The adoption of sound change : synchronic and diachronic processing of regional variation in Dutch

Voeten, C.C.

Citation

Voeten, C. C. (2020, October 13). *The adoption of sound change : synchronic and diachronic processing of regional variation in Dutch*. *LOT dissertation series*. LOT, Amsterdam. Retrieved from https://hdl.handle.net/1887/137723

Author: Voeten, C.C.
Title: The adoption of sound change : synchronic and diachronic processing of regional variation in Dutch
Issue Date: 2020-10-13

CHAPTER 5

Noticing the change: misrepresentation, not misperception, of allophonic variants in sound change

This chapter has been submitted.

Abstract

Linguists have posited that phonological change arises through misperception. This is evaluated using a longitudinal EEG experiment. Three on-going changes in Dutch are studied: the diphthongization of /eː,øː,oː/ to [ei,øy,ou], the blocking of diphthongs before coda /l/, and the gliding of coda /r/ to [ɹ]. These changes have essentially completed in the Netherlands, but have not taken place in Flanders, the Dutch-speaking part of Belgium. A passive-oddball task was performed with Flemish-Dutch speakers (plus Netherlandic controls) who migrated to the Netherlands to start their university studies. Previous work has shown that such sociolinguistic migrants readily adopt the local phonology, and hence on-going changes. Participants did the experiment four months after arrival and again four months later. Results show that, initially, the Flemish participants have stronger mismatch negativities to a deviant [ei] within a stream of standard [eː]s, but four months later this difference has disappeared: they have learned the vowel shift. The gliding of /r/ continues to elude them: they show an MMN, but with a less frontal topographical distribution—they find the glide less salient. This is interpreted as showing successful phonological learning, but not yet sociolinguistic learning. The results argue against misperception, favoring misrepresentation instead.

5.1 Introduction

It was proposed by Ohala (1981) that sound change originates when a listener under- or overapplies rules compensating for coarticulation during the transmission of the speech signal. In this scenario, the listener has processed the incoming speech signal incorrectly, either by phonologically encoding accidental details ("intrinsic variation", Wang & Fillmore 1961) or by failing to encode information that was linguistically relevant ("extrinsic variation", ibid.). Thus, while the listener hears the speech signal without problem, they make an error in *processing* it. This type of categorization error is commonly referred to as "misperception", and is said to lead to sound change if the listener subsequently adjusts their phonological representation to match (Hyman 1976). Theoretical analyses of historical sound changes have found (indirect) support for this mechanism of sound change, claiming it the most likely scenario in various case studies (e.g. the infamous [k]>[ʧ]>[s] change in Proto-Romance, giving Latin *caelum*, Italian *cielo*, and French *ciel*; Guion 1998).

An alternative account by Hamann (2009) suggests that these types of sound changes do not in fact take place in the listener's processing of the speech signal, but rather in the grammar they use to perform this processing with. In this view, the speaker and the listener have acquired slightly different mappings of phonetic cues to phonological categories, and as a result the listener understands a different phonological category than the speaker intended, because they attach different cue weights to the same auditory information. A similar proposal was made by Beddor (2009), which differs from Hamann (2009) only in not requiring that the grammatical innovation take place in childhood. In these accounts, the critical difference between speaker and listener is not located in the listener's *perception* of the speech signal, but rather in their grammatical *representation* of the relevant phonological and phonetic features.

The present chapter aims to contribute to the debate surrounding these two alternatives—broadly speaking, the misperception account by Ohala (1981) and others versus the misrepresentation account by Hamann (2009) and Beddor (2009)—using neurolinguistic evidence. The chapter draws strongly on Grosvald & Corina (2012), who used a mismatch-negativity paradigm to show that listeners are able to perceive and *encode* sound change—in their case, long-distance vowel-to-vowel coarticulation of up to three syllables away—automatically. Grosvald & Corina (2012) demonstrated this encoding using a mismatch-negativity ("MMN") experiment, which is also the paradigm used in the present paper. MMN studies have a long history of use in phonology (see e.g. Cornell, Lahiri, & Eulitz 2011, Grosvald & Corina 2012, Hestvik & Durvasula 2016, Lahiri & Reetz 2010, Lanwermeyer et al. 2016, Mitterer & Blomert 2003, Scharinger & Lahiri 2010; for accessible introductions to the

MMN in general, see Näätänen 1990, Näätänen & Alho 1997, Sussman et al. 2014). Generally speaking, the MMN is an event-related potential evoked in EEG experiments using the passive-oddball task. In these experiments, participants listen to a stream of "standard" sounds, which is sometimes interrupted by a "deviant" sound. If the standard has a neurally-encoded feature that the deviant does not, this difference triggers an MMN (but the reverse difference does not; Cornell, Lahiri, & Eulitz 2011, Lahiri & Reetz 2010). To disentangle phonological encoding from phonetic encoding (i.e. to ensure that the MMN probes a phonological feature rather than the obvious acoustic differences between standards and deviants), MMN studies typically use multiple, different, tokens of the same surface allophone, called the "varying standards" paradigm (Hestvik & Durvasula 2016). In this paradigm, the presence of an MMN reflects the phonological encoding of the standard stimulus, with the size of the MMN in microvolts proportional to the phonological distance between the standards and the deviant (Näätänen et al. 2007).

MMNs have been used to study sound change, ranging from phonemic mergers to allophonic shifts. Lanwermeyer et al. (2016) show that a phonemic merger resulting in lexical confusion elicits an MMN that is much earlier (100–200 ms) than the allophonic MMN found by Grosvald & Corina's (2012) (275–325 ms). In addition, they find a P600, which reflects the semantic reintegration and reevaluation of an initially misanalyzed phoneme category (see also Kung, Chwilla, & Schriefers 2014 and Chapter 6). However, in a contrasting condition where only allophonic differences were manipulated (similar to Grosvald & Corina 2012), the MMN was reduced and temporally shifted and the P600 disappeared. The absence of the P600 in this condition is not surprising, as an allophonic difference cannot result in lexical overlap, and hence no reanalyis was necessary. Similarly, the reduction of the MMN is logically explained as allophonic switches being less salient than phonemic switches, as is argued by Kazanina, Phillips, & Idsardi (2006), who failed to find an allophonic MMN. However, these authors presented their allophonic condition (Korean [t,d]; these are allophones of the same phoneme, with /t/ becoming [d] intervocalically) without providing the requisite phonological context (both variants were presented word-initially, which does not trigger the allophony). This may explain why they did not find an allophonic MMN, whereas other studies (Jacobsen 2015, Lanwermeyer et al. 2016, Steinberg, Truckenbrodt, & Jacobsen 2010a, 2010b, 2011) did. The temporal shifting of the allophonic MMN observed by Lanwermeyer et al. (2016) brings it exactly in line with the window where Grosvald & Corina (2012) found their strongest effect (Lanwermeyer et al.: 250–350 ms, Grosvald & Corina: 275–325 ms), which inspires confidence that the allophonic MMN is indeed later than the phonemic MMN. Note that the mentioned allophonic MMNs are really responses to the phonological allophone, and do not simply reflect acoustic differences in the phonetic signal: both Grosvald & Corina (2012) and Lanwermeyer et al. (2016) used the varying-standards paradigm.

More specific than research into allophonic variation is research into allophonic *violations*. In this strand of research, one does not investigate allophonic differences in realization *per se*, as in Kazanina, Phillips, & Idsardi (2006) and Lanwermeyer et al. (2016), but rather the grammatical knowledge that is a prerequisite for processing such differences. An example, and the phenomenon studied in this chapter, is the rise of a new allophonic rule due to phonotactically-conditioned regular sound change. Previous studies (Jacobsen 2015, Steinberg, Truckenbrodt, & Jacobsen 2010a, 2010b, 2011) have shown that phonotactic violations result in MMNs. However, these papers are about violations of well-established allophonic rules in the standard variety of a language. It might be the case that novel allophonic rules involved in on-going sound change are less salient (and hence encoded less strongly) than well-established allophonic rules of the type studied in Jacobsen (2015) and Steinberg, Truckenbrodt, & Jacobsen (2010a, 2010b, 2011). Hence, the present study integrates and extends the aforementioned findings by studying a currently-on-going sound change that involves the genesis of a new allophone distinction. The study uses a combined cross-sectional and longitudinal design, aimed at providing a window into the processing of sound change as it unfolds in real time.

The language used for the investigation is Dutch, in which the tense mid vowels /eː,øː,oː/ have changed into upgliding diphthongs [ei,øy,ou]. This regular sound change is blocked before coda /l/, leading to novel allophone pairs: monophthongs before coda /l/, diphthongs elsewhere. Independently of these changes, Dutch has also undergone an allophone split in the rhotic, changing /r/[1] to [ɹ] in coda position. These three distributional changes are regionally stratified. They have all but completed in the Netherlands, but the Dutch spoken in Flanders (the northern part of Belgium) has not undergone them, leading to salient sociolinguistic differences between Netherlandic Dutch and Flemish Dutch (Sebregts 2015, Van de Velde 1996). This is particularly true for the rhotic, which is perhaps the most-well-known sociolinguistic variable in the Netherlands and Flanders (Sebregts 2015). Table 5.1 provides a complete overview of the relevant allophonic rules in both varieties. The present study

---

[1] The phonetic implementation of /r/ is highly variable between different regions of Dutch (see Sebregts 2015 for a thorough overview), including alveolar as well as uvular trills and fricatives. However, the phonological allophone split between onset and coda variants is restricted to Netherlandic Dutch, and is also only implemented by means of the [ɹ] realization; in addition, this realization can never occur in onset position in either Netherlandic Dutch or Flemish Dutch. The phonetic implementation of the onset allophone in this experiment was the alveolar trill, as this is the standard variant in Flemish Dutch and is one of the major variants in Netherlandic Dutch, and shares its place of articulation with the [ɹ] allophone, which helps keep the difference between standards and deviants to the minimum required.

Table 5.1: The relevant allophonic rules involved in the on-going sound changes and their regional differences.

Underlying form	**Netherlandic realization**	**Flemish realization**
/eː/ followed by coda /l/	[eː]	[eː]
/eː/ elsewhere	[ei]	[eː]
/ər/	[əɹ]	[ər]

uses an MMN experiment to investigate the processing of these allophonic rules in two populations: a control group of Netherlandic students, and an experimental group of Flemish students in their first year of study at a university in the Netherlands. It is expected, and has been shown experimentally in similar research (Evans & Iverson 2007), that this experimental group will adapt to the Netherlandic realizations as part of the normal process of accent accommodation, paralleling the adoption of these historical sound changes. To investigate such adaptation over time, the cross-sectional comparison is performed two times, with four months in between.

The task is an oddball task, the same task used by Grosvald & Corina (2012) and Lanwermeyer et al. (2016). The two accounts of sound change under discussion make different predictions on the degree of encoding of the allophones in question, and hence on the degree to which the oddball task should yield MMN ERPs. Under the misperception account, the prediction would be that the Flemish participants will not grammatically encode the difference between the allophones for each of [eːɫ∼eiɫ], [eː∼ei], and [ər∼əɹ], as these differences are not relevant in their own grammars and hence fall under the purview of intrinsic variation. This would preclude MMN effects from showing up. In turn, the misrepresentation account predicts that the Flemish participants do *encode* the allophonic distinctions, but subsequently *evaluate* them in a different way (e.g. through the P600, as observed by Lanwermeyer et al. 2016 and Chapter 6; a separate experimental paradigm would be required to evaluate this possible mechanism). In this case, the participants will perceive a mismatch between the deviant and the standards on an extrinsic property, which is visible as the MMN. A second prediction for the present experiment, which holds for both accounts of historical sound change equally, is that in the four months between the two sessions of the experiment, the Flemish participants have begun to adopt the Netherlandic system, such that the differences between the groups will have become smaller. There is evidence for this type of post-adolescent adjustment of the perception of vowel categories from both sociolinguistics (e.g. Bowie 2000, Evans & Iverson 2007, Nycz 2011, Ziliak 2012, Chapter 4) and the related field of second-language acquisition (Flege 1987, Flege & Wayland

2019, *inter alia*); it seems reasonable to hypothesize that those findings reflect a general ability that is also relevant here.

The pool of participants suitable for this experiment is small, because the experiment calls for a specific and special population: participants in the experimental group must be Flemish, must have migrated to the Netherlands post-adolescence (and not before), must be measured relatively shortly after arrival (no later than a couple of months; compare Chapter 3 and Chapter 4), and must be willing to commit to a two-part experiment with some time in between. Because the GDPR was not yet in effect when recruitment for this experiment was initiated, it was possible to obtain a list of recently-arrived Flemish individuals who had just begun their studies at two universities in the Netherlands: Leiden University and the University of Amsterdam. This made it possible to recruit eight participants, resulting in fourteen obtained repeated-measures datasets. Both are typical sample sizes for similar sociolinguistic studies on the adoption of phonetic variation along the lifespan (e.g. Alshangiti & Evans 2011, Bauer 1985, Carter 2007, Cedergren 1987, Chambers 1992, De Decker 2006, Evans & Iverson 2007, Harrington 2006, Harrington, Palethorpe, & Watson 2000, Hinton 2015, Nahkola & Saanilahti 2004, Nycz 2011, Nycz 2013, van Oostendorp 2008, Prince 1987, Sankoff 2004, Sankoff & Blondeau 2007, Sankoff, Blondeau, & Charity 2001, Trudgill 1988, Wagner 2008, Yaeger-Dror 1994, Ziliak 2012). However, small sample sizes raise concerns about the power of the experiment. This issue of power is explicitly taken into account in the present chapter by using appropriate statistical methods, particularly the generalized additive mixed model ("GAMM"; Wood 2017). Contrary to ANOVA, GAMMs do not require data from experimental trials to be averaged over both the time and space dimensions, thus achieving more precision. At the same time, the GAMM analyses used in this chapter make it possible to model the topographical distribution of the MMN (as was the focus in Grosvald & Corina 2012) without incurring the "curse of dimensionality", by not requiring that electrodes be coded using many-leveled factors for hemisphere and anteriority, thus permitting parsimonious models. The Bayesian approach adopted in this paper, which is explained below, provides a natural measure of the power of the analysis, by defining power as the degree to which the experimental goal of rejecting the null or the alternative hypothesis was reached (Kruschke & Liddell 2018). The Bayes factors used in this chapter provide a direct measure of the degree to which this goal was attained, and hence whether the sample size was sufficient to detect the presence or absence of group differences.

5.2 Method

5.2.1 Participants

Participants were eight Flemish-Dutch first-year students in the Netherlands and nine Netherlandic-Dutch controls. The participants were measured in two sessions: one approximately four months after the start of the academic year, and once again approximately four months later; with the exception of two Flemish students, all participants took part in the second session. This yielded 14 datasets for the Flemish-Dutch students and 18 datasets for the Netherlandic-Dutch students.

The experiments followed the Ethics Code for linguistic research in the faculty of Humanities at Leiden University, which approved its implementation. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

5.2.2 Stimuli

Stimuli were realizations of [eː] versus [ei], [eiɫ] versus [eːɫ], and [ər] versus [əɹ]. The stimuli were produced in a carrier word and sentence by a trained phonetician. Of all stimuli, five different tokens were selected. Praat (Boersma & Weenink 2016) was used to extract the relevant segment(s), to equalize all F0s to the average of all tokens used, to equalize all amplitudes to 60 dB SPL, and to fix the durations of the stimuli [eː,ei,ər,əɹ] at 200 ms and those of the stimuli [eːɫ,eiɫ] at 300 ms. For each of the six stimulus types, all five tokens were included as varying standards (68 presentations each, constituting 85% of the experiment when taken together) and one token was included as the deviant (60 presentations, or 15%. This resulted in a total of six experimental conditions, which are summarized in Table 5.2. In the remainder of the chapter, these conditions will be referred to by the corresponding deviant stimulus, such that "the [əɹ] condition" is the condition where [əɹ] was the deviant and [ər] were five varying standards. As an illustration of the stimuli, Figure 5.1 shows waveforms, spectrograms, and F3 trajectories[2] of the stimuli used in this condition; note how the [əɹ] stimulus has a much lower F3 than the others.

---

[2] The jittery F3 track in the trilled part of the [r] is not erroneous; this is inherent to the nature of this trill.

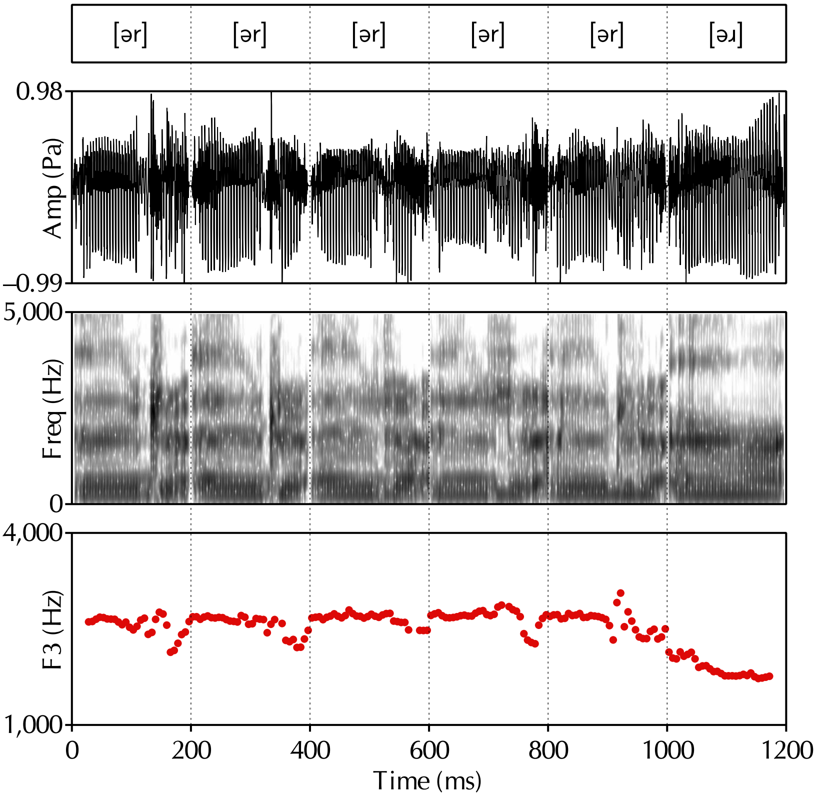

Figure 5.1: Example waveforms, spectrograms, and F3 trajectories (the critical difference between the two types of rhotic realization) of all five tokens of [ər] used as standards and one token of [əɹ] used as deviant, which together make up Table 5.2's [əɹ] condition.

Table 5.2: Design of the six conditions used in the experiment.

Standard (68×5 tokens)	**Deviant (60×1 token)**
[eː]	[ei]
[ei]	[eː]
[eiɫ]	[eːɫ]
[eːɫ]	[eiɫ]
[ər]	[əɹ]
[əɹ]	[ər]

5.2.3 Procedure

The experiment took place in a dimly-lit sound-attenuated booth. Participants were seated in a chair in the center of the room in front of a computer monitor, which was located behind an electrically-shielding glass pane. Two loudspeaker boxes were placed in the corners of the room at a distance of approximately 80 cm from the participant. During the experiment, the computer monitor was used to display a silent movie, so as to occupy the participant's attention. The sound stimuli corresponding to the six experimental conditions were presented over the loudspeaker boxes at a volume that was comfortable to the participant. The experiment was administered using PsychoPy (Peirce 2007) on a PC running Windows 7. EEG activity was recorded using a BioSemi ActiveTwo system with a sampling rate of 512 Hz. 32 AgCl electrodes were used, arranged according to the 20/10 system. Six additional electrodes recorded the left and right mastoids and the left and right horizontal and vertical extra-oculograms. Raw data were collected and were referenced off-line to the linked mastoids. Previous research has shown that this is the optimal reference choice for the auditory MMN (Mahajan, Peter, & Sharma 2017), because this is where the MMN effect achieves the highest signal-to-noise ratio even when it is small in magnitude (Kujala, Tervaniemi, & Schröger 2007, Mahajan, Peter, & Sharma 2017, Picton et al. 2000).

Before the start of the experiment, participants were instructed by the researcher to try to sit still and to try to keep blinking to a minimum. When the researcher started the experiment, an audio file spoken by a male speaker of Netherlandic Standard Dutch was played, which provided the participant with instructions. Participants were instructed to focus their attention on the silent movie and to ignore the auditory stimuli, and again to try to sit still and keep blinking to a minimum. A transcript of the spoken instructions was also shown on the screen. Participants could initiate the experiment of their own accord by pressing any one of two buttons located on either armrest of the chair. The six conditions were then presented to the participants in pseudorandomized order. There were no breaks in the experiment, which lasted exactly 28 minutes.

5.2.4 Data analysis

The raw EEG data were processed using R (R Core Team 2020) package `eegUtils` (Craddock 2019) by detrending the referenced data and applying a band-pass filter with a low cut-off of 1 Hz and a high cut-off of 30 Hz. Data were epoched from −100 ms to 450 ms post-stimulus-onset, where the first 100 ms served as baseline. Eyeblinks were removed from the epochs using least-squares regression. Trials contaminated by artifacts were rejected automatically. The data analysis focused on the six sounds used as deviants, compared to when these same six sounds were used as one of five varying standards. As such, trials of standards that were not also used as deviants were removed from the data.

The temporal window for the analysis was set at 275–325 ms post-stimulus-onset. This is the same window that was used by Grosvald & Corina (2012), and a narrower version of the 250–350-ms window used by Lanwermeyer et al. (2016). Other 50-ms windows were also investigated, but results from other possible MMN windows (e.g. 175–225 ms) were qualitatively similar enough that arbitrarily selecting a different window from the established 275–325 ms was not warranted. The data were averaged over time within this interval. Following the approach by Grosvald & Corina (2012), the data were not subsequently averaged over a specific region of electrodes, but the 32 electrodes were instead explicitly involved in the analysis. Compared to Grosvald & Corina (2012), a slightly more modern approach is used by modeling the electrodes as measurement sites on the surface of a 3D sphere. This removes the need to fit complex interaction terms of the "Hemisphere × Anteriority × Electrode" type, while retaining their advantages of specifying a precise model that achieves sufficient statistical power despite the relatively small sample size.

The statistical analysis was implemented by means of generalized additive mixed models, using function `gam` from R package `mgcv` (Wood 2017). The EEG amplitude was modeled using a smooth spline of the 32 EEG electrodes, which were mapped to a sphere based on their latitude and longitude coordinates using at most fifteen basis functions. This "spline on the sphere" informs the model that the data sampled from nearby electrodes are correlated to one another in a way that corresponds to data collected from the surface of a sphere. Difference smooths were included by the factors "Group" (coded such that Netherlandic = 0 and Flemish = 1), "Session" (coded such that the first session = 0 and the second session = 1), and "Deviant" (coded such that the sound used as standard = 0 and the sound used as deviant = 1) and all interactions. Random smooths by participants were added for the reference condition and by the factors "Deviant", "Session", and "Deviant × Session". Thus, each model contains a reference smooth, which models the topographical distribution of the EEG activity of the Netherlandic listeners, in session 1, presented with standards. Separate terms then model the difference in activity between this reference condition and the various factors manipulated in the experiment. Separate models were run for each of the six vowels. All models were fitted to scaled-t errors.

To test for possible asymmetrical effects, significance was not established via p-values but rather using Bayes factors. These make it possible to argue not just that an MMN is *present*, but also that it is *absent*, which is expected if the MMNs to be obtained are indeed asymmetrical (Cornell, Lahiri, & Eulitz 2011, Lahiri & Reetz 2010). For each of the eight smooth terms present in the model, a model with this term removed was fitted using maximum likelihood. The BIC (Schwarz 1978) of this model was compared to that of the full model (refitted using maximum likelihood). The difference between the two BICs was converted into a Bayes factor using equation (5.1), which is due to Wagenmakers (2007).

$$\mathrm{BF}_{10} = \exp\left(-\frac{1}{2}(\mathrm{BIC}_{\text{full model}} - \mathrm{BIC}_{\text{reduced model}})\right) \tag{5.1}$$

Bayes factors larger than 1 indicate support for the alternative hypothesis (the full model providing a better fit than the reduced model) and values smaller than 1 indicate support support for the null hypothesis. Section 5.3 reports these on the $\log_{10}$ scale instead, in which case the interpretation is symmetrical around zero: a $\log_{10}$ Bayes factor of zero indicates no support, positive values indicate support for the alternative hypothesis, and negative values indicate support for the null hypothesis.

5.3 Results

Table 5.3 shows the $\log_{10}$ Bayes factors corresponding to the the terms included in the statistical analyses. These are considered to be significant if their magnitude exceeds 0.5; this corresponds to what Jeffreys (1961) calls "substantial" evidence. Bayes factors with smaller magnitudes than this criterion indicate that there was insufficient evidence to be confident in a conclusion; this is indicative of an insufficiency in statistical power (Kruschke & Liddell 2018). Within each vowel, the critical effect is the difference between the vowel used as standard and the same vowel used as deviant, encoded by the factor "Deviant" and its interactions with the other factors in the design. Hence, of interest for the hypotheses are the effects for "Deviant", "Deviant × Group", and "Deviant ×

	Condition					
Factor	[ei]	[eː]	[eːɫ]	[eiɫ]	[əɹ]	[ər]
Reference smooth	0.88	1.34	−0.01	2.75	0.34	31.28
Deviant	0.58	−0.26	−0.11	−0.10	34.70	24.46
Group	0.01	−0.08	2.94	7.09	−1.71	0.17
Session	−0.07	15.61	3.42	7.41	3.72	−1.18
Deviant × Group	0.52	2.05	−0.67	−1.42	20.73	−12.30
Deviant × Session	11.18	−0.41	4.31	−0.05	5.94	−0.39
Group × Session	−0.04	0.00	−0.04	0.92	3.16	−12.29
Deviant × Group × Session	0.53	15.14	−0.42	0.43	−0.07	−3.56

Table 5.3: Results of the statistical analyses, reported as Bayes factors on the $\log_{10}$ scale.

Group × Session". Where these terms' Bayes factors provide significant support for the alternative hypothesis, Figure 5.2 visualizes the marginal effect mapped onto a stereographic projection of the head. The topographical plots in this figure thus directly correspond to the significant differences in EEG amplitude across the scalp.

In the [ei] condition, where [ei] is the deviant and varying [eː]s are the standards, there is substantial evidence for a difference between the [ei] used as standards versus used as deviant, i.e. an MMN. Figure 5.2 shows that this corresponds to a very small MMN, which reaches a minimum of −0.21 μV at Fp1/Fp2 (compared to a maximum of +0.01 μV near PO3). There is substantial evidence that this MMN differs for the Flemish students: their MMN is more negative by −0.67 μV frontally (to −0.47 μV near PO3). There is also substantial evidence that this between-groups difference changes over the sessions: the decrease in Session 1 of the experiment is counteracted by at least +0.53 μV near F7 and at most +1.01 μV around PO4, bringing their MMN in line with that of the Netherlandic controls.

In the reverse condition, with [eː] as the deviant and [ei] as standards, there is insufficient evidence to warrant claims about differences between [eː] used as standard versus as deviant. Following Kruschke & Liddell (2018), this can be rephrased as a lack of statistical power. There is, however, "decisive" (Jeffreys 1961) evidence for a different response to the deviants by the Flemish students, as well as decisive evidence that this difference changes over the two sessions. The second row of plots in Figure 5.2 shows that the Flemish students have a less negative MMN ERP to the [eː] deviants than the Netherlandic controls, by as much as +0.83 μV around Fz. In contrast, in the second session they have attained a strong MMN, which differs from the first session's MMN by

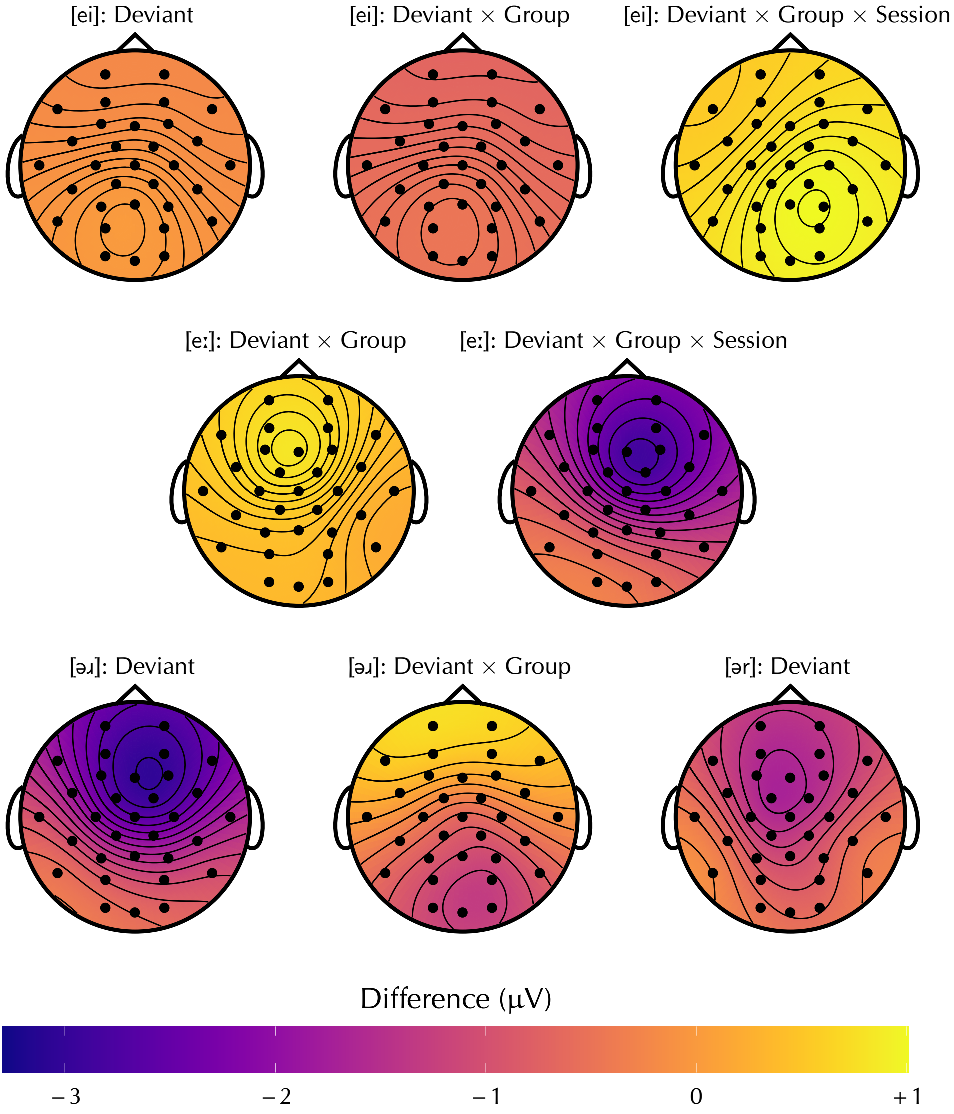

Figure 5.2: Topographical plots of the marginal effects of interest whose Bayes factors indicated at least substantial support for the alternative hypothesis. The baseline is the sounds used as standards, heard by the Netherlandic controls, in the first session of the experiment.

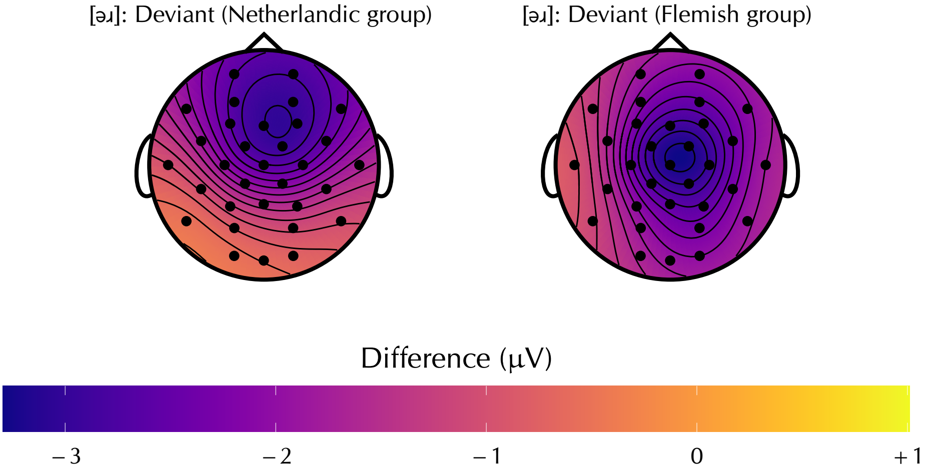

Figure 5.3: Side-by-side comparison of the "[əɹ]: Deviant" effect (left) and the sum of this effect and the "[əɹ]: Deviant × Group" effect (i.e. the MMN difference between the baseline Netherlandic and the contrasting Flemish group; right). The magnitudes of the MMNs are very similar between the groups, but the Flemish group has the effect shifted significantly towards the midpoint of the scalp.

up to −2.88 µV near Fz.

Finally, in the [əɹ] condition, decisive evidence is found for an MMN response to the [əɹ] deviants. The effect reaches a magnitude of up to −3.04 µV between Fz and F4. The evidence for a between-groups difference in this MMN is also decisive, such that the Flemish group's MMN is significantly less pronounced at the frontal pole (with a maximal difference of +0.76 µV) and more negative near occipital sites (by up to −1.31 µV). When this effect is added on top of the main effect for "Deviant" (see Figure 5.3), the result is an MMN similar in magnitude to the one for the Netherlandic group (with a largest negativity of −3.29 µV), but shifted away from the frontal pole and closer towards the center of the head for the Flemish group. No evidence was found that this shifting of the MMN in the Flemish group changed over the two sessions. In the reverse condition, where [əɹ] formed the standards and [ər] was the deviant, there is again decisive evidence for an MMN response to the [ər] deviants, but this MMN is smaller in magnitude (−1.70 µV around Fz). There is decisive evidence that there are no differences between the groups in this MMN, and that this did not change for the Flemish group over the two sessions.

5.4 Discussion

The results from the present oddball experiment show differences in aspects of phonotactic knowledge of Netherlandic Dutch between the Flemish experimental group and the Netherlandic control group. The sound change diphthongizing [eː] to [ei] has left its mark, in that both the Flemish students and the Netherlandic students exhibit a small MMN ERP when the monophthongal allophone is changed to a diphthongal one. This MMN is much larger in the Flemish group, where it reaches a peak negativity of $-0.90\,\mu V$. The sizes of these MMNs are on the same order of magnitude as the one reported by Lanwermeyer et al. (2016) in their allophonic condition, making the results credible as reflections of allophonic knowledge related to sound change. The [ei] condition additionally shows that the Flemish students are learning over the course of their stay in the Netherlands: approximately four months after the first session, their MMN to the [ei] allophone has reduced in size, bringing it to the same small level as the Netherlandic controls. The reverse condition, where [eː] is the deviant amidst [ei] standards, shows a similar learning effect: in the first session, the Flemish group has a significantly attenuated MMN response compared to the Netherlandic controls, but in the second session they attain a strong MMN at the expected topographical location.

The Flemish students' increased MMN to the [ei] realization in the first session of the experiment shows that, already in the first session, this difference is encoded by the Flemish students, and is in fact represented more strongly than it is by the Netherlandic controls. After the Flemish students have spent more time in the Netherlands, and have become more used to the diphthongal allophone, they are observed to attenuate their MMN response, coming in line with the Netherlandic controls. The same is observed in the reverse condition, where [eː] is the deviant. Here, the Flemish students' MMN is attenuated in the first session, indicating that they do not find the [eː] realization as salient as the Netherlandic controls do, but they reverse this difference in the second session of the experiment.

The results for the [eː∼ei] allophones provide evidence that sets apart the misperception and misrepresentation accounts of sound change. The pattern of results for the [ei] allophone is incompatible with misperception: not only did the Flemish participants perceive the difference at all, they encoded it even more strongly than the Netherlandic controls did, already in the first session of the experiment. The former result can be explained by either account, but the latter cannot be explained by making reference only to sound perception. The misrepresentation account has no problem with this result, and might speculatively attribute the stronger response in the first session of the experiment to an on-going learning effect (analogous to that found in second-language and second-dialect acquisition). The return of this between-groups difference to the baseline, Netherlandic, level in the second session is fully in line with the second prediction made in the Introduction: by the second session of the experiment, the Flemish participants have acquired the Netherlandic pattern. The result for the [eː] allophone can be explained by either theory. The Flemish participants find the switch from [ei] to [eː] less noteworthy in the first session, but have gotten more attentive to it by the second session; under the misperception account, this is because in the first session, they have not yet learned to neurally encode the difference between these stimuli as strongly as the controls, but by the second session they have. On the other hand, under the misrepresentation account, the reason is that after they correctly perceive the [eː] sound, they impart less sociolinguistic salience to this switch, which by the second session of the experiment they have managed to acquire. The results for the rhotic, described further down, will lend more credence to the latter interpretation.

The results for the [eːɫ] and [eiɫ] realizations are quite different from those for their single-vowel counterparts. No MMN-related effects were found, and for the most important term "Deviant × Group", there was (very) strong evidence that there was no difference between the groups. This result is surprising, given the positive findings in the single-vowel conditions and the findings by Jacobsen (2015) and Steinberg, Truckenbrodt, & Jacobsen (2010a, 2010b, 2011), who also used vowel-consonant sequences to demonstrate allophonic knowledge in an oddball task. The major phonological difference between the latter authors' experiment and the present one is the type of allophonic rule: in their publications the vowel determined the realization of the following consonant, whereas in the present experiment the opposite was true. However, this cannot be the full explanation, as this was not the case for the comparable null findings by Kazanina, Phillips, & Idsardi (2006). One possible scenario is that there are in fact MMNs in the baseline condition, but that the present sample was not sufficient to detect them: in both the [eːɫ] and the [eiɫ] conditions, the Bayes factors indicated that the statistical power was too low to draw any firm conclusion one way or the other. Further research is necessary.

The results for the rhotics are partially similar to those for the single vowels. Both [əɹ] and [ər] generate MMNs when presented as deviants, but the MMN to [əɹ] is twice the size of that to [ər]. The difference between these two sounds compared to the two vowel conditions is twofold: first, the critical allophone difference is in the consonant rather than in the vowel; second, the rhotic condition is significantly more salient sociolinguistically. The [ɹ] realization, though Netherlandic-Dutch, is a highly salient sociolinguistic variable in both the Netherlands and Flanders (Sebregts 2015), whereas the [r] realization is just as sociolinguistically demarcative, but does not come with the strong public awareness of its counterpart. With prior research not making a strong case for an explanation of the rhotic results in terms of the C/V distinction (recall the results by Jacobsen 2015 and Steinberg, Truckenbrodt, & Jacobsen 2010a, 2010b, 2011), the sociolinguistic explanation remains. The MMNs show that both [əɹ] and [ər] deviants elicit a mismatch, and that the [əɹ] elicits a larger MMN, which shows that this sound is more salient (per Scharinger, Monahan, & Idsardi 2016).

The [əɹ] deviant additionally elicits a difference between the Flemish group and the Netherlandic controls. This difference is topographical in nature: the lowest MMN value is approximately the same for both groups, but the Flemish group shows less activity at frontal sites (for comparison, Figure 5.3 shows the groups side by side, with the effect in the Netherlandic group on the left and the difference with the Flemish group added on top of it on the right). Due to the inverse-mapping problem (computing how electrical signals transmitted from a certain dipole in the brain are distorted by the surrounding brain areas, the skull, and the skin tissue is straightforward; computing the reverse starting from the voltage measured at the scalp is an unsolved problem), the difference in EEG signal at this location does not necessarily reflect differential activity of specifically the frontal brain areas in the Flemish participants. However, we know from prior literature that, among others, frontal areas are involved in MMN generation (e.g. Baldeweg et al. 2002, Giard et al. 1990, Rinne et al. 2000) and that these areas are also responsible for attention (Deouell 2007, Rinne et al. 2000), which is the primary component of sociolinguistic salience (Rácz 2013). As the only remaining difference between the two rhotic allophones is the increased sociolinguistic salience of the [ɹ] allophone, a speculative explanation in sociolinguistic terms could be as follows: the Flemish students did not grow up with this [ɹ] allophone, and hence do not have its sociolinguistic salience ingrained in their representations, hence the reduced contribution from frontal areas to the MMN. However, this needs to be tested thoroughly by future research; as sociolinguistic salience was not the primary manipulation in this study, any explanation in terms of salience can only be exploratory here. Future research should investigate effects of salience on the MMN directly.

In conclusion, the results found in the present chapter do not support Ohala's (1981) account of the actuation of sound change, and do support the views by Hamann (2009) and Beddor (2009). The group differences in the MMNs show that the Flemish students are perfectly able to perceive and categorize the diphthongal [ei], monophthongal [eː, and gliding [əɹ] allophones, but process them differently. In the [ei] case, they even have a *stronger* MMN than the Netherlandic control group. These results do not make sense if the Flemish participants were not able to appropriately perceive or encode these sounds. The results *do* make sense, however, with reference to phonological and sociolinguistic knowledge (the latter by process of elimination, although a neurophysiological basis was suggested). Such knowledge operates on a higher and more abstract level than bottom-up phonetic processing, and effects of such knowledge in the process of sound change are therefore incompatible with Ohala's (1981) view. If, however, sound change happens not during the transmission of the speech signal but during its grammatical evaluation, the results follow naturally as a result of differences in the phonological and sociolinguistic representation of the stimuli in the present study.

5.5 Conclusion

The present study built on previous work by Grosvald & Corina (2012) and others to investigate the listener's role in sound change: is sound change due to differences in low-level perceptual processing (Ohala 1981) or due to differences in higher-level representation in the grammar (Beddor 2009, Hamann 2009)? The results showed evidence against the former but in favor of the latter: the experimental group in this experiment perceived the diphthongal [ei], monophthongal [eː], and glided coda [ɹ] just fine, but processed them differently compared to the control group. While this in and of itself does not speak against Ohala (1981), and in fact would be predicted by him, the specific differences that were found are not easily amenable to an explanation in terms of misperception. The Flemish participants' MMN to the [ei] vowel implies that their perception of it is more than adequate, and is in fact even stronger than it is for the Netherlandic controls. On the rhotics, the Flemish group displayed a less frontal MMN to the [ɹ] allophone, which shows that their perception of this sound is again fine, but that they do not find this sound as salient as the Netherlandic controls do (a sociolinguistic observation which has support in neurophysiological findings, but should be subjected to future research). A possible explanation of the group difference in the [eː] allophone, which was present in the first session and inverted in the second session, was along similar lines. The aforementioned findings reflect sources of grammatical knowledge that are of a higher order than Ohala's (1981) distinction between intrinsic and extrinsic variation (Wang & Fillmore 1961), on which his proposal is founded. In contrast, the results follow naturally if the necessary information for the processing of sound change is evaluated as a normal component of the grammatical system as a whole, and thus if sound change corresponds to a change in the linguistic *grammar*. This is the proposal by Hamann (2009) and Beddor (2009). The observation that the Flemish participants became more Netherlandic-like in their perception of the [ei] vowel after four months' time is thus an observation of grammar learning, not of changes in perception.

Acknowledgments

I thank my dissertation supervisors, Claartje Levelt and Yiya Chen, for steering me into the direction of the experiment reported in this chapter. I additionally thank Daan van Renswoude, who introduced me to Wagenmakers's (2007) method for calculating Bayes factors, and Janet Grijzenhout, for helping me make the link with the work by Flege. I am also grateful for the helpful feedback I received at the University of Amsterdam and at the fifth Sociolinguistics Circle conference, where I presented early versions of this chapter.