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

Voeten, C.C.

CHAPTER 1

Introduction

In a seminal paper that determined the agenda of sociolinguistics for decades to come, Weinreich, Labov, & Herzog (1968) identified five major problems that need to be solved in order to explain the phenomenon of linguistic change. The *actuation problem* concerns the initiation of change: why does a certain change occur in a certain language at a certain point in time, but not in another language or at another point in time? The *constraints problem* seeks to identify the sets of possible and impossible changes, and their structural conditions. The *embedding problem* situates an individual change within its larger linguistic and social context, and the closely-related *evaluation problem* discusses the social meaning of a change. The final problem, the *transition problem*, is the focus of this dissertation. In its original formulation, the goal of the transition problem is to identify the pathway through *linguistic structure* by which a change progresses. An example, taken from Scheer (2014), is the change from [l] to [ʁ] in intervocalic position in the Genoni dialect of the Italian peninsula of Sardinia. Synchronically, this change is "crazy" (Scheer 2014), in the sense that it is phonetically unmotivated and hence appears unnatural. The same is true diachronically: a historical change l → ʁ / V_V is "crazy". However, Scheer (2014) argues that this change has actually arisen quite plausibly, via an intricate chain of l > *ɬ > w > g^w > *$ɣ^w$ > ʁ. The transition problem formulated in Weinreich, Labov, & Herzog (1968) deals with establishing these types of plausible historical derivational chains.

There is, however, a second type of transition that Weinreich, Labov, &

Herzog (1968) do not discuss in detail. Just as the embedding problem is (correctly) split into "embedding in the linguistic structure" and "embedding in the social structure" (Weinreich, Labov, & Herzog 1968:185), the idea I wish to put forward is that the same distinction should be made for the transition problem. That is, a distinction should be made between the transition throughout *linguistic structure* (corresponding to Weinreich, Labov, & Herzog's 1968 original formulation of the transition problem) and the transition throughout *social structure*, i.e. the speech community. The concept of such a split was present in Weinreich, Labov, & Herzog (1968): situated within the context of the 1960s, the transition of a change throughout social structure was in fact wholly incorporated in Weinreich, Labov, & Herzog's (1968) evaluation problem. With the present-day knowledge and (statistical) methods available, however, an addendum is necessary. Linguistics in the 21st century is much more concerned with individual differences than it was in the 1960s, and recent advances in statistics such as mixed-effects models or the more flexible generalized additive mixed models make it possible to explicitly take into account such individual variation at no additional methodological cost. These advances in the field make it possible to look at the transition of novel linguistic structure not just throughout the social community, but also throughout the *individual members* of that community. This individualized version of the transition problem is concerned with the individual language user as a processor of spoken language: how do you pick up a language change, and what is the time course involved? This is the question this dissertation aims to answer.

Throughout the remainder of this introductory chapter, I will refer to this narrowly-scoped variant of the transition problem as the *adoption problem*. This dissertation investigates the adoption problem by taking advantage of a set of sound changes that are currently on-going in Dutch. These are introduced in Section 1.1. Section 1.2 then discusses the origin and actuation of sound change within a single individual and the large-scale propagation of sound change throughout the community, two large issues between which the adoption problem is perfectly sandwiched. Section 1.3 discusses the extent to which these theoretical notions are in line with psycholinguistic and computational models of human speakers and listeners, and also goes into some important methodological innovations that make it possible to perform the psycholinguistic experiments used in this dissertation. Finally, Section 1.4 concludes the introduction of this dissertation with an overview of the remaining chapters.

1.1 The Polder shift: an on-going vowel shift in Dutch

A sound change that has been on-going in Dutch for approximately a hundred years by now is the diphthongization of the tense mid vowels /e:,ø:,o:/ towards present-day [ei,øy,ou]. Because this sound change is relatively recent, it has been well-described (Adank, van Hout, & Smits 2004, Adank, van Hout, & Van de Velde 2007, Van de Velde 1996, Van de Velde, van Hout, & Gerritsen 1997, Van de Velde & van Hout 2003, Zwaardemaker & Eijkman 1924). Of particular note are Zwaardemaker & Eijkman (1924), who were the first to notice (and express their disapproval of) a small offglide developing in the vowel /e:/, probably realized as something approaching [e:^{j}]. A detailed account of subsequent events that took place between 1935 and 1995 is provided by Van de Velde (1996) on the basis of historical radio recordings. He shows that the adoption of the diphthongization resembles a logistic curve—the S shape that is typical of on-going language change. In 1935, the vowels are diphthongized only negligibly, in 1965, the average diphthongization is about 50%, and in 1995, nearly all realizations are fully diphthongized. Later synchronic measurements by van der Harst (2011) confirm that, as of the twenty-first century, the tense mid vowels have indeed become genuine diphthongs. This is typical for processes of language change: in the beginning of a change, only a few individuals share a linguistic innovation; at some point, these individuals spread the innovation to their peers and, perhaps most importantly, their children, causing a rapid ascent in the adoption of the change; finally, the (usually older) people who have not acquired the change either manage to acquire it, or pass away. This yields an S-curve with exponential growth at the beginning and exponential decay at the end. The fact that the diphthongization of the Dutch tense mid vowels follows exactly such a curve suggests that this is, indeed, a language change in progress.

The diphthongization of the tense mid vowels went hand in hand with a second change, namely the lowering of the original diphthongs /ɛi,œy,ɔu/, particularly /ɛi/ (Blankestein 1994, Gerritsen & Jansen 1980, Gussenhoven & Broeders 1976, Jacobi 2009, Mees & Collins 1983, Stroop 1992, Stroop 1998, Van de Velde 1996, Voortman 1994), commonly referred to as "Polder Dutch" (Stroop 1998). The relatedness of these two changes is obvious, but the causal connection is debatable. Stroop (1998) suggests that the phonetic lowering of /ɛi/ initiated a drag chain that attracted the tense mid vowels' nuclei, causing the intrinsic tendency towards slight diphthongization of these vowels to become extrinsicized. Another option, however, would be to postulate a push chain, where the diphthongization of /e:,ø:,o:/ was the original innovation, which then pushed the lax diphthongs /ɛi,œy,ɔu/ out of the way (vi-

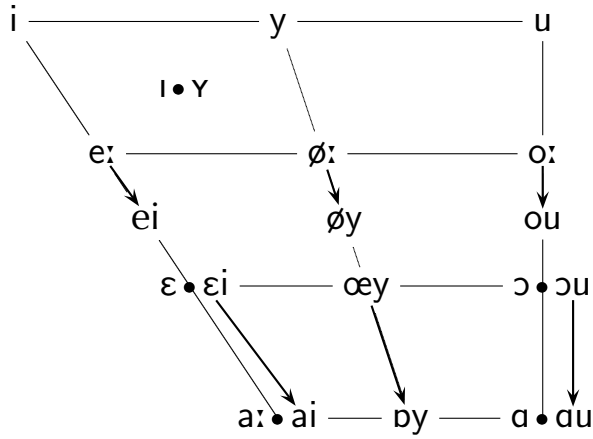

Figure 1.1: Vowel diagram showing the changes constituting the on-going vowel shift. The arrows indicate the diachronic changes; secondary arrows to indicate upgliding diphthongization are not included to prevent cluttering the diagram.

sualized in Figure 1.1). This latter option seems more credible for three reasons. First, the diphthongization of the original /eː,øː,oː/ was first observed in 1924 (Zwaardemaker & Eijkman 1924), while the lowering of the original diphthongs was only put to paper in 1990 (Jacobi 2009). Secondly, diphthongization of tense mid vowels is a natural phonetic development (Labov, Yaeger, & Steiner 1972, Watt 2000), which makes this change a plausible initiator of a chain shift, whereas the lowering of pre-existing diphthongs is more often the consequence of an earlier change in a chain shift (Labov 1994). Finally, Labov (1994:116) notes that "in chain shifts, the nuclei of upgliding diphthongs fall". If the diphthongization of [eː,øː,oː] was the first step in this chain shift, then the second step would be their nuclei falling, which would cause them to merge with the original, lax, diphthongs. Preservation of phonemic contrast would then naturally require these original diphthongs to move "out of the way", resulting in a push chain, as opposed to a drag chain; thus, Labov's observation indirectly makes a push chain more likely than a drag chain.

For as long as they have existed, the original diphthongs /ɛi,œy,ɔu/ have been subject to distributional restrictions. The canonical reference is Booij (1995), who notes that these diphthongs are realized as long monophthongs [ɛː,œː,ɑː] before coda /l/, presumably, Booij argues, for articulatory reasons. It might then come as no surprise that this very same restriction was also noted by Zwaardemaker & Eijkman (1924): according to them, the synchronic process of /eː/ → [eːʲ] was also blocked before coda /l/; later authors added /r/

(Gussenhoven 1993) and /ʊ,j/ (Collins & Mees 1999) to the list, completing the natural class of Dutch approximant consonants. The precise set of distributional restrictions is somewhat complicated; a comprehensive description is given in Voeten (2015). These distributional restrictions complicate the picture of the vowel shift described thus far, by adding a *phonological* dimension to the phonetic changes affecting these vowels.

The resulting vowel shift I will call the "Polder shift", in hommage to Stroop's (1998) "Polder Dutch" term. This shift seems to be an ordinary vowel shift, albeit one with a specific, known, historical track record. This is a rare fortune in historical linguistics. In addition, the Polder shift can be shown to recruit phonological knowledge rather than being a plain phonetic vowel shift (Voeten 2015). A final point, which makes the Polder shift especially suitable for the investigation undertaken in this dissertation, is that the Polder shift has resulted in significant sociolinguistic variation. Previous research has shown significant differences, particularly in the realizations of the tense mid vowels and diphthongs, between the Dutch spoken in the Netherlands and the Dutch spoken in Flanders, the northern half of Belgium; in particular, see Adank, van Hout, & Smits (2004), Adank, van Hout, & Van de Velde (2007), Van de Velde (1996), and Van de Velde, van Hout, & Gerritsen (1997), and Van de Velde & van Hout (2003), among others. Chapter 2 is devoted to a thorough investigation of the way the Polder shift has spread throughout these two communities of spoken Dutch. It will be shown that there are robust synchronic differences between Netherlandic Dutch and Flemish Dutch, which parallel the diachronic differences between modern-day Netherlandic Dutch and its state before the Polder shift. This observation forms the foundation for the remainder of the dissertation, in which these synchronic differences are exploited to perform a *synchronic* investigation of the adoption of *diachronic* change.

This dissertation investigates these on-going sound changes in Dutch on the basis of psycholinguistic experiments with speakers of Netherlandic Dutch and speakers of Flemish Dutch. In doing so, the dissertation aims to describe and explain the phonetic and psycholinguistic mechanisms underlying the adoption of sound change using a variety of methods. The following sections of this introductory chapter show that this is a particularly necessary enterprise: currently, historical phonologists and psycholinguists are at odds with one another in their explanations of how humans process variation, and therefore how sound change can be actuated, adopted, and transmitted.

1.2 Theories about the lifecycle of sound change

1.2.1 Misperception as a source of sound change

Historical approaches to sound change tend to focus on the role of the listener in their interaction with the speaker. Thus, the most oft-cited source of the origin of sound change is *misperception*. Bermúdez-Otero (2007) calls a misperception event a "coordination failure" between a speaker (henceforth "S") and a listener (henceforth "L"). A coordination failure occurs when L perceives S's realization of a hypothetical category /A/ as [A′] rather than [A]; L is then assumed to reanalyze his representation of /A/ as /A′/ when he himself becomes the speaker in future communication events. A couple of remarks on this core proposal are in order. First of all, the present formulation of this proposal leaves ambiguous whether S actually *realized* [A′], or whether this is merely what L *heard* (sensorily) or *perceived* (after parsing the sensory input as a phonetic category). Classic accounts such as Hyman (1976) or Ohala (1981) assume that [A] and [A′] are actually the same speech signal, but that this speech signal is misparsed by L. To reuse the example used in Hyman (1976): if S plans to realize a syllable /bā/, S's F_0 will be slightly lowered at the onset of the /ā/ vowel for unavoidable reasons of human anatomy, resulting in [bā]; if L incidentally fails to perceptually compensate for this intrinsic effect, they will perceive /bā/. The theoretical account, however, does not require that what L perceived is actually what S *produced*; a simple mishearing on the part of L will produce the same result. In fact, there are at least three different types of misperception, as summarized by the three pillars of Blevins's (2004) Evolutionary Phonology framework. Misperception of the Ohalian kind, whereby L mis-partitions the intrinsic and extrinsic sources of variation in the speech signal, is considered "CHANCE" in Evolutionary Phonology. This classic type of misperception is more generally called "hypocorrection", following Ohala (1989) and Lindblom et al. (1995). The second type of misperception in Evolutionary Phonology terminology is "CHANGE", which corresponds to Ohala's (1989) and Lindblom et al.'s (1995) "hypercorrection". This situation is the opposite of CHANCE, in that rather than incorrectly attributing intrinsic variation to an intended gesture by S, L now thinks that part of the intended variation by S was actually *not* intended, thus "overparsing" the speech signal. Finally, sound change due to "CHOICE" takes place when L hears S produce multiple phonetic variants of the same word, and reconstructs a different underlying form for this word than S had intended. This is a type of drift in the statistical distribution of word tokens: the mode of the distribution shifts slightly towards a different variant.

The view that sound change has its basis in misperception is attractive, because it is obvious that a language-acquiring child must perceive before it can

produce. If authors like Labov (2007) or Hamann (2009) are correct that non-superficial reanalyses can only be made in first-language acquisition, then the driving force of sound change must be one of two things. On the one hand, it is possible that sound change is due to literal misperception (Ohala 1981), which leads to reanalysis by the language-learning child (Labov 2007). However, because it is this reanalysis that is the critical step in the acquisition, it can also suffice to single out this step of reanalysis. In this case, which is the interpretation by Beddor (2009) and Hamann (2009), the language-learning child does not make an error in perception, but rather uses a grammar with different cue weights than those of the adult speaker, thus arriving at a different phonetic interpretation of the speech signal produced by the adult. Chapter 5 provides evidence that this is more likely to be correct than Ohala's (1981) misperception account, although it will be shown *not* to hinge on children as the actors (as already argued by Bybee & Slobin 1982). This is problematic for generative theories of sound change, since these assume that grammatical restructuring after the critical period is not possible due to the inaccessibility of UG. Recent evidence by Pinget, Kager, & Van de Velde (2019) (also available in Pinget 2015) demonstrates that this assumption is unwarranted. Their results show not only that adults are capable of changing their sound systems just as much as children are (a well-known fact outside of generative linguistics, which will be revisited in Chapter 4), but also that there are different roles for perception and production. On the basis of two on-going mergers in Dutch, Pinget, Kager, & Van de Velde (2019) show that when changes are incipient, perception goes first: individuals need to perceive a change (via, e.g., a coordination failure of the Ohalian kind) before they will produce it. Depending on each individual's perception–production link, they may subsequently continue to spread the change among their peers, thereby pushing the change from the incipient into the on-going stage; see Coetzee et al. (2018) for detailed discussion of this phase of the process. Finally, when the sound change reaches the advanced stage and nears its completion, Pinget, Kager, & Van de Velde (2019) show that the perception–production relationship reverses, such that individual speakers come to produce sound systems in which the change has completed, but are still able to draw on any remaining fine phonetic detail in perception left in place by the old system. Curiously, this last step has also been observed to reverse: Labov, Yaeger, & Steiner (1972) report on the phenomenon of "near-mergers", where individuals fail to perceptually differentiate between two sounds that are involved in an on-going merger, yet consistently produce such differentiation in their own productions.

A corollary of a central role for misperception must be that "mini sound changes" are actuated all the time. In fact, if misperception is the sole mechanism behind sound change, Weinreich, Labov, & Herzog's (1968) actuation problem has been solved: a certain sound change actuates at a certain point in

time due to a misparsing by L of a certain kind, which subsequently becomes entrenched in L's grammar and is then spread to other individuals. The bigger question is then that of transmission: how is this incidental reanalysis transmitted, by L turned speaker, to the other members of their community? Yang (2009) provides a mathematical model of what is required, and solidifies his claims on the basis of real-world data by Johnson (2007). His answer is very simple: transmission takes place if at least 21.7% of the input consists of novel forms. However, what is going on between 0 and 21.7%? If the transmission of sound change beyond those 21.7% is reliant on the listener (as Ohala 1981 would claim), it could be that the first 21.7% have to be actuated by the speaker. Section 1.2.2 discusses this.

1.2.2 Speaker-induced sound change and the role of the representation

Research on the speaker as a possible source of sound change is, perhaps surprisingly, rather scarce. This can be understood when viewed in light of the structuralist tradition of phonology in which the lion's share of prior work on sound change has been done: under structuralist views, interlocutors both possess discrete categories, and an incidental pronunciation change by a speaker in continuous ("analog") acoustic space cannot initiate a sound change at the discrete ("digital") level. The listener, on the other hand, can initiate sound change of the Ohalian kind by erroneously creating a novel category or merging existing categories during a serendipitous misperception event. If the speaker is to play a substantive role in sound change, it must be on the continuous level of phonetics, a domain which is implicitly neglected in structuralism-inspired generative views on phonology.

Various attempts have been made to integrate continuous phonetics with discrete phonology. In the context of historical phonology, the theoretical necessity of an integration of both levels into theories of sound change was most clearly argued by Hyman (2013). Later authors, such as Bermúdez-Otero (2015) and Ramsammy (2015) provide explicit models taking this into account. These papers appear to have been implicitly influenced by Boersma's (2011) BiPhon model, as they incorporate the same major building blocks, save for the distinction between the Auditory Form and the Articulatory Form. As a framework for modeling sound change not just by the listener, but also by the speaker, BiPhon is a particularly relevant candidate, because it is *bidirectional*, as opposed to most models of sound change, which are (often implicitly) feedforward. This is a problem, because there is psycholinguistic evidence (see Section 1.3) that, for example, a misperception by L can be counteracted by L's lexical knowledge of what S is likely to have said given the discourse context. Accomplishing this either requires the ability to move back and forth between

different levels of representation in the model, i.e. feedback, or requires that the various sources of information are accumulated and a decision is only made in a single evaluation at the very end of the model. The latter follows directly from the parallel-OT architecture of BiPhon, but by definition cannot be obtained by feedforward models operating on ordered transformational rules.

BiPhon's parallel architecture offers another advantage over traditional feedforward models, because it models the listener and the speaker in exactly the same way: to change the role of the listener into that of the speaker, simply traverse the model in reverse. Thus, where the listener starts from a speech signal as input and interprets the optimal underlying form, the speaker starts from an underlying form and outputs the optimal speech signal, without requiring any alterations to be made to the grammar or its evaluation. As noted by Boersma (2011), this automatically bidirectional architecture generates specific predictions concerning sound change. If a BiPhon listener encounters variation of the Ohalian kind (i.e. variation due to anatomical differences or otherwise speaker- or condition-specific variation), termed "transmission noise" by Boersma (2011), said listener will acquire a broader distribution of possible auditory values for the sound(s) in question. This may over time lead to drift of the means of these distributions towards the novel realizations. However, because the listener will eventually also need to speak, this drift is counteracted by Boersma's (2011) mechanism of *prototype selection*. Prototype selection is the process of finding the optimal auditory form given a certain phonological surface form. As shown in simulations by Boersma & Hamann (2008), the increased variance in auditory values caused by transmission noise is counteracted by the fact that more extreme values, and hence large deviations from an established mean value, are more difficult to realize. The reason is partly in universal anatomical restrictions, but also lies in the fact that an individual's repertoire of motor programs will have been optimized for the original values. This causes the adoption of Ohalian variation to be maladaptive from the perspective of the speaker. Therefore, given that BiPhon models speaking and listening using the same grammar, a sound change can only be obtained if it is acquired by the listener via, e.g., Ohalian misperception *and* does not present problems for the speaker.

1.2.3 Exemplar Theory

An altogether different approach to integrating the speaker and the listener into the same model is to abandon the classic structuralist model of phonology and branch out to a new type of models. This step has been taken by Exemplar Theory (Pierrehumbert 2001). This brings us back to the possible types of sound change discussed in Bermúdez-Otero (2007), who argues that exemplar models are in fact the only theoretical device available that can predict sound

change that is both phonetically gradual and lexically gradual. The main claim of exemplar models is that our minds do not contain discrete phonological categories, but rather store (in full detail) the raw acoustic data with which we are provided as listeners. As this information passes through short-term memory, working memory, and long-term memory, it gradually decays (unless reactivated by, for instance, an attempt on the part of the speaker to reproduce the utterance). Decay can be considered the loss of fine phonetic detail, and the subsequent consolidation of different exemplars into a single prototype. Subsequent examples entering into long-term memory become absorbed into this prototype, and influence it by being averaged with the prototype exemplar; thus, many realizations which are accepted as belonging to a certain prototype but are not exactly equal to it will slowly result in drift. When the listener turns into a speaker, this same prototype will be used to generate speech, and thus new examples. Hence, slow drift in perception begets slow drift in production.

Bybee (2002) provides a general overview of how Exemplar Theory can be used to describe a specific type of sound change, viz. *reductive* sound change (reduction of segments or of individual articulatory gestures within segments, leading to lenition). In the case studied by Bybee (2002), /t,d/ deletion in American English, the reductive sound change is both phonetically and lexically gradual, and hence requires a framework such as Exemplar Theory to be modeled successfully. Bybee's (2002) idea is along the following lines. Words that are predictable, i.e. high-frequency words, are particularly good candidates for reduction: their inherent predictability means that for a correct interpretation, they are less dependent on the redundancy that is inherent to a full pronunciation. If, following Levelt, Roelofs, & Meyer (1999) and Wheeldon & Levelt (1995), words are viewed as highly-trained motor programs, and if humans desire to minimize their articulatory effort whenever possible (Passy's 1890 "Economy" principle), then reduction can be expected to take place particularly in these high-frequency words. By the tenets of Exemplar Theory, these words will then have a high proportion of reductions stored within their corresponding exemplar clouds. The average of all exemplars will then naturally shift towards the more reduced variants, until an equilibrium is reached between the desire for articulatory reduction ("Economy") and the need for functional communication (Passy's 1890 "Emphasis"). Note that no separate mechanism is required to spread these exemplars from one speaker to the next: because Exemplar Theory does not assume a unidirectional feedforward model of phonological processing, the more-reducing speaker will naturally provide more-reduced exemplars to his or her peers, without the need for any specific theoretical machinery.

1.2.4 Which comes first: perception or production?

While the preceding sections could be taken to suggest that pure speaker-caused sound change is rare, there are studies that have found change in production before change in perception, in which case it must have been the speakers who have made the first move. One such study is Evans & Iverson (2007), who studied the accents of British-English high-school students who were about to enter university. While at the end of high school these students had distinctly local English accents, after a year in university their productions had measurably changed to be more in line with Standard Southern British English. In perception, however, Evans & Iverson (2007) did not find significant changes, although they did find a correlation between their perception measures and the degrees to which their participants had changed in production. These results are incompatible with a view in which sound change starts in the listener and only later spreads to speakers.

Other work, however, has obtained precisely the reverse findings. In a study of the devoicing of Dutch fricatives and bilabial plosives, Pinget (2015) found strong evidence that it is the listener who has to initiate a sound change, which is in line with classic misperception-based accounts of sound change. Harrington, Kleber, & Reubold (2008), in a study of Standard Southern British English /u/-fronting (whereby /u/ is changing into /ʉ/), come to the same conclusion: their data are compatible with an Ohalian account of sound change, and not compatible with a speaker-initiated account of sound change.

Why do Evans & Iverson (2007) find change starting in production, and do Pinget (2015) and Harrington, Kleber, & Reubold (2008) find change starting in perception? The critical difference between the studies is the level of representation at which the change is playing out. For Evans & Iverson (2007), the sound changes to be acquired by their subjects are changes in surface realizations: a vowel /V/ changes its realization $[V_1]$ to $[V_2]$. In the case of the other two studies, the sound changes that are on-going are not phonetic, as in the case of Evans & Iverson (2007), but phonological. Harrington, Kleber, & Reubold (2008) make the case that their sound change started by listeners undercompensating for coarticulation of /u+t/ sequences, where the [u] becomes more front due to coarticulation with [t], leading them to a reanalysis of /u/ as /ʉ/. In this case, differently from Evans & Iverson's (2007), the realization [ʉt] was already in existence and the actual sound change is the reanalysis of the underlying form /u/ as /ʉ/. For Pinget (2015), the same is true: her study investigates the merger of the *phonemes* /f/ and /v/ and of the *phonemes* /p/ and /b/. Her conclusion that these sound changes needed to be perceived by listeners before they would become produced by speakers is in line with a hypothesis that change at the underlying level must be initiated by listeners, but change at only the surface level starts with the speakers.

The idea that underlying-form change starts in perception and surface-form change starts in production is supported by evidence from psycholinguistics. Research in psycholinguistics has shown that listeners are exceptionally skilled at compensating for variation in the phonetic input they receive. Thus, a listener presented with a subtle difference in a phonetic realization (as is the object of change in Evans & Iverson 2007) will perceptually compensate for this difference, preventing the actuation of an Ohalian sound change. However, as will be extensively discussed in Section 1.3, this ability to compensate for changes has limits. In particular, Witteman et al. (2015) have shown that listeners fail to accommodate on-line to changes by which a sound is realized as a member of another phoneme category—in this case, Dutch /i/ realized as [ɪ], which also exists as a separate phoneme in Dutch. This suggests that a listener cannot compensate for sound changes affecting underlying forms, and hence, for these sound changes, it has to be the listener who performs the crucial reanalysis. The implications for sound change are that if initiated by a speaker, sound change involves a reanalysis of the concrete realization corresponding to an abstract phonological category, i.e. of the phonology-phonetics mapping. In contrast, if a sound change is initiated by a listener, the reanalysis is one of the abstract category system itself, i.e. the phonetics-phonology mapping.

1.2.5 Types of change

After an individual has come into contact with a sound change, it needs to spread through their linguistic system. This is exactly Weinreich, Labov, & Herzog's (1968) original transition problem; recent literature (e.g. Bermúdez-Otero 2007) prefers to speak of *implementation*. Implementation is generally recognized to take place in one of four ways. The two most-well-known are Neogrammarian change (Osthoff & Brugmann 1878) and change by lexical diffusion (Wang 1969). Neogrammarian changes are those that start out as gradual phonetic innovations, which are then grammaticalized according to the lifecycle discussed previously. It follows that if sound change starts out in the realization of a phonetic category, all words in the lexicon are affected by the change at the same rate at the same time; thus, Neogrammarian change is phonetically gradual but lexically abrupt (Bermúdez-Otero 2007). By contrast, in the case of change by lexical diffusion, the change is a phonetically abrupt substitution of one phonetic category for another, which takes place in individual lexical items. Here, the locus of change does not lie in the realizations of phonetic categories, but rather in the realizations of individual lexical items: some words will have implemented the change, other words will not have. Thus, classic lexical diffusion is lexically gradual, but phonetically abrupt.

As a third option, Bybee (2002) notes that there are some sound changes

that appear to be both phonetically and lexically gradual. The dichotomy between Neogrammarian change and lexical diffusion presented above explicitly disallows this third mechanism of spread: a change is Neogrammarian, which is phonetically gradual but lexically abrupt, or it is lexically diffuse, in which case it proceeds through the lexicon slowly, but when a word changes, it immediately does so fully. Bybee's (2002) observation that there *are* changes that can be shown to be lexically diffuse, but where the individual lexical items are not changing at the same rate, shows that a third option is needed, which I provisorily term "change by exemplar", with "exemplar" referring to Pierrehumbert's (2001) Exemplar Theory. Bybee's (2002) analysis of these kinds of sound changes is that the exemplar clouds that ultimately give rise to phonological representations are slowly undergoing a regular sound change. Since exemplar clouds are formed separately for each word in the lexicon, the diffusion throughout the lexicon of this phonetically gradual regular sound change then naturally obtains.

The fourth and final mode of implementation in which a sound change can be implemented is phonetically and lexically abrupt. These changes arise due to (un)conscious choices, such as in accommodation. What changes here is what Janson (1983) calls the *norm*, which in his view is a conscious sociolinguistic allophone choice (such as which of a large variety of rhotics to use for a single /r/ category), which, upon transmission to a new generation of speakers only (Janson 1983), will result in a change in the underlying form for this category. In terms of the lifecycle, these changes follow the same steps as Neogrammarian changes, with the single difference that the original phonetic change is too large to have arisen gradually. In these changes, a central role is to be played by sociolinguistic factors such as accommodation by listeners to speakers they evaluate positively (Auer & Hinskens 2005, Chambers 1992, Janson 1983, Pardo 2006, Sonderegger, Bane, & Graff 2017, Trudgill 1986). However, this reveals a deficit in Janson's (1983) account that it shares with that of generative views on misperception. If adults can only change the underlying forms associated with phonemes and not the number and distribution of these phonemes themselves, as explicitly claimed by Janson (1983), then any categorical reanalysis can be performed only by their children. However, if adults can only change their surface realizations, the only system that they can transmit to their children is one in which changes in surface representations alone will do, and, due to the subset principle (Berwick 1985), no incentive for children to change will *ever* arise.

1.2.6 Summary

The received, generative, view of sound change claims that sound change originates in a coordination failure between a speaker and a listener, which involves

over- or underparsing of phonetic cues as phonological and vice versa. While coordination failures may occur at any place and point in time, reanalyses beyond the surface level can be performed by children only (Section 1.2.1). Sound change tends to be caused by the listener, but it can also be caused by the speaker, if and only if this comes at no additional cost for them, i.e. the change does not result in a system that is harder to produce or more difficult to understand. Specifically, the speaker selects prototypes that are optimal according to both production and perception (Section 1.2.2); this may result in the actuation of sound change if the distribution of available tokens or exemplars is skewed towards a better variant that is not currently the norm (Section 1.2.3). This type of speaker-induced sound change may be more likely to start as a physical change in the phonetic realization, whereas listener-induced sound change may be more likely to start as an abstract reanalysis of the underlying form (Section 1.2.4). Change may be categorized along phonetic and lexical abruptness/graduality, and these different modes of implementation may operate on different principles (Section 1.2.5).

This view is well-established, convenient, and largely plausible. However, Section 1.2.5 ended with a critical remark: if adults can only *produce* grammars that are consistent with surface-level changes, they can also only *transmit* grammars that are consistent with surface-level changes. It is then up to their children to reanalyze such phonetic changes as being part of the phonology, but the subset principle predicts that they do not generally do so. More importantly, however, is that neither of these predictions are completely in line with reality. As the Polder shift demonstrates (Voeten 2015), adults *can* in fact perform truly phonological reanalyses and *can* transmit these to their children. In addition, the same is true for children: as part of the normal process of phonological acquisition, children often come up (at least temporarily) with incorrect phonological analyses. What makes all this possible? The answer is probably in the way speakers and listeners cope with variation.

The individual's processing of variation falls under the purview of psycholinguistics, to which due attention is paid throughout the remainder of this dissertation. Empirical research in this field has provided additional challenges for the received view on sound change, and findings by psycholinguists will provide important stepping stones in the synthesis offered in Chapter 7. Section 1.3 provides a brief summary of the ways in which pycholinguistic empirical research is and (mostly) is not compatible with the received view of sound change. This provides the empirical background for an important component of this dissertation: methodology.

1.3 The psycholinguistics of variation in perception

1.3.1 Perceptual learning as the antagonist of misperception

The misperception-based account of sound change faces fundamental obstacles from a field of linguistic research that can from time to time be underappreciated by historical phonologists. Decades of work in psycholinguistics (summarized in major works such as Cutler 2012) indicate that human speakers and listeners are, in fact, extremely skilled at compensating for variation in the speech signal. In fact, psycholinguistic evidence shows that *compensating* for variation is a misnomer: variation is actually *used* in talker-specific processing strategies (Nygaard, Sommers, & Pisoni 1994, Palmeri, Goldinger, & Pisoni 1993). A specific problem for the misperception account of sound change is posed by the existence of lexically-guided perceptual learning (Norris, McQueen, & Cutler 2003): the phenomenon that listeners adapt to persistent deviations from their own expectations by individual speakers.

Perceptual learning is a process that primarily operates on the basis of lexical knowledge. The original Norris, McQueen, & Cutler (2003) paper was based on knowledge of individual words. Their experiment divided Dutch listeners into two groups, presenting them with words containing either /f/ or /s/ phonemes realized as an ambiguous intermediate sound [?]. In a training phase, these words were selected such that only one of these two interpretations was possible (e.g. "witlo[?]" can only be "witlo[f]" meaning "chicory", as there is no Dutch word *"witlo[s]"). In a subsequent categorization task, the /f/-familiarized participants showed expanded /f/ categories (an ambiguous sound needed to be more [s]-like for them than for the other group before they would categorize it as /s/), and the /s/-familiarized participants showed the reverse. In a later study on the same phenomenon, McQueen, Cutler, & Norris (2006) found that this effect generalized not just to phoneme categorization, but also to word recognition. In their experiment, participants were familiarized in the same way as in Norris, McQueen, & Cutler (2003), but then did a lexical-decision task with ambiguous words (e.g. "doo[?]", which can make either "doof" meaning "deaf", or "doos" meaning "box") with a cross-modal priming component. For both groups, the results from this experiment showed facilitatory priming effects for prime–target pairs congruent with training (e.g. auditory "doo[?]" paired with the picture of a box in the /s/-familiarized condition), and inhibitory priming effects for incongruent prime–target pairs. These results show that listeners use lexical information to retune their sound categories, and that this retuning generalizes to new items, and hence takes place at an abstract phonological level.

Further evidence demonstrates that this retuning of categories is not limited to lexical words nor to phonemes, but is also obtained for more surface-

like phonetic categories. Cutler et al. (2008) replicated Norris, McQueen, & Cutler's (2003) findings on the [f∼s] continuum, this time based not on lexical knowledge (participants trained on words like "witlof"), but based on phonotactics. In their study, British-English listeners trained on "[?]rul" learned that the ambiguous sound must have been /f/ (*[sr] being an illicit word onset in English), whereas a second group of British-English listeners trained on "[?]nud" learned /s/ (*[fn] being an illegal English word onset). In a later categorization task, /f/-trained participants gave more /f/ responses, and /s/-trained participants gave more /s/ responses along an [f∼s] continuum. These results show that perceptual learning does not necessarily rely on individual words inside the lexicon: phonological knowledge of static lexical patterns also suffices to trigger the process.

For misperception-based sound change to take place, the mechanism for perceptual learning needs to be impaired somehow. An obvious candidate is the amount of exposure. For instance, Maye, Aslin, & Tanenhaus (2008) have shown that listeners can adapt to entire vowel shifts (all vowels lowered by one degree, so "wicked witch" becomes "weckud wetch"), but these participants received twenty minutes of consistent exposure in a laboratory setting. On the other hand, results by Witteman et al. (2015) show that participants can adapt after as little as 3.5 minutes, and that such adaptation is even long-lasting (see also Gaskell & Dumay 2003, who suggest a critical role for sleep in such long-term accommodation). If adaptation to different speakers and their sound systems is this rapid, amount of exposure might not be a viable contender for bypassing perceptual learning. Witteman et al. (2015) suggest one possible failure mode, namely that their participants were unable to adapt to realizations that crossed phoneme boundaries (which probably has a neurolinguistic correlate in the P600; Chapter 6), but since not all sound changes are phonemic mergers, this cannot be a general explanation. Results by Witteman, Weber, & McQueen (2014) implicate the consonant-vowel asymmetry as contributing to sound change, having found that adaptation to vowels (as in Maye, Aslin, & Tanenhaus 2008) is easier than adaptation to consonants, but again, this applies only to a portion of all sound changes.

Neurolinguistic research, mostly centered around the mismatch negativity ("MMN") ERP component, provides some more perspective. In a study of long-distance coarticulation as a possible source for sound change actuation, Grosvald & Corina (2012) found that the brain was sensitive to long-distance coarticulation. That is, a vowel colored by coarticulation from a vowel that was one or two (but not more) syllables away elicited a significant MMN in an oddball task, showing that the brain was capable of detecting the phonetic difference from a stream of non-coarticulated vowels. Other work has shown that the MMN is not necessarily acoustic (and Grosvald & Corina indeed argue that theirs is not), but can also reflect phonological knowledge. Four publi-

cations by the same researchers on the same data (Jacobsen 2015, Steinberg, Truckenbrodt, & Jacobsen 2010a, 2010b, 2011) report significant MMNs to the difference between a correct allophone (the German realization [ɛç]) and an incorrect allophone (the German realization *[ɛx]). These results show that the brain does not compensate for all types of variation, even at the more abstract level of phonology. This ties in with results from the field of regional-and-foreign-acccent processing, which report cumulative interference effects in reaction times (Floccia et al. 2009, Floccia et al. 2006) and the N400 (Goslin, Duffy, & Floccia 2012). This suggests that while a human being in a conversation is very adept at compensating for variation, "under the hood" there are problems that the brain needs to actively resolve. It is highly probable that the degree to which this succeeds is subject to a significant degree of individual variation.

As detailed in the introduction to Chapter 4, various studies of adaptation to phonetic differences across 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) have found that while some individuals adopt such differences (which include sound change) with relative ease, others do not. Similarly, psycholinguists have warned for a very long time that analyses of psycho- and neurolinguistic data need to properly take into account variation between individual participants, due to obvious variation in psychophysiological makeup leading to equally obvious variation in measures such as response latencies in RT experiments. Psycholinguists have additionally realized that the incorporation of merely participants as a random factor in statistical models of language processing is not enough; language items are equally random, leading Clark (1973) to formulate his "language-as-a-fixed-effect fallacy". While the remedy—the mixed-effects model—had already been formulated by statisticians as far back as 1950 (Henderson), it was only through the effort of authors like Baayen, Davidson, & Bates (2008) that mixed-effects models finally became popular in (psycho-) linguistics. Section 1.3.2 reflects on these and other methodological innovations that made the research in this dissertation possible.

1.3.2 Methodological innovations for psycholinguists

Since Barr et al. (2013), psycholinguists have scrambled to incorporate differences between participants and items into their statistical models to the absolute fullest extent, citing Barr et al.'s (2013) slogan to "keep it maximal". It is

nowadays believed that the advice by Barr et al. (2013) to unconditionally fit the maximal random-effects model was somewhat overzealous. Beyond the computational expense to fitting models with random slopes up to the error term, the resulting models often converge to a solution that fails the KKT criteria (Karush 1939, Kuhn & Tucker 1951) or converges to a boundary solution for which these criteria do not even apply. The former failure mode is reported by most statistical software as "failure to converge", whereas boundary solutions are currently only reported by recent versions of R (R Core Team 2020) package `lme4` (Bates et al. 2015b), which reports them as "singular fit"s. The cause for either failure is the same: the model contains too many, probably (multi)collinear, unknown terms, for which no (statistically felicitous) global optimum can be found. Bates et al. (2015a) argue that these problems with such overparameterized models make them unsuitable for routine use, and in some cases even lead researchers to incorrect conclusions. Even if the maximal model can be made to "work", that is, fit nonsingularly without convergence warnings, the inverse relationship between the number of free parameters and statistical power means that such models are costly not just in terms of CPU time, but also in terms of the number of participants and items necessary to be able to detect a true effect (Judd, Westfall, & Kenny 2017, Matuschek et al. 2017).

Bates et al. (2015a) propose to tackle these issues by starting with an infelicitous (nonconverged or singular) maximal model, and then manually identifying the extraneous random effects. They have developed an R function (now incorporated into package `lme4`) called `rePCA` which assists in this process. However, this is a laborious and not necessarily straightforward task, as the θ parameters on which `rePCA` operates do not always correspond directly to individual terms in the researcher's design matrix. For example, one offending θ parameter may correspond to the correlation of two levels of two different categorical variables with many levels—should the researcher then drop all correlations between all of these combinations, or find some way to convince the software to hold only this problematic parameter at a value of zero? Even software that allows more flexible covariance structures than `lme4`, such as R package `glmmTMB` (Brooks et al. 2017), cannot easily accommodate such a request.

An alternative approach is suggested by Matuschek et al. (2017). They propose to use backward stepwise elimination, a well-established technique in the field of psycholinguistics, to identify which of the terms included in a maximal model are truly required. As the backbone of this technique is a simple likelihood-ratio test, this approach to random-effects selection is principally motivated and straightforward to use. The only requirement that may be difficult to meet is a converged maximal model from which to start backward elimination. This procedure thus finds the balance between Type I error rate and

power that Barr et al.'s (2013) approach lacks, but still requires a feasible maximal model from which to start. One way to define a feasible maximal model is as the model that includes the *most important* random-effect terms that can still converge. To find such a model, this dissertation relies on R package `buildmer` (Voeten 2020). This is an R package that builds up a mixed-effects model by starting with only the fixed effects, and adding random-effect terms one by one as long as the model is still able to converge. Random effects are added to the model in order of their contribution to the likelihood-ratio-test statistic or an information criterion, such that when the model eventually fails to converge, the most important random effects have made it in. From this maximal *feasible* model, backward elimination is then used to identify which of the included random and fixed effects significantly improve the model fit.

Beyond the mixed-effects model for linear regression (which includes ANOVA) and generalized linear regression, linguists have added another methodological notch onto their toolbelt in the past few years. The generalized additive model, known since Hastie & Tibshirani (1987) but popularized in linguistics by recent papers such as Baayen et al. (2017), makes it possible to perform regression analysis with predictors that are not linearly related to the response variable, but have effects that take arbitrary forms. Chapter 2 of this dissertation uses this type of model to deal with a longstanding and particularly vexing problem in phonetics, namely the problem of segmenting VC sequences where the consonant is very vowel-like. Dutch coda /l/, realized as [ɬ] in the Netherlands and also beginning to vocalize there (van Reenen & Jongkind 2000), is an example of such a problem-creating consonant. The transitions between VC sequences like [eːɬ] are smooth and continuous rather than discrete, and hence the concept of an *a priori* acoustic segmentation simply does not apply. However, it is nonetheless perfectly possible for a phonetician to formulate hypotheses about the temporal dynamics of vowels followed by coda /l/—in fact, Chapter 2 will demonstrate that coda /l/ indeed plays a major role in the Polder shift. Using the second formant as an example, one such hypothesis could be that the F2 will remain relatively high throughout the course of the vowel and fall as the articulation transitions into the [ɬ]. The modern implementation of the generalized additive mixed model (henceforth "GAMM") in R package `mgcv` (Wood 2017) makes it possible to model this nonlinear trajectory, including random effects, without additional methodological cost by the experimenter beyond a powerful computer. By using GAMMs, Chapter 2 does not require explicit segmentations of these highly gradient [Vɬ] transitions, but simply models the entire VC trajectory as a smooth nonlinear function of time. This makes it possible to compare hard-to-segment [Vɬ] sequences to unproblematic sequences of the same vowels followed by a nonapproximant consonant, dispensing with manual segmentation of the former but not the latter.

Other situations where linear mixed-effects models fare poorly are cases involving many categorical predictors. If multiple many-leveled categorical predictors interact to produce meaningful differences, a regression model ends up becoming very complex to interpret due to including all combinations of all factor levels of the highest interaction and all lower-order terms. In these situations, regression *trees* (see Tagliamonte & Baayen 2012, who discuss the closely-related conditional-inference trees) are more appropriate. These models operate on the basis of recursive partitioning, which results in very intuitive tree diagrams of the relative importance and effects of each variable given the variables that were of higher importance. R package `glmertree` (Fokkema et al. 2018) extends the basic principle of the regression tree by making it possible to incorporate random effects, using a very simple quasi-likelihood algorithm that iterates between building the tree given the random effects and estimating the random effects given the built tree until convergence. Chapter 3 relies on this technique to quantify the degree to which individuals adopt the Polder shift in their perception over a period of nine months.

The mixed-effects model has more uses than controlling for differences between participants and items, in which case the random effects are just nuisance terms. It is also possible for these random effects to be of interest in and of themselves. Psycho- and neurolinguistics have recently begun to realize the potential these models have of offering insight into individual differences (Eekhof et al. submitted, Kliegl et al. 2011, Mak & Willems 2019); the same is true of sociolinguistics (Drager & Hay 2012, Tamminga to appear). Chapter 4 uses the by-participants predicted random effects to classify individuals who have been exposed to the Polder shift for varying numbers of years as "adapted" or "non-adapted". It will be shown that the individual-level differences provide a more nuanced view than the aggregate group differences.

The aforementioned statistical techniques all rely on the existence of a null hypothesis that an effect to be tested is equal to zero, a philosophy known as "null-hypothesis significance testing" ("NHST"). A p-value $<.05$ means that the probability of observing the measured outcome variable y, given that this null hypothesis is true, is smaller than 5%. Therefore, either the null hypothesis is false, or the data are improbably unrepresentative. It is strange to think about statistical models in this way, making inferences about the value of a parameter β by arguing that the probability of the *data*, given the parameter being zero, is very low, i.e. $p(y|\beta = 0) < .05$. What we really want to know is the probability of the *parameter* having a certain value taking the data as given, i.e. $p(\beta = \hat{\beta}|y)$. This is the difference between frequentist statistics and Bayesian statistics. The former is an incoherent hybrid of the original views of Fisher (1955, 1956) and Neyman & Pearson (Neyman 1950, 1957; see Gigerenzer 2004 for details), while the latter is philosophically more sensible, but not the standard scientific practice (see Kruschke 2010a, 2010b for commentary).

In practice, null-hypothesis significance testing is a useful tool to have available for testing point hypotheses about model parameters, particularly because the frequentist maximum-likelihood estimates can be computed efficiently, which is not true for the Bayesian maximum-*a-posteriori* estimates except in special cases. It is also safe to use as long as one is aware of the pitfalls, the most important of which are that a p-value $<.05$ does not constitute absolute, categorical, evidence that the alternative hypothesis can only be true, and that a p-value $>.05$ does not constitute *any* kind of evidence that the null hypothesis is true. If hypotheses are framed within these constraints, it is unlikely that a Bayesian analysis would result in a different substantive conclusion than a frequentist analysis, unless the data were prepared to be excessively pathological. For most of the studies reported in this dissertation, these caveats are acceptable, except for the EEG experiment reported in Chapter 5. This chapter investigates the mismatch-negativity ERP, henceforth "MMN". The MMN is an automatic brain response that is generated when the brain detects a change in sensory stimulation—in the case of this dissertation, when a syllable like [ei] is replaced by one like [e:]. This ERP is almost always asymmetric (Lahiri & Reetz 2010), which means that a switch such as [ei]→[e:] will generate an MMN, but the reverse switch [e:]→[ei] will not. The presence of an MMN can be argued using NHST, but its absence cannot; thus a Bayesian approach is needed. Chapter 5 takes the approach by Wagenmakers (2007), in which Bayes factors are computed based on the difference in BIC (Schwarz 1978) between two candidate models. If these model comparisons are set up such that a full model is compared to a model in which a single focal term has been removed, the corresponding Bayes factor quantifies the odds of that term being equal to zero. Given the *a priori* assumption that both models are equally likely, this Bayes factor makes it possible to quantify the evidence against this assumption (similar to the NHST p-value) as well as *in favor of* this assumption (thus providing evidence that the models are indeed equivalent, i.e. that the effect being tested *is* zero).

All of the aforementioned statistical methods assume that the researcher knows what they are looking for. For example, when analyzing the data of an EEG experiment, the researcher needs to specify *a priori* what combination of electrodes is of interest, and at which moments in time. This information is not available beforehand when the research is exploratory. In this situation, permutation testing can be used (Maris & Oostenveld 2007), and this is done in Chapter 6 of this dissertation. This chapter reports an exploratory investigation of EEG differences that are related to the Polder shift. The chapter will reveal a phonological P600 modulated by factors unrelated to the Polder shift, of which the precise nature is not yet fully known. The permutation tests made it possible to identify a window of statistically-significant differences that corresponded precisely to a P600, which, combined with the correct direction for the

observed difference, corroborate the sparse scientific literature on the phonological P600, and made it possible to take first steps to further qualifying the conditions under which this effect can be obtained. The analysis used in Chapter 6 led to the development of R package `permutes` (Voeten 2019c), the code of which has also seen use outside of the Polder shift by authors such as Ruijgrok (2018).

1.4 This dissertation

The literature discussed thus far paints the picture of a field that is internally divided. On the one hand, we have historical phonologists, who claim that production and perception are especially prone to sound change. In this view, perception is fallible and may result in perceptual reanalyses, and production favors articulations that are more familiar or in other ways "easier", and may therefore lead to speaker-driven sound change. On the other hand, psycholinguistic evidence has shown that both the production and the perception apparatus are extremely flexible, and can cope with incidental variation without any problems. The overarching question of this dissertation is: **what factors influence the adoption of sound change?**

The currently-ongoing Dutch Polder shift provides an opportunity to investigate this major question. There are three properties that make this vowel shift particularly well-suited for this purpose. First and foremost, the Polder shift is currently on-going, which means that, for once, we are *not* too late (cf. Pinget 2015). Secondly, the Polder changes are phonologically conditioned (Voeten 2015). This makes it possible to disentangle phonetic changes from phonological changes, and thus provides a unique test case for the claims from Section 1.2 that adults are not able to adopt phonological changes, but *can* simulate them using more superficial accommodation rules. Finally, the Polder shift is geographically stratified, such that both conservative and innovative individuals are available for psycholinguistic experiments. This brings us to the first research question. For practical reasons of needing to select participants suitable for a psycholinguistic investigation of the on-going Polder shift, a clear picture of the current state of affairs of the specific sound changes subsumed under the Polder shift is necessary. The first research question in this dissertation can therefore only be: **what is the synchronic diatopic diffusion of the sound changes involved in the Polder shift?** This is discussed in Chapter 2.

Based on the findings in Chapter 2, the most suitable participants for further empirical investigation of the adoption of the Polder shift were found to be *sociolinguistic migrants*, specifically speakers of Flemish Dutch who have migrated to the Netherlands, and have hence come into contact with the Polder shift. In order to generally investigate the adoption problem and specifically

test the claims of the generative view on sound change outlined in Section 1.2, two related research questions present themselves. This brings us to the second research question—(**how**) **do sociolinguistic migrants adopt the Polder shift?**—as well as the third—**which individuals, after how much time, are more likely to adopt the Polder shift?** These two questions are superficially similar, but call for different methodological approaches. The former question calls for a *longitudinal* investigation at the *group* level. By contrast, the the latter question is more advantageously defined by a *cross-sectional* study at the *individual* level. Naturally, this dissertation takes these considerations into account. Chapter 3 answers the former question using a small-scale group-level investigation, whereas Chapter 4 deals with the latter issue using a large-scale individual-level study.

It was established in Section 1.3 that phonological variation is handled by specific psycho- and neurolinguistic mechanisms. Two specific correlates of phonological variation were identified in ERPs: the MMN and the P600. These ERPs were shown to be sensitive specifically to phonological-rule violations, and hence may inform us about individuals' adoption of the Polder shift. This brings us to the fourth and final research question: (**how**) **is the adoption of the Polder shift reflected in ERPs?** Chapter 5 focuses on the MMN component, using the specific changes in the phonological rules involved in the Polder shift, whereas Chapter 6 focuses on the P600 component and ends with a more general claim about the types of variation in which the P600 is involved.