r\hat{o}\hat{a}s\eta/r\hat{o}\hat{u}s\eta$  is suggested to be discriminated more easily than the deviant  $r\hat{o}\hat{u}s\eta/r\hat{o}\hat{a}s\eta$ , since in the former condition the direction of change is from the less peripheral vowel  $/\hat{o}\hat{u}/$  to the more peripheral  $/\hat{o}\hat{a}/$ . This easier discriminability is reflected by the earlier negativity effect for  $r\hat{o}\hat{a}s\eta/r\hat{o}\hat{u}s\eta$  in contrast to  $r\hat{o}\hat{u}s\eta/r\hat{o}\hat{a}s\eta$ , indicating a facilitated discrimination of the vowel  $/\hat{o}\hat{a}/$  due to its more peripheral position in the vowel space and thus its salience. The higher saliency of the  $/\hat{o}\hat{a}/$ -diphthong might also be due to the fact that it is a falling diphthong which is a rather uncommon phenomenon in German varieties. Since asymmetries in vowel perception can be observed even in native listeners when lexical processing is involved in the task (cf. Polka and Bohn, 2011), the peripheral position of the  $/\hat{o}\hat{a}/$ -diphthong might be an influencing factor concerning the relative earliness of the negativity effect for  $r\hat{o}\hat{a}s\eta/r\hat{o}\hat{u}s\eta$  in contrast to  $r\hat{o}\hat{u}s\eta/r\hat{o}\hat{a}s\eta$ . Although the negativities found display different latencies, it is important to keep in mind that both lexemes elicited an error detection mechanism induced by the semantic priming, leading to a problematic lexical retrieval for both forms.

## Cross-Dialectal Comprehension and Phoneme Change

Data from dialectology studies show that a change of the  $/\hat{o}\hat{a}/$ -phoneme can be detected in several lexemes in BA which is thought to be due to communicative interactions between speakers of BA and CB who synchronize their different linguistic competencies. Synchronization in our view refers to the calibration of competence differences in the performance act, which results in a stabilization or modification of the active and passive competencies involved (cf. Schmidt, 2010; Schmidt and Herrgen, 2011). Desiring to be understood, listeners actively and interactively synchronize their individual competencies. In every speech production, speakers try to match their individual competences to their interlocutor's communicative expectations and abilities in order to be understood. Negative feedback from the interlocutor signaling a lack of comprehension or only partial comprehension leads to modification of the applied language production strategy effecting a reconstruction of the speaker's individual competence (so-called microsynchronization; cf. Schmidt, 2010; Schmidt and Herrgen, 2011). In long-term contact situations (e.g., school classes, youth peer groups, regional groups), subjects interact over an extended period of time. This leads to a series of parallel synchronizations implying repetitive positive or negative feedback. The result is a congruent part of linguistic knowledge, which enables the communication partners to communicate successfully. Such a series of parallel acts of synchronization performed by individuals in personal contact situations, which lead to the establishment of common



context-dependent linguistic knowledge, are designated as so-called mesosynchronizations (cf. Schmidt, 2010; Schmidt and Herrgen, 2011)<sup>2</sup>.

The lexically gradual change of the /ōā/-diphthong in the Bavarian-Alemannic transition zone may be due to such mesosynchronizations. In this regionally defined speaker group, a constant stream of parallel synchronizations led to parallel optimization strategies resulting in a convergence to the dominant Central Bavarian variant. Thus, the synchronizations led to a gradual modification of the group's linguistic competence, which is reflected by a word-by-word phonological redistribution of the /ōā/-diphthong. This gradual change can also be explained by the interactional approach. During the adaptation process, speakers modify their pronunciation of particular words rather than the complete phonological system at once. Some words are thus affected earlier by the change than others. This is caused by the speakers' motivation to pronounce individual lexemes just as speakers from the target variety (cf. Trudgill, 1986).

The current study focuses on the question whether this slow but continuous change in the usage of the dialectal /ōā/-phoneme leads to increased neural costs during sentence processing. In fact, the results display clear processing differences

between the deviants  $rōasn̩/rōʊsn̩$  and  $lōas̩/lōʊs̩$  in contrast to  $lōū/lōū$ . The aforementioned forms elicit early negativities (N2b) associated with an early error detection mechanism, which is absent for  $lōū/lōū$ . Furthermore,  $rōasn̩/rōʊsn̩$  and  $lōas̩/lōʊs̩$  elicit LPCs reflecting an evaluation process of the previously detected mismatch, which is also absent for  $lōū/lōū$ .

The effects found for /rōasn̩/ and /lōas̩/ can be interpreted within the scope of the synchronization theory, as the detection and evaluation of erroneous usage of the native form /rōasn̩/ and the unknown form /lōas̩/ lead to negative feedback. Our data suggest that both misunderstanding as well as incomprehension induce similar neural costs during cross-dialectal comprehension. In both cases, the non-native variant leads to increased costs at an early stage of processing reflecting the detection of an error in the sentential context. Thus, these cross-dialectal competence differences do lead to cross-dialectal comprehension difficulties, which are reflected in both neural responses as well as in behavioral data. These difficulties can be interpreted as a trigger for phoneme change, since the error detection could be proven for  $rōasn̩/rōʊsn̩$ —a lexeme involved in language change in BA, amongst others.

However, the conditions misunderstanding and incomprehension differ partially with regard to the early negativity's amplitude. Surprisingly, the amplitude of the N2b is less pronounced for  $rōasn̩/rōʊsn̩$ , although the behavioral data show a significant effect between standard and deviant.

<sup>2</sup>The accommodation theory developed by Howard Giles focuses on questions about the convergence in face-to-face interaction, as well. It is not possible to discuss theoretical differences between the theories of accommodation and synchronization within the scope of this paper.

Moreover, the behavioral data show that, although there is a significant difference in the evaluation of deviant and standard, the unacceptability rates for the reversed misunderstanding contrast are higher. It thus seems as if the deviant form /rōasŋ/ was not as easily rejectable as its counterpart /rōūsŋ/. This might be attributed to the close contact situation of the participants to the merged community, as the exposure to regional variability influences the perception and processing of vowel contrasts in their own accent as well. Hence, although the respective vowel contrast is still preserved in CB, the less clear rejection of rōasŋ/rōūsŋ indicates a level of uncertainty during lexical and evaluation processing caused by the exposure to regional variability. These results are in line with Brunellière et al. (2009) who show that listeners who preserve a particular phonological contrast, but are often exposed to merged variants, discriminate this contrast less easily than regional stable contrasts. Conrey et al. (2005) could further show that unmerged dialect speakers are still worse at distinguishing between merged vowels than other vowels unaffected by a merger.

This influence of the close dialect contact is also visible in the results for the incomprehension condition. As a deviant, the well-formed pseudo-word /lōas/, which is non-existent in the listeners' lexicon, was presented. However, /lōas/ is indeed an existing word in the adjacent dialect of BA. This might be a further reason for the fact that we did not find an N400 effect, which is normally elicited by unknown novel word forms, as they require additional processing resources in order to classify them as non-existing words (cf. Bentin et al., 1999; Domahs et al., 2009). Further evidence comes from the behavioral data: Although the pseudo-word is less acceptable in comparison to the real word form /lōūs/, no significant effect between the match and mismatch condition could be found for primed sentences. The wide dispersion of the standard deviation in this condition suggests uncertainties when evaluating the sentences, indicating a more extensive evaluation of its lexical status and thus semantic classification. This process could be impeded by the circumstance that the pseudo-word is the correct phonemic form for the same semantic concept in the neighboring dialect area. Thus, especially in the priming context, the classification of the deviating form is especially problematic for the participants. This fact is also reflected in the high dispersion of the standard deviation (cf. Section ERP Data in Condition 1: Misunderstanding).

Unlike the conditions 1 and 2, neither ERP nor behavioral data show comparable results for the potential comprehension condition. Although the deviant /lōūs/ is also a non-existing pseudo-word, it is phonetically very similar to the native lexeme. The absence of an LPC and the similar acceptability rates for both forms indicate that no reanalysis-mechanism was required and that the evaluation of both forms was equally easy to categorize. This suggests that both forms are associated with the meaning 'wage', i.e., the phonetically deviant variant /lōūs/ is perceived as an allophonic variant of /lō:/, even though it is not an appropriate form in the listener's competence. It does thus not elicit a mismatch between expectation and perceived input, as both forms were equally categorized as existing word forms. Henrich et al. (2015) support the unproblematic evaluation of both critical

and control conditions illustrated by a similar absence of a P300 in the evaluation process of rhythmically well-formed and subtle deviant structures. The fact that allophonic variability does not lead to enhanced processing costs, while phonological contrasts (rōasŋ/rōūsŋ and lōas/lōūs) elicit clear ERP effects is in line with results from Kazanina et al. (2006), who found an MMNm for the phonemic [t]-[d] contrast for Russian listeners, which is absent for Koreans for whom the consonants are purely allophonic.

Thus, the usage of the /ōūs/-diphthong in lexemes going back to MHG *ō* does not lead to comprehension difficulties and the change from /ōa/ to /ōūs/ seems to be a worthwhile strategy to reduce both misunderstandings/incomprehension and neural costs in processing during cross-dialectal communication.

## CONCLUSION

The aim of the present ERP study was to examine the speech processing of Central Bavarian speakers, who are exposed to certain lexemes common to the Bavarian-Alemannic transition zone. The speaker groups differ with regard to their linguistic competence insofar as a phonemic merger of MHG *ō* and MHG *ei* to /ōa/ took place in the Bavarian-Alemannic transition zone. By contrast, in the Central Bavarian dialect two phonemes are distinguished (/ōa/-/ōūs/). It is assumed that these competence differences lead to communication problems between the dialect areas.

To approach this particular cross-dialectal communication setting, a more natural design in which listeners are exposed to complex sentences was created in combining an oddball design with whole sentences and a semantic rating task. This advanced and innovative design is quite promising, as single minimal pairs can be investigated during sentence processing.

Results show robust effects despite the low variance of critical items and recurrent sentence contexts. Indeed, the Bavarian-Alemannic lexemes /lōas/ 'sow' and /rōasŋ/ 'roses' evoke enhanced neural costs during sentence comprehension, reflected by biphasic ERP patterns consisting of an N200 and an LPC. These results demonstrate a conscious mismatch detection and evaluation process with regard to the non-native lexemes. Thus, it seems likely that the usage of these lexemes cause problems during cross-dialectal communication, leading to negative feedback and subsequently to competence modifications. In contrast, the results for the lexeme /lōūs/ show only reduced negativity effects and no LPC suggesting that in this case allophonic deviations are perceived but do not prevent successful communication.

To conclude, the results reveal first ERP evidence for cross-dialectal misunderstanding when speakers from the merged area use the /ōa/-diphthong stemming from MHG *ō*. No similar effects could be detected for the phonetically related /ōūs/-phoneme. The empirically observable change of /ōa/ to either /ōūs/ or /o:/ in the Bavarian-Alemannic transition zone can thus be interpreted as a strategy to avoid costly communication difficulties in close dialect contact settings. Insofar, dialectal competence differences resulting



in enhanced neural processing costs can indeed trigger dialect change in order to facilitate successful cross-dialectal communication.

## AUTHOR CONTRIBUTIONS

ML conception of experimental stimuli and experimental paradigm; phonetic analysis of auditory stimuli; data acquisition; interpretation of results; preparation of the manuscript's first draft, revision of the manuscript. KH conception of experimental stimuli and experimental paradigm; data acquisition; data analysis; interpretation of results; preparation of the manuscript's first draft, revision of the manuscript. MR conception of experimental stimuli and experimental paradigm; phonetic analysis of auditory stimuli; data acquisition; interpretation of results; proofreading. HS conception of experimental stimuli and experimental paradigm; phonetic analysis of auditory stimuli; data acquisition; interpretation of results. AW conception of experimental stimuli and experimental paradigm; phonetic analysis of auditory stimuli; data acquisition; interpretation of results; proofreading. JH conception of experimental paradigm; interpretation of results; proofreading. JS conception of experimental paradigm; interpretation of results; proofreading.

## REFERENCES

- Bendixen, A., Scharinger, M., Strauß, A., and Obleser, J. (2014). Prediction in the service of comprehension: modulated early brain responses to omitted speech segments. *Cortex* 53, 9–26. doi: 10.1016/j.cortex.2014.01.001
- Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., and Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *J. Cogn. Neurosci.* 11, 235–260. doi: 10.1162/089892999563373
- Bohn, K., Knaus, J., Wiese, R., and Domahs, U. (2013). The influence of rhythmic (ir)regularities on speech processing: evidence from an ERP study on German phrases. *Neuropsychologia* 51, 760–771. doi: 10.1016/j.neuropsychologia.2013.01.006
- Boulenger, V., Hoen, M., Jacquier, C., and Meunier, F. (2011). Interplay between acoustic/phonetic and semantic processes during spoken sentence comprehension: an ERP study. *Brain Lang.* 116, 51–63. doi: 10.1016/j.bandl.2010.09.011
- Brown, C., and Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *J. Cogn. Neurosci.* 5, 34–44. doi: 10.1162/jocn.1993.5.1.34
- Brunellière, A., Dufour, S., and Nguyen, N. (2011). Regional differences in the listener's phonemic inventory affect semantic processing: a mismatch negativity (MMN) study. *Brain Lang.* 117, 45–51. doi: 10.1016/j.bandl.2010.12.004
- Brunellière, A., Dufour, S., Nguyen, N., and Frauenfelder, U. H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition* 111, 390–396. doi: 10.1016/j.cognition.2009.02.013
- Buchwald, J. S., Guthrie, D., Schwafel, J., Erwin, R. J., and Van Lancker, D. (1994). Influence of language structure on brain-behavioral development. *Brain Lang.* 46, 607–619. doi: 10.1006/brln.1994.1033
- Connolly, J. F., and Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *J. Cogn. Neurosci.* 6, 256–266. doi: 10.1162/jocn.1994.6.3.256
- Conrey, B., Potts, G. F., and Niedzielski, N. A. (2005). Effects of dialect on merger perception. ERP and behavioural correlates. *Brain Lang.* 95, 435–449. doi: 10.1016/j.bandl.2005.06.008

## FUNDING

The research presented here is supported by the Hessen State Ministry of Higher Education, Research and the Arts, Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz (LOEWE), Research Focus: Exploring fundamental linguistic categories grant, research project Phonological word-Constituents of the Phonological Word under the leadership of JH, JS, and AW.

## ACKNOWLEDGMENTS

The authors thank Lydia Riedl and Corinna Schwarz for their help in preparing the stimuli, Rupert Schamberger for recording the stimuli, Alexander Dröge for his support in programming, Jeffrey Pheiff for language proofreading as well as the participants in Isen.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00739>

- Coulson, S., King, J. W., and Kutas, M. (1998). ERPs and domain specificity: beating a straw horse. *Lang. Cogn. Process.* 13, 653–672. doi: 10.1080/016909698386410
- Domahs, U., Kehrein, W., Knaus, J., Wiese, R., and Schlesewsky, M. (2009). Event-related potentials reflecting the processing of phonological constraint violations. *Lang. Speech* 52, 415–435. doi: 10.1177/0023830909336581
- Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2008). The processing of German words stress: evidence for the prosodic hierarchy. *Phonology* 25, 1–36. doi: 10.1017/S0952675708001383
- Eichinger, L. M. (ed.). (2011). *Sprachatlas von Oberbayern. Band 1: Phonologie I. Langvokale und Diphthonge*. Heidelberg: Universitätsverlag Winter (Bayerischer Sprachatlas. Regionalteil 6).
- Eulitz, C., and Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *J. Cogn. Neurosci.* 16, 577–583. doi: 10.1162/089892904323057308
- Folstein, J. R., and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170. doi: 10.1111/j.1469-8986.2007.00602.x
- Henrich, K., Alter, K., Wiese, R., and Domahs, U. (2014). The relevance of rhythmical alternation in language processing: an ERP study on English compounds. *Brain Lang.* 136, 19–30. doi: 10.1016/j.bandl.2014.07.003
- Henrich, K., Wiese, R., and Domahs, U. (2015). How information structure influences the processing of rhythmic irregularities: ERP evidence from German phrases. *Neuropsychologia* 75, 431–440. doi: 10.1016/j.neuropsychologia.2015.06.028
- Herold, R. (1997). Solving the actuation problem: merger and immigration in eastern Pennsylvania. *Lang. Var. Change* 9, 165–189. doi: 10.1017/s0954394500001861
- Huynh, H., and Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *J. Educ. Stat.* 1, 69–82. doi: 10.2307/1164736
- Kazanina, N., Phillips, C., and Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11381–11386. doi: 10.1073/pnas.0604821103
- Knaus, J., Wiese, R., and Janßen, U. (2007). “The processing of word stress: EEG studies on task-related components,” in *Proceedings of*

- the 16th International Congress of Phonetic Sciences (Saarbrücken), 709–712.
- Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the Event-Related Brain Potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657
- Kutas, M., and Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161–163. doi: 10.1038/307161a0
- Kutas, M., and Van Petten, C. K. (1994). “Psycholinguistic electrified. Event-related brain potential investigations,” in *Handbook of Psycholinguistics*, ed M. A. Gernsbacher (San Diego, CA: Academic Press), 83–143.
- König, W. (ed.). (1998). *Sprachatlas von Bayerisch-Schwaben. Band 5: Lautgeographie III. Qualität und Quantität der mittelhochdeutschen Langvokale und Diphthonge*. Heidelberg: Universitätsverlag Winter (Bayerischer Sprachatlas. Regionalteil 1).
- Labov, W. (2010). *Principles of Linguistic Change. Volume 3. Cognitive and Cultural Factors*. Oxford: Wiley-Blackwell (Language in Society 39).
- Lahiri, A., and Reetz, H. (2002). “Underspecified recognition,” in *Laboratory Phonology 7*, eds C. Gussenhoven and N. Warner (Berlin: Mouton de Gruyter), 637–677.
- Lahiri, A., and Reetz, H. (2010). Distinctive features: phonological underspecification in representation and processing. *J. Phon.* 38, 44–59. doi: 10.1016/j.wocn.2010.01.002
- Lameli, A. (2013). *Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland*. Berlin; Boston, MA: De Gruyter (Linguistik – Impulse & Tendenzen 54).
- Larsson, J. P., Constán, F. V., Sebastián-Gallés, N., and Deco, G. (2008). Lexical plasticity in early bilinguals does not alter phoneme categories: I. Neurodynamical modeling. *J. Cogn. Neurosci.* 20, 76–94. doi: 10.1162/jocn.2008.20004
- Lau, E., Almeida, D., Hines, P. C., and Poeppel, D. (2009). A lexical basis for N400 context effects: evidence from MEG. *Brain Lang.* 111, 161–172. doi: 10.1016/j.bandl.2009.08.007
- Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933. doi: 10.1038/nrn2532
- Menning, H., Zwitserlood, P., Schöning, S., Hihn, H., Bölte, J., Dobel, C., et al. (2005). Pre-attentive detection of syntactic and semantic errors. *Cogn. Neurosci.* 16, 77–80. doi: 10.1097/00001756-200501190-00018
- Miglietta, S., Grimaldi, M., and Calabrese, A. (2013). Conditioned allophony in speech perception: an ERP study. *Brain Lang.* 126, 285–290. doi: 10.1016/j.bandl.2013.06.001
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434. doi: 10.1038/385432a0
- 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. doi: 10.1016/j.clinph.2007.04.026
- Partanen, E., Vainio, M., Kujala, T., and Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. *Psychophysiology* 48, 1372–1380. doi: 10.1111/j.1469-8986.2011.01214.x
- Patel, S. H., and Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the event-related potential. *Int. J. Med. Sci.* 2, 147–154. doi: 10.7150/ijms.2.147
- Pritchard, W. S., Shappell, S. A., and Brandt, M. E. (1991). “Psychophysiology of N200/N400: a review and classification scheme,” in *Advances in Psychophysiology: A Research Annual*, eds J. R. Jennings, P. K. Ackles, and G. H. Coles, Vol. 4 (London: Jessica Kingsley Publishers), 43–106.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: an MEG mismatch study. *J. Cogn. Neurosci.* 12, 1038–1105. doi: 10.1162/08989290051137567
- Picton, T. W. (1992). The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456–479. doi: 10.1097/00004691-199210000-00002
- Polka, L., and Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *J. Acoust. Soc. Am.* 100, 577–592. doi: 10.1121/1.415884
- Polka, L., and Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Commun.* 41, 221–231. doi: 10.1016/S0167-6393(02)00105-X
- Polka, L., and Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: an emerging view of early phonetic development. *J. Phon.* 39, 467–478. doi: 10.1016/j.wocn.2010.08.007
- Polka, L., and Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 421–435. doi: 10.1037/0096-1523.20.2.421
- Rein, K. (ed.). (2005). *Sprachregion München. Ergänzungsband zum Sprachatlas von Oberbayern*. Heidelberg: Universitätsverlag Winter (Bayerischer Sprachatlas. Regionalteil 6).
- Renoult, L., Wang, X., Mortimer, J., and Debrulle, J. B. (2012). Explicit semantic tasks are necessary to study semantic priming effect with high rates of repetition. *Clin. Neurophysiol.* 123, 741–754. doi: 10.1016/j.clinph.2011.08.025
- Ritter, W., Ford, J. M., Gaillard, A. W. K., Harter, M. R., Kutas, M., Näätänen, R., et al. (1984). Cognition and event-related potentials. I. The relation of negative potentials and cognitive processes. *Ann. N.Y. Acad. Sci.* 425, 24–38. doi: 10.1111/j.1749-6632.1984.tb23521.x
- Roehm, D., Bornkessel-Schlesewsky, I., Rösler, F., and Schlewsky, M. (2007). To predict or not to predict: influences of task and strategy on the processing of semantic relations. *J. Cogn. Neurosci.* 19, 1259–1274. doi: 10.1162/jocn.2007.19.8.1259
- Scharinger, M., Monahan, P. J., and Idsardi, W. (2012). Asymmetries in the processing of vowel height. *J. Speech Lang. Hear. Res.* 55, 903–918. doi: 10.1044/1092-4388(2011/11-0065)
- Scharinger, M., Monahan, P. J., and Idsardi, W. J. (2011). You had me at “Hello”: rapid extraction of dialect information from spoken words. *Neuroimage* 56, 2329–2338. doi: 10.1016/j.neuroimage.2011.04.007
- Schmidt, J. E. (2010). “Language and space: the linguistic dynamics approach,” in *Language and Space. An International Handbook of Linguistic Variation, Vol. 1: Theories and Methods (Handbook of Linguistics and Communication Science 30.1)*, eds P. Auer and J. E. Schmidt (Berlin: de Gruyter), 201–225.
- Schmidt, J. E. (2011). Formation of and change in regiolects and (regional) dialects in German. *Taal Tongval* 63, 143–173. doi: 10.5117/TET2011.1.SCHM
- Schmidt, J. E., and Herrgen, J. (2011). *Sprachdynamik. Eine Einführung in die moderne Regionalsprachenforschung*. Berlin: Erich Schmidt Verlag (Grundlagen der Germanistik 49).
- Schmidt, J. E., Herrgen, J., and Kehrein, R. (eds.). (2008). *Regionalsprache.de (REDE). Forschungsplattform zu den Modernen Regionalsprachen des Deutschen*. Marburg: Forschungszentrum Deutscher Sprachatlas.
- Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. *J. Mem. Lang.* 52, 240–255. doi: 10.1016/j.jml.2004.11.001
- Sendlmeier, W. F., and Seebode, J. (2007). *Formantkarten des deutschen Vokalsystems*. Available online at: [http://www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten\\_des\\_deutschen\\_Vokalsystems\\_01.pdf](http://www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf)
- Trudgill, P. (1986). *Dialects in Contact*. Oxford: Basil Blackwell (Language and Society 10).
- Wang, W. S.-Y., and Lien, C. (1993). “Bidirectional diffusion in sound change,” in *Historical Linguistics. Problems and Perspectives*, ed C. Jones (London; New York, NY: Longman), 345–400.
- Wiesinger, P. (1983). “Die Einteilung der deutschen Dialekte,” in *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung. 2. Halbband*, ed W. Besch (Berlin; New York, NY: de Gruyter) (Handbücher zur Sprach- und Kommunikationswissenschaft 1.2), S. 807–900.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Lanwermeijer, Henrich, Rocholl, Schnell, Werth, Herrgen and Schmidt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.