Voicing in transition

Laryngeal characteristics in West-Germanic and Italo-Romance dialects

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: Buttered Cat Paradox, by Alla Krikun

ISBN: 978-94-6093-305-9
NUR: 616

Voicing in transition

Laryngeal characteristics in

West-Germanic and Italo-Romance dialects

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
maandag 10 december 2018
om 14.30 uur precies

door

Nina Yasmin Ouddeken

geboren 10 augustus 1990
te Leiderdorp

Promotoren:	Prof. dr. M. van Oostendorp
	Prof. dr. R.A.G. D'Alessandro (Universiteit Utrecht)
Manuscriptcommissie:	Prof. dr. J.P.M. Fikkert
	Dr. B. Alber (Università di Verona, Italië)
	Prof. dr. J. Grijzenhout (Universiteit Leiden)
	Dr. S.R. Hamann (Universiteit van Amsterdam)
	Prof. dr. R.W.N.M. van Hout

For Roy
For my parents

Contents

Acknowledgements

Writing this dissertation has been an adventure, and many people have been so very important in the process of completing it.

First, I thank my supervisors, Marc van Oostendorp and Roberta D'Alessandro. Thank you for making this great adventure happen, thank you for your comments and advice, and your support and guidance. I have learned so much from you.

Thank you to the members of the committee, Paula Fikkert, Birgit Alber, Janet Grijzenhout, Silke Hamann and Roeland van Hout, for reading and approving my manuscript.

I would also like to thank my colleagues at the Meertens Instituut. Hans, Sjef, Gertjan, Leonie, Ben and Frans: thank you for your comments and advice. Varun, Mirella, Erik, Etske, Edoardo and Lotte deserve a very special mention. You guys have been amazing colleagues, and made working on this dissertation so much fun!

A huge thank you goes to everyone who has helped me in my data collection. Piet and Anke van Reenen: thank you for making the data so much more easily accessible and for your enthusiasm. Thanks to the Forschungszentrum Deutscher Sprachatlas and Regionalsprache.de, for allowing me to use their data for this research. I also thank Giovanna Marotta, for taking the time to assist me during my stay in Italy. To all my informants in Italy, and to everyone who has helped me in my search for them: thank you. Without your help and time, this dissertation would not have been what it is now.

Thank you to my new colleagues at the NRO. In particular Jelle, Ruud, Liesbet and Renée: thank you for allowing me to finish this project, and helping me start a whole new adventure at the same time. Your support really means a lot to me.

Finally, I am very grateful to my friends and family. Mum and dad, thank you for always believing in me and for teaching me to be true to myself. Thank you for that tiny bit of stubbornness I get from both of you, and that helped me finishing this project. Timo and Job, thank you for always knowing how to make me laugh. Pepé and Memé, thank you for being the most amazing grandparents anyone could want.

Johan, Mike, Wilma and Ingrid, thank you for being so welcoming, and always making me feel at home.

Meliantha, Hilde, Benjamin and Saskia: thank you for the tea. Malou, Tamara, Laurie and Sonja: thank you for having been a part of my life for so long. Also, thank you for taking my mind of my dissertation when I needed to.

Last, but certainly not least, I want to thank my fiancé Roy. Thank you for being there for me, and for making me smile when I needed to. I could not have done this without you.

CHAPTER 1

Introduction

1.1 Language variation

Language variation has been one of the central topics of study in linguistics for many years. Quite trivially, the field of linguistics emerged partly from the fact that languages differ from each other. Generative linguistics (a term first used by Chomsky (1957)) assumes that even though all languages are different on the surface, they are all guided by the same linguistic faculty that generates the languages' phonology, morphology, syntax etc. If one assumes that all language speakers make use of the same linguistic faculty, even though on the surface their languages may be very different, the study of language variation can (and should) be used when making assumptions on the nature of the language faculty: in general, it should be able to produce any sentence that is grammatical in any language. In other words, if one assumes all speakers make use of the same language faculty, one must also consider every human language when making assumptions on this faculty. Ever since Chomsky, the field of linguistics not only had to describe the differences between languages, but also had to be able to account for these differences.

Over the last few decades variation has become a more central topic of linguistics. Variation is not only studied between different languages, but also within individual languages or even in the speech of individual speakers. At a theoretical level, the study of variation can be divided into three different subfields. First, variation is studied at the level of theoretical linguistics: how can a formal theory of linguistics account for the presence of variation? Grammars must be formulated so that all different structures found in the languages of the world can be accounted for, while at the same time it must be restrictive in some way, as we know that the presence of variation in language does not simply mean that anything goes: there are limits on variation. Individual languages and speakers do not show unlimited amounts of variation, but instead show a lot of consistency in their linguistic output. In other words, theoretical linguistics searches for one formal framework that can produce all possible human languages, that does not create an impossible human language (i.e., it should not overgenerate) and that does not exclude any possible human language (i.e., it should not undergenerate). This can be rephrased as a question that plays an important role in the study of language variation: what are the limits on variation?

Second, variation is studied at a social level. Here, the focus lies on differences between speakers that are explained by extra-linguistic factors, such as age, gender or social class. The study of sociolinguistics led to the insight that a speaker's language may differ depending on social context: in more formal settings a speaker's language is likely to be different from that same speaker's language in more informal settings. These various insights led to the realisation that one speaker cannot have access to only one, invariable, grammar. Instead, speakers must have access to multiple grammars, or to one grammar that allows for variation.

Finally, variation can be studied at a geographical level, i.e. dialectology. It studies geographical patterns in variation, and tries to group different linguistic varieties together in so-called dialect families. Although dialectology is mentioned here as a separate branch in the field of language variation, it differs slightly from the theoretical and social focus points: it is very common for dialectologists to use geographical variation to answer questions regarding either theoretical or social variation.

As variation has become a more important question in linguistics, more questions regarding the study of variation have been identified. With respect to phonetics and phonology, two different fields of study can be distinguished. The first is related to the modelling of phonetic variation (and is thus concerned with the interface between phonetics and phonology), while the second is related to the modelling of phonological variation.

The interface between phonetics and phonology has become a popular field of study in the last three decades (Keating (1988), Ohala (1990), Blumstein (1991), Kingston (2007), Romero and Riera (2015)). Although it is clear that the relationship between phonetics and phonology must be specified, and although the interface is well-studied, the different approaches put forward in the literature could not be more different from each other. On one end of the spectrum, there is the proposal that phonetics and phonology should be strictly separated (substance-free phonology; e.g. Hale and Reiss (2000), Hale and Reiss (2008), Reiss (2017)). In this approach, the phonetic realisation of an (innate) segment is regulated via "complex transduction processes" (Reiss (2017, p. 2)), and the phonetic realisation of a feature does not have any influence on phonological processes. On the other end of the spectrum, there is the proposal that phonetics and phonology should not be seen as two different disciplines, but rather as one (e.g. Ohala (1990), Exemplar Theory (e.g. Pierrehumbert (2001))). Many different, intermediate approaches exist as well. An early example is Keating (1984), who proposes that the phonetic realisation of a phoneme is guided by phonetic implementation rules. Every phoneme is associated with a phonetic category (e.g., in the case of consonant voicing consonants are phonologically [+voice] or [−voice], but [+voice] can belong to the phonetic category {voiced} or to the phonetic category {voiceless unaspirated}, while [−voice] can belong to the phonetic category {voiceless unaspirated} or {voiceless aspirated}). The precise phonetic realisation of a phonetic category is guided by the implementation rules. Both the phonetic category of a phonological feature, and the precise phonetic realisation of a phonetic category are language-specific. A similar approach is found in BiPhon (Boersma (2009), Hamann (2009), Boersma (2011), Hamann (2011), Hamann (2014)). In this approach, the phonetic realisation of a phoneme is guided by Optimality Theory (OT) constraints, the ranking of which has to be acquired by the language-learning child. Chapter 5.4 will discuss this approach in more detail.

Besides approaches where the interface is specified as a set of rules or constraints, several frameworks exist that propose a bigger interplay between phonetics and phonology. A first example is Articulatory Phonology (Browman and Goldstein (1991), Browman and Goldstein (1992), Goldstein and Fowler (2003)), which has as its main tenet that phonemes are represented by articulatory gestures. These are movements made by the different articulators (upper lips, lower lips, jaw, tongue tip, tongue body, velum and glottis). Each articulator can make different gestures. For example, the lips can be 'protruded' or 'spread', the velum can be 'open' or 'closed', etc. Different phonemes are a combination of different simultaneous gestures of the articulators. This "means that gestures are basic units of contrast among lexical items as well as units of articulatory action" (Browman and Goldstein (1992, p. 23)). Changing the timing of individual gestures may cause phonetic or phonological change. For the articulation of a nasal consonant, for example, the velum has to be opened. During the articulation of a preceding vowel the opening of the velum is anticipated, so that part of the vowel will be realised with an open velum and will sound slightly nasalised. When the timing of the velum opening is moved even more forward, the vowel will be fully nasal. This can be compared to a phonological process of nasal assimilation. This example shows the interplay between phonetics and phonology in Articulatory Phonology.

A second example is Evolutionary Phonology (Blevins (2004), Blevins (2006)), a model that seeks to explain the difference between common and uncommon sound patterns by looking at patterns of sound change. Some patterns are more likely to be caused by sound change (e.g. final devoicing of obstruents) while other patterns are very unlikely to be caused by sound change (e.g. final voicing of obstruents). Frequency patterns in phonology are thus related to patterns in sound change. The approach differs from traditional approaches to phonology (such as generative phonology and OT) in the sense that the traditional approaches use phonology to explain phonology (e.g. final devoicing occurs frequently because a final voiced obstruent causes a marked phonological structure). In Evolutionary Phonology, the burden of explaining phonological patterns is put on phonetics more than phonology: only when phonetics cannot explain phonological patterns is phonology invoked as an explanation.

Another important question in phonological theory is how the presence of variation should be modelled. Several processes have been identified for which speakers show variation with respect to the application of a phonological process. That is, one speaker might sometimes choose to implement the phonological process, while at other times the same speaker might choose not to implement this process. Whether or not the speaker applies the process does not necessarily depend on the linguistic environment; it may depend on the speech register the speaker uses at that moment but it can also be purely a matter of chance. However, whether the speaker's choice depends on the speech register or on chance (i.e., when the output cannot be predicted by the linguistic context), the speaker must have access to both the form where the process has applied and the form where the process has not applied. Several approaches to modelling the presence of variation in the grammar of a speaker exist. Roughly, they can be divided into approaches assuming the speaker has access to multiple (different) grammars and approaches assuming the variation must be modelled within one grammar. While the different approaches will be discussed in more detail in Chapter 5, a rough explanation of the differences between the approaches will be discussed in this section.

An early approach to modelling variability is couched in a rule-based framework. In this framework, all phonological processes are the result of the application of a (set of) rules. Variation in phonology must thus be the result of a rule that applies variably: if we consider variation that is conditioned by extra-linguistic rather than intra-linguistic factors, the variation cannot be built into the rule. In Rule-based phonology (RBP), variable application is formalised by assigning every rule an application probability: a rule that always applies, regardless of the extra-linguistic contexts, is assigned a probability of 100%, while a rule that applies only half of the time is assigned a probability of 50%. In the latter scenario, the rule is applied in 50% of the contexts in which it could theoretically apply, while the rule is not applied in the other 50% of the contexts in which it could theoretically apply.

In OT, several strategies can be invoked to account for variation in the phonology. In OT four different components can be identified: the input, the constraints, GEN (which generates an infinite number of output candidates based on the input), and EVAL (which evaluates all output candidates based on the constraint ranking). OT does not place any restrictions on the input, and GEN generates an infinite number of possible outputs. Consequently, variation cannot be modelled there, but must be modelled in the constraint ranking or EVAL. In the first scenario, one can assume that individual constraints are floating (either throughout the entire grammar or within a smaller range of constraints), or unranked with respect to each other. A floating constraint is an individual constraint that is not assigned to a specific position in the ranking. Instead, it 'floats' over all the other constraints, which are assigned a fixed position in the ranking. Only when the speaker produces an utterance does the floating constraint receive a fixed position in the ranking. However, this position may differ for every individual utterance. This is how phonological variation is accounted for: in the speaker's final ranking floating constraint A may be ranked over constraint B (which is fixed with respect to all other constraints except A) or below it.

If constraints are unranked, the position of one constraint can be switched relative to the other. If only two constraints are involved (e.g. constraint A is unranked with respect to constraint B, or vice versa), two grammars are possible: A >> B and B >> A. This is identical to a situation where floating constraint A ends up in a position between constraint B and the constraint directly above B, or between B and the constraint directly below B. If more constraints are involved (e.g. A, B and C), the grammar gets more complicated. If all constraints are unranked with respect to each other, six different grammars are possible: A >> B >> C, A >> C >> B, B >> A >> C, B >> C >> A, C >> A >> B, C >> B >> A. If, on the other hand, only constraint A is unranked with respect to constraints B and C (which are ranked with respect to each other), only three grammars are possible: A >> B >> C, B >> A >> C, B >> C >> A. This situation is identical to a situation where constraint A is floating over constraints B and C. In both this approach and the approach with floating constraints, the chances of encountering one particular output in the speaker's speech are supposed to be correlated to the predictions made by the grammar: if two out of six grammars predict output X_1, two predict output X_2, one predicts X_3 and one X_4, outputs X_1 and X_2 are each expected to surface in 33% of the cases while outputs X_3 and X_4 are expected to surface in 16.7% of the cases.

A slightly different approach, Stochastic OT (Boersma (1998), Boersma (1999)), does not assume an absolute constraint ranking but rather a relative ranking of the constraints. This means that the distance between different constraints may vary: constraints C_1 and C_2 may be ranked relatively close to each other while constraints C_2 and C_3 are ranked relatively far apart. Boersma (1998), Boersma (1999) models this by assigning each constraint a ranking value. The higher the ranking value, the higher the position of the constraint in the ranking. Variation is modelled by adding noise to the ranking: constraints with a ranking value relatively close to each other may be reversed as a result of the noise in the system, while constraints ranked relatively far apart will probably never switch positions in the ranking.

Instead of modelling variation in the constraint ranking, it can also be modelled in EVAL. In the standard approach, EVAL only selects the best (most optimal) output candidate. In other words, there is only one winner. Instead, one could assume that EVAL selects a larger number of candidates as possible outputs, so that there are more possible winners. It is then up to the speaker to select one of those winners as the eventual output.

A final, framework-independent approach assumes that variability is not modelled within a speaker's grammar, but instead speakers have access to multiple, slightly different grammars.

In this study, the question of language variation will be addressed once more. While this study certainly will not provide an answer to all questions, it will help to shed some light on the topic and yield new insights. The different focus areas of language variation that have been discussed above will all play a role in this study, although some fields will play a bigger role than others. There will be a large focus on regional variation, (data for) the phonetics-phonology interface and variation in theoretical linguistics. While social variation and the modelling of variation will also be discussed, they will play a smaller part in this study than the other fields. Further, while regional variation plays a large role here, it is not the object of study itself, but instead it is used as a tool for studying variation at the different levels mentioned.

If one wants to study language variation, from whichever perspective, a necessary condition is of course that the linguistic variety studied shows a certain amount of variation. A situation that may very likely lead to the presence of variation is language contact: if speakers of two or more different varieties are in contact with each other, their languages are likely to be influenced by each other. In the present study, language variation will therefore be studied from a geographical perspective. One phenomenon that has been described in dialectology is a very promising starting point for this type of research: Chambers and Trudgill (1980) have described the existence of so-called transition zones, areas where a geographically gradual change between linguistic varieties can be found. The gradual change of one linguistic variety into another is bound to lead to linguistic variation. In this study, transition zones will thus be used as a tool for studying variation. Two different language families, the Germanic and Romance language family, will be included in the study: the language families are different enough to provide interesting research topics, and for both language families an abundance of data is already accessible so that interesting research topics can be identified relatively easily.

The remainder of this chapter is devoted to a short overview of the dialectological literature and the presentation of the research topic. In Sections 1.2 and 1.3, the existing literature on dialectology, dialect geography and transition zones will be presented, in order to get a better understanding of the phenomenon and to possibly identify several interesting transition zones. The precise research topic will be presented in Section 1.4.

1.2 Maps and Grammar

In the 19th century, linguists started to use maps in their studies of dialects. Early examples do not show individual linguistic features on the map, but are overview maps of where different languages or dialects are spoken. The first language map is argued to be the map by Ten Kate (1723), showing the 'Volk- en tael-verspreiding over Europa' (the spread of people and languages across Europe). Other early works include a map of the Bavarian dialect area by Schmeller (1821) and the collection of dialect phrases by Georg Wenker (the Wenkerbogen from 1876 and later, which form the basis of the Deutscher Sprachatlas (Wrede, Mitzka and Martin (n.d.))). The first example of dialect geography for the Dutch language area is the map by Jellinghaus (1892), which is based on the dialect translations of the parable of 'de Verloren Zoon' (the Lost Son) by Winkler (1874), as well as on several collections of dialect words. Other examples of dialect maps of the Dutch language area are Te Winkel (1901) and Van Ginneken (1913).[1]

In 1925, Blancquaert and Pée (in cooperation with numerous other dialectologists) started the Reeks Nederlandse Dialectatlassen (RND): the series of Dutch dialect atlases. The series consists of 16 volumes, each describing a different part of the Dutch language area,[2] and was completed only in 1982. Instead of indicating which dialect is spoken in which area, the atlases contain detailed information about the phonetics of 1956 dialects spoken in the Dutch language area. For each location several dialect speakers were asked to translate a questionnaire,[3] containing 141 test items (139 prespecified items and two additional questions asking informants about local names for surface measures and for local water names), into their local dialect. Recordings of these translations were later transcribed using a phonetic alphabet. The atlases of the RND display the realisations of individual words or groups of words, and give a very detailed view of the different isoglosses present in the Dutch language area.

During the final stages of the RND project, the feeling arose that a new project was necessary, one that would better connect with new theoretical insights from structuralist grammar and sociolinguistics. A new project, the Goeman-Taeldeman-Van Reenen-Project (GTRP), was started, which contains the phonetic realisation of slightly less than 1900 items (individual lexical items as well as short sentences) in 622 dialects. This database forms the basis of two dialect atlases: the Fonologische Atlas van de Nederlandse Dialecten (FAND) and the Morfologische Atlas van de Nederlandse Dialecten (MAND) (Phonological and Morphological Atlas of the Dutch Dialects). The FAND consists of three volumes, published in 1998 (Goossens, Taeldeman and Verleyen (1998)), 2000 (Goossens, Taeldeman and Verleyen (2000)) and 2005 (De Wulf, Goossens and Taeldeman (2005)) respectively; the MAND consists of two volumes published in 2005 (De Schutter, Van den Berg, Goeman and De Jong (2005)) and

[1] See `https://www.meertens.knaw.nl/projecten/mand/CARThistorisch.html` for an overview of the history of linguistic maps.

[2] The volumes discuss, in respective order, the following areas: Klein-Brabant ('small Brabant'); South-East-Flanders; North-East-Flanders and Zeelandic Flanders; Flemish Brabant; the Zeelandic Islands; West Flanders and French Flanders; Antwerpen; Belgian Limburg and South-Netherlandic-Limburg; North-Brabant; East-North-Brabant, the River Area and North-Netherlandic-Limburg; South-Holland and Utrecht; Gelderland and South-Overijssel; North-Holland; South-Drenthe and North-Overijssel; Fryslân; Groningen and North-Drenthe.

[3] The questionnaire was, where possible, kept constant between the different dialects.

2008 (Goeman et al. (2008)). In the first years of the 21st century, fieldwork for a syntactic dialect atlas was carried out, which resulted in a two-volume Syntactische Atlas van de Nederlanse Dialecten (SAND) (Syntactic Atlas of the Dutch Dialects), published in 2005 (Barbiers, Bennis, De Vogelaer, Devos and Van der Ham (2005)) and 2008 (Barbiers et al. (2008)). The SAND contains dialect data for 267 dialects (158 in the Netherlands, 102 in Flanders and 7 in France).

For the Italian language area, several different atlases have been published. Two of them, the Sprach- und Sachatlas Italiens und der Südschweiz (AIS) and Atlante Linguistico Italiano (ALI), cover the entire country including, in the case of the AIS, the Italian varieties spoken in southern Switzerland. Fieldwork for the AIS was carried out between 1919 and 1925, covering 407 locations in Italy (including Sicily and Sardinia) and southern Switzerland. Between 1928 and 1940 eight different volumes were published including an introductory book and an index (see `http://www3.pd.istc.cnr.it/navigais/navigais-index`). The ALI was published between 1995 and 2008 (Lameli, Kehrein and Rabanus (2011)). Data collection took place in 993 locations in Italy (including Sicily and Sardinia), Slovenia, Croatia and France. The atlas consists of nine different volumes.

Besides the atlases covering the entire country, regional atlases for, a.o., Sicily, Tuscany, Umbria, Molisano, Campania, Puglia and Lucania have been published (Lameli et al. (2011)). As methodological approaches differ between the atlases, and the focus point differs between several atlases as well (some focussing on phonology and others on phonetics), the data are not well comparable.

The plotting of individual dialect features, which shows the location of individual isoglosses instead of borders between different dialects, can reveal interesting links between language and geography. An example of such a pattern is a linguistic border that corresponds to a geographical border (e.g. a mountain range, a large lake, etc.). The Frisian dialects, for instance, are spoken in the province of Fryslân in the north of the Netherlands. In the east, these dialects are situated next to the Groningen dialects, such that the precise location of the border between the two is not clear. In the west, however, these dialects are separated from the North-Hollandic dialects by the IJsselmeer. This lake functions as the border between Frisian and Hollandic dialects.

Besides these relatively obvious geographic patterns in language, more intricate patterns can be found. Personal pronouns in Dutch, for example, can surface either as a phonologically strong (unreduced) form, or as a phonologically weak (reduced) form, which is usually encliticised and usually has a reduced vowel (Noske (2005), Van Oostendorp (2012)). The personal pronoun *ik* 'I', for example, can be realised as either /ɪk/ or as /k/. Reduction cannot apply equally in different contexts; the chances of applying depend on the linguistic environment of the pronoun. When the pronoun appears before an auxiliary, chances of reduction applying are bigger than when the pronoun appears before a lexical verb. In the phrase *ik heb* 'I have', reduction is thus more likely to occur than in the phrase *ik hoor* 'I hear'. Second, when vowel reduction would create a structure with a simplex onset (which it would in the case of *ik heb* and *ik hoor*, as the word-initial /h/ of the verb is deleted as well) it is more likely to occur than when reduction would create an onset cluster consisting of a plosive followed by a fricative (e.g. *ik zal* 'I will' or *ik zeg* 'I say'). However, reduction is still more likely to occur in the latter environment than in an environment where it would lead to an onset cluster with two plosives (e.g. *ik ben* 'I am' or *ik beloof* 'I promise'). This hierarchy is not only visible in individual speakers' behaviour, but is also visible on the map. A small core area exists in which speakers can use the reduced form in all contexts. Around this core area an outer layer is found, in which speakers can use the reduced forms in all contexts except *ik ben* and *ik beloof.* The next layer allows reduction only in *ik heb*, *ik hoor* and *ik zal*; the layer after that only in *ik heb* and *ik hoor*, and the outside layer only allows reduction in *ik heb* (Van Oostendorp (2012)). Van Oostendorp (2012) shows that this geographical inclusion pattern can also be found in individual speakers' judgments: when forced to rank the six forms from best to worst, this ranking corresponds to the pattern found on the map. In *ik heb* reduction is judged to be most grammatical, while *ik hoor* is judged to be slightly less grammatical. The form *ik zal* is judged even more ungrammatical, followed by *ik zeg* and then *ik ben.* Finally, *ik beloof* is judged to be the least grammatical form.

Transition zones are another example of geographical patterns in language: dialects which combine characteristics of two different dialect groups. Technically, the concept appears to be rather clear and self-evident, referring to a continuum of points where the two endpoints (A and B) of the continuum are different in nature and where the change between the two endpoints is gradual; every step (in distance) away from A and closer to B involves a small change (in characteristics) away from A and closer to B. Logically, the items between A and B have characteristics in common with both A and B; but the points closer in distance to A have more characteristics in common with A while the points closer in distance to B have more characteristics in common with B. If one wants to classify all the points in the continuum as belonging to either A or B, this will be relatively easy for the items closest to either one of the groups but very difficult if not impossible for the items in the middle, as they share characteristics with both A and B.

In linguistics, a transitional dialect is a dialect that shares linguistic characteristics with two (or more) other dialects to such an extent that it is impossible to group the transitional dialect with one of the dialects it shares characteristics with. Often (although not necessarily) the transitional dialect is part of a bigger transition zone: an area where a gradual change between two (or more) different linguistic systems can be found (often referred to as a dialect continuum). This means that there is no discrete boundary between the different systems (contrary to, e.g., political boundaries), but rather, when moving from one end of the continuum to the other, the linguistic varieties slowly change from the variety found at one end of the continuum to the variety found at the other end of the continuum. In other words, when moving from one end of the continuum to the other, the number of differences between a location and one of the endpoints of the continuum accumulates (Chambers and Trudgill (1980)). It is thus difficult to classify the dialects in the middle of the continuum as belonging to group A or group B. As the transitional dialects have characteristics in common with more than one other dialect, it is impossible that these areas originate outside a situation of language contact: if there were no language contact, the presence of characteristics shared with two different dialect groups would be completely coincidental.

Chambers and Trudgill (1980) distinguish two different patterns in transition zones: mixed and fudged lects. They discuss the transition between southern and non-southern English varieties, which differ in the realisation of the Middle English short vowel ⟨u⟩. In the south, this vowel changed into [ʌ] in some words (but remained [ʊ] in other words), while in the north the vowel did not undergo any change and remained [ʊ] in all words. The transition zone between these two areas is characterised by several mixed lects on the one hand, and several fudged lects on the other. In the mixed lects, the vowels that occur in the two stable regions, [ʊ] and [ʌ], are both present. In some words, the speaker uses [ʊ], and in other words the speaker uses [ʌ]. The two vowels are thus allophones of the same phoneme, and there is variation in the usage of these vowels: different speakers may have different distributions of the two vowels (Chambers and Trudgill (1980) do not discuss whether speakers themselves can use the two vowels interchangeably in one word or if speakers tend to have a fixed lexical distribution for the vowels). Speakers of a mixed lect thus always have both realisations in their linguistic system. In the case of a fudged lect, speakers do not have both vowels present in the system simultaneously, but rather the two vowels [ʊ] and [ʌ] are phonetically merged into a new vowel [ɤ] (this vowel is the phonetic mid point between [ʊ] and [ʌ]). There is thus no competition between two allophones. The new vowel does not occur in any of the neighbouring dialects, but it does occur in any position where [ʊ] and [ʌ] would occur in the stable areas. A third type of lects distinguished by the authors are scrambled lects, which are a combination of both mixed and fudged lects. They do not appear in the case study on Middle English short ⟨u⟩, but they would be dialects where the speakers would use the entire vowel space between [ʊ] and [ɤ]. In a case study on the contemporary realisation of Middle English short ⟨a⟩, the three different types do occur together, showing that transition zones need not be characterised by only mixed lects, only fudged lects or only scrambled lects, but that the three can occur all together.

These contact situations might create interesting patterns, as speakers receive linguistic input from two different systems. This might create unstable situations, where speakers might be torn between two linguistic variables or where they have to combine input from two incompatible linguistic systems. These areas may thus shed light on linguistic variation and how speakers deal with variation: are there limits on variation and if so, can we define these limits? Further, variation in transition zones may shed light on the question of how variation should be represented in the grammar of a speaker. In Section 5.4, different approaches to this question will be discussed.

An example of a transition zone in the Dutch language area can be found on the famous map by Kloeke (1927). This map shows the realisation of Old Dutch *oe* ([u]) in the words *huis* ('house') and *muis* ('mouse'). In Old Dutch, the realisation of this phoneme started to change toward *uu* ([y]) in most dialects (Van Bree (1987)). However, some dialects were not affected by this change, and up to this day retain the old *oe* pronunciation. In some of the dialects where *oe* changed to *uu*, a further change of the phoneme to *ui* ([œy]) took place "at the boundary between Middle Dutch and New Dutch" (Van Bree (1987)).[4] This is in itself not a remarkable process, but Kloeke's maps show a very unexpected pattern. Contrary to what was stated by the Neogrammarians,[5] Kloeke (1927) showed the existence of dialects where the pronunciation of *oe* in *huis* ('house') had changed to *uu* ([y]), while the pronunciation of *oe* in *muis* ('mouse') was unaffected. The map thus shows stable areas with *oe* in both lexical items, *uu* in both lexical items or *ui* in both lexical items, but it also shows an area where two different phonemes, *oe* and *uu*, are found in two different lexical items. This area can be argued to form a transition between the stable *oe* area and the stable *uu* area. In the next section, an overview of the literature on transition zones in the Dutch-speaking language area will be presented.

Other transition zones have been described in the literature on Dutch dialects besides the area described by Kloeke (1927) and Chambers and Trudgill (1980). In the following sections, different approaches to dialectology, and the different divisions of the Dutch and Italian dialect areas they have distinguished, will be discussed.

1.3 Transition zones

In trying to divide the Dutch and Italian language areas into different dialect groups, several different transition zones have been identified. In this section, the findings from different approaches to dialectology will be discussed.

1.3.1 Classical dialectology

The main focus of classical dialectology has been to classify dialects in bigger dialect groups. Traditionally, the dialectologist selects a set of linguistic characteristics, and bases their classifications on the differences between dialects with respect to these characteristics. In the Dutch speaking language area, Van

[4]Translation of the original Dutch "op de grens van Mnl. [Middelnederlands, NO] naar Nnl. [Nieuwnederlands, NO]".

[5]Sound change always applies regularly: if a phoneme changes its pronunciation in a specific context, every instance of that phoneme occurring in that particular context will be affected.

Ginneken (1913) identifies Frisian, Saxon, Hollandic-Franconian, Brabantic-Franconian and Limburg-Franconian, while Weijnen (1966) identifies eight different dialect areas: Frisian, “German to the east of the Benrath line”, south-eastern dialects, north-eastern dialects, south-western dialects, southern-central dialects, north-western dialects and northern-central dialects. Most dialect classifications of the Dutch speaking language area are similar to these two, identifying a Frisian and Low Saxon area plus several different Franconian areas (Hollandic, Brabantic and Limburgian Franconian are commonly identified regions). Most dialectologists classify dialects that do not clearly belong to one dialect group as ‘transitional’ dialects. In the literature on Dutch dialectology alone, many different dialects have been classified as transitional. In this section, several such transitional dialects will be discussed, followed by a discussion of the problems with this method of classification.

A region that has been classified as transitional by different authors, and that forms a very clear example of a dialect continuum, is the border area between the Franconian and Low-Saxon dialects (Berns (2002, p. 20), Heeringa (2004), Heeringa and Nerbonne (2001), M. Jansen et al. (2011, p. 61) Scholtmeijer (2003, p. 79), Spruit (2008), Wieling, Heeringa and Nerbonne (2007)). Many isoglosses are located in that area:

- the *old*-*oud* isogloss;
- umlaut of old West-Germanic *au*;
- umlaut of old West-Germanic *â*;
- realisation of final -/n/ in nominal and verbal morphemes;
- realisation of old West-Germanic *î* (diphthong or long monophthong);
- presence vs. absence of the verbal plural morpheme -*t*.

This area, located in the Veluwe region, has been called a ‘terraced landscape’ in the work of Daan and Blok (1969), referring to the fact that here one can find an accumulation of linguistic differences: in the south-west of the region one finds mostly linguistic traits that are characteristic of Franconian dialects, but when moving to the north-east one gradually finds more and more linguistic features that are characteristic of Low-Saxon dialects, and fewer and fewer Franconian characteristics (examples of these characteristics are the diphthongisation of Old Dutch long /u/ and long /i/ and the vocalisation of /l/, both of which show a change in the Veluwe region (Scholtmeijer (2011)).

Another region that is mentioned by different authors is the area between Ripuarian and East-Limburgian dialects (Hinskens (1998), Cornips (1998), Hinskens and Taeldeman (2013, p. 131)). This area is part of the Rhenish Fan: a geographical region where a bundle of related isoglosses can be found. These isoglosses all coincide to the north-east of the Rhenish Fan, but approximately 40 kilometres east of the river Rhine the isoglosses start to diverge (Bloomfield (1933)). These isoglosses all relate to the High German Consonant Shift, a phonological process that changed Old High German plosives into affricates

(G. W. Davis (2008), Schrijver (2011)). The isoglosses divide the dialect area as follows:

- dialects in which no consonant has shifted its pronunciation;
- dialects in which word-final /k/ shifted to /x/ (/ɪk/ → /ɪx/) (Ürdingen line);
- dialects in which intervocalic /k/ shifted to /x/ (/makən/ → /maxən/) (Benrath line);
- dialects in which word-final /p/ in a cluster shifted to /f/ (/dɔrp/ → /dɔrf/);
- dialects in which a single word-final /p/ shifted to /f/ (/ɔp/ → /auf/);
- dialects in which word-final /t/ shifted to /s/ (/dɑt/ → /dɑs/);
- dialects in which intervocalic /p/ shifted to /pf/ (/ɑpəl/ → /ɑpfəl/).

These changes accumulate from north-west to south-east, so the dialects in the north-west have undergone none of these changes; a bit more south-eastward the dialects have undergone the first change; more south-eastward the dialects have undergone the first and second change; etc.

The Ripuarian dialects can be found at the south-eastern side of the Benrath line; they are the north-westernmost dialects where a regular application of the High German Consonant Shift can be identified (not necessarily all phonemes (/p, t, k/) underwent the change, but the ones that did, did so in a regular manner) (Hermans (2013)). The East-Limburgian dialects are located to the north-west of the Benrath line; they still display some traces of the Second German Consonant Shift but its application is highly irregular (Hermans (2013)).

The transition zone between the Ripuarian and the East-Limburgian dialects is, according to Hermans (2013) (citing Goossens (1965)), demarcated by the Benrath line in the south-east, and the *sagen*-line in the north-west. This latter isogloss separates [zɛɡə] dialects in the west (East-Limburgian) from [zaːɣə] dialects in the east (the transitional dialects). These dialects are thus distinguished from each other based on their realisation of the word *zeggen* 'to say' (Hermans (2013, p. 337)). However, Hermans does not specify any linguistic characteristics of the region that would make it transitional. Hinskens (1998) mentions that the area is a terraced dialect landscape: from west (East-Limburgian dialects) to east (Ripuarian dialects), the number of dialect features gradually increases. The East-Limburgian dialects, which are the westernmost dialects of the three, have the smallest amount of dialect features; the transitional dialects have all the dialect features that are present in the East-Limburgian dialects plus some more; the Ripuarian dialects, finally, have all the dialect features that are present in the transition zone plus some more. The transition zone is thus transitional in the sense that the number of dialect features present lies between the number of features in East-Limburgian and in Ripuarian.

Other regions that are sometimes proposed in the literature, but without much supporting evidence or data:

- a transitional area between Brabantine and West-Limburgian dialects;
- a transitional area between Brabantine and East-Flemish dialects (by some authors East-Flanders is recognised as a transitional area itself, connecting the West-Flemish and Brabantine dialects);
- the Stellingwerf dialects, which form a transitional area between Frisian and Low Saxon dialects (Bloemhoff, De Haan and Versloot (2013)).

In Italian dialectology, one isogloss bundle famously plays a large role: the La Spezia-Rimini line, which links the city of La Spezia on the west coast to the city of Rimini on the east coast (some refer to this line as the Massa-Senigallia line, arguing that it should link the cities of Massa on the west coast and Senigallia on the east coast), and roughly follows the provincial borders between Emilia-Romagna and Tuscany and between Emilia-Romagna and Marche (Savoia (1997)). The La Spezia-Rimini line is considered to be the border between the northern Italian varieties and the central and southern Italian varieties, and is the southern border for, a.o., the following linguistic characteristics (Savoia (1997), Cravens (2002), Benincà, Parry and Pescarini (2016):

- degemination;
- lenition of intervocalic /p, t, k/ (voicing, spirantisation or even deletion);
- deletion of final unstressed vowels (except /a/; Benincà et al. (2016));
- loss of unstressed vowels.

Savoia (1997) gives a broad overview of linguistic differences present within the Italian language area. In his words, "[a] major typological distinction between the dialects is constituted by the three principal vowel systems": a western, a Sicilian and a Sardinian vowel system. The three systems differ in how they developed from the Latin vowel system, which distinguishes ten different vowels (/i(ː), e(ː), a(ː), o(ː), u(ː)/. A rough division places the Sardinian system, which entirely lost the Latin length distinction, in part of southern Basilicata and northern Calabria; the Sicilian vowel system, which merged /i/, /iː/ and /eː/ into /iː/, /a/ and /aː/ into /aː/ and /u/, /uː/ and /oː/ into /uː/ (while /e/ and /o/ changed their quality into /ɛ/ and /ɔ/), is found on Sicily, in southern Calabria and Salento; the western system, which merged /i/ and /eː/ into /eː/ while /e/ is realised as /ɛ/, and which merged /u/ and /oː/ into /oː/ while /o/ is realised as /ɔ/, is (roughly) found in the rest of the country.

With respect to consonants, a first division can be made by looking at lenition processes (Savoia (1997)). In the north, lenition of intervocalic consonants is abundantly present, whereas south of the La Spezia-Rimini line the process is much less frequent. Several other consonantal processes are present in the language area (e.g. assimilation of postnasal obstruents, palatalisation of [ll] before high vowels, or absence of palatalisation of velars before front vowels), but they do appear to affect different geographical regions (Savoia (1997) links most processes to the Italian provinces).

With respect to morphosyntax, Savoia (1997) argues that the northern varieties form a region separate from the non-northern (central and southern) varieties, the two regions separated by the La Spezia-Rimini line. Some distinguishing features are, a.o., the use of subject clitics and the "lack of agreement between intransitive verb and postposed NP".

An overview of the literature reveals that within the Dutch dialect area numerous transition zones have been identified, while within the Italian dialect area the focus has been more on describing individual processes in smaller regions. Several transitional areas can probably be found, however, around the La Spezia-Rimini line and the Rome-Ancona line. However, the dialects that have been classified as transitional in the literature do not form an ideal starting point for studying language variation. First, it is possible that not all of the dialects classified as transitional are in fact transitions between dialect groups; they might be interpreted as such because they cannot be classified as belonging to one of the larger groups. As dialectologists base their classifications on a subset of isoglosses, there may be small differences between the set of isoglosses used by individual dialectologists. This may in turn cause differences in their classifications, leading to different dialects that have to be classified as transitional. In other words, the set of isoglosses used by dialectologists might affect the classification of individual dialects. A related point is that a precise number of isoglosses that have to bundle together before they can be considered a dialect boundary has never been specified. As a consequence, transitional dialects in classifications based on a smaller number of isoglosses might show less variation than transitional dialects in classifications based on a larger number of isoglosses. This poses a problem for the present study, as the smaller the number of transitional characteristics of a dialect, the less interesting the dialect becomes for the purpose of the present study.

A second problem is that isoglosses vary hugely in their nature. Some isoglosses may refer to the exact phonetics of, for example, the low-mid front unrounded vowel /ɛ/ (for example as [æ] or [ɛ]), whereas other isoglosses may refer to different underlying phonemes in a particular phonological environment (e.g. the presence of Umlaut in several Dutch dialects, which fronts the stem vowel when the stem is followed by a diminutive suffix (Sloos and Van Oostendorp (2010))). The latter isogloss seems to be of more strucural influence than the former, but in dialectology both are treated as equally important. This problem has been perfectly described by Daan and Blok (1969): "The linguist can identify all isoglosses and compare their density and direction using statistical methods. However, in his objectivity he will miss which isoglosses are important, which are unimportant, he does not gain insight in the relevance of the isoglosses as dialect borders".[6]

Finally, the concept of isogloss is in itself problematic, since it implies that a sharp border between two linguistic phenomena can be found. In reality, this is often not the case (cf. Chambers and Trudgill (1980)). It is usually possible

[6] "De taalkundige kan alle isoglossen vaststellen en hun dichtheid en richting volgens statistische methoden vergelijken. Maar in zijn objektiviteit zal het hem ontgaan welke isoglossen belangrijk, welke onbelangrijk zijn, hij krijgt geen inzicht in de relevantie van de isoglossen als dialektscheiding."

to define two core areas, but in the border area one can expect to find a lot of variation, with (for example) both realisations of the relevant phenomenon being present in one location. The sharp border that is implied by an isogloss is in many cases an idealised representation of reality; rather, the boundary must be seen as an area where there is a lot of variation between the two realisations. In conclusion, transitional dialects identified in dialectological literature are not the most promising regions to focus on in the present study. This is not to say that its results have to be ignored: isoglosses do in fact indicate where cumulative differences between linguistic varieties can be found. Classifications based on these isoglosses should, however, be treated with care.

Coming back to the transition zones identified in the literature, it has to be noted that in many cases the dialectologist had problems classifying the linguistic variety as belonging to either variety A or B. Keeping in mind the problems with the isogloss method outlined above, these classification problems may not be the result of the variety's truly ambiguous nature, but may rather be the result of the set of isoglosses used by the dialectologist. If the set of isoglosses would have been larger, or if different isoglosses had been used, a different picture might have emerged, in which the transitional variety may not have been classified as such but rather as clearly belonging to one of the dialect groups.

This brings us to one of the main problems with the concept of transition zones: a proper definition has never been given, and is certainly not used in the different dialectological studies. In their article, A. L. Davis and McDavid (1950) give the following definition: "[a] transition area is one which has undergone influence from two or more directions, so that competing forms exist in it side by side. [...] The speech patterns of a transition area are likely to be more complex than those of the other two types [focal and relic areas, NO]". In the literature on Dutch dialectology, this definition is, however, not used. In practice, a transition zone is simply identified as a region that poses difficulties for classification, so any dialect that the dialectologist cannot group together with other dialects can be argued to be a transitional variety. As argued, this is not the best way to identify transition zones, but the definition put forward by A. L. Davis and McDavid is not ideal either. Although their definition does capture the idea that a transitional linguistic variety has to have characteristics in common with both core areas, nothing is said about the geographical location of the core areas and the transition zone: in their definition, mixed linguistic varieties that can be found at a large distance from (one of) the source languages (e.g. mixed linguistic varieties resulting from extensive language contact between different groups of migrants) would also count as transitional varieties, at least in the earlier stages of the language when there would still be variation and competing forms. When dealing with varieties that are no longer in a contact situation with their source languages (or core areas), it can be expected that eventually these varieties will reach a stable situation and display no more variation than linguistic varieties found in core areas. Though interesting, these transitional linguistic varieties that are not or no longer in a contact situation with their source languages or core areas will have to be excluded from the definition, by adding a geographical and a language contact part (cf. for example Britain (2004)).

However, even when using this new definition, the transition zones identified in the literature still face some problems. First of all, it is mostly assumed that transition zones form a transition between two core areas: regions that do not display a lot of variation and clearly behave as one dialect (group). A proper linguistic definition of dialects has, to my knowledge, never been given; dialect classification often seems influenced by geopolitical factors (with many dialects being named after the city, village or geographical region where they are spoken). Furthermore, dialect groups identified in the literature are separated by isogloss bundles of varying numbers, causing differences in the eventual classifications. As a result, there is a large amount of different dialect classifications. All in all, the results of the traditional approach to dialect classification must be treated with caution. However, they should not be overlooked entirely either; even though many different classifications exist, one can find many similarities among these classifications. Furthermore, even though the classifications are based on arbitrarily defined sets of isoglosses, the dialectologist's intuition plays an important role as well, leading them to use (bundles of) isoglosses and identify regions that are not unrealistic.

In the next section, a potentially more objective method for dialect classification will be discussed.

1.3.2 Dialectometry

Dialectometry (e.g. Séguy (1973)) is an approach to dialect classification that may overcome the problems that traditional dialectology faces. This method quantifies the differences between dialects, making it less arbitrary than traditional dialectology. An example of its application is the Levenshtein distance, which calculates how many segments have to be changed to get from the pronunciation of lexical item X in linguistic variety A to the pronunciation of that same lexical item in linguistic variety B. Calculations are made for each lexical item in a pre-selected sample for each pair of locations, resulting in a matrix of distances between all possible pairs of dialects. An advantage of this approach is that it is less arbitrary than the isogloss method. All differences that are found are weighted (whether all differences are weighted equally or some differences are given more importance than others depends on the exact application of the method), so that the regions that can eventually be identified do not depend on the isoglosses the researcher has chosen to work with, but rather on all differences found in the sample of lexical items.

The methods for calculating dialect distances can vary between different approaches to dialectometry. A classical approach is, as discussed, to count the number of segments that have to be changed. How the number of changes is counted might vary. Some approaches will treat insertion, deletion and substitution of a segment all as one change, while other approaches might treat substitution of a phoneme as a case of deletion followed by insertion (or vice versa), and thus consider substitution to be two changes instead of one. Other approaches count the differences between dialects in terms of phonological features (the substitution, addition or deletion of every feature is seen as one change). The difference between /sɔr/ and /zɔr/, for example, involves one change in both systems (in the phoneme difference approach the only change is a change from /s/ to /z/, while in the feature approach the only change is a change from [∅] to [voice][7] in the first phoneme), but the difference between /sɔr/ and /vɔr/ involves one change in a phoneme approach (/s/ is changed into /v/) and two changes in a feature approach ([∅] is changed into [voice]; [alveolar] is changed into [labiodental]). Other different approaches involve the nature of compared lexical items (a possibility is to only compare cognates with each other; another option is to also compare different lexical items with the same meaning with each other) and the weighting of different types of changes (an example is the difference between the phoneme and feature approach, but the length of the compared items may also play a role, changes in the beginning of the item may weigh more than changes in the end of the word as speakers pay more attention to the beginning of the word, or changes in stressed syllables may weigh more than changes in unstressed syllables as speakers rely more heavily on stressed than unstressed syllables in word recognition (see e.g. Wieling (2007) for a discussion of these different methods); a very different approach is used by Hoppenbrouwers and Hoppenbrouwers (2001) who compare the frequency of phonological features across dialects). As a dialectometrical approach quantifies the differences between linguistic varieties, the linguist can objectively determine which varieties are very similar and which very different. The method thus appears to be less arbitrary than classical dialectology.

Although dialectometrical studies have been carried out for other languages than Dutch (e.g. Kessler (1995) for Irish Gaelic, Heeringa, Johnson and Gooskens (2009) for Norwegian, Scherrer and Stoeckle (2016) for Swiss German), in this section the findings of several different dialectometrical studies of the Dutch language area are discussed. Most of them deal with phonological and phonetic characteristics, but one study focusses on morphosyntax.

Nerbonne et al. (1996)

Nerbonne et al. (1996) is the first dialectometrical study of Dutch dialects, based on data from the RND. It compares 100 lexical items for twenty dialects (including two Flemish and three Frisian dialects). As these data were not digitally available (the RND contains only handwritten phonetic transcriptions of dialect recordings), transcriptions were translated into a set of ASCII symbols. Subsequently, an algorithm calculated the Levenshtein distance for every lexical item in each pair of dialects. The dialects were clustered based on their linguistic distances: linguistically more similar dialects were clustered

[7] Or from [−voice] to [+voice] in a bivalent framework.

closer together, and linguistically more different dialects were clustered further apart. The clustering reveals patterns that are very similar to those that have been found in classical dialectology: the Frisian varieties are very similar to each other, the Low Saxon varieties are very similar to each other, the Flemish varieties are very similar to each other, the Brabantish varieties are very similar to each other, the Limburg variety is similar to the Brabantic varieties.

Hoppenbrouwers and Hoppenbrouwers (2001)

Hoppenbrouwers and Hoppenbrouwers (2001) is a classification of 156 Dutch dialects based on the Feature Frequency Method (FFM). This method calculates the presence of phonological features in a phonologically transcribed text, using a feature matrix for every phoneme. The authors used the RND for their study, as this database consists of a translation of the same set of sentences for 1956 dialects spoken in the Netherlands and Flanders; it is thus possible to compare the frequency of phonological features in a consistent way. The RND transcriptions were first translated into a phonological representation that could be understood by a computer program. For every phoneme used in the transcriptions a feature matrix was specified, thereby correcting for redundancy (a phoneme specified for [front] and [high] is automatically also specified as [vocalic], as the features [front] and [high] are never used in the representations of consonants. The redundant feature [vocalic] is deleted from the feature matrix). Subsequently, the computer program calculated the presence of each phonological feature for every individual dialect. Based on the feature frequencies, the authors compare the similarities between the dialects: dialects with comparable feature frequencies are more closely related than dialects with different feature frequencies. Based on these similarities a clustering of all dialects is made, represented in a dendrogram. Varieties that are very similar cluster together at a relatively low level, while varieties that are very dissimilar cluster together at a relatively high level.

Based on their calculations, Hoppenbrouwers and Hoppenbrouwers came to a classification of the 156 linguistic varieties used. They distinguish 9 larger dialect groups: Limburgish, Flemish, Zeelandic, Belgian Brabantian, Northern Brabantian, Southern Hollandic, Northern Hollandic, Frisian and (Low) Saxon. Each of these groups is subdivided in a core area and peripheral varieties, but only if a core area was clearly visible in the clustering. Peripheral varieties are found in the Low Saxon, Northern Brabantian, Belgian Brabantian and Flemish dialects. The peripheral varieties may classify as transition zones, since they clearly do not cluster directly with the core areas, and they are argued to have characteristics in common with more than one dialect group. Hoppenbrouwers and Hoppenbrouwers do, however, not specify on which aspects these peripheral varieties are different.

Apart from the 9 core areas with their periferal varieties Hoppenbrouwers and Hoppenbrouwers also identify a group of transitional dialects (‘overgangsdialecten’), located in the Veluwe region. Hoppenbrouwers and Hoppenbrouwers mention that these dialects have always been difficult to classify for dialectologists, and in their own classification the transitional varieties form a region of their own, rather than being classified as belonging to a larger dialect group. Traditionally these transitional varieties are considered to form a transition zone between the Franconian and the Low Saxon dialects.[8] Note that not only Hoppenbrouwers and Hoppenbrouwers (2001) classify this group of dialects as transitional; the group corresponds to the group of dialects classified as transitional by traditional dialectologists (as discussed in Section 1.3.1).

In conclusion, four different transition zones can clearly be identified based on this classification. First there is the group of dialects that are classified as transitional varieties and cannot be grouped together with any of the larger dialect groups. Then there are the peripheral varieties: the periphal varieties between Flemish and Belgian Brabantian together form a transition zone between these two varieties, and the peripheral varieties between Belgian Brabantian and Northern Brabantian together form a transition zone between those two varieties. Finally, there is a transitional area between Low Saxon and Frisian varieties.[9]

Heeringa and Nerbonne (2001)

Heeringa and Nerbonne (2001) used the Levenshtein distance to calculate the linguistic distance between 27 Dutch dialects, ranging from Bellegem (Flanders) in the south-west to Scheemda (Netherlands) in the north-east. Using a sample of 125 words, they calculated the number of mutations necessary to come from a phonological representation of a lexical item in one dialect, to the phonological representation of the same item in another dialect (the authors calculated the number of feature mutations necessary instead of the number of phoneme mutations). All lexical items were compared in this way for every pair of dialects.

When clustering the different dialects based on their linguistic distance, Heeringa and Nerbonne find that the entire continuum can be divided into three larger subgroups: Saxon, Dutch Franconian and Flemish Franconian. They show that the border between Saxon and (Dutch) Franconian is not one clear 'line' found between two adjacent villages, but rather there appear to be three borders[10] spread across four different locations (there is a border between

[8]However, in Hoppenbrouwers and Hoppenbrouwers' classification they seem to be more similar to the Frisian varieties.

[9]It is possible to argue that the peripheral Flemish dialects are a transition zone between Flemish and Zeelandic dialects, but since there is no Zeelandic core area it is unclear what status Hoppenbrouwers and Hoppenbrouwers would like to give to these linguistic varieties. The status of the southern peripheral Low Saxon dialects is another problem, since they are adjacent to the transitional varieties, and it seems illogical to assume a transition zone between the transitional varieties and the Low Saxon core area.

[10]The presence of three borders rather than any other number depends partly on the choice of lexical items, as the authors mention themselves (i.e., they mention that three lexical items correspond to the borders they found).

location 1 and 2, one between location 2 and 3 and one between location 3 and 4).

Heeringa (2004)

Heeringa (2004) carried out a dialectometrical study of Dutch dialects. The basis of this study is the same as in Heeringa and Nerbonne (2001): a sample of 125 lexical items is taken from the RND, but instead of 27 dialects this study includes 360 dialects in the Netherlands and Flanders. Most likely, the 125 items used in Heeringa (2004) are the same as the items used in Heeringa and Nerbonne (2001), but not all items in Heeringa and Nerbonne (2001) are specified (although the items that are mentioned in Heeringa and Nerbonne (2001) do correspond to items in Heeringa (2004)).

The dialects included in the study (360 out of 1956 dialects included in the RND) are chosen based on the size of the village and of the eventual spread of dialects over the entire linguistic landscape. In general, villages with less than 5,000 inhabitants were judged too small to be included in the study, as their dialects might be unstable because of people migrating in and out of the community, deaths in the community etc. Villages with more than 10,000 inhabitants, on the other hand, were judged to be too big to be included in the study, as more than one linguistic variety might be spoken in these villages.[11] The smaller (but not the larger) villages were only included when excluding their dialects led to an uneven spread of dialects in the study.

The 125 items were selected in such a way that they "represent (nearly) all vowels (monophthongs and diphthongs) and consonants. Also the consonant combination [sx] is represented, which is pronounced as [sk] in some dialects and as [ʃ] in some other dialects". As the RND only contains handwritten transcriptions of recordings, the transcriptions were digitised before analysis. The method of digitisation is not discussed (Heeringa refers to a website where the digitised forms should be available, but this website is no longer online). Because the RND only contains sentences but not single lexical items, individual items may be subject to assimilation. The data were corrected for this by comparing assimilated items with their realisation in a non-assimilating context in the same dialect, or with their realisation in a non-assimilating context in a neighbouring dialect. Based on the digitised data, the Levenshtein distance between every pair of dialects is calculated using an algorithm that finds the smallest number of changes needed (Levenshtein distances are normalised for the length of the alignment).

The results were visualised by drawing lines between every pair of dialects. The darker the line, the smaller the distance between the dialects.[12] The vi-

[11] Why these specific numbers are chosen as the limits is not specified.

[12] In case the Levenshtein distance yielded an irregular pattern, a 'bigger' location was sometimes replaced by a 'smaller' location. Definitions of bigger, smaller and irregular patterns are not given, nor is it specified how often this procedure was done. Also, this step is not really justified; an irregular pattern might indicate a dialect border. It is unclear why the

sualisation revealed several different dialect areas. The Frisian varieties form a clear dialect group (although the town Frisian varieties, which have been influenced by dialects from Holland, are not as closely related to the Frisian varieties as they are related to each other); the Low Saxon varieties are split into a Groningen group, a Drenthe group and an Overijssel group; there is a Zeelandic variety; a French variety; a West Flemish variety; an East Flemish variety; an Antwerp variety; and a (Belgian) Brabantic variety. The other Franconian varieties do not show clearly different dialect groups: there are just a few dialect pairs with a small linguistic distance between them. For most dialect pairs the linguistic distance is bigger. There is, however, a clear border visible between Low Saxon and Franconian varieties.

To reveal the patterns in the Franconian area, the dialects are clustered based on their linguistic distances: dialect pairs with smaller distances are clustered close together, while dialect pairs with bigger distances are clustered further apart. The 360 dialects are grouped into 13 clusters: this number corresponds with the number of groups identified in the visualisation of distances. Choosing a larger number of groups led to a clustering with several groups consisting of only one dialect.[13] The clustering reveals the following clusters: Frisian, Frisian mixed varieties (town Frisian and Stellingwerfs), Groningen, Overijssel, south-west Limburg, (Belgian) Brabant, central Dutch varieties, Urk, east Flanders, west Flanders, Zeeland, Limburg and north-east Luik.

In the final clustering, some of the Stellingwerf varieties are clustered with the Frisian varieties, while others are clustered with the Low Saxon varieties. Heeringa argues that this may be partially explained by transcriber differences,[14] but another part of the explanation might be that the varieties have characteristics in common with both the Frisian and the Low Saxon varieties and can thus be considered transitional areas. The dialects of Nunspeet, Vaassen and Bronkhorst are found to form a transitional area between the Low Saxon and the Franconian dialects, and the dialect of Groesbeek is argued to be intermediate between the Gelderland varieties (belonging to the central Dutch cluster) and the Limburg varieties. The Flemish and French varieties in west Flanders are separated from each other by the dialects of Nieuwkerke, Veurne and Alveringem, which form a transition zone between them. Finally, Heeringa identifies the cluster consisting of the dialects of 's-Gravenvoeren, Aubel, Baelen and Eupen to belong to the transition zone between east Limburg and the Ripuarian dialects, which corresponds to the location of the Benrather line.

A final step taken by Heeringa involves mulitdimensional scaling of the data. In the words of Heeringa, "[T]he purpose of multidimensional scaling (MDS) is to provide a visual representation of the pattern of distances among a set of elements". In this case, the method can be used to visualise linguistic distances

smaller location would be better than the bigger one.

[13]The eventual clustering with 13 groups has one dialect group consisting only of the dialect of Urk, however.

[14]The Stellingwerf varieties belong to different parts of the RND, which are transcribed by different transcribers.

between locations. The distance between locations can be represented by two or more dimensions: the more dimensions are used, the larger the percentage of variance that is explained. Visualisation is easy when two dimensions are used (distances can be plotted in a simple graph), but when three or more dimensions are used visualisation becomes more difficult. In that case, distances are not visualised in a graph, but rather the coordinates for every dimension (for example the coordinates on the x, y and z axis) are translated into different colour gradations. Every axis corresponds to a different colour, so that the coordinates of every location are translated into a specific colour. The linguistic distance between locations with similar colours is thus smaller than the linguistic distance between locations with different colours. Heeringa uses three dimensions, which together explain 88% of the variance in the data. Adding a fourth dimension would explain 92%, which is "a rather small improvement [...]".

The final results of the multidimensional scaling do not reveal patterns different from the ones revealed by hierarchical clustering or visualising linguistic distances. However, each dimension used in multidimensional scaling corresponds to a different lexical item, which means that the items corresponding to the three dimensions used in this study are the items that most strongly divide the linguistic landscape into different areas (i.e. the variation found in those lexical items correlates very well with the boundaries between different dialect groups). Heeringa shows that the first dimension corresponds to the item *waren* 'were'. This item divides the linguistic area into an area where the final nasal is realised (Fryslân, Low Saxon and some extremely south-western (West-Flemish) varieties) and an area where the final nasal is not realised (the rest of the linguistic area). The second dimension divides the area into Frisian varieties on the one hand, and Low Saxon and Franconian varieties on the other. It correlates to the lexical item for the word 'father': this is *heit* (with varying pronunciations) in Frisian and *vader* (with varying pronunciations) in the other varieties. The third dimension divides the southern part of the area into an eastern and a western area, and correlates to the realisation of the trill: either uvular ([ʀ], in the east) or alveolar ([r], in the west).

Wieling (2007)

Wieling (2007) (and also Wieling et al. (2007)) is the first dialectometrical study of the GTRP data. The GTRP data are more recent than the RND data, so that Wieling's data can be diachronically compared with Heeringa's data. Out of the 613 locations in the GTRP, 224 are shared with the 360 locations from the RND used by Heeringa (2004), and the two databases share 59 lexical items. These dialect locations and lexical items are used in Wieling's study. The Levenshtein distance is calculated for the 224 GTRP dialects, comparing phonemes instead of features (diacritics are ignored as there are transcriber differences both between the RND and GTRP, and within the GTRP; the Levenshtein distances are not normalised for the length of the alignments).

The results reveal patterns similar to the patterns found by Heeringa: a Frisian variety, several Low Saxon varieties, a west Flemish variety, an east Flemish variety and an Antwerp variety can be identified. Clear dialect areas within the Dutch part of the Franconian area are not found. However, a clear boundary between the Franconian and Low Saxon varieties is revealed. To compare the GTRP data with the RND data, a correlation test was carried out (for the Netherlandic and Flemish dialects separately). This showed a correlation of r = 0.83 for the distances between Netherlandic dialects, and a correlation of r = 0.82 for the distances between the Belgian dialects. The correlation is not perfect, but this is not unexpected considering the differences between dates of recording of the two databases.

Spruit (2008)

Spruit (2008) used the SAND data for a study that "investigates how to adequately measure syntactic variation in Dutch dialects". This is the first study on syntactic variation in Dutch dialects, as the SAND is the first collection of syntactic variation in the Dutch language area. The distances between every pair of dialects is calculated using the Hamming distance: "[...] the distance between dialect A and dialect B is increased by 1 for each variant that is observed in dialect A but not in dialect B, and vice versa" (Spruit (2008, p. 36)). The Hamming distance was calculated for 510 feature variants in (267*266)/2 = 35511 dialect pairs. Linguistic distances, relative to geographical distances, are plotted using multidimensional scaling: "Neighbouring dialect locations will have corresponding colours if there is a correlation between geographical distance and syntactic distance. In other words, a perfect correlation will result in a colour continuum, whereas a low correlation will result in a mosaic-like map" (Spruit (2008, p. 38)). The eventual map does not reveal very striking patterns: the areas that are visible on the map have been described in the previous literature.

Comparison of Spruit's map to the map created by Heeringa (2004) does not reveal very striking differences. Some minor differences include the absence of a separate group of Town Frisian varieties from Spruit's map, and the similarity of several Frisian varieties with south-western Flemish varieties. The first dissimilarity may be explained partly by the fact that the SAND data only include 3 Town Frisian locations while the RND includes 13 Town Frisian varieties, and partly by the fact that Town Frisian is a variety of Hollandic dialects once acquired by speakers of Frisian. As such, the syntax of these dialects may be similar to the (Franconian) Hollandic dialects. The second dissimilarity does not seem to be explainable by extra-linguistic facts.

Comparison of Spruit's map to Daan and Blok's map mainly shows that the large dialect areas identified by Daan and Blok are not fine-grained enough, as Spruit identifies many smaller regions within the bigger dialect areas. Second, Daan and Blok found a relatively large number of independent dialects areas within the Low Saxon area, while Spruit (p.53) "reveals only a few relatively subtle dialect areas in the north-eastern area". Finally, the Frisian area does not show as clearly on the map by Spruit as on the map by Daan and Blok.

Italian dialectometry

While the field of dialectometry has published a relatively large amount of work on Dutch dialects, less attention has been devoted to the Italian language area. A small number of studies, however, has been published using data from either the AIS or the ALI (see for example Goebl (2007) or Goebl (2008)). These works focus on the entire Italian language area and include linguistic varieties spoken in 383 locations. Instead of creating maps showing similarities between all varieties, following e.g. Wieling (2007) or Spruit (2008), Goebl (2007, 2008) chooses the linguistic variety of a single location as the point of reference for every individual map. These different maps often show the presence of the La Spezia-Rimini line in the linguistic landscape, and have also roughly found the same geographical divisions of the area as the divisions based on dialectology. One other study focuses on the process of *gorgia toscana* (a process that lenites intervocalic plosives in several different linguistic environments, and is active in many Tuscan varieties), and identifies several linguistic varieties that form a transition between varieties with and varieties without *gorgia* (Montemagni, Wieling, De Jonge and Nerbonne (2013).

Conclusions dialectometry

By making use of this method, different authors have identified different transition zones. The region that very clearly stands out as a transitional region is the border area between Franconian and Low-Saxon (Heeringa (2004), Heeringa and Nerbonne (2001), Spruit (2008), Wieling et al. (2007)).

However, the selection of this sample might be problematic, as there is no clear method for selecting the items in the sample. This method therefore depends on the differences that are found: differences that are not found simply cannot be used in the calculations. Which differences will be found therefore depends on the selection of lexical items by the researcher, and a poor choice of items may result in differences not being found. This problem can be overcome in two ways: either by selecting a sample based on the researcher's foreknowledge, or by using a large enough sample of testing items. The former is less informative than the latter: when choosing a sample based on one's foreknowledge, one will inevitably create dialectometrical patterns that are consistent with the existing classifications. Choosing a large enough sample thus turns out to be the best way to go.[15]

By using a dialectometrical approach the differences across linguistic varieties are quantified, so that the researcher can easily identify which varieties are highly similar and which varieties are very different. An advantage of this method is that its workings are less arbitrary than the workings of traditional

[15] According to Séguy (1973), a sample of 100 items should be large enough.

dialectology, since all differences that are found are taken into account, so that the classification is not based on an arbitrarily defined set of isoglosses. However, the differences that are found do depend on the sample of items that is used, so the sample will have to be large enough to be able to find all relevant differences across dialects (as mentioned above). The method is, however, not an ideal one. In a computational approach the historical factor is completely left out. Dialects that have different phoneme inventories or phonotactics are argued to be more distant than dialects that have highly similar phoneme inventories or phonotactics, even if in the latter case the similarities are merely a coincidence (being the result of two completely unrelated linguistic varieties evolving in the same direction) rather than very meaningful (when the two linguistic varieties are historically related). Factors like these might be relevant, so it will be useful to include them in the analysis. Statistics can of course help in distinguishing coincidental from meaningful similarities, but another, significant problem with the approach remains: there appear to be many different approaches to calculating dialect distances. A standard implementation is to count the number of substitutions, deletions and additions of phonemes necessary to go from the phonemic representation of a lexical item in one linguistic variety to the phonemic representation of that same lexical item in another linguistic variety. Another possibility is to count the number of substitutions, deletions and additions of phonological features necessary to go from the phonemic representation of a lexical item in one linguistic variety to the phonemic representation of that same lexical item in another linguistic variety. Both these approaches can be divided into an *all-word* and a *same-word* approach. In the first example all comparisons are made, even if the two linguistic varieties use different lexical items for the same linguistic concept, whereas in the second approach the comparison is only made if the two linguistic varieties use the same lexical item for the linguistic concept (Kessler (1995)). The existence of all these different approaches can lead to differences in dialect classifications.

1.3.3 Perceptual dialectology

A final approach to dialectology that deserves attention is perceptual dialectology (Weijnen (1966)). In this approach dialect speakers (without any linguistic knowledge) are asked to indicate how similar they judge other dialects to be to their own dialect. Based on these judgments a map of the linguistic landscape is made. Daan and Blok (1969) created such a map for the Dutch language area. They asked informants[16] to indicate in which locations in their neighbourhood[17] speakers spoke the same dialect or a dialect similar to their own

[16] For the Netherlandic part of the area the authors were able to find enough informants, as the Dialectenbureau had access to a large enough group of dialect speaking correspondents. For the Flemish part of the area, however, they could not find enough dialect speakers. The map for the Flemish area is based on the judgments of Flemish dialectologists, but as all of them were dialect speakers themselves the authors did not consider this to be a large problem.

[17] How ‘neighbourhood’ was defined is not clear from the text.

dialect, and in which locations in their neighbourhood speakers spoke a dialect very different to their own. The last question also asked informants to indicate some of the differences between their own dialect and the dialect they judged to be different. Based on these judgments the linguistic landscape is divided into different dialect areas: if a speaker of dialect A indicates that dialect B is similar to his or her own dialect, the two dialects are joined by an arrow on the map. If speakers of dialect B judge dialect C to be similar to their own dialect, dialects B and C are joined by an arrow as well. Although speakers of dialects A and C did not indicate their dialects to be similar to each other, they are indirectly joined together because both are judged to be similar to dialect B. When speakers judged a dialect to be different from their own dialect, the two dialects are separated by a dialect border as they are not joined together by an arrow.

On the map by Daan and Blok (1969), 28 different dialect groups are distinguished. The Franconian, Frisian and Low Saxon varieties are clearly judged to be different from each other. Within the Frisian area there is a division between Frisian and Town Frisian dialects, but the Low Saxon and Franconian areas are divided into 9 and 17 dialect areas respectively. In the Belgian area three transitional dialect groups are identified: the 'dialect of the area between Brabantic and Limburgian', the 'dialects of the area between west and east Flemish' and the 'dialect of the area between east Flemish and Brabantic'. In the Netherlandic area no transitional varieties are identified. Whether this is simply a matter of labelling or an effect of different types of informants is unclear. Although Daan and Blok (1969) do not refer to any of the Netherlandic varieties as transitional, they do mention the idea that linguistic boundaries are never sharp. In fact, the book starts with the statement that when crossing the country from west to east, a listener will notice a gradual change in dialects.[18]

The classification of the Flemish dialects could not be based on informants' judgments, as the project could not find enough informants in Flanders. Instead, classification of the Flemish part is based on the linguists' knowledge. As the approach is the same approach as is used in classical dialectology, the outcomes of the two do not differ from each other. The division of the Netherlandic part of the language area, however, is also very similar to a division based on isoglosses.

Within the Italian language area this approach has not been applied extensively. Romanello (2002) has used the approach to study several southern Italian urban dialects (investigating whether there are dialectal differences within and between the cities), and Gally (2015) studies the dialects of ten villages in west Piedmont. A study covering the entire language area has, to my knowledge, never been published.

[18] "Wie van Den Haag naar het oosten wandelt en hier en daar een praatje maakt, hoort de plattelandsdialekten geleidelijk veranderen." (p.9) ("If you walk from The Hague to the east and have a small chat here and there, you can hear the dialects of the countryside gradually change.")

1.3.4 International transition zones

Apart from transitional areas in the Dutch language area, transition zones have also been identified for other language areas. One of the most striking examples is the transitional linguistic variety identified by A. L. Davis and McDavid (1950): a linguistic variety spoken in northwest Ohio is a mixed linguistic variety, which therefore shows characteristics of both source languages. This does indeed sound like a prototypical example of a transitional linguistic variety; however, this mixed variety is the result of two different groups of immigrants from the New York area and Pennsylvania transferring features from their local English variety to the new, mixed variety. In this example the geographical factor is completely absent: the source languages of the mixed language are not geographically neighbouring varieties, but rather are spoken in another, distant, area. Since there no longer is an (extensive) contact situation between the source languages (which can be compared to the core areas in the case of a dialect continuum) and the mixed language (the transition zone), the expectation is that the mixed language will undergo stabilisation and can, eventually, be considered a linguistic variety on its own, independent of the source languages.

Other examples of transition zones which better fit the idea of a dialect continuum, are the two areas described in Chambers and Trudgill (1980) (see Section 1.2). Dialectometrical work has identified transition zones in many other different language areas. See for example Scherrer and Stoeckle (2016) for transition zones in Switzerland.

1.3.5 A methodological problem

When studying the literature on transition zones, a few things have come to light. A first problem is that so far, no one has given a clear definition of linguistic transition zones. The term appears to be applied freely to dialects that cannot be easily classified, but there are no other requirements on transitional dialects. No author has mentioned that, in order to be classified as a transitional dialect, the dialect must have characteristics in common with the dialects or dialect groups between which it forms a transition, or that a situation of language contact is necessary (see for example A. L. Davis and McDavid (1950), where there is no contact situation between the different dialects). Even though such a stabilised transitional dialect will be able to tell us how speakers have dealt with the presence of variation, the absence of a contact situation is likely to cause stabilisation of the linguistic situation, leading to less variation in the stabilised variety. A contact situation between two or more dialects or dialect groups will thus provide a more interesting test case for studying linguistic variation. In this study a contact situation between two (or more) different dialect groups will thus be a prerequisite for the postulation of a transition zone. This implies that the dialects must be geographically adjacent.

In Section 1.3.1 a more fundamental problem of the classical dialectological approach has already been briefly discussed. This problem concerns the fact that dialect classification is not as straightforward as it is commonly presented. It has already been discussed that different dialectologists use different sets of linguistic characteristics, and therefore all have different classifications of dialects. The problem that dialect classification faces, however, is not merely caused by the set of linguistic characteristics that the dialectologist chooses to use; there is a much deeper and fundamental cause of the problem. Not a single dialect is linguistically identical to another dialect.[19] Further, there is no objectively defined threshold by which two different linguistic varieties can be judged as 'different'. Considering these two facts, it has to be concluded that dialect classification is in fact an impossible task. Consequently, if it is impossible to group dialects into larger families, it is impossible to identify dialects that are a transition between these families.[20]

When considering the literature on Dutch dialects, several transition zones are mentioned by many (if not all) dialectologists: the border area between Franconian and Low Saxon dialects, and the Rhenish Fan. The identification of these areas does not seem to depend on the method that is chosen. These areas do therefore seem to be the perfect areas to study. However, the dialectal changes that are present in those two areas do not involve deep changes in the grammar, but are rather quite superficial. An example is the presence of the final nasal in the suffix *-en* (indicating infinitives or plural in verbs, and indicating plural in nouns): in Franconian dialects the final nasal is often deleted (although it may surface in liaison contexts such as *we lopen op straat* 'we walk on the street', where the final nasal of the verb *lopen* can be realised to avoid a hiatus between the vowels of the verb and the following preposition), while in Low Saxon dialects it is still present. Another example is the vocalisation of /ol/ to /ou/ in syllables where the cluster is followed by /d/ or /t/ (e.g. *woud* vs. *wold* (forest): the vocalisation has taken place in most of the Franconian dialects, but not in most of the Low Saxon dialects. Similar examples can be given for the Rhenish Fan area: the main difference between the dialects is the contexts in which a sound change did apply and in which it did not, creating variation in the lexicon. At the level of the grammar, however, the differences are not appreciable.

Several other examples can be given for both transition zones, but the examples all relate to relatively superficial language characteristics: none of them involves a systematic difference in the grammar. The variation that can be found in the transition zones does not concern true phonological variation, but rather lexical: some items may be affected by the change that vocalises /ol/ to /ou/ while other items may not be affected. Lexical variation most likely will

[19] In fact, not a single speaker has a linguistic system that is constant in all different social contexts.

[20] Chambers and Trudgill (1980) point out that the isogloss is not particularly suited for identifying linguistic variation, as it is often based on differences between two speakers (one speaker per location) who would both be consulted only once. Variation between speakers at the same location is thus not easily spotted.

not reveal very interesting patterns and most likely does not shed much light on the question how speakers deal with variation: it may reveal a pattern in which items are affected by a phonological change and which items are not (e.g. a frequency pattern), but it does not show how speakers deal with grammatical variation. In this study the focus will therefore not be on one of the aforementioned transition zones. Instead the focus will be on one linguistic phenomenon per region, because it is possible to identify the approximate boundary of one linguistic characteristic or the approximate region where the implementation of an individual characteristic is changed. The expectation is that if the phonological systems are different enough (i.e., if they do not merely involve the presence or absence of a final nasal, the vocalisation of a vowel-sonorant cluster, etc.), the contact area can potentially display interesting variational patterns in the phonology. These variational patterns might contribute to two important questions in linguistic theory: the nature of the phonetics-phonology interface, and the locus of variation. These two topics will be discussed in the following section. The final section describes the two regions that will be studied in this dissertation in more detail.

1.4 The research topic

The present study looks at phonetic and phonological variation in two individual transition zones. To avoid the problem discussed in Section 1.3.5, the study focusses on individual phonological phenomena rather than on transitions between different dialect groups as a whole. In the latter type of transition zones mostly lexical variation can be found. While this is interesting, it does not shed much light on the question of the phonetics-phonology interface; it will only be helpful in a study on the locus of variation. This study therefore focusses on two individual voicing phenomena, one relating to the phonetics-phonology interface and the other relating to the locus of variation. Voicing, or laryngeal contrasts, are a well-studied topic in both phonetics and phonology. A fundamental theoretical discussion continues to be held to this day, however. The most outstanding discussion relates to the phonological representation of two different types of laryngeal contrasts: should both be represented in the same way (assuming that differences are merely phonetic), or should they be represented differently (assuming that differences are not only phonetic but also phonological)? This discussion is thus very relevant for the discussion about the phonetics-phonology interface. A contact situation between the two language types will be likely to cause both phonetic and phonological variation, which can shed light on both the question of the representation of the different contrast types and on the nature of the phonetics-phonology interface. Within the Germanic language family both contrast types are present, but a direct contact situation between the two different contrast types can only be found between different dialects spoken in the Netherlands and different dialects spoken in Germany.[21] More specifically, within the Low Saxon language area is an ideal testing location, providing a contact situation between the two laryngeal systems within the language area. As this area forms a dialect continuum, the risk of other linguistic characteristics interfering with the laryngeal contrast is negligible.[22]

The second transition zone is located at the border between northern and central Italian varieties, more specifically between Emilian-Romagnol and Tuscan varieties. In the first group of dialects there is a neutralisation of the voicing contrast in intervocalic position, affecting only the alveolar fricative (Intervocalic /s/-voicing (ISV)). This neutralisation is restricted to a number of linguistic environments; outside these environments it cannot apply. This neutralisation is not present in the second group of dialects. This can lead to interesting phonological situations, as speakers will have to decide whether or not ISV is active in their language. Possibly some speakers will show variability in the application of the rule, which may lead to interesting grammatical situations. The area is therefore an ideal region to help answering the question on the locus of variation in the grammar. Besides the existence of a contact area, several authors have claimed that linguistic varieties in central Italy are currently undergoing a change towards the northern varieties, and speakers from central Italy are implementing ISV. The change is most visible in the speech of younger speakers, meaning that their grammar will likely differ from the grammar of older speakers. As these younger speakers cannot have learned the distribution of voiced and voiceless fricatives in intervocalic fricatives from their parents, they must have borrowed it from speakers of northern varieties. The transition zone thus forms an ideal testing case for the hypothesis put forward by Thomason and Kaufman (1988), who argue that it should be possible for speakers to borrow not only superficial language characteristics (e.g. a phonetic characteristic), but the borrowing of more structural characteristics such as phonemes, morphemes, phonological rules or syntactic constructions should also be possible.

The two regions are very different with respect to their geographical, geological and political characteristics. A comparison between the two regions might therefore yield interesting insights in the nature of linguistic variation. It is commonly assumed that language contact is a necessary condition for the presence of language variation, but the question whether it is also a sufficient condition is not frequently asked. The differences in the extra-linguistic characteristics might have an effect on the variation found in the areas, thereby

[21] Dutch is one of the languages with a for the Germanic languages atypical voicing contrast. Yiddish and Afrikaans also have deviating voicing contrasts, but they are not spoken in a dialect continuum with another Germanic language.

[22] Within the Franconian dialects in the Netherlands and Germany the same contact situation between two different laryngeal systems can be found, but several Franconian varieties have a tonal distinction (e.g. Köhnlein (2011)). As tone is related to laryngeal characteristics it might interfere with the laryngeal contrasts. The Low Saxon dialect continuum does not have a tonal contrast. The present research therefore studies laryngeal contrasts in the Low Saxon continuum.

giving a (partial) answer to the latter question. Further, the differences in the political situations might yield new insights in the study of language variation. The influence of the national political situation on languages and dialects (e.g. dialect levelling or standardisation) is not a new topic in the study of variation. However, much less is known about the effects of politics across country borders or within a single country. Finally, this study includes two areas that are linguistically different from each other. The processes and characteristics that will be studied differ in such ways that variational patterns may vary between the two areas. The study might thus shed new light on the discussion on the limits of variation.

This dissertation will be structured as follows. Chapters 2 and 3 discuss, respectively, the Dutch-German and the Italian transition zones. A thorough theoretical background of the two phonological topics is given first, followed by an overview of the two geographical regions and a discussion of the findings of the study. In Chapter 4 different analyses of the findings are discussed, followed by a discussion of their implications for the different approaches to the phonetics-phonology interface in Chapter 5. Chapter 6 will be the concluding chapter.

CHAPTER 2

Voicing distinctions in the Dutch-German dialect continuum

2.1 Voicing distinctions

Ever since the discovery by Lisker and Abramson (1964) (see also Abramson and Lisker (1965)) that the phonetic realisation of a voicing contrast may differ across languages, voicing distinctions have been one of the most well-studied phenomena in both phonology and phonetics. Lisker and Abramson found that, with respect to voicing, languages can generally make use of three different plosive categories. Phonetically, the main difference between these plosives is one of Voice Onset Time (VOT): the onset of vocal cord vibration relative to the moment of plosive release. When vocal cord vibration initiates before plosive release, VOT is negative (prevoicing); when vocal cord vibration initiates after plosive release VOT is positive. Positive VOT's can be further divided into two different categories: plosives where VOT is initiated shortly after the plosive release (short-lag VOT or plain voiceless plosives) and plosives where the onset of vocal cord vibration is delayed (long-lag VOT or (voiceless) aspirated plosives).[1] Most languages with a two-way laryngeal contrast either contrast a prevoiced plosive with a plain voiceless plosive (this is commonly referred

[1] A fourth category that, to the ear of the listener, seems to combine prevoicing and aspiration in one plosive exists. Phonetically, it is of course impossible to combine prevoicing with a delay in voicing. The apparent 'voiced aspiration' of breathy voice plosives results from combining periodic voicing with substantial airflow through the glottis (Gordon and Ladefoged (2001)). Breathy voice plosives are not very common across languages (e.g. again Gordon and Ladefoged (2001)), and they will not be considered any further.

to as a voicing contrast), or a plain voiceless plosive with an aspirated plosive (this is commonly referred to as an aspiration contrast). A third category contrasting prevoiced plosives with aspirated plosives also exists, but appears to be very infrequent (but see Beckman, Helgason, McMurray and Ringen (2011) for Swedish or Hunnicutt and Morris (2016) for Southern American English).

Before Lisker and Abramson (1964) no distinction between voicing and aspiration contrasts was made. The terms 'voiced' and 'voiceless' were not only used to refer to phonetically voiced (prevoiced) and voiceless plosives, but also referred to any plosive involved in a laryngeal contrast, regardless of the specific phonetic realisation of this contrast: in voicing languages, 'voiced' refers to the phonetically voiced and 'voiceless' to the phonetically voiceless plosive, but in aspiration languages 'voiced' was used to refer to the phonetically plain voiceless and 'voiceless' to the phonetically aspirated plosive. Presence of phonetic voicing was thus not a requirement for a plosive to be referred to as 'voiced'. This approach was visible in the phonological representation of laryngeal contrasts, as several phonologists have proposed to represent all two-way laryngeal contrasts as a [±voice] distinction (e.g. Keating (1984), Wetzels and Mascaró (2001)). The plosive with the shortest VOT (where prevoicing is considered to be shorter than short-lag VOT) is marked [+voice], and the plosive with the longest VOT is marked [−voice].[2] Phonetic differences between the languages (VOT differences) are argued to be the result of language-specific implementation (phonetic implementation rules).

After Lisker and Abramson's study, the phonetic realisation of laryngeal contrasts has been studied for many different languages. See for example Slis and Cohen (1969), Jansen (2004), Van Alphen and Smits (2004), Van Alphen (2007) for Dutch, Jessen (1998), Braunschweiler (1997) and Jessen and Ringen (2002) for German, Moosmüller and Ringen (2004) for Austrian German, Kallestinova (2004) and Ögüt, Kiliç, Engin and Midili (2006) for Turkish, Helgason and Ringen (2008) and Beckman et al. (2011) for Swedish and Lisker and Abramson (1964) for English. Besides the specific phonetic realisation of the laryngeal contrasts in many different languages, different authors investigated the phonological representation of the different contrast types. Different authors have argued to transfer the phonetic differences between voicing and aspiration languages to the phonology, putting forward several reasons. First, there is a physical difference between the three plosive categories. Prevoiced plosives require the presence of vocal cord vibration, which require the vocal cords to be slack and adducted. The air pressure below the vocal cords must be higher than the air pressure above the vocal cords for the Bernoulli effect, which is required for voicing, to take place.[3] Aspirated plosives require the glot-

[2]Monovalent approaches also exist. The phoneme with the shortest VOT value (where negative VOT is considered the lowest, and positive VOT the highest) would then be represented as [voice] (whether it be a prevoiced or a plain voiceless plosive), and the phoneme with the longest VOT value would then be represented as [∅] (whether it be a plain voiceless or an aspirated plosive) (e.g. Cho (1990)).

[3]For phonetic voicing to occur there has to be a difference in air pressure between the

tis to be in an open, abducted position (Kim (1970)): because of the abducted state of the vocal cords during the plosive release, there is a delay in voice onset (long-lag VOT) caused by the time needed to reach an adducted state of the vocal cords. Plain voiceless plosives, however, can be seen as neutral plosives. This physical difference is ignored in a system that uses only one phonological feature to describe all contrasts.

Second, plain voiceless plosives are most common cross-linguistically. Maddieson (1984) shows that out of 50 languages with only one plosive series, 49 of them have a plain voiceless realisation of this series. Further, approximately 89% of languages with a two-way laryngeal contrast in their plosive inventory include the plain voiceless series, which is contrasted with either prevoiced or with aspirated plosives (Maddieson (1984)). Languages contrasting prevoiced with aspirated plosives, however, are rare (but see Beckman et al. (2011) and Hunnicutt and Morris (2016)). If [+voice] plosives can be realised as both prevoiced and plain voiceless, and if [−voice] plosives can be realised as both plain voiceless and aspirated, the infrequency of languages contrasting prevoiced [+voice] plosives and aspirated [−voice] plosives should be explained as a mere coincidence, and cannot be predicted on phonological grounds.

Third, language-acquiring children show an asymmetry with respect to the plosive series they acquire first: in voicing languages children acquire the fortis plosives before they acquire the lenis plosives, while in aspiration languages children acquire the lenis plosives before they acquire the fortis plosives. Phonologically, this means that all children acquire the unmarked series before they acquire the marked series. Further, if a child makes a production error the direction of the error is always from marked to unmarked: in voicing languages a lenis (prevoiced) plosive will be realised as fortis (plain voiceless), while in aspiration languages an (aspirated) fortis plosive will be realised as a (plain voiceless) lenis plosive (Kager, Van der Feest, Fikkert, Kerkhoff and Zamuner (2007)).

Finally, there are phonological differences between languages contrasting prevoiced with plain voiceless plosives and languages contrasting plain voiceless with aspirated plosives. In voicing languages, intervocalic lenis plosives always have full voicing during closure, while in aspiration languages intervocalic lenis plosives can have either full or partial voicing during closure. In the former case, voicing is systematic and argued to be a phonological property of the plosives, while in the latter case voicing is optional and argued to be a phonetic effect (intervocalic voicing, e.g. Beckman, Jessen and Ringen (2013)). In both language types fortis plosives can only have partial voicing

sinus above and the sinus below the glottis. The air below the glottis will cause the vocal folds to be abducted, so that the velocity of the air flow through the glottis increases. This increase in velocity causes a drop in the pressure at the glottis, so that the vocal folds will adduct. The increase in velocity causing the air pressure to drop is called the Bernoulli effect. The process of abducting and adducting the vocal folds as a result of changes in air pressure can be repeated as long as there is enough air pressure below the glottis (Rietveld and Van Heuven (1997)).

during closure. Another difference regards a difference in the active feature in, and the direction of voicing assimilation in plosive clusters. In languages where the [+voice] plosive is realised with prevoicing, a cluster with a [+voice] second plosive always surfaces with full voicing during the entire closure of the cluster, regardless of the feature value of the first plosive. The first plosive is thus assimilated in voicing to the second plosive. Because most voicing languages have a devoiced realisation of syllable-final plosives, a [+voice] first plosive undergoes assimilation just like a [−voice] first plosive: if the second plosive were [−voice], a first [+voice] plosive would surface as voiceless as the result of final devoicing. An [αvoice][+voice] cluster thus always surfaces as if both plosives are [+voice]. [+voice] can never spread to a following consonant, however. In a [+voice][−voice] cluster assimilation does not occur: the entire cluster will surface as voiceless. The first, [+voice] plosive devoices as a result of final devoicing, and the second plosive is already voiceless. Because assimilation can only be from right to left, but never from left to right, this process is referred to as Regressive Voice Assimilation (RVA). Typically, voicing languages only show assimilation to the voiced member of the cluster, but never to the plain voiceless member.[4] Languages with an aspiration contrast typically show assimilation towards the fortis member of the cluster instead of the lenis (Iverson and Salmons (1999)): in clusters with a lenis second plosive, this plosive does not show phonological activity. Clusters with a first fortis obstruent, however, show phonological activity of this first plosive. In English, for example, the default realisation of the plural morpheme is lenis, but when this (monophonemic) morpheme is preceded by a fortis obstruent it is realised as fortis (e.g. *cat-[z]* $\rightarrow$ *cat-[s]*[5]). In clusters with a [+voice] second plosive, there is no spreading of this feature. Vocal cord vibration seizes during the closure phase of the first plosive in the cluster, and is only resumed after the release of the second plosive. However, in clusters with a [−voice] first plosive, the second plosive can assimilate to the first plosive (Progressive Voice Assimilation (PVA)). A [−voice][αvoice] cluster thus surfaces without any voicing during the closure. The presence of certain phonological processes thus appears to depend on the phonetic realisation of the voicing contrast. In a [$\pm$voice] analysis this can only be explained as a coincidence, but it cannot be attributed to phonological differences between the two types of contrast.

A common approach is to distinguish a phonological voicing contrast (for voicing languages) from a phonological tenseness contrast (tense vs. lax plo-

[4] There is, however, a clear counterexample to this point: the case of Dutch past tense formation. The default past tense morpheme is -de, which surfaces as -te when it follows a voiceless plosive. After a sonorant or a voiced obstruent the suffix surfaces as -de, as in *hij noemde* 'he named' and *hij krabde* 'he scratched'. After a voiceless obstruent the suffix surfaces as -te, as in *hij raakte* 'he hit'. If the default suffix is assumed to be -te, assimilation to -de in the case of a preceding sonorant cannot be explained, as sonorants are not underlyingly specified for voice. If one assumes the suffix is underlyingly -de, the suffix surfaces with a devoiced initial plosive if the previous obstruent is not voiced either.

[5] The phonetic realisation of the fricative is here represented as a voicing contrast, for reasons of typographical ease.

sives, for an aspiration contrast) (e.g. Jakobson, Fant and Halle (1952)). In this approach, "[i]n contradistinction to the lax phonemes the corresponding tense phonemes display a longer sound interval and a larger energy [...]. In consonants, tenseness is manifested primarily by the length of their sounding period, and in stops, in addition, by the greater strength of their explosion". Chomsky and Halle (1968) also propose the existence of a feature [tense], but propose that it is accompanied by 'heightened subglottal pressure' in aspirated plosives. Another common option is to distinguish a phonological voicing contrast from a phonological fortis-lenis contrast (e.g. Kohler (1984)). This feature "may be associated with an articulatory timing and with a laryngeal power/tension component. The former relates to the speed of stricture formation and release, and is probably a language universal, the manifestation of the latter (aspiration, voicing, glottalization) is language-specific" (p.168). The feature fortis can be associated with greater articulatory power, an increase in vocal fold tension, a shorter closure duration (which is a language-specific characteristic) and f_0 perturbations. Based on the findings by Lisker and Abramson (1964), Halle and Stevens (1971) propose a framework of articulatory features, which refer to the position and status of the glottis at the moment of articulation. Four binary features in total are used: [±spread (glottis)], [±constricted (glottis)], [±stiff (vocal cords)] and [±slack (vocal cords)]. The features [+spread] and [+constricted] can obviously not be combined, nor can the features [+stiff] and [+slack]. All other feature combinations can be made. [+constricted] plosives (ejectives and implosives) will not be further considered here. Prevoiced sounds require slack vocal cords (stiffness of the vocal cords inhibits vibration), so they would be represented with the feature [+slack]. The glottis is neither spread ([−spread]) nor constricted ([−constricted]). Plain voiceless plosives have stiff vocal cords ([+stiff]) so that no vocal cord vibration occurs, but the glottis is neither spread (so that no aspiration occurs either) nor constricted. Aspirated plosives have stiff vocal cords to inhibit voicing, and a spread glottis that causes aspiration. Although these proposals do transfer the phonetic distinction between voicing and aspiration languages to the phonology, they are rather complicated as they require four different features (Halle and Stevens (1971)), or do not link laryngeal characteristics to timing differences (Kohler (1984)).

An approach that does link laryngeal characteristics to timing differences, and does not require the presence of many features, is the approach by Iverson and Salmons (1995) (cf. Iverson and Salmons (1999), Iverson and Salmons (2003), Avery (1996)). The authors propose the existence of two, monovalent, features: [voice] and [spread glottis]. The feature [voice] represents plosives with prevoicing, while the feature [spread glottis] represents plosives with aspiration. As both features are monovalent, the plosive series contrasting with either prevoiced or aspirated plosives (plain voiceless plosives in both voicing and aspiration languages) is unmarked for laryngeal characteristics ([∅]). Several arguments have led the authors to this proposal. First, the differences in timing of voicing onset are linked to the position of the glottis. As Kim (1970) has shown for Korean, the glottis assumes a wide, open position ('spread') for an aspirated plosive. Because the vocal cords are far apart at the time of plosive release, and have to be brought together for voicing to commence, a delay in voicing onset occurs naturally in these plosives. Second, the link between the nature of the voicing contrast (voicing or aspiration) is now linked to the presence of certain phonological and phonetic processes:

- the consistent full voicing during closure of intervocalic lenis plosives in voicing languages is the result of the presence of the feature [voice], which requires phonetic (pre)voicing;
- the inconsistent full voicing during closure of intervocalic lenis plosives in aspiration languages is the result of a phonetic process: lenis plosives in aspiration languages are laryngeally unmarked ([∅]). This means that there is no phonological feature requiring phonetic voicing of that plosive in any position. However, as there is no contrasting [voice] plosive series, unmarked lenis plosives are free to vary between full voicing during closure and partial voicing during closure. In other words, a phonetic process of intervocalic voicing is free to apply to these plosives, resulting in some fully voiced and some partially voiced plosives. The very same argument can be made for post-pausal lenis, unmarked plosives: they have no phonological feature that requires a specific phonetic realisation, so these plosives are free to show phonetic variation, as long as they do not merge with the relevant contrasting plosive series. In aspiration languages this contrasting plosive series is marked [spread glottis], which requires aspiration, so the lenis plosive cannot surface with a long-lag VOT, but nothing prevents the plosive from surfacing with prevoicing;
- the presence of either RVA or PVA has no longer an (apparently random) link with the phonetic realisation of the [±voice] contrast, but is instead linked to the presence of either [voice] (RVA) or [spread glottis] (PVA). This is a rather pleasing approach, as it clearly explains why assimilation is always in the direction of the voiced plosive series in voicing languages, and to the aspirated plosive in aspiration languages: in voicing languages the feature [voice] is the only present feature, marking voiced (lenis) plosives. Voiceless unaspirated (fortis) plosives are unmarked: they lack a laryngeal feature. Obviously, a phoneme cannot actively spread a feature that it does not have: so, as fortis plosives in voicing languages lack a laryngeal feature, they consequently cannot spread a laryngeal feature to an adjacent plosive. Voicing languages are therefore not expected to show assimilation to the voiceless member of the cluster.[6] In aspiration languages the only active laryngeal feature is the feature [spread glottis], marking aspirated (fortis) plosives. Plain voiceless (lenis) plosives are laryngeally unmarked. Consequently, lenis plosives cannot spread a laryngeal feature as they do not have one, but fortis plosives can spread their feature [spread glottis]. Aspiration languages are thus not expected to show assimilation to the lenis member of the cluster.

[6] Remember, however, the case of Dutch past tense formation.

An additional advantage of this approach is that the distribution of aspiration can easily be explained. In aspiration languages, fortis plosives in the onset of stressed syllables are realised with a long-lag VOT (e.g. *pin* [pʰɪn]), but when these plosives are preceded by a voiceless fricative in the same syllable they are realised with a short-lag VOT (e.g. *Spinne* 'spider' [spɪ.nə], not [spʰɪ.nə]).[7] In traditional, [±voice] approaches the plain voiceless realisation of fortis plosives preceded by a fricative in the same onset cannot be easily explained: the distribution of aspiration is guided by phonetic implementation rules. However, in the approach by Iverson and Salmons (1995) (cf. also Kehrein and Golston (2004)), the lack of aspiration can be explained on phonological grounds. Not only aspirated plosives are marked with the feature [spread glottis]; voiceless fricatives are marked with the same feature. In clusters consisting of /s/ and an aspirated plosive, both obstruents are thus marked with the feature [spread glottis]. Instead of both obstruents being marked for [spread glottis] independently, the two segments share one feature [spread glottis] (Iverson and Salmons (1995, p. 2)). This assumption is based on the observations by Kim (1970, p. 113-4):

> It has been hypothesized [. . .] that the minimal unit of motor commands is a syllable, not a phoneme, and, accordingly, that segments within a syllable receive a simultaneous package of instruction for articulation [. . .]. If we accept this hypothesis, it follows that the effector organs already possess information concerning the second consonant in, say, /sp/ at the same time the articulation of the first consonant /s/ is being accomplished. And if we further assume that in the coordination of the articulations of the segments within a syllable, there is no requirement to delay a certain articulatory movement of the second segment until a certain movement of the first segment has been completed, as long as the two movements are not simultaneously incompatible due to some inherent physiological constraints, then it follows that the glottal movement for /p/ of /sp/ will start during /s/, i.e., the glottis will begin to widen. This means that, if the glottis is instructed to open to the same degree and for the same period for /p/ of /sp/ as it would for initial /p/, the glottis will begin to close by the time the closure for /p/ is made, and consequently, by the time /p/ is released, the glottis will already have become so narrow that the voicing for the following vowel will immediately start, and thus we have unaspirated /p/ after /s/. Note that the notion of simultaneous compatibility is crucial here, i.e., since /s/ is voiceless and does not require the closing of the glottis, the opening of the glottis for /p/ does not have to wait for the completion of /s/ but can proceed simultaneously with the oral articulation of /s/.

[7] In both Dutch and German, the only fricative that is allowed in the same onset as a fortis plosive is /s/.

So, in short, fortis plosives in aspiration languages are realised without aspiration when preceded by a voiceless fricative because the open position of the glottis is already assumed for the articulation of /s/, and closing of the glottis can be initiated before the release of the plosive. During the articulation of a fortis plosive that is not preceded by /s/, closing of the glottis commences only after the release of the plosive. The lack of aspiration on fortis plosives preceded by /s/ is thus explained on physiological as well as phonological grounds, rather than by invoking language-specific rules for the distribution of aspiration. Because this approach does justice to phonological differences between voicing and aspiration languages, and "makes clear the true nature of possible laryngeal contrasts", it has been named Laryngeal Realism (LR) by Honeybone (2002, p. 123).

Avery and Idsardi (2001) have developed the LR framework further, by grouping several features under laryngeal dimensions. Instead of using only the features [voice], [spread glottis] and [constricted glottis], like in traditional LR approaches, Avery and Idsardi (2001) use six different (privative) features: [spread], [constricted], [stiff], [slack], [raised] and [lowered]. The features [spread] and [constricted] are grouped under the Glottal Width dimension; [stiff] and [slack] under the Glottal Tension dimension and [raised] and [lowered] under the Larynx Height dimension. The features "are to be interpreted as motor instructions to the articulators. Thus, they share much in common with the *gestures* of Brownman and Goldstein (1989), [...]". Instead of the features, "only dimensions can be contrastive in obstruents". A voicing contrast is thus not represented as a contrast between [stiff] (plain voiceless) and [slack] (voiced) plosives, but rather as a contrast between [∅] (plain voiceless) and Glottal Tension (GT; voiced) plosives; an aspiration contrast is represented as a contrast between [∅] (plain voiceless) and Glottal Width (GW; aspirated) plosives. The dimensions are, on the phonological level, not specified for a gesture. The gestures are added on a later level (via a process called completion), and every dimension has a universal, default gesture that is added during completion. For GT this gesture is [slack], for GW this gesture is [spread]. [stiff] and [constricted] can also be added to GT and GW respectively, by means of a process called enhancement. As this process is not relevant to the present discussion, it will not be further discussed.

An approach that is very similar to LR is an element-based approach (Kaye, Lowenstamm and Vergnaud (1985), J. Harris and Lindsey (1995), Backley (2011)). Rather than using phonological features, the approach represents phonemes by combining different elements. While an individual phonological feature is not pronouncable without the support of another feature, an element is. Dependency Phonology (Anderson and Ewen (1987), Van der Hulst (2006)) is an early example of a framework using elements rather than features (although Anderson and Ewen (1987) refer to elements as 'components'). In this approach the element |V|[8] is "a component which can be defined as 'rel-

[8] Elements are always placed between vertical brackets.

atively periodic'" (Van der Hulst (2006, p. 455)); in other words, |V| is the element indicating (contrastive) voicing. The element |O| represents 'glottal opening' (Van der Hulst (2006, p. 457)), and is used "in three-way oppositions of phonation-type, in a voicing opposition amongst sonorants and in the distinctive use of aspiration" (Anderson and Ewen (1987)). The contrast between voicing and aspiration languages is thus visible in the presence of a different element: |V| in voicing languages and |O| in aspiration languages.[9]

In a slightly more recent variant of Element Theory (as commonly used in Goverment Phonology, cf. Kaye, Lowenstamm and Vergnaud (1990); see also Backley (2011)), the two relevant elements for laryngeal distinctions are |N| (or |L|) and |H|. In isolation they are realised as a voiced nasal[10] and a voiceless glottal fricative respectively. Just like features, elements can be combined into complex representations, but unlike features they do not have to be combined. In the case of plosives, a feature |ʔ| is added to indicate an "oral or glottal occlusion" (Backley and Nasukawa (2010)), and an element |I|, |U| or |A| represents the place of articulation (in the case of an oral occlusion; for glottal stops no place element is present in the representation). A contrastively voiced plosive has the element |N| (or |L|) in its representation; a contrastively aspirated plosive has the element |H| in its representation. Plain voiceless plosives (whether contrasting with voiced plosives, aspirated plosives, both plosives or neither plosive) do not have a laryngeal element in their representation. The representation in Element Theory is thus very similar to a featural representation using the features [voice] and [spread glottis].

2.2 The Low Saxon dialect continuum

In LR the representation of voicing languages is structurally different from the representation of aspiration languages. In a contact situation between the two different languages, this may lead to interesting linguistic patterns as speakers may try to combine characteristics from both systems into one. This can yield new insights in linguistic variation. In this chapter an area with contact between the two language types will be studied, showing how the transition between the two systems is implemented. Ideally, this transition is studied within a dialect continuum, so that other grammatical characteristics are as similar as possible.

The Germanic languages are mostly characterised by aspiration languages, but Dutch, Afrikaans and Yiddish are exceptions to this and have a voicing contrast (e.g. Iverson and Salmons (1995), Jansen (2004)). Afrikaans is mostly spoken outside of Europe (Ethnologue (2005)), and can thus not be found in a dialect continuum with an aspiration language. The Yiddish language came into existence when part of the Jewish population migrated to the Rhineland

[9]The authors do not discuss how the elements would be realised in isolation.

[10]|L| is argued to represent only voicing in isolation, not nasality (Kaye et al. (1990)). The authors do not propose a specific realisation of the element in isolation (in a consonant position), but one can imagine it would be realised as a lateral because of the label itself.

area: these speakers already spoke both Hebrew and Laaz ("a Jewish language of Romance stock"; Weinreich (1959)) and came into contact with speakers of Germanic varieties. The Yiddish language might thus show remnants of other, very different languages, and cannot be found in a dialect continuum with an aspiration language either. Around the Dutch-German border, however, Dutch is in a contact situation with German. Dialects spoken around this border do not form clearly separated dialect areas, but rather form a continuum. The area of the Rhenish Fan, which has been discussed in the previous chapter, is such an example. Another example, which has not yet been discussed, is the Low Saxon dialect continuum. The Low Saxon dialects (which belong to the West-Germanic branch of the Germanic languages) are located in the north of the Netherlands and Germany, ranging from the provinces of Overijssel, Groning and Drenthe in the Netherlands to the provinces of Niedersachsen, Schleswig-Holstein, Bremen, Hamburg, Mecklenburg-Western Pommerania, Nordrhein-Westfalen and Brandenburg in Germany (Lindow et al. (1998), Matras and Reershemius (2003), Prehn (2011)). A distinguishing characteristic of these dialects is the absence of the second High German consonant shift (Lindow et al. (1998)), which changed word-initial plosives, geminated plosives and plosives following a liquid or nasal consonant into affricates. An example is the word for 'horse', which is *Pferd* /pfɛrt/ in High German varieties but is realised as /pɪːrt/, i.e. without affrication of the word-initial plosive, in Low Saxon.

The absence of affrication does not distinguish the Low Saxon varieties spoken in the Netherlands from other dialects spoken in the Netherlands, as the second High German consonant shift did not take place in the dialects spoken in the Netherlands. A feature that distinguishes the Low Saxon dialects spoken in the Netherlands from other dialects spoken in the Netherlands is the vocalisation of /ɔl/ and /ɑl/ to the diphthongs /ou/ and /au/ when the vowel-lateral cluster was followed by an alveolar plosive. This change did occur in the Franconian dialects, but the Low Saxon dialects were not affected.

In this study the transition between voicing and aspiration languages in the Low Saxon dialect continuum will be studied. Different studies have shown that the two standard languages involved, standard Dutch and standard German, differ in their realisation of the laryngeal contrast: in standard Dutch this distinction is realised as a voicing contrast (Slis and Cohen (1969), Van Alphen (2007)), while in standard German it is realised as an aspiration contrast (Jessen (1998)). This difference is not restricted to the standard languages, but also distinguishes the different dialects in the Low Saxon dialect continuum. The varieties spoken in Germany are reported to have an aspiration contrast (Braun (1996)), while the varieties spoken in the Netherlands are reported to have a voicing contrast (Jansen (2004), but see Pinget (2015a)). As the Low Saxon dialects form a continuum (Heeringa, Nerbonne, Niebaum, Nieuweboer and Kleiweg (2000)), and as both contrast types are present in the continuum, it forms an ideal region to study the transition between the two systems.

One might argue that the country border, which divides the continuum into two parts, marks the presence of two different standard languages and has its origins already in the early Middle Ages (De Vrankrijker (1946)), is likely to have an influence on the realisation of the contrast: the dialects spoken in the Netherlands might be oriented towards standard Dutch, and the dialects spoken in Germany might be oriented towards standard German. Further, dialect speakers from the two countries might not be in contact with each other because of the country border. However, there is a lot of contact and movement of speakers across the border. The presence of the country border does thus not necessarily have a problematic influence on the existence of the continuum (cf. Giesbers, Van Hout and Van Bezooijen (2005), who argue that the country border between the Netherlands and Germany in the Kleverland area is mostly a political border).

A final important remark needs to be made about Frisian. The present study includes several Frisian dialects, so that the Netherlandic part of the continuum would contain more villages. As the Frisian dialects have been reported to have a voicing contrast (Jansen (2004)) instead of an aspiration contrast, these language varieties will be included in the present study. The results in Section 2.5 will show that there is indeed no difference between Frisian and Low Saxon varieties.

2.3 Possible scenarios

Although the behaviour of individual voicing and aspiration languages have been very well studied, contact areas between the two language types have not. The behaviour of languages in such contact areas, where speakers receive input from both language types, can be of much interest to the question of the phonological representation of the contrast, as the different approaches make different predictions.

In the traditional, [(±)voice], approach, there is no phonological distinction between the two language types. The only difference is one of phonetic implementation: language-specific phonetic rules guide the realisation of [+voice] as either prevoiced or plain voiceless, and of [−voice] as either plain voiceless or aspirated. The features are thus not universally linked to specific phonetic values, and any phonetic implementation rule should in theory be possible. Although this does not exclude an abrupt switch between voicing and aspiration languages, it is not unrealistic to expect a phonetically gradual transition between the two systems if the relevant feature distinction is [±voice] in both language types. From west (voicing) to east (aspiration), VOT values are then expected to gradually increase, without showing a clear boundary between the two systems.

If phonological features (or elements) are directly linked to specific phonetic implementations, as is assumed in LR and Element Theory, phonetic variation within each category is possible, but a consistently gradual phonetic transition is less likely. Rather, one would expect one of three things: a) an intermediate area where characteristics of both voicing and aspiration languages are found[11]; b) an intermediate area where characteristics of neither system are found; c) an abrupt change between the two systems (i.e., no clear transitional area exists). Because of language contact between the two systems, an intermediate area between the two systems does not seem unlikely. This intermediate area might then show characteristics of both systems, such that some lenis plosives are prevoiced (as in a voicing systems) while others are plain voiceless (as in an aspiration system), and that some fortis plosives are plain voiceless (as in a voicing system) while others are aspirated (as in an aspiration system), or an intermediate area where lenis plosives are prevoiced and fortis plosives aspirated (cf. Beckman et al. (2011)).[12] Another possible effect of contact between the two systems, however, is an intermediate area without any characteristics of either system. Such a transition zone would lack an laryngeal opposition in its plosive inventory. This is not an unexpected scenario and has been described by Vietti, Alber and Vogt (2018) for Tyrolean dialects, where the laryngeal contrast between voiced and plain voiceless plosives is neutralised in initial position.

The link between phonological features and specific phonetic values does not exclude phonetic variation, not even in the three scenarios outlined above. If features and elements are linked to specific phonetic values, there must be a cut-off point between phonetic values that count as prevoiced and phonetic values that count as aspirated, but between the cut-off points of each category phonetic variation should be freely possible. Although it seems likely that if phonetic variation were present, it would be between the cut-off points of each category, a phonetically gradual increase in VOT values is not entirely impossible. If there is phonetic variation around the cut-off points, so that one speaker will have a phonetic realisation that only just falls in one category (e.g., prevoiced), while the next speaker will have a phonetic realisation that only just falls in the other category (e.g., plain voiceless), the increase in VOT values might still appear gradual.

The predictions about assimilation patterns made by LR are quite clear: only the phonological feature that is present in the language can be active in phonological processes. In a voicing language, assimilation can thus only

[11]Note that this scenario shows similarities to patterns found in the wave model (e.g. Chambers and Trudgill (1980, p. 166)), where innovations spread from one place to another in a wave ("The 'wave model' visualised innovation diffusion as a pebble-in-pond effect, with a centre of influence (the point of impact of the pebble) sending ripples outwards in all directions (the movement of the wave)." (Chambers and Trudgill (1980, p. 166)).

[12]An example of dialects that exhibit characteristics of two different systems are several dialects spoken in Twente (a region in the east of the Netherlands). In several dialects of Twente the first person singular of verbs is realised with a final schwa (*ik gelo[və/]* 'I believe'). In another group of dialects this final schwa is deleted, creating a context in which Final Devoicing can (and does) apply (*ik gelo[f]* 'I believe'). A small subset of the dialects in which the final schwa is deleted, however, does not apply Final Devoicing in that context. Speakers will thus realise the final fricative in these words as voiced, thereby combining the characteristics of the two systems (Schoemans and Van Oostendorp (2004), Van Oostendorp (2007)).

be spreading of [voice], while in an aspiration language assimilation can only be spreading of [spread glottis]. A segment which is unmarked for laryngeal characteristics, i.e. [∅], can never spread its laryngeal feature as it does not have one. In voicing languages assimilation can thus never be to the fortis member of the cluster, while in aspiration languages assimilation can never be to the lenis member of the cluster. In a [±voice] approach, however, both [+voice] and [-voice] can be active, so the prediction is that both will be, regardless of the realisation of the contrast.

2.4 Methods

The present study focusses on the realisation of laryngeal distinctions in the Dutch-German dialect continuum. This area was chosen for several reasons. First of all, the presence of both voicing and aspiration systems in the continuum has been confirmed in the literature (Braun (1996) for German dialects, Jansen (2004) for Dutch dialects). Second, the standard languages of the two countries have different laryngeal distinctions: standard Dutch is a voicing language (Van Alphen and Smits (2004), Van Alphen (2007)), while standard German is an aspiration language (Jessen (1998), Jessen and Ringen (2002)). Thirdly, many dialect studies have been conducted in the West Germanic language area, including recordings of spoken dialect. The data used in this study have been taken from several of these databases. The data for the Dutch side of the continuum come from the database for the GTRP, which forms the basis for both the Phonological (Goossens et al. (1998), Goossens et al. (2000), De Wulf et al. (2005)) and Morphological (De Schutter et al. (2005), Goeman et al. (2008)) Atlases of the Dutch Dialect. The database contains recordings of informants from over 600 locations in the Netherlands and Flanders, translating a list of words and sentences into their local dialect. The recordings are all available at the Meertens Instituut, and for most interviews transcriptions are available. Both the transcriptions and recordings are easily searchable online (`http://www.meertens.knaw.nl/mand/database/`). The German data come from several different databases (Regionalsprache.de (REDE), Niedersächsisches Dialektarchiv (NSD), MR Phonetisch-phonologischer Atlas von Deutschland (MRPAD) and MR Deutsche Dialekte (MRDD)), and can be found at `https://regionalsprache.de/`.[13] All German data are based on the Wenker sentences (see Appendix B), and contain recordings of informants translating the phrases into their local dialect. The questionnaire used for the GTRP database can be found at `http://www.meertens.knaw.nl/projecten/mand/GTRPlijstsec.html`.

The data from the GTRP have been collected between 1981 and 1990, the German data between 1956 and 1987. Speakers' age ranges between 47 and

[13] I want to express my gratitude to the Forschungszentrum Deutscher Sprachatlas and Regionalsprache.de, who kindly made their data available for my research.

82 years old (Netherlands) and between 21 and 77 years old (Germany).[14] All informants (both male and female speakers were included) were dialect speakers, and most likely also spoke standard Dutch resp. standard German to at least a certain degree as the questionnaires were written in the respective standard languages. However, standard language proficiency is not documented, and because most interviews were conducted in the local dialect, standard language proficiency is not known with certainty. The recording location of the interviews is not documented, but interviews appear to be conducted at informants' houses.[15] Interviews in the Netherlands were recorded on tape (for most locations a Sennheiser microphone was used; interviews were recorded with a Nagra recorder in earlier stages and with a Uher recorder in later stages), and later digitised to a BWF file (Broadcast Wave Format) with a sampling frequency of 96 kHz, 24 bit. These files were subsequently converted to MP3-files, which are available for research at the Meertens Instituut.[16]

The choice of locations has been influenced primarily by data on the Netherlandic part of the continuum. Standard Dutch, and most dialects spoken in the Netherlands and Flanders, are voicing languages. However, dialects in the north-east of the country, the province of Groningen, are commonly known as aspirating dialects: when non-dialect speakers try to imitate the dialect of Groningen they typically use heavily aspirated plosives (see e.g. Bloemhoff, der Kooi, Niebaum and Reker (2008), p.163). The dialect has also been described as aspirating by e.g. Jansen (2007). Because the dialects in Germany were expected to show aspiration (see for example Braun (1996)), and no mention of voicing in those dialects is made in the literature, the presence of aspiration in Dutch dialects was leading in the choice of locations: voicing can quite certainly be expected in the Netherlands, so if the transition zone includes Dutch dialects with phonetic aspiration, it is quite certain that the entire transition zone will include both voicing and aspiration languages. Further, if some traces of aspiration can already be found in the Netherlands, it is unlikely that the linguistic border will simply coincide with the country border. A final argument is that the transcriptions of the GTRP are transcribed in a very systematic manner, with aspiration being indicated with a ⟨h⟩ following the plosive the transcriber interpreted as being aspirated, while such transcriptions are not available for the recordings from Germany. The precise location of the transition zone was based on a first investigation of the Dutch transcriptions.

On the website `http://www.meertens.knaw.nl/mand/database/` it is possible to search in the transcriptions of the interviews using the International Phonetic Alphabet (IPA) or regular expressions. Using the regular expression ⟨p⟩, ⟨t⟩ or ⟨k⟩ followed by ⟨h⟩, the search function returns any transcription

[14] Not all recording dates and informants' ages are documented.

[15] In some cases domestic sounds like the ticking of a clock, a pet, someone doing the dishes, etc., can be heard in the background, indicating that the interviews were indeed conducted at the informants' house.

[16] I want to thank Kees Grijpink from the Meertens Instituut who gave me the relevant information on the digitisation procedures.

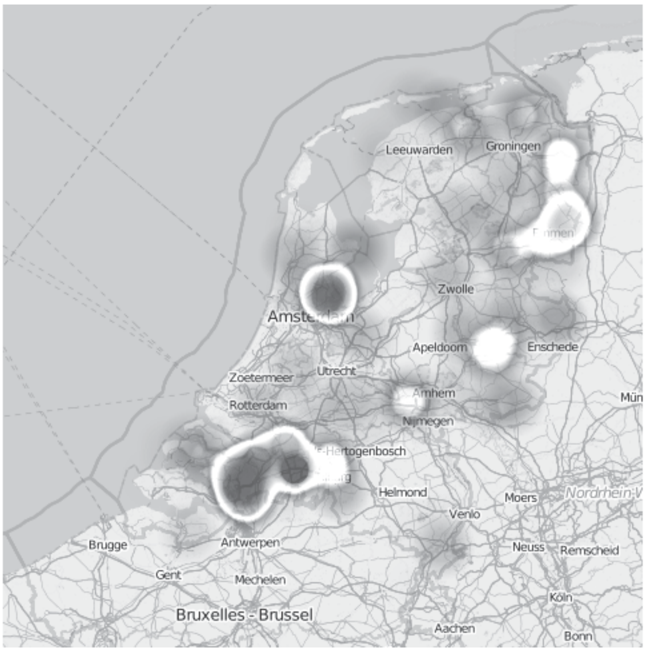


Figure 2.1: Concentration of aspirated ⟨p⟩, highest percentage (Hollandsche Veld): 83.58%. Darker colours represent higher concentrations of aspiration. Map created by Erik Tjong Kim Sang, maps ©OpenStreetMap contributors.

where a fortis plosive is followed by an ⟨h⟩. The expression has to be preceded as well as followed by an asterisk to indicate that any number of graphemes (including zero) can precede and follow ⟨ph⟩, ⟨th⟩ or ⟨kh⟩. The command thus returns any transcription containing ⟨ph⟩, ⟨th⟩ or ⟨kh⟩, regardless of the plosive's position in the transcription. From the list of results, syllable-final plosives and plosives followed by the phoneme h (i.e. where ⟨h⟩ does not represent aspiration but rather a phoneme itself, as in *het huis* 'the house') were excluded. Subsequently the ratio of aspirated syllable-initial plosives to the total number of syllable-initial plosives per location was calculated, giving an overview of which locations show more aspiration than others. These ratios were then plotted in heat maps, which show where concentrations of locations with relatively many aspirated plosives can be found. The lowest ratios of aspiration are represented by the blue shades; the highest ratios are represented by red shades.[17]

Figures 2.1 to 2.3 reveal that the dialects spoken in Groningen indeed have relatively much aspiration, but so do the dialects spoken in the province of

[17] As the heat maps represent concentrations of locations with high numbers of aspirated plosives, the reddest colour spot on the map does not immediately correlate to the location with the highest percentage of aspiration.

Figure 2.2: Concentration of aspirated ⟨t⟩, highest percentage (Ulft): 71.875%. Darker colours represent higher concentrations of aspiration. Map created by Erik Tjong Kim Sang, maps ©OpenStreetMap contributors.

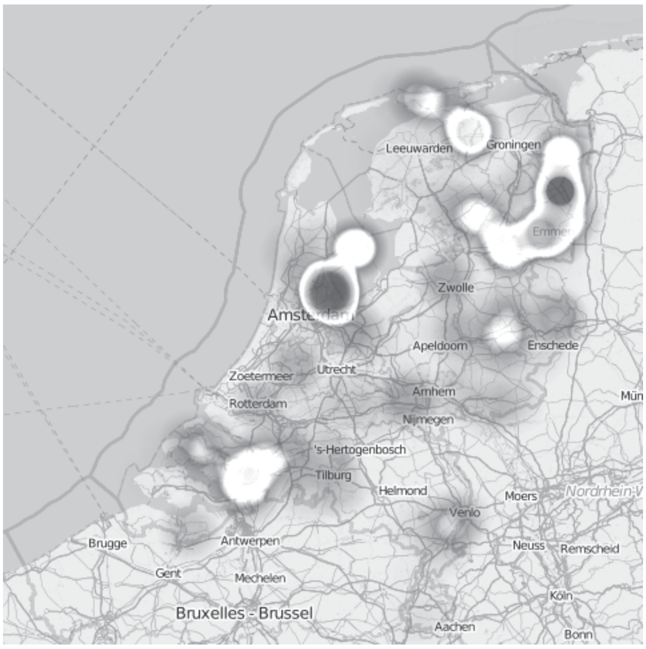


Figure 2.3: Concentration of aspirated ⟨k⟩, highest percentage (Hollandsche Veld): 61.76%. Darker colours represent higher concentrations of aspiration. Map created by Erik Tjong Kim Sang, maps ©OpenStreetMap contributors.

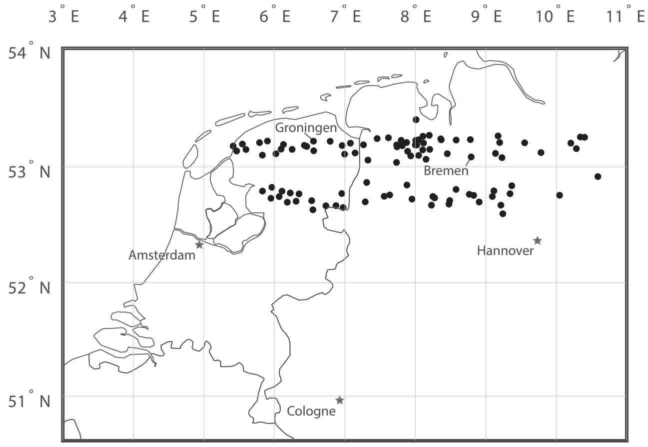


Figure 2.4: Test locations

Drenthe (to the south of Groningen).[18] Because aspiration is, to some extent, present in the north-east for all three plosives, this region was chosen as the start of the transition zone. The areas more to the west (Fryslân and the west of Drenthe) are also included, to make sure that languages with a voicing system are present in the continuum as well. Within this entire area locations on two horizontal lines were selected (including all locations in the entire area would be too time-consuming): one in the very north of Groningen and Fryslân and one in the south of Drenthe. The choice of locations was based on the locations that were included in the GTRP database. Subsequently, the two regions were extended into Germany (where, again, the choice of locations depended on the locations that were included in the databases). Figure 2.4 shows all locations included in the study (a list of all locations is included in Appendix A).

In this study the five plosives ⟨b, d, p, t, k⟩[19] have been included. ⟨g⟩ is not included because it is not part of the phoneme inventory in all included dialects: it is absent from many varieties spoken in the Netherlands. For every location, per plosive at least one item with that plosive in word-initial position has been selected. The database for the Netherlands contained enough material to select two items per plosive (one where the plosive is followed by a vowel and one where the plosive is followed by a sonorant consonant (a trill, lateral

[18]Note that the areas around Amsterdam and the province of Zeeland show much aspiration as well. While this is very interesting, these regions do not form a transition zone with true aspirating languages, so they are not included in the present study.

[19]The plosives are represented with their orthographic labels here instead of their phonemic labels: in all dialects, the bilabial and alveolar lenis plosive are represented by, respectively, the graphemes ⟨b, d⟩, and the bilabial, alveolar and velar fortis plosives are represented by, respectively, the graphemes ⟨p, t, k⟩. However, the phonology of the plosives differs within the continuum, as there is an (expected) change from prevoiced vs. plain voiceless (/b, d/ vs. /p, t, k/ respectively) to plain voiceless vs. voiceless aspirated (/p, t, k/ vs. /p^h, t^h, k^h/ plosives.

or glide)), but the databases for Germany were quite scarce, so that only one item per plosive could be selected (the plosive is followed by either a vowel or a sonorant consonant (a trill, lateral or glide)). Further, for every location and for every plosive an item with the plosive in intervocalic position was selected. For the Dutch language area enough data were found in the GTRP database, but for the German area it was not possible to collect many data. Finally, for every location at least one item with a potential assimilation context (either a lenis or a fortis first plosive, and a lenis second plosive) was selected. Where possible another item with a fortis second plosive was selected.[20] All measurements are carried out in Praat (Boersma and Weenink (2016)).

A small remark on why past tense formation is not used to study the laryngeal characteristics of the dialects included is in order here. In many Germanic varieties this suffix is either (orthographic) -de or -te, with a voicing value corresponding to the voicing value of the last segment of the stem: if this segment is a voiceless obstruent, the suffix is -te; if this segment is a voiced obstruent, a sonorant or a vowel, the suffix is -de. However, in many Low Saxon dialects (Groningen, Drenthe, Salland, Achterhoek, Twente, East Fryslân, Emsland, Münsterland (Roos (2009))) only the suffix -de is used. Instead of the suffix changing voicing values based on the final segment of the stem, the final segment of the stem changes voicing values under the influence of the past tense suffix: if the final segment is a vowel, sonorant or lenis obstruent no changes take place, but a final fortis obstruent undergoes RVA and surfaces as lenis. As this type of assimilation can only occur in languages with a voicing contrast, as argued in Section 2.1, the past tense formation seems a good indication of the phonological contrast underlying the laryngeal distinctions in the continuum. However, the varieties spoken in Bremen-Oldenburg, Hamburg, Holstein and Mecklenburg-West Pomerania no longer have a past tense suffix, as this suffix was deleted (Roos (2009)). In these dialects the past tense formation thus cannot be used to study laryngeal contrasts. Although for the more western dialects past tense formation could be insightful, the process cannot be used across the entire continuum. It will therefore not be used in this study.

For word-initial plosives VOT is measured as the difference between the moment of plosive release and the moment of voice onset. When voice onset precedes plosive release the resulting value is negative, indicating prevoicing; when voice onset follows plosive release the resulting value is positive. Voice onset is placed at the first visible wave in the waveform, corresponding with the presence of the voice bar in the spectrogram. Every wave, however small, is accepted as voicing. Plosive release is placed at the moment of a sudden change in amplitude in the waveform, corresponding with the presence of the release burst in the spectrogram. When two releases were visible, the first release was used as the reference point.

For intervocalic plosives the percentage of voicing during closure is mea-

[20] 201 tokens with a lenis C2 were selected for 111 speakers in 101 locations; 118 tokens with a fortis C2 were selected for 90 speakers in 82 locations.

sured. The onset of the closure is placed at the point where the waveform showed a decrease in amplitude of the preceding vowel, combined with the absence of the higher formants in the spectrogram (Rietveld and Van Heuven (1997), Jessen (1998), Kulikov (2012)). The offset of the closure was placed at the point where a sudden increase in amplitude was visible in the waveform, again combined with the presence of the black vertical bar in the spectrogram. The onset of voicing during the closure corresponded with the onset of the closure, while the offset of voicing was placed at the last visible voicing cycle during the closure. If voicing recommenced during the closure, the second voicing onset was placed at the beginning of the first visible voicing cycle, while the second voicing offset corresponded with the closure offset (plosive release). Voicing in the waveform always corresponded with the presence of the voicing bar in the spectrogram. The percentage of voicing during closure is calculated by dividing the duration of voicing by the total duration of the closure (if voicing was interrupted, i.e. voicing ended during the closure but recommenced later, the percentage of voicing was based on the duration of both voicing periods together).

The assimilation measurements are identical to the measurements for intervocalic plosives, as they measure the percentage of voicing during a closure. The onset of the closure is thus placed at the point where the waveform showed a decrease in amplitude of the preceding vowel, combined with the absence of the higher formants in the spectrogram; the offset of the closure is placed at the time of the release of the second plosive (usually only one plosive release was visible; but this release always corresponded to the second plosive. If two plosive releases were visible, closure offset was placed at the second plosive release). Voicing offset again corresponded to the last visible cycle in the waveform combined with the offset of the voice bar in the spectrogram, and voicing onset (if voicing recommenced during the closure) corresponded to the first visible cycle in the waveform combined with the presence of the onset of the voice bar in the spectrogram. The percentage of voicing was again calculated by dividing the total duration of voicing by the total duration of the closure.

2.5 Results

2.5.1 Initial plosives

Voice Onset Time

Figures 2.5 to 2.9 plot the VOT values of word-initial <b, d, p, t, k> (respectively) against the geographical longitude coordinates (the west-east dimension) of each location. All plots show a clear distribution of VOT values, where lower values are concentrated in the west and higher values in the east. The plots for <b> and <d> show an abundance of prevoicing in the western part of the continuum and scarcity of prevoicing in the east. The middle area, for both

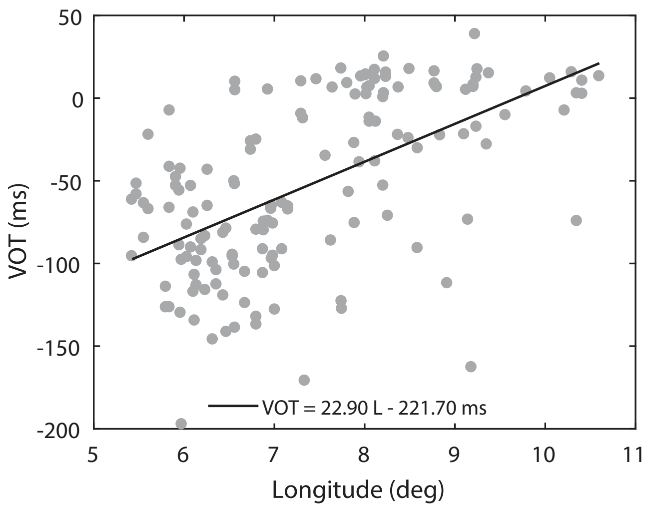


Figure 2.5: VOT values of <b>, r = 0.575, p=0.000*

plosives, shows both prevoiced and plain voiceless plosives.

The picture for the fortis plosives differs slightly from the lenis plosives. When the boundary between plain voiceless and voiceless aspirated plosives is set at a VOT of around 30 ms (following Keating (1984)), the east shows mostly aspirated fortis plosives, while the rest of the continuum shows both plain voiceless and voiceless aspirated plosives. A middle and western area can thus not be clearly distinguished.

All plots seem to show a gradual increase in VOT values, with lower values concentrated in the west and higher values concentrated in the east. To test if this seeming gradualness can be confirmed, the statistical correlation between VOT values and longitude coordinates is calculated using a Pearson product-moment correlation coefficient. The results are given in the captions to each figure, showing for each plosive a significant positive correlation between longitudes and VOT values.

Chambers and Trudgill (1980) discussed the distinction between mixed, fudged and scrambled lects. A mixed lect in the transition zone between voicing and aspiration systems would use both systems simultaneously. More specifically, speakers would thus use both prevoiced and plain voiceless realisations for lenis plosives, and both plain voiceless and aspirated realisations for fortis plosives. In the case of fudged lects, speakers would have to find a midway realisation in VOT values for both lenis and fortis plosives. The difficulty in the latter scenario is that a realisation midway between prevoiced and plain voiceless is still a prevoiced plosive, only with shorter prevoicing than the ‘original’ plosive (i.e., the plosive upon which the midway realisation is based). Also, VOT is a highly variable characteristic, as the length in milliseconds (ms) is

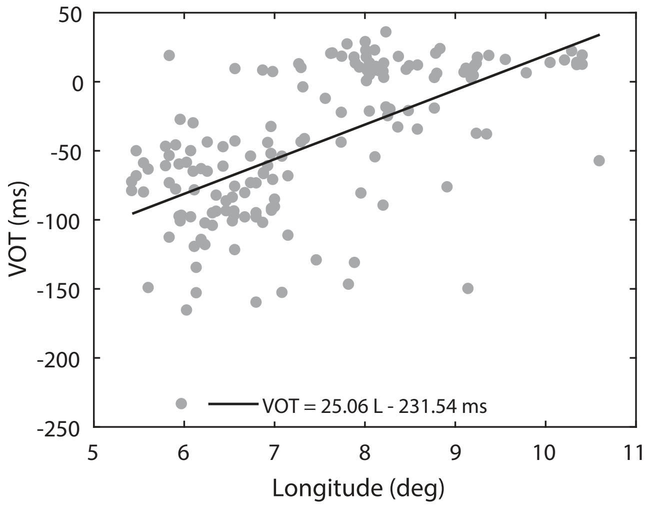


Figure 2.6: VOT values of <d>, r = 0.604, p=0.000*

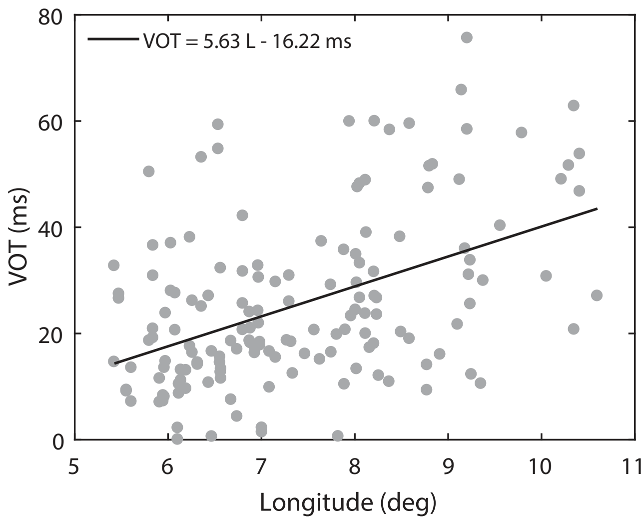


Figure 2.7: VOT values of <p>, r = 0.464, p=0.000*

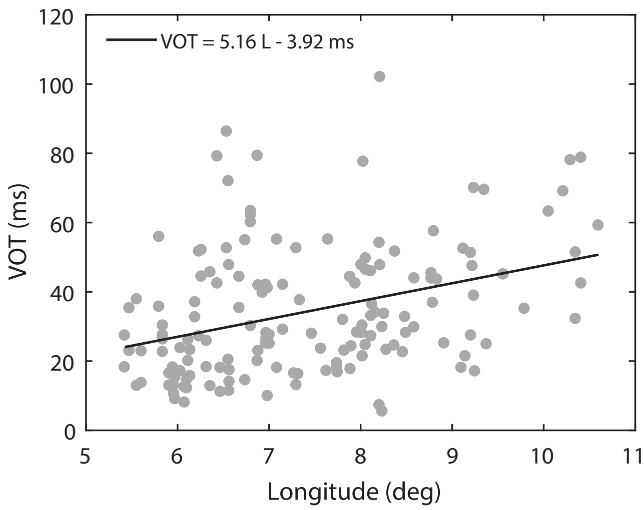


Figure 2.8: VOT values of <t>, r = 0.356, p=0.000*

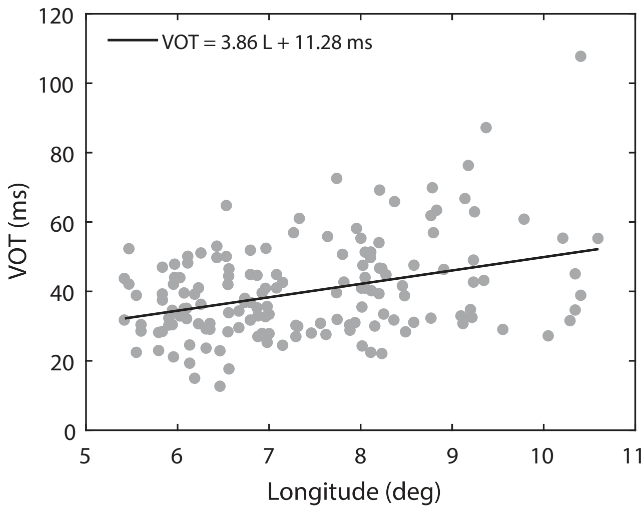


Figure 2.9: VOT values of <k>, r = 0.357, p=0.000*

not perfectly controllable for a speaker. It is thus difficult to determine if a speaker uses a mixed or a fudged system (or of course a scrambled system). An intermediate VOT value for a lenis plosive (i.e. prevoiced, but not with as much prevoicing as a lenis plosive in the stable voicing area) might be the result of fudging, but it might also be a value occurring in a mixed lect which is shorter because of physiological reasons or it might simply be shorter by coincidence. The same can be argued for fortis plosives. Because the boundary between plain voiceless and aspirated plosives is quite unclear, it is not very well possible to determine to which category a voiceless plosive belongs, and thus it is difficult to determine if a plosive is a mix of plain voiceless and voiceless aspirated (and thus characteristic for a fudged lect), or if it should belong to either one of the two categories rather than being an intermediate realisation (and thus characteristic of a mixed lect). However, clearly prevoiced, plain voiceless or aspirated plosives can still be of some value, as they can show the existence of a mixed lect. Lenis plosives with more prevoicing than found in the stable area, fortis plosives with more aspiration than found in the stable area, and plain voiceless plosives with only very small positive VOT values can be seen as evidence of a mixed lect, as they are clearly not intermediate values.

Looking at the data, it is evident that a number of the dialects are mixed. One speaker can realise lenis plosives with either prevoicing or short-lag VOT, and can realise fortis plosives with either short-lag VOT or aspiration. For example, there are 25 instances in which one speaker realises a <b> with prevoicing and a <d> with short-lag VOT or vice versa.[21] The existence of fudged lects is more difficult, as explained above. However, the data show that in the transition zone a number of speakers have a prevoiced realisation of the lenis plosives, but the duration of the prevoicing is quite short. As the occurrences of a shorter duration of prevoicing are all clustered in the same region (the middle area), this may be an indication of the presence of fudged lects in the transition zone.

Fundamental frequency

Another characteristic that is linked to VOT values is the f_0 value of the vowel following the plosive. As the position of the larynx at the time of plosive release differs depending on the intended VOT, the fundamental frequency of the following vowel varies along with it. Voiced plosives commonly cause a lower fundamental frequency on the following vowel, while voiceless plosives commonly cause a higher fundamental frequency on the following vowel. This is accompanied by a difference in f_0 movements: the lower f_0 immediately following a prevoiced plosive is followed by a rise in the f_0, while the higher f_0 immediately following a (plain or aspirated) voiceless plosive is followed by a

[21] The locations in which speakers have different realisations of lenis plosives are: Barssel, Blockwinkel, Eelde, Filsum, Hellwege, Jeddeloh II, Kuinre, Lauenbrück, Leer, Lohne, Mallinghausen, Neuenkruge, Nieuw Schoonebeek, Ocholt, Osteressen, Ostertimke, Putensen, Reppenstedt, Saterland, Schwanewede, Stederdorf, Süddorf, Veendam, Wagenborgen and Warpe.

fall in the f_0. This link between the two phonetic variables can be the cause of tonogenesis: when the phonetic distinction between voiced and voiceless plosives undergoes neutralisation, the phonetic distinction between higher and lower fundamental frequency of the following vowel may be phonologised (see Yip (2002)). Yip (2002) does not specify the phonetics of the voicing contrast, but since she mentions that the position of the larynx is lower and the vocal folds are slacker, it can be assumed that this is not only a phonological voicing distinction but also a phonetic voicing distinction. The effects of an aspiration contrast on fundamental frequency are not very well-studied, but see Lai, Huff, Sereno and Jongman (2009) for an overview of the relevant literature: "It is generally assumed that f_0 after voiceless aspirated stops is higher than after voiceless unaspirated stops [...]" (p. 2). It remains unclear, however, what the effects of an aspiration contrast on f_0 are.

Even though VOT values of word-initial plosives show a lot of variation, f_0 values might show a more consistent pattern and shed more light on laryngeal distinctions. For every item the onset of the vowel or sonorant consonant following the plosive was marked in a Praat TextGrid. The onset of the vowel or sonorant-vowel was placed either at the onset of voicing (in case of a voicing lag) or at the closure release (in case of a voicing lead). Vowel offset was placed at the point where the formants were no longer visible. A Praat script was used to extract f_0 values for the entire duration of the vowel, per millisecond.

Several problems were encountered in the f_0 analysis. First, different authors have used different time frames for the pitch analysis (pitch analysis per millisecond or per 10 ms). In this study pitch was analysed per millisecond because of the short duration of some vowels: analysis per millisecond was a better representation of pitch movements than analysis per 10 ms. Second, pitch movements were often much more complicated than pictured in the literature. It often occurred that the consonant was immediately followed by a relatively short rising movement, followed by a falling movement that lasted much longer (or vice versa). It remains unclear which of these movements is more important: the first, short, rising movement, or the second, longer, falling movement? The literature is not informative on this question either.

The results of these f_0 measurements were analysed in several ways, but none of them proved to be very insightful. Both the absolute f_0 values at the onset of the vowel as well as the movement of the fundamental frequency in the vowel (rising or falling) are set out against the longitude coordinates of each location, and both values (absolute f_0 values and f_0 movements) are compared to the VOT values of each plosive.

As mentioned, a link between f_0 values and VOT values is expected. Prevoiced plosives are expected to have a lower f_0 value at the vowel onset, while plain voiceless plosives are expected to have higher f_0 values at the vowel onset. Voiceless aspirated plosives are expected to show even higher f_0 values than plain voiceless plosives. The graphs in Figures 2.10 to 2.14, however, do not clearly show higher f_0 values being linked to higher VOT values. An overlap plot, combining the f_0 value at the vowel onset for lenis and fortis plosives plot-

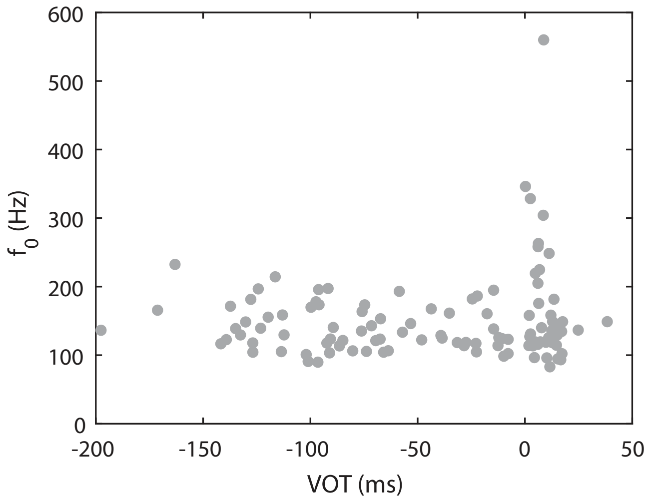


Figure 2.10: f_0 values at vowel onset linked to VOT (<b>)

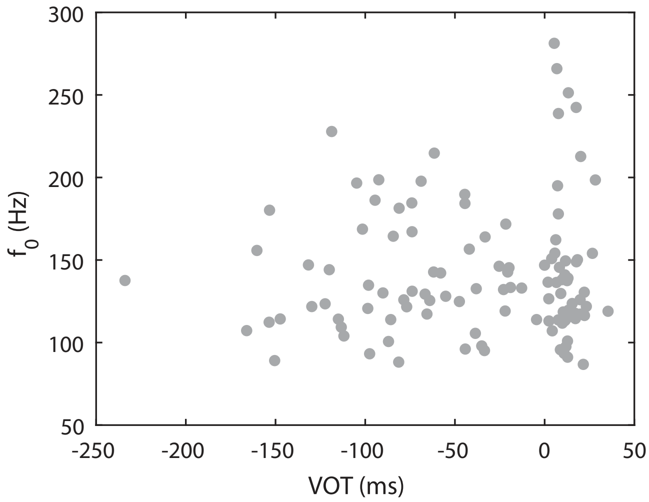


Figure 2.11: f_0 values at vowel onset linked to VOT (<d>)

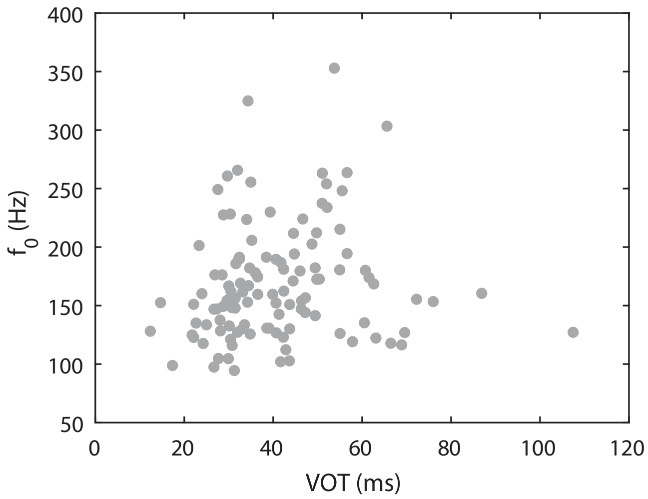


Figure 2.12: f_0 values at vowel onset linked to VOT (<k>)

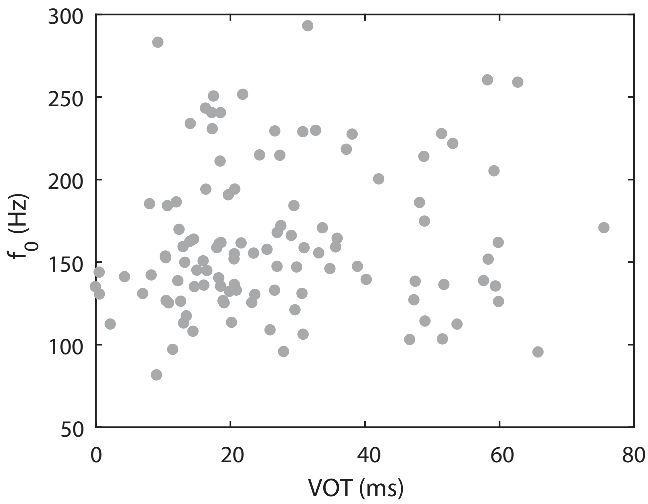


Figure 2.13: f_0 values at vowel onset linked to VOT (<p>)

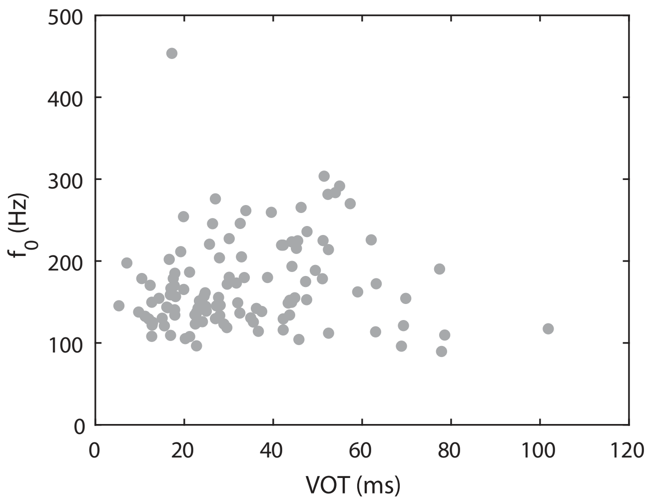


Figure 2.14: f_0 values at vowel onset linked to VOT (<t>)

ted against the longitude coordinates of every location, might show a pattern: as fortis plosives are expected to have higher f_0 values after release than lenis plosives, a separation of f_0 values for lenis and fortis plosives is expected.[22]

The graphs show that speakers in the entire continuum have similar f_0 values for lenis and fortis plosives. A clear distinction between the two categories is not made. The absolute f_0 value at vowel onset is thus not a reliable identifier for laryngeal characteristics, at least not for these data.[23]

The graphs in Figures 2.17 to 2.21 show the first f_0 movements of every vowel or sonorant consonant plotted against VOT values. A falling f_0 movement (negative value) is expected for a positive VOT value, while a rising f_0 movement (positive value) is expected for a negative VOT value. The plots show that the unexpected patterns, falling f_0 movements for negative VOT values and rising f_0 movements for positive VOT values, are much more frequent than expected. Similar to the absolute f_0 values at vowel onset, f_0 movement is an unreliable identifier for laryngeal characteristics.[24]

[22] Note that an overlap plot cannot combine f_0 values and VOT values, as the range of VOT values differs between lenis and fortis plosives. The overlap in f_0 values between <b> and <p> would then not be visible as the two categories would be separated on the axis representing VOT values.

[23] Absolute f_0 values have been shown to be reliable measures in other studies. Why it is unreliable here is unclear.

[24] Again, earlier studies have shown that f_0 movements in the vowel can be used reliably to identify laryngeal characteristics of the preceding plosive. Why f_0 is an unreliable identifier in this study is unclear.

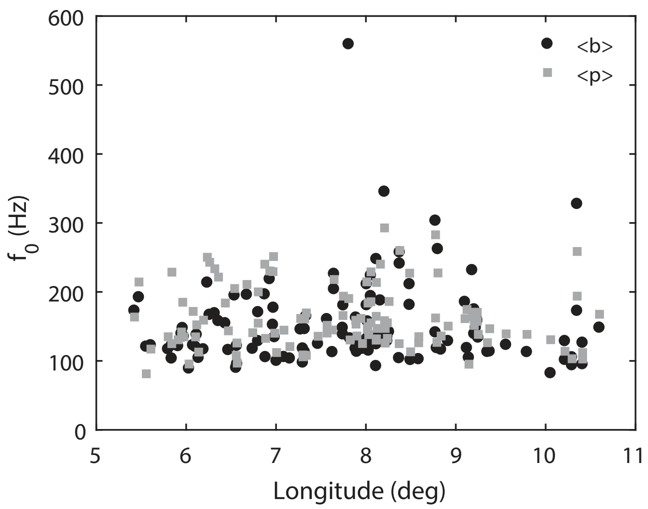


Figure 2.15: f_0 at vowel onset, lenis and fortis bilabial combined

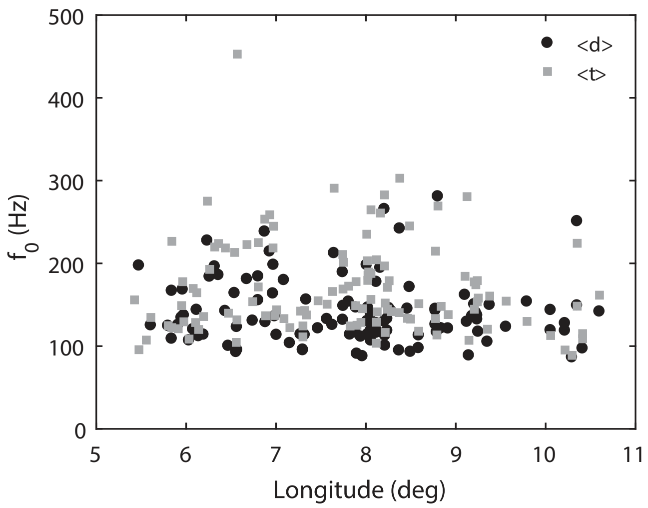


Figure 2.16: f_0 at vowel onset, lenis and fortis alveolar combined

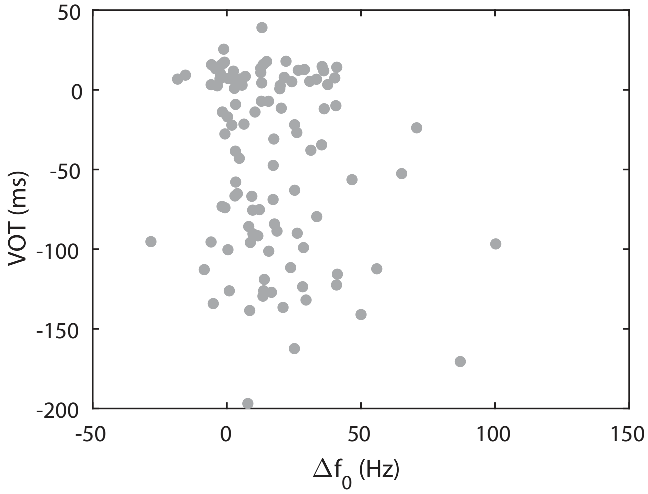


Figure 2.17: f_0 movement during the vowel linked to VOT (<b>; outliers removed)

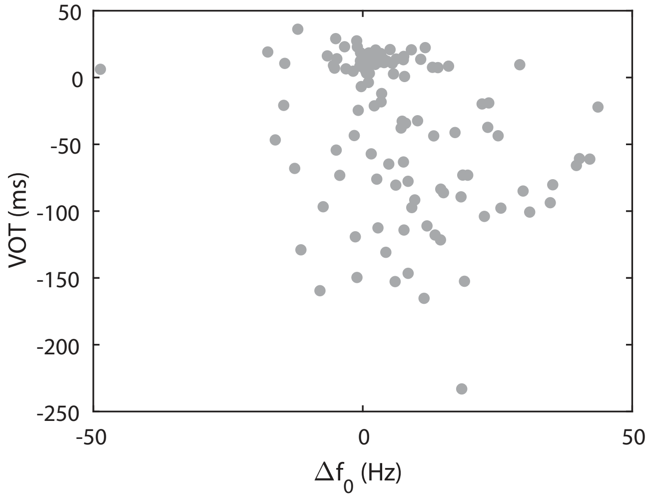


Figure 2.18: f_0 movement during the vowel linked to VOT (<d>)

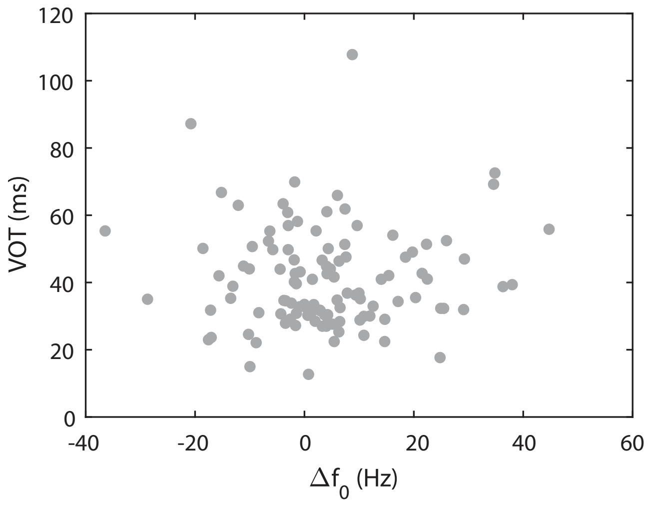


Figure 2.19: f_0 movement during the vowel linked to VOT (<k>; outliers removed)

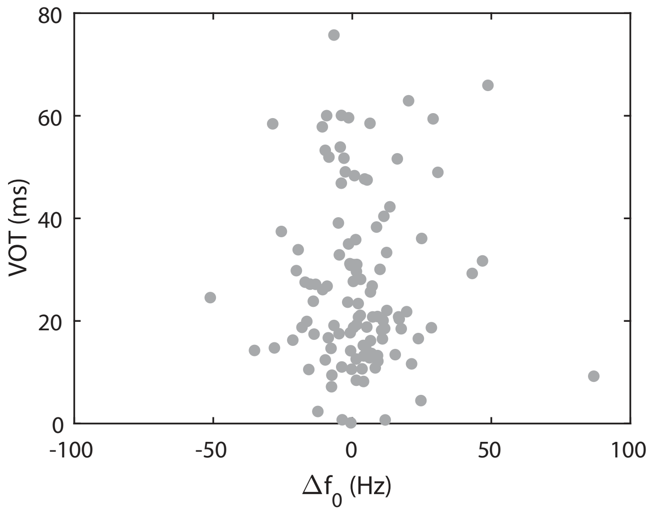


Figure 2.20: f_0 movement during the vowel linked to VOT (<p>)

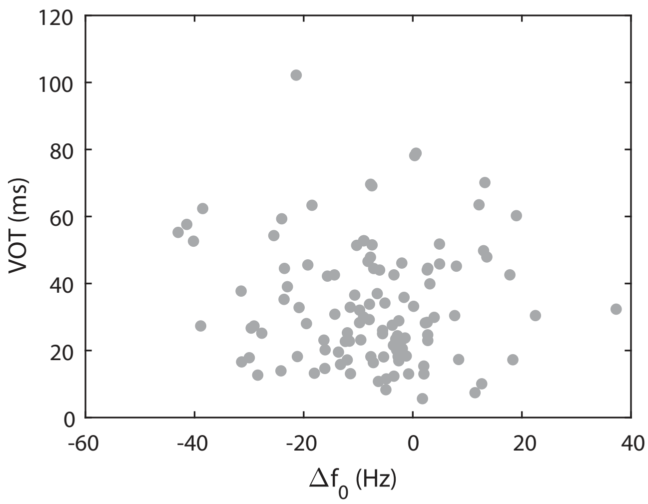


Figure 2.21: f_0 movement during the vowel linked to VOT (<t>; outliers removed)

2.5.2 Intervocalic plosives

The number of intervocalic plosives in the German database is very small. For two plosives, the lenis bilabial (<b>) and fortis velar (<k>) plosive no items could be found in the database. The other plosives could only be selected for a small number of locations. The dialects spoken in Germany are thus underrepresented in the study of the realisation of intervocalic plosives. For the Dutch dialects only 5 items are missing.[25] Because of the small amount of intervocalic plosives in the study they cannot be used to form an analysis, but in the following chapters it will be shown that they do in fact support the proposed analysis.

2.5.3 Assimilation

In Sections 2.1 and 2.3 the different assimilation patterns in voicing and aspiration languages, and the different predictions made by [±voice] approaches and LR have been discussed. In this section I will present the results of voicing patterns in plosive clusters. The onset of the closure has been placed at the point where the higher formants are no longer visible; the release of the second plosive is marked as the release of the entire closure. Voicing onset coincides with the onset of the closure, voicing offset is placed at the final visible periodic wave in the waveform, combined with the offset of the voice bar.

The assimilation data show two stable areas: an area with and an area without voicing assimilation. The middle area seems unstable, however. First,

[25] For Kuinre and Slagharen no item with intervocalic <b> could be found, for Rottevalle no <p> could be found, and for Coevorden and Schoonebeek no <k> could be found.

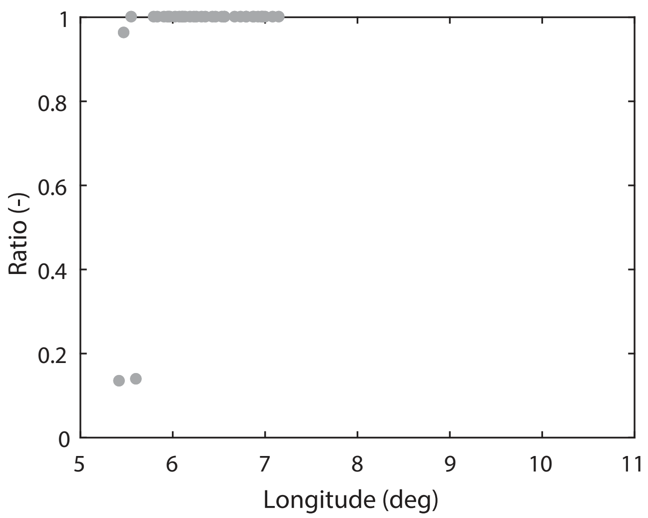


Figure 2.22: Percentage of voicing of intervocalic <b>

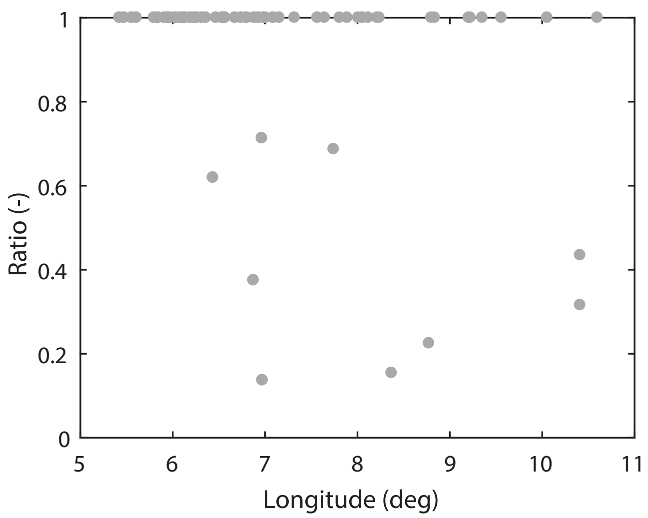


Figure 2.23: Percentage of voicing of intervocalic <d>

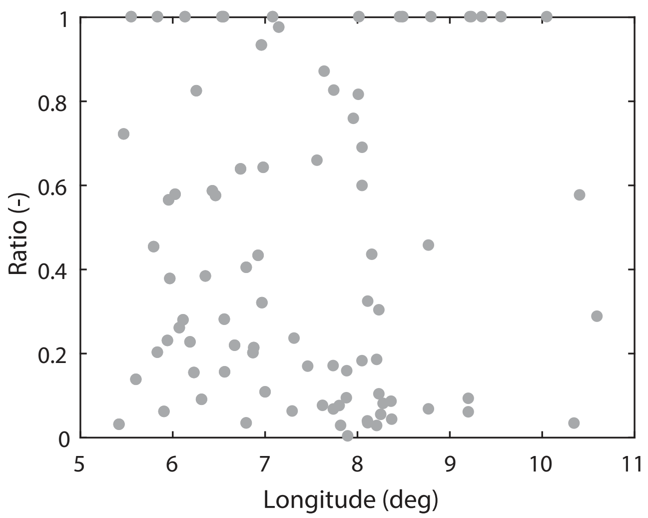


Figure 2.24: Percentage of voicing of intervocalic <p>

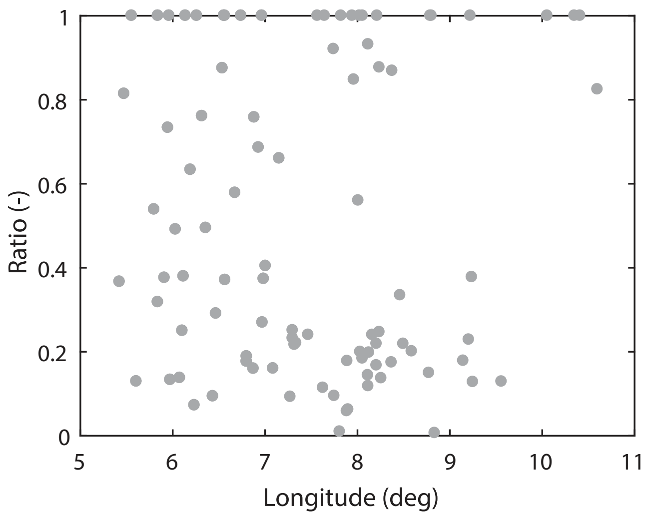


Figure 2.25: Percentage of voicing of intervocalic <t>

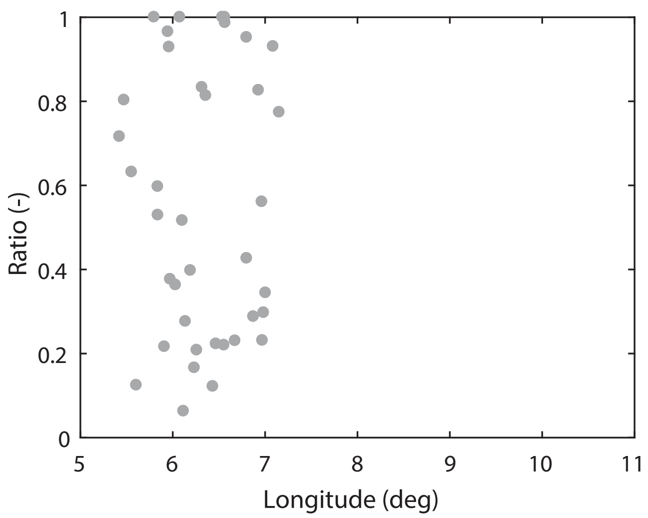


Figure 2.26: Percentage of voicing of intervocalic <k>

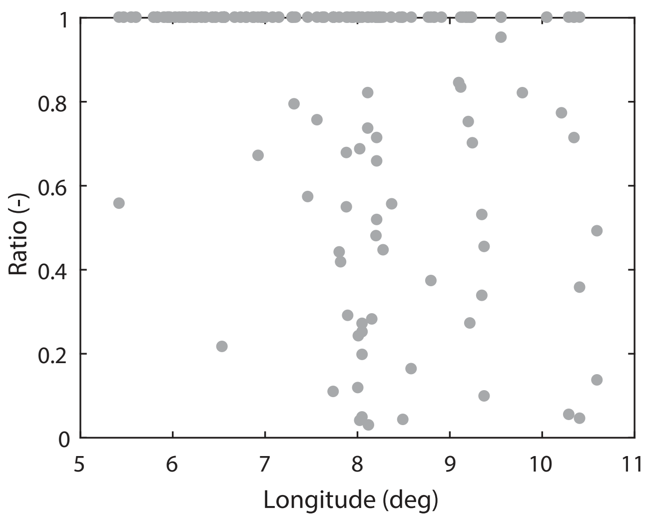


Figure 2.27: Voicing during closure in clusters - lenis C2

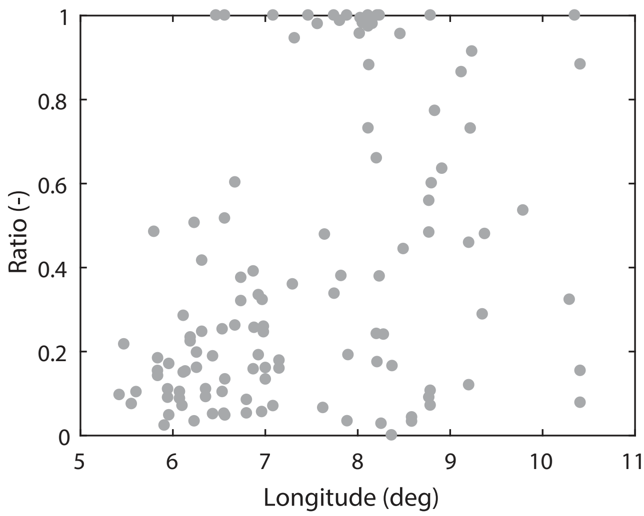


Figure 2.28: Voicing during closure in clusters - fortis C2

voicing assimilation triggered by a lenis C2 does appear to occur, but not consistently: it is applied less frequently than in the west, but more frequently than in the east. When comparing this to the mixed and fudged lects of Chambers and Trudgill (1980) the pattern seems to be compatible with patterns found in mixed lects. In fudged lects one would predict that the percentage of voicing in clusters is the average of the percentage found in the west and the percentage found in the east. Voicing in clusters would thus be intermediate between the two areas. This is not borne out by the data: a large number of clusters with a lenis C2 is fully voiced, while the clusters without full voicing have voicing percentages similar to the percentages found in the east.

The clusters with a fortis C2 show a pattern that does not appear to be compatible with either mixed or fudged lects. Neither the west nor the east shows full voicing of the cluster, but in the transitional area a large number of these clusters surfaces with full voicing. This can neither be classified as a mixed lect nor as a fudged lect, because both these types of lects assume that the transitional variety combines characteristics of the stable regions of the continuum. In the case of clusters with a fortis C2, the transitional area shows characteristics of its own: full voicing of these clusters occurs neither in the west nor in the east. With respect to clusters with a fortis C2, these dialects cannot be classified as either mixed or fudged. It thus appears that there is another type of dialects; one that Chambers and Trudgill did not discover in their data. This type of dialects may be referred to as 'imitating lects', as on the surface the dialects are very similar to both surrounding dialect groups, but deeper in the grammar, in the phonology, they do not share any characteristics with them. The similarities are thus purely superficial. The current data are, however, not the first data to show the existence of such dialects. Revithiadou, Van Oostendorp, Nikolou and Tiliopoulou (2006) show the existence of a type of vowel harmony in several Asia Minor dialects of Greek, which are argued to display a type of vowel harmony based on but not identical to Turkish vowel harmony. As the Asia Minor dialects have been in contact with Turkish for a long time while no longer being in contact with other Greek varieties, the authors argue that this type of vowel harmony is the effect of language contact.

2.6 Concluding remarks

In Section 2.5.1 I have shown that for all plosives a gradual increase in VOT values from west to east can be found. The two plots for the lenis plosives show a similar picture. In the west prevoicing is abundantly present, while short-lag VOT values are almost completely absent. The eastern end of the continuum shows mostly short-lag VOT values for lenis plosives, while prevoicing is all but absent. The middle area, contrary to the two outer ends of the continuum, shows a large spread of VOT values; both positive and negative values are found. The presence of this variation might hint at the presence of a transition zone between voicing and aspiration systems.

The plots for the fortis plosives show a less clear picture. The eastern end of the continuum shows mostly long-lag VOT values (with the boundary between short-lag and long-lag values at approximately 30 ms; Lisker and Abramson (1964) find that the lowest VOT values of aspirated plosives lie around 30 ms), while the rest of the continuum is characterised by a large amount of variation with both short-lag and long-lag VOT values being present.

If, as mentioned above, the transition zone is characterised by the presence of phonetic variation, the lenis and fortis plosives show a different pattern. For the former series the west and east of the continuum are relatively stable (no phonetic variable is fully stable so variation is always present), but for the latter series only the east of the continuum displays stable behaviour, while the west shows phonetic variation. If the presence of phonetic variation is used to identify and locate the transition zone, the task is impeded by the behaviour of VOT values in the west. It is difficult, if not impossible, to decide if the presence of phonetic variation in the realisation of fortis plosives in the west (but the absence of phonetic variation in the realisation of lenis plosives in this region) is enough to label the west as a transitional area. Furthermore, it is difficult to decide the onset and offset of the transition zone based on phonetic variation, because for the German locations only one plosive per Place of Articulation (PoA) is included in the study. The patterning of lenis and fortis VOT values with respect to each other is likely to be more revealing. In Figures 2.29 and 2.30, the VOT values for bilabial (Figure 2.29) and alveolar (Figure 2.30) plosives are plotted against the longitudes of each location.[26]

[26] An overlap plot for velar plosives is obviously missing as <g> is not included in the study.

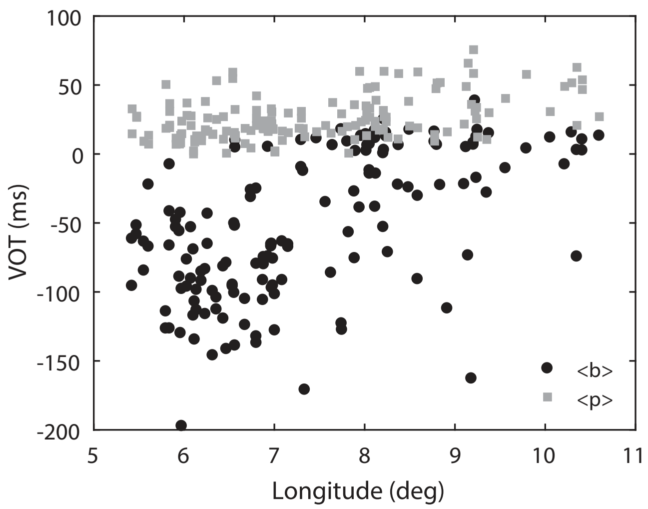


Figure 2.29: Overlap in VOT values for bilabial plosives

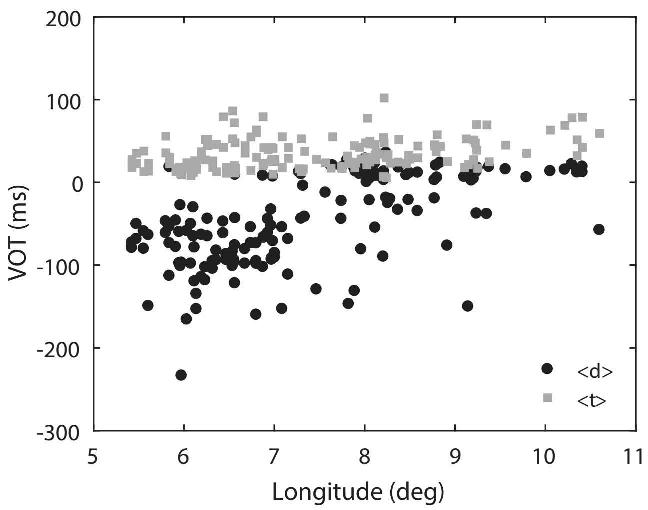


Figure 2.30: Overlap in VOT values for alveolar plosives

The plots in Figures 2.29 and 2.30 show a pattern similar to that in the plots in Figures 2.5 to 2.9, but the existence of the three areas (west, middle, east) can be shown with more certainty. In the west and east VOT values are clearly separated from each other; in the west around a VOT of 0 ms, in the east around a VOT of around 20-30 ms (the lowest VOT values found for the fortis plosives). In the middle VOT values of fortis and lenis plosives show overlap in the short-lag region. If all speakers in a region would choose for the same system, VOT values would be expected to be clearly distinguished for the two series (as is the case in the west and the east). If there is phonetic overlap between VOT values, correct perception of voicing categories will be problematic. The pattern found in the middle area is thus a highly unexpected one. Considering the presence of values compatible with a voicing system as well as values compatible with an aspiration system, it can be argued that this middle area shows a transition between the two systems.[27]

A point worth mentioning here is that the presumed onset of the transition zone roughly coincides with the political border between the Netherlands and Germany (around a longitude of 7°). It might therefore be argued that the presence of the transition zone is not merely a linguistic phenomenon but rather induced by (several) other factors, such as the different standard languages (standard Dutch is a voicing system, while standard German is an aspiration system) or the different databases. The Dutch database contains word list translations, while the German database contains sentence translations. This triggers different speech rates (with lower speech rates for reading a word list than for reading sentences), which has an effect on VOT values: Beckman et al. (2011) have shown for Swedish that lower speech rates tend to cause exaggeration of VOT values of phonologically marked plosives. Plosives marked [voice] will thus have longer prevoicing values, and plosives marked [spread glottis] will have longer aspiration values. A possibility is that the prevoicing values found in the Netherlands are longer than they would be in slightly faster speech, causing a difference between Dutch and German VOT values for lenis plosives. However, looking at not only the Dutch but also the German data, it does not seem likely that the western boundary of the transition zone is only the result of a database effect. First, VOT values for lenis plosives would be expected to be much more exaggerated in the Netherlands (Lisker and Abramson (1964) found an average prevoicing duration of 85 ms for /b/ (VOT values ranging between 145 and 50 ms of prevoicing) and of 80 ms for /d/ (VOT val-

[27]In Section 2.4 it was argued that the choice of the location depended on the presence of aspiration in the Netherlandic varieties. This may give the impression that the onset of the transition between voicing and aspirating varieties must be located within the Netherlands, instead of at the political border between the two countries. However, the aspiration present in the Netherlandic varieties can be assumed to be the result of phonetic variation within a voicing system (as only a feature [voice] is present for the lenis series, the fortis series is free to be realised as either a plain voiceless or a voiceless aspirated plosive), while the phonetic overlap in the middle area shows that the phonological onset of the transition zone must be located at the political border. The two statements are thus not contradictory, even though they may seem to be at first glance.

ues ranging between 115 and 45 ms of prevoicing) in standard Dutch). Second, if there would be a database effect a sudden change in VOT values between the Netherlands (with exaggerated lenis VOT values) and Germany (with normal lenis VOT values) would be expected. The graphs do not show this pattern: in Germany similar VOT values for lenis plosives as in the Netherlands can be found. Finally, while the Dutch VOT values follow a clear pattern, the German VOT values do not: some are compatible with a voicing system, others with an aspiration system. This cannot be the effect of the different databases, as all German data have been collected from the same database.

There might, however, be a database effect caused by differences in recording date: most dialects in Germany were recorded in the 1950's, while most dialects in the Netherlands were recorded in the 1980's. It might be possible that the Netherlandic dialects had already undergone more standardisation than the German dialects at the time of recording. Auer and Hinskens (1996) mention the fact that standardisation in the Low Saxon continuum is guided by the two different standard languages: dialects in the Netherlands are standardised towards standard Dutch while dialects in Germany are standardised towards standard German. It is thus possible that this has an effect on the different dialects in the continuum, with the dialects spoken in the Netherlands having voicing patterns compatible with those of standard Dutch and the dialects spoken in Germany having voicing patterns compatible with standard German. The differences in recording dates may explain why the onset of the transition zone coincides with the political border. This might be argued to be problematic for the analysis, but it must be noted that the transition zone extends well into Germany (past the city of Soltau). Therefore, even if the beginning of the transition zone is partly an effect of standardisation or of the different databases, the end of the transition zone is not. The phonetic and phonological patterns found in Germany still show characteristics compatible with both voicing and aspiration systems, and can be regarded as transitional regardless of the location of (one of) the boundaries of the transition zone.

The assimilation data shown in Section 2.5.3 show a division similar to the VOT data. There is a western area where clusters with a lenis second plosive all have full voicing during the closure, while clusters with a fortis second plosive all have partial voicing during closure. This pattern is fully compatible with the presence of a feature [voice]: only if the second plosive is marked with that feature, it can spread to the first plosive of the cluster, so that the entire cluster surfaces with full voicing. The presence of consistent assimilation to the voiced member of the cluster, combined with consistent prevoicing, can only be explained by the presence of the feature [voice]. An eastern area, with full voicing during closure for some items with a lenis C2, and partial voicing during closure for other items with a lenis C2, but consistent partial voicing for items with a fortis C2,[28] is also clearly visible. This pattern hints at the presence of the feature [spread glottis]: full voicing of clusters with a lenis C2 is

[28] Notice that there is one item with a fortis C2 that has full voicing during closure.

possible but not obligatory,[29] but it is clearly blocked in clusters with a fortis C2. If [spread glottis] were not present in the system, these clusters would be expected to show full voicing (the result of a phonetic process) as well.

Where the west and east of the continuum show stable linguistic patterns, the middle area shows an abundance of variation. Both clusters with lenis and fortis second consonants surface with full voicing during closure. While this is not unexpected for clusters with a lenis C2, it should be impossible for clusters with a fortis C2 in both voicing and aspiration languages. Similar to word-initial VOT values, voicing patterns in plosive clusters are equally incompatible with either a voicing or an aspiration language.

[29] There is no phonological process spreading [voice] as [voice] is not present in the system, but phonetic voicing of the cluster is possible, leading to the occasional full voicing of clusters with a lenis C2.

CHAPTER 3

Intervocalic /s/-voicing in Tuscany and Emilia-Romagna

3.1 Introduction

3.1.1 The languages of Italy

The linguistic situation in Italy is a complicated one. The many dialects and language varieties spoken in Italy belong to many different subbranches of the Romance languages. Standard Italian is thus not the variety of Italian from which all dialects and other varieties derive, but instead standard Italian itself derives from the variety of Italian as spoken in Florence (Tuscany). The 'Italian dialects' (or rather, the dialects of Italy) are local varieties deriving from Latin (Maiden and Parry (1997)). The linguistic area in Italy can roughly be divided into four major regions: northern, central, upper southern and lower southern varieties (Harris and Vincent (1988)). The northern varieties, subsequently, are separated from the central, upper southern and lower southern varieties by the La Spezia-Rimini line (e.g. Von Wartburg (1936), Vignuzzi (2010)). This line consists of a bundle of isoglosses that roughly coincide geographically. It does not only divide the Italian dialect area into two, but rather the entire Romance language area: to the north-west of the line we find the western Romance languages, and to the south-east of the line we find the eastern Romance languages (Kabatek, Pusch, Kortmann and Van der Auwera (2011)). Among the many differences that distinguish the northern dialects from the central and southern dialects is the isogloss that divides dialects that have lost the consonantal length contrast that was present in Latin (northern dialects), from the dialects that have preserved this contrast (central and southern dialects; e.g. Maiden and Parry (1997)); the isogloss that separates dialects that have voiced Latin intervocalic voiceless plosives (northern dialects) from the dialects that have not (central and southern dialects; Harris and Vincent (1988)); the isogloss that separates dialects without (northern dialects) and dialects with (central and southern dialects) Raddoppiamento Fonosintattico from each other (Harris and Vincent (1988), Rogers and D'Arcangeli (2004)).

Until quite recently many people in Italy spoke either the local language or both the local language and standard Italian. More recently, however, the use of the local varieties has significantly decreased while the use of standard Italian has increased in both informal and more formal situations, a result of the increase in people migrating across the country, the media and the social status of the local varieties (Tosi (2004)).

3.1.2 Intervocalic /s/-voicing

One of the many linguistic characteristics that distinguishes the northern from the central and southern Italian dialects is the realisation of a stem-internal intervocalic alveolar fricative (Loporcaro (1999)). In Latin this fricative was realised voiceless (Rohlfs (1966)), but the realisation of this fricative has undergone some change in the different languages spoken in Italy. In northern dialects and in standard Italian spoken in the north, this fricative is consistently realised as fully voiced (see e.g. Maffei Bellucci (1977) for Lunigiana varieties, Massariello Merzagora (1988) for Lombardian varieties, Zamboni (1974) for Venetan varieties or Frau (1984) for Friulian varieties), while in central and southern varieties (and standard Italian spoken in the centre and south) there is a phonemic opposition between /s/ and /z/ in intervocalic, stem-internal position (see for example Giannelli (1976) for Tuscan).

In central and southern dialects, voiced and voiceless fricatives in intervocalic, stem-internal position are lexically distributed: in some words the fricative is voiced, while in others it is voiceless (see Example 1). The phonological contrast between the fricatives is visible in the minimal pair in Example 2 (taken from Krämer (2005)):

(1) a. *la ca[s]a* 'the house'
 b. *la co[z]a* 'the thing'

(2) a. *fu[s]o* 'spindle'
 b. *fu[z]o* 'melted'

The northern dialects show a neutralisation of the contrast in intervocalic position:

(3) a. *fu[z]o* 'spindle'
 b. *fu[z]o* 'melted'

While the central and southern dialects have developed a phonological contrast between /s/ and /z/ intervocalically, the northern dialects are assumed to have changed intervocalic /s/ to /z/. This process is referred to as Intervocalic /s/-voicing (ISV).

Although there is a neutralisation of /s/ and /z/ in northern Italian varieties to the voiced member of the pair, a contrast between these two fricatives can still be found because of the later neutralisation of the consonantal length contrast: geminate /sː/ is degeminated to /s/. Northern varieties thus show an opposition between /s/ and /z/ in the same context, albeit with a different origin:

(4) a. *ca[s]a* 'cash register'
b. *ca[z]a* 'house'

In central and southern varieties this degemination did not take place, so these languages have three different phonemes in intervocalic, stem-internal position: /s/, /z/ and /sː/. Although there is no minimal pair contrasting all three phonemes, the differences can be seen in the following two examples:

(5) a. *ca[sː]a* 'cash register'
b. *ca[s]a* 'house'

(6) a. *fu[s]o* 'spindle'
b. *fu[z]o* 'melted'

The neutralisation of /s/ and /z/ in northern varieties does not apply in every context, but is limited to stem-internal fricatives (as in Example 3) and prefix-final fricatives (Example 7). In the latter context central and southern varieties also have a consistently voiced realisation, just like northern varieties:

(7) *di[z]-onesto* 'dis-honest'

In other linguistic environments, e.g. at a morpheme or word boundary, ISV is not active. Although there is no neutralisation of voiced and voiceless fricatives, there is no contrast between voiced and voiceless fricatives in these other linguistic environments. Rather, only voiceless fricatives occur in this position (both in northern, central and southern variaties). The following examples show the realisation of fricatives at, respectively, a morpheme boundary (stem-initial fricative in Example 8), a word boundary (Example 9), word margins (Example 10 for word-final fricatives; Example 11 for word-initial fricatives; note that the non-application of ISV in the word-final and word-initial categories is already implied by the non-application of ISV at word boundaries), and a clitic boundary (Example 12).

(8) *a-[s]ociale* 'a-social'

(9) *le cose [s]ono* 'the things are'

(10) *anana[s]* 'pineapple'

(11) *[s]ono andato* 'I am gone'

(12) *vende-[s]i* 'for sale'

There is one context in which the fricative is not voiceless: stem-initial fricatives at an opaque morpheme boundary are realised as voiced (Example 13).

(13) *re-[z]istenza* 're-sistance'

Although a voiceless realisation is expected (cf. the realisation of fricatives in Example 8), the voiced realisation can be explained by the opacity of the morphological structure: because speakers no longer analyse the word as morphologically complex (prefix-stem), but as morphologically simplex, the fricative no longer occurs at a morpheme boundary but rather in intervocalic, stem-internal position, where a voiced realisation is possible in central and southern varieties, and obligatory in northern varieties.

Because the application of ISV depends on the morphological and syntactical context of the fricative, it can be concluded that it is not a phonetic process, which is expected to apply to every intervocalic context, but rather a phonological process.[1]

3.1.3 Previous analyses

In this section several different analyses of ISV will be discussed. An early analysis is Kenstowicz (1996), who proposes an analysis in terms of a base-identity constraint: derived forms should be identical in form to the base they are derived from (unless another constraint that creates a different outcome is ranked higher). This means that a stem-internal intervocalic fricative (e.g. *a[z]ola 'button hole'*) can freely undergo ISV because there is no other base form with a voiceless realisation of the fricative to which it must be identical; the entire stem is itself the base. ISV is not blocked by the base-identity constraint in the case of prefix-final fricatives (as in Example 7) and stem-initial fricatives at an opaque morpheme boundary (see Example 13), as there is no independently occurring base to be identical in form to: the /s/-final prefix *dis-* can never occur independently from a stem, and the stem *sistenza* (from *re-sistenza*) never occurs without the prefix *re-*. ISV is, however, blocked in all other contexts, where the base can occur independently. Stem-initial fricatives at a transparent morpheme boundary cannot be voiced because the fricative is realised as voiceless if the stem surfaces without a prefix (e.g. *a-[s]ociale* 'asocial' is identical in form to *[s]ociale* 'social'); at word boundaries the fricative is part of a word that already occurs as the base itself, and the same is true for fricatives at word margins; fricatives at a clitic boundary are voiceless because

[1] In the remainder of this dissertation the archiphoneme |S| will be used to refer to the alveolar fricative if no distinction between the voiced and voiceless variant can or needs to be made.

the clitic can surface in a context where it does not undergo voicing (if the fricative is part of the clitic) or because the fricative is part of a stem, which can of course occur independently.

A different approach to ISV is Bertinetto (1999), who analyses ISV in terms of boundary strength: "Natural languages exhibit, as a rule, some reluctance to phonological change at morpheme boundaries. Not all boundaries, however, have the same strength". The stronger the boundary, the less likely phonological processes are to apply. In the case of ISV there is a clear effect of morpheme boundaries, as the blocking of the process's application depends on the nature of the intervening morpheme boundary. Stem-internally there is no boundary between the fricative and either one of the vowels, so the process can freely apply. The same is true for morphologically opaque boundaries: because speakers no longer analyse the word as morphologically complex but rather as morphologically simplex, there is no boundary between the fricative and either one of the vowels. ISV can thus freely apply. The weakest boundaries are the boundaries between a stem and an inflectional suffix, followed by the boundary between a stem and a derivational suffix. As Italian does not have /s/-initial suffixes, there are no examples of this morphological environment with the boundary preceding the fricative, and no data on the realisation of the fricative in this context exist. However, there are /s/-final stems[2] followed by a vowel-initial suffix (both inflectional and derivational examples exist). In Standard Italian (Bertinetto (1999) focusses only on Standard Italian, not on northern varieties) an alveolar fricative in this context can be realised as either voiced or voiceless. The author does not address the question further, but it appears to show that these two boundaries do not block ISV. The boundary that follows the boundary between a stem and a suffix in strength is the boundary between a prefix and a stem. When the boundary follows the fricative (see Example 7), the fricative surfaces as voiced in both northern and central varieties (and in Standard Italian, which Bertinetto focusses on), but when the boundary precedes the fricative (see Example 8), the fricative surfaces as voiceless in all varieties. All other boundaries (between two stems, between a clitic and a host, and between two words) are considered stronger than the boundaries just discussed, and indeed ISV does not apply across these boundaries.[3] Although the data do appear to fit in the boundary strength pattern at first sight, there are several problems. First, it is unclear why ISV is not active at boundaries between a stem and a suffix (it is not blocked in Bertinetto's terms, as some fricatives surface as voiced in this context, but considering the fact that there

[2]Noun stems can be considered consonant-final if the final vowel of the stem, which indicates gender and number, is considered a suffix. There is, however, no general consensus on the morphological status of the final vowel: according to some linguists it is part of the stem (e.g. Scalise (1986)), but according to other linguists it is an inflectional suffix (e.g. Peperkamp (1995)). Note that, if the final vowel belongs to the stem instead of being an inflectional suffix, the examples given by Bertinetto (1999) do not count as examples of the realisation of /s/ at a morphological boundary between a stem and an inflectional or derivational suffix, but are simply examples of stem-internal fricatives.

[3]Boundary strengths are not only based on the Italian but also on French data.

is a phonological opposition between /s/ and /z/ in this context it does not seem right to call this ISV), while it is active at a prefix-stem boundary (with the boundary following the fricative). Second, it is unclear why ISV is active at a prefix-stem boundary if the boundary follows the fricative, but not if it precedes the fricative. Bertinetto (1999) explains this as an effect of differences in boundary strength: the prefix-stem boundary is stronger when /s/ precedes the boundary than when /s/ follows the boundary. The author does not explain, however, how such differences arise in the first place.

A different approach explains ISV on the basis of phonological rather than morphological structure (Nespor and Vogel (1986), Peperkamp (1995, 1997), Van Oostendorp (1999)). Although different authors have slightly different views on the approach, in short the assumption is that ISV applies within and not across the phonological word. A stem with all suffixes is incorporated in one phonological word, while prefixes are incorporated if both prefix and stem have a specific phonological structure: the prefix has to be monosyllabic and has to end in a consonant, while the stem has to begin with a vowel. A monosyllabic prefix ending in a consonant will not be incorporated into the prosidic word if the following stem starts with a consonant, and a monosyllabic prefix ending in a vowel will not be incorporated into the phonological word either, regardless of the whether the stem starts with a vowel or consonant. This asymmetry between prefixes ending in a vowel and prefixes ending in a consonant is explained on the basis of syllabification: Italian requires all syllables to start with an onset. The final consonant of a prefix can be incorporated into the following stem in order to adhere to the ONSET constraint (assuming that resyllabification is more optimal than the epenthesis of an onset consonant or the deletion of the vowel). Alignment constraints further require that the left edge of a prosodic word and the left edge of a syllable are aligned; so either the prefix is not at all incorporated (but this violates ONSET) or the entire prefix is incorporated. Intervocalic /s/ is thus only voiced if it appears inside a phonological word.

Krämer (2001, 2003a, 2003b, 2005 and 2009) analyses ISV in OT. The process is modelled by the ranking of two constraints, *(VC̥V)PWD and F-CONTIGUITY. The latter constraint is violated by a change in feature values of an internal segment, while the former is violated by a voiceless intervocalic consonant within a prosodic word.[4] When F-CONTIGUITY is ranked over *(VC̥V)PWD, ISV does not apply, as changing the voicing value of a prosodic word-internal segment incurs a worse violation than having a voiceless intervocalic consonant within the prosodic word. The inverse ranking, however, ensures that ISV applies, as a voiceless intervocalic fricative within the prosodic word incurs a worse violation than having a voiceless fricative in that position. This latter constraint ranking thus ensures that fricatives within a stem and fricatives at an opaque morpheme boundary surface fully voiced. Voicing of

[4] Only /s/ is affected by this constraint as the voice identity constraints for the other consonants are ranked higher than *(VC̥V)PWD.

prefix-final /s/ is in fact a different process. Following Peperkamp (1995) and Van Oostendorp (1999), Krämer assumes that the prefix-final fricative is incorporated into the onset of the following vowel-initial stem (and thus into the prosodic word containing the stem): every syllable must start with an onset because of a high-ranked constraint ONSET, but because of an even higher ranked constraint DEPIO, epenthesis incurs a worse violation than having no onset. A constraint against resyllabification must be ranked lower than ONSET, so that the prefix-final fricative is incorporated into the following stem. Contrary to Peperkamp and Van Oostendorp, however, Krämer does not assume that the rest of the prefix is incorporated into the prosodic word as well. Instead, he assumes that this fricative is voiced as a result of an alignment constraint working on the featural level: the final fricative of the prefix is incorporated in the onset of the stem (and also incorporated in the prosodic word) because of a high-ranked constraint ONSET. As a result, the alignment constraint ALIGNL (stating that the left edge of every stem must be aligned with the left edge of the prosodic word) is violated. Because incorporating the entire prefix into the prosodic word incurs a worse violation than incorporating only the fricative into the prosodic word, the latter option is selected (i.e., in /di.(s#o.nes.to)/ the alignment constraint is violated only by /s/; in /d(i.s#o.nes.to)/ it is violated by /is/ and in /(di.s#o.nes.to)/ it is violated by all three segments). However, because ALIGNL works on the featural level, deleting the laryngeal feature [-voice][5] yields a more optimal output than the output where the laryngeal feature is not deleted. The prefix-final fricative subsequently surfaces as voiced as the result of feature spreading from the following vowel. Note that the assumption of a bivalent feature [±voice] is not crucial to this analysis: if a privative feature [voice] is assumed, the fricative would not have a laryngeal feature at all so that the feature of the following vowel can still spread.

3.1.4 Language change

Several authors (Galli de' Paratesi (1984), Bertinetto and Loporcaro (2005), Nocchi and Schmid (2007), Nocchi and Filipponio (2011)) have argued that non-northern varieties are currently undergoing a levelling process. Although standard Italian is based on (but not identical to) Tuscan varieties, northern varieties are abundantly present in the media. Northern Italian also has higher prestige because of the prosperity of the area. As a consequence, non-northern varieties are undergoing a change towards the northern varieties. ISV is one of the northern characteristics that is entering the non-northern varieties. Especially the Tuscan varieties show a change towards a system with ISV, so that the region will show not only a diatopic, but also a diachronic change between varieties with and varieties without ISV.

[5] Krämer assumes a bivalent feature [±voice].

3.1.5 Possible scenarios

Two sets of hypotheses can be identified, one with respect to language change (diachronic change) and one with respect to the transition zone (diatopic change). The following diachronic hypotheses can be made:

- ISV is consistently present in the dialects and regional Italian of Emilia-Romagna (for both younger and older speakers), in those contexts where it is expected to apply;
- ISV is consistently absent from the dialects and regional Italian of Emilia-Romagna (for both younger and older speakers), in those contexts where it is not expected to apply;
- older speakers in Tuscany are not expected to show evidence of ISV, except for some items which are always realised with a voiced fricative;
- younger speakers in Tuscany are expected to show more evidence of ISV, at least in the contexts where it is expected to apply;
- in the transition zone, fricatives that are not expected to undergo ISV might undergo it after all because of 'failed' rule borrowing.

The following diatopic hypotheses can be made:

- the transition between the two systems is sudden;
- the transition between the two systems is phonetically gradual (a gradually decreasing average voicing percentage from north to south, for all categories at the same time);
- the transition between the two systems is phonologically gradual (in the north all categories are affected by ISV, but going to the south less and less categories are affected).

3.2 Fieldwork and methods

As there are no data available for the present study, they were collected during a fieldwork session in September and October 2015. In this section an outline of the fieldwork process will be provided.

3.2.1 Location selection and recording

The process of ISV is present in the north but not in the centre and south of Italy, so the transition zone can be found between the north and the centre. As the central varieties are further in the language change process (Galli de' Paratesi (1984) shows that younger speakers of Florentine Italian have more voiced realisations of intervocalic /s/ than younger speakers of Roman Italian), none of the locations for the study was chosen further south than Tuscany. For this study location selection could not be based on previous results, as there are no dialect studies where both northern and central dialects are compared:

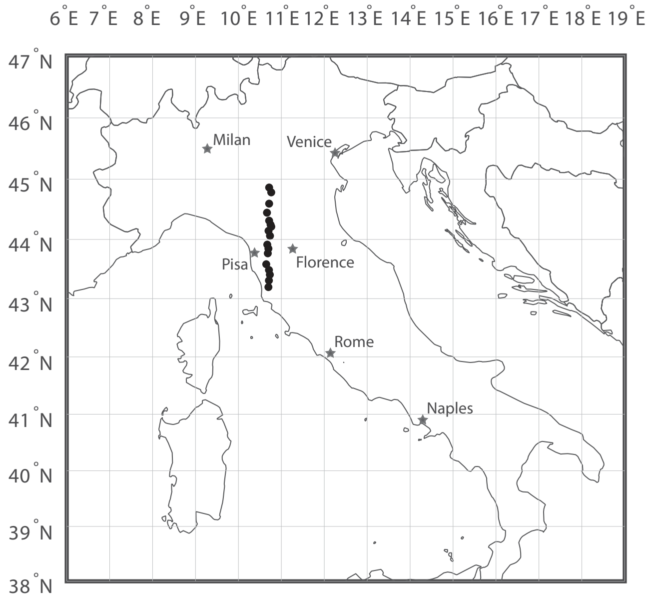


Figure 3.1: Recording locations

although the Tuscan varieties are studied in the Atlante Lessicale Toscano (ALT), the northern and other central varieties are not described in such detail. The location selection was therefore based on findings of earlier literature. All northern varieties are expected to consistently apply ISV, so the location selection was based on the central varieties. Because Galli de' Paratesi (1984) showed that younger Florentine speakers have more ISV than older speakers, and because Nocchi and Filipponio (2011) showed that younger speakers in Tuscany have more ISV than older speakers, Tuscany was included in this study. As there are no significant differences between the northern varieties reported, the northern locations were all chosen in Emilia-Romagna. All villages were located on a straight line ranging from north to south. As the Tuscan varieties are argued to be changing into varieties with ISV, all Tuscan locations were also present in the ALT in order to be able to compare the new language situation with the old language situation if the Tuscan varieties would not show large differences from the Emilian-Romagnol varieties. Because ISV is expected to be present in the entire northern region (there are no reports on the absence of ISV from the south of Emilia-Romagna), the number of locations in Emilia-Romagna is kept smaller for reasons of feasibility. In total 17 locations are included in the study, 10 in Tuscany and 7 in Emilia-Romagna (see Appendix C).

For every location four informants were looked for: two older speakers (50 years or older) and two younger speakers (30 years or younger). This division was made to compare younger and older speakers with each other, to test the hypothesis of language change. All speakers had to be dialect speakers or speakers of the local variant of Italian, born and raised in the village. Because of this restriction it was often difficult to find enough younger informants, as most of them had moved at some point in their life. The restriction on informants being born and raised in the village could not be loosened, however, because a process like ISV could potentially be influenced by the linguistic behaviour of speakers from another area (as Galli de' Paratesi (1984) showed in her study on the changing pronunciation of intervocalic |S|). Besides the generation difference, the experiment also tested for differences between standard Italian and the local dialect or vernacular. In the north no difference between the two was expected, but the situation in Tuscany is slightly more complex. Most speakers are not aware of the fact that there is a difference between the local Italian as spoken in different regions in Tuscany, and standard Italian. To test if speakers themselves do make a difference between the two, both the local vernaculars and standard Italian were tested. In both categories one younger and one older speaker were included. For some locations, the number of informants is smaller than 4. In that case the dialect or vernacular was given priority over standard Italian.

All municipalities were contacted and asked for help with finding the right informants. In case help was offered, an appointment for the fieldwork session was made with all informants together. The informants were informed of the general goal of the fieldwork (a study on Italian dialects), but not on the precise topic of investigation. All informants gave their signed consent for the recording (see Appendix E), and were informed that they could stop the interview at any point in time without further explanation (see Appendix F).[6] After the entire interview the informants were informed of the precise topic of investigation.

The recording session started with a conversation between all informants, either in the local dialect (Emilia-Romagna) or the local vernacular (Tuscany). After that, two informants (one of the older and one of the younger generation) translated the questionnaire to their local dialect or vernacular. The two other informants (again one of the older and one of the younger generation) read the questionnaire in standard Italian. Several municipalities were unable to offer their help. In those cases informants were found in the village itself, but in most cases this did not allow for a dialect conversation between informants before the translation of the questionnaire.

Recordings were all made on an Edirol R-09 HR. Sampling rate was set at 44,100 kHz (16 bits per sample). All recordings were made in a quiet location, but some background noise could not always be excluded. In case of background noise informants were asked to repeat the items from the questionnaire that were interrupted.

[6]Only one interview was broken off before the end, as the informant was too tired to continue.

3.2.2 Questionnaire

The questionnaire (see Appendix D) consisted of 68 sentences written in standard Italian. 38 sentences contained test items (words with an alveolar fricative), the other 30 sentences were filler items (in fact these sentences did not contain any alveolar fricative at all). In addition to this the questionnaire contained 33 adjectives of country names, as a native speaker of Italian pointed out that there is lexical variation in the realisation of this suffix. 23 of these items were test items; because of the small amount of adjectives of country names without |S| only 10 filler items were used.

3.2.3 Phonetic analysis

All interviews were annotated using Praat, marking sentence number and the words relevant for the analysis. These words were subsequently extracted from the sound file, and saved in a separate file. For every file an accompanying TextGrid was created. In several cases speakers showed heavy lenition of the fricative, to the extent that it was no longer audibly nor visibly present. These items were excluded, as they are not relevant for fricative voicing.[7]

Following Pinget (2015a), the onset and offset of the fricatives were identified using the spectral Centre of Gravity (CoG). Van Son and Pols (1996) define the CoG as “the mean frequency” of a spectrum (p.1530). The CoG of fricatives and plosives is higher than the CoG of vowels and sonorants (Pinget (2015a)), so that CoG values are a reliable measure for fricative identification: the onset of the fricative can be identified as the point in time where CoG values start rising, while the offset of the fricative can be identified as the point in time where the CoG values stop falling. Two Praat scripts[8] were used in the procedure of calculating the CoG: one script opened the sound file with accompanying TextGrid, and one calculated the CoG. Fricative onset and offset were marked manually as the technique does not allow for automatic fricative detection[9]: the onset of the fricative was placed at the point where the CoG values started to increase, while the offset was placed at the point where the CoG values stopped falling. These values were always compared to the spectrogram, which had to show corresponding presence of noise. For the entire duration of the fricative, the presence or absence of voicing was marked manually. Voiced intervals receive the label ‘voiced’, voiceless intervals receive the label ‘voiceless’. Voicing is said to be present in the parts of the fricative that show periodicity in the waveform and a voice bar in the spectrogram.

[7] Of course lenition of intervocalic fricatives is an interesting study, but as the present study focuses only on fricative voicing lenition with respect to manner of articulation is regarded as irrelevant.

[8] I am very grateful to Anne-France Pinget (Pinget (2015b)) for offering me her scripts.

[9] Multiple fricatives might be present in a single utterance, but the script cannot be told which fricative is the relevant one for that utterance; CoG values constantly rise and fall which makes it impossible to identify which rises and falls mark the fricative.

A Praat script was used to extract the durations of the fricatives as well as the durations of all voiced and voiceless parts of the fricatives. The results were transported to an Excel file, where the relative voicing duration was calculated by adding the durations of all voiced parts of the fricative and dividing this value by the total duration of the fricative.

It is interesting to see what different authors have said about the proportion of |S| that has to be voiced in order for it to be interpreted as a voiced fricative. Several authors have studied the phonetics of (alveolar) fricatives in Italian, with quite different outcomes. The earliest study is Baroni (1998), who investigates the phonetics of alveolar fricatives in a set of prefixed words (where the author previously found high amounts of intraspeaker variation), as well as in two sets of prefixed nonce-words (one set containing words with a prefix followed by a stem-initial /s/, the other containing words with a prefix-final /s/ followed by a vowel-initial stem) using data from an electroglottograph (EGG) study. The EGG allows the researcher to visualise the state of the glottis by measuring impedance: a low impedance indicates that the vocal folds are closed, a high impedance indicates that the vocal folds are open. Phonetic measurements were based on a combination of the waveform, the EGG and the spectrogram. The author reports that every single fricative where a voiced realisation was expected, was indeed realised as fully voiced. Not a single fricative appears to deviate from this pattern, as the average voicing ratio (duration of voicing relative to the duration of the fricative) is exactly 1 for each category, all with a standard deviation of 0.

Pape and Jesus (2015) studied the realisation of the intervocalic obstruents in Veneto Italian, a region where ISV is active. Using data from six speakers, they show that these fricatives are highly likely to be realised with full voicing, similar to the findings of Baroni (1998). However, it must be noted that Pape and Jesus did not use filler items in their questionnaire, which consisted of a total of 792 test sentences that had to be read by the informants. The topic of the investigation might therefore have been clear to the participants during the experiment, causing them to exaggerate the differences in voicing between voiced and voiceless fricatives. Furthermore, participants were given specific instructions on the pronunciation of the alveolar fricative. As all speakers have an active rule of ISV, the voiceless realisation of an alveolar fricative had to be 'forced' by instructing informants to pronounce orthographic ⟨s⟩ "as in the word[s] *sano* (healthy)" and orthographic <z> "as in *sbarbato* (shaved)". It therefore seems likely that the overall high voicing ratios of /z/ are an artificial effect.

Although the results of Baroni (1998) and Pape and Jesus (2015) do confirm the voiced realisation of intervocalic fricatives, their findings are highly unlikely. The findings by Baroni (1998) imply that voicing is never, not even once, interrupted. The author does not give a specific number of test items realised with a voiced fricative, but from his article it can be calculated that this number is probably 32, all realised by the same speaker. The number of test items as well as the number of speakers is very small, so that it is difficult to make reliable theoretical assumptions. Further, although this is a small number of test items, it is still unlikely that all of them surface fully voiced. As Ohala (1983) has pointed out, it is difficult to maintain voicing during the production of a fricative, because the latter requires a high oral pressure while the production of voicing requires low oral pressure.

A production study by Rivas (2006), comparing Dutch and Italian fricatives, found very different voicing ratios. All Italian informants had to be from the north of the country.[10] The study only included real words, so that informants did not have to be instructed on the pronunciation of <s>. The results of her study are quite different from the studies by Baroni (1997), Baroni (1998) and Pape and Jesus (2015). Instead of finding only fully voiced realisations for /z/, a large number of voiced fricatives is realised with partial voicing. Unfortunately, however, the minimal voicing ratio required for a fricative to be interpreted as voiced is unclear.

Nocchi and Schmid (2007) postulate a rule that lenites postvocalic fricatives and "affects all postvocalic contexts (both word-internal and across word boundaries)." This rule applies variably rather than categorically. The authors also hypothesise the existence of three fricative categories: voiceless fortis, voiceless lenis and voiced. They hypothesise "that lenes display a shorter duration, lower overall intensity, and partial sonorisation." However, precise voicing (sonorisation) percentages are not given.

It can only be concluded that production data on Italian fricative voicing are inconsistent. The results of studies reporting full voicing of voiced fricatives are highly unlikely, but the only study reporting both full and partial voicing of voiced fricatives does so very inconsistently. The present study therefore does not label each fricative as 'voiced' or 'voiceless', but rather compares average voicing ratios per speaker. As the aim of the study is to find out whether or not ISV is active in the different locations, it is impossible to determine the underlying status of the fricative (at least in intervocalic position for Tuscan speakers).

3.3 Results

For all items the exact pronunciation has been checked so that they could be placed in the correct morphosyntactic or phonological category; for example, in many northern Italian varieties final vowels are truncated. As a result, intervocalic fricatives followed by a single vowel before the word-final margin surfaced as word-final fricatives. All items were manually placed in the correct category, depending on the exact pronunciation. Subsequently, the average percentage of voicing per fricative category per speaker has been calculated, as well as

[10] However, Rivas (2006) also included speakers from Florence. She argues that speakers from Florence apply ISV, but unfortunately does not give a reference for this. Whether or not this is true remains an open question.

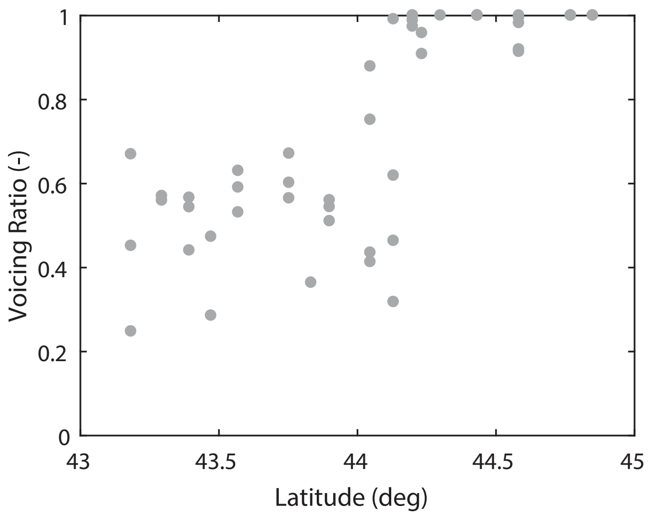


Figure 3.2: Average ratio of voicing in intervocalic fricatives

the percentage of fully voiced fricatives per speaker per category.[11] The results are plotted in the following graphs, where per phonological or morphosyntactic category, the x-axis shows the latitude coordinates of each recording location, and the y-axis shows the average ratio of voicing or the ratio of voiced items per speaker. The northernmost location in the study is thus found on the right of the graph, while the southernmost location is found on the left of the graph.

3.3.1 Intervocalic fricatives

The intervocalic fricatives in Figures 3.2 and 3.3 show a clear division between north and south. In the northern region no speaker has an average voicing percentage of less than 90%, while in the southern region only one speaker has an average voicing percentage of more than 90%. The transition zone between Emilia-Romagna and Tuscany seems to be a very narrow one, and seems largely to be the effect of the voicing realisations of three speakers: one speaker in Rivoreta and two in Popiglio. These two villages are both located in Tuscany. It thus appears that ISV is a stable process in Emilia-Romagna, while it is all but absent from the Tuscan varieties of Italian.

As mentioned in Section 3.1.4, several authors have argued that ISV is spreading into Tuscany. If this is true, there should be a difference in voicing realisations between older and younger speakers. The graphs in Figures 3.4 to 3.7 show that there is indeed a difference in their realisations.

[11] Only fricatives with 100% voicing were counted as fully voiced, because the literature does not give conclusive evidence for the proportion of the fricative that has to be voiced in order for it to be interpreted as voiced.

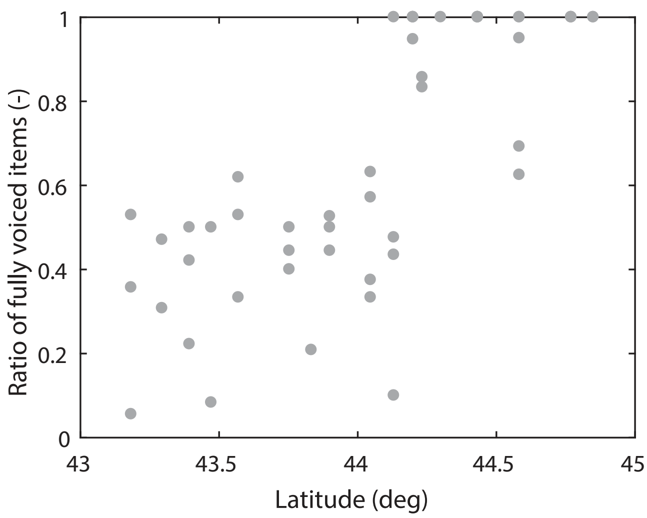


Figure 3.3: Ratio of fully voiced realisations of intervocalic fricatives

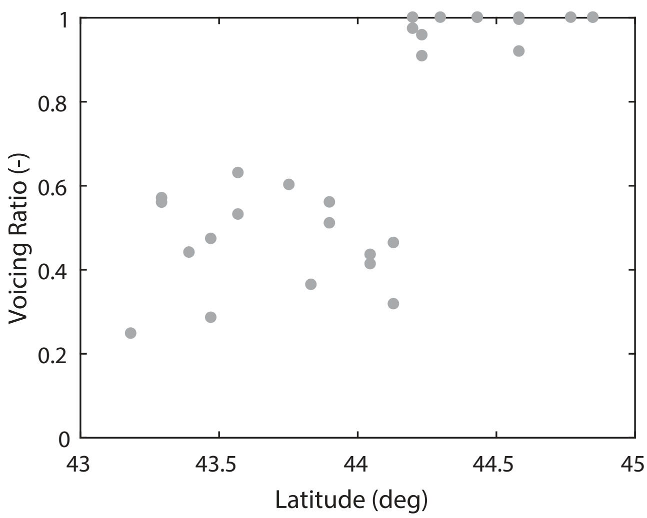


Figure 3.4: Average ratio of voicing in intervocalic fricatives (older speakers)

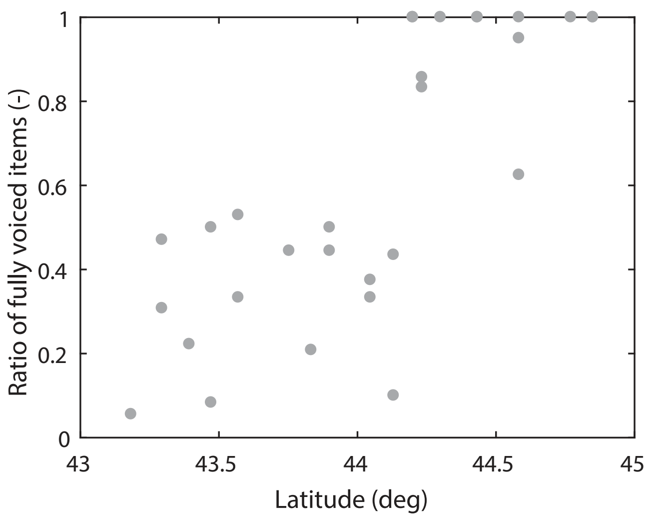


Figure 3.5: Ratio of fully voiced realisations of intervocalic fricatives (old speakers)

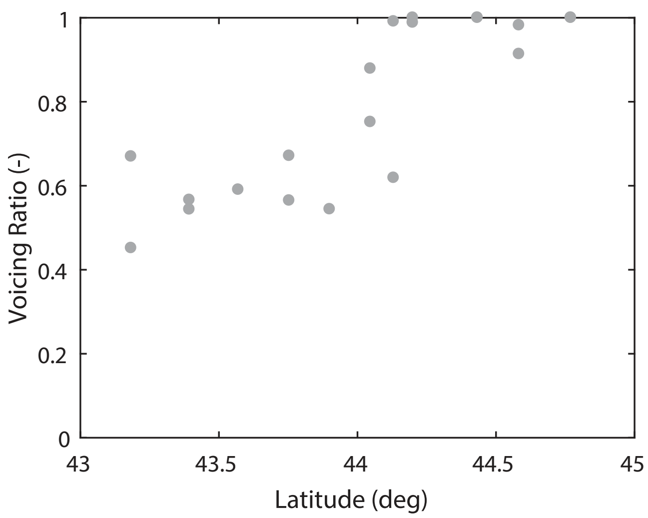


Figure 3.6: Average ratio of voicing in intervocalic fricatives (younger speakers)

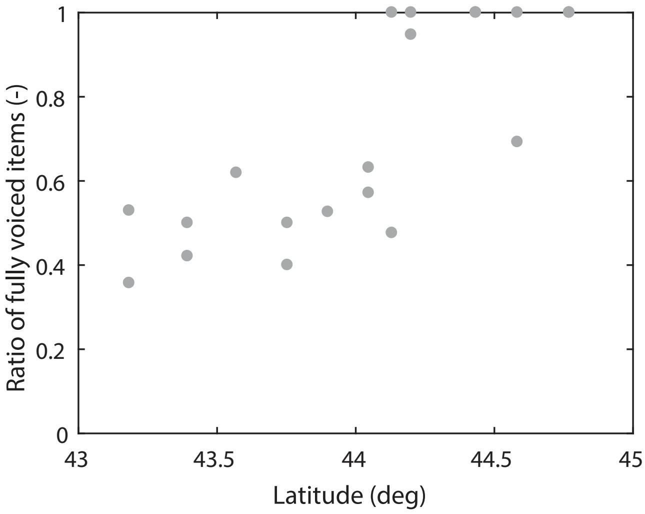


Figure 3.7: Ratio of fully voiced realisations of intervocalic fricatives (young speakers)

Indeed, for older speakers (50 years or older) the highest average percentage in Tuscany is only slightly more than 60% (63.01%, San Pietro Belvedere). The lowest percentage in Emilia-Romagna, on the contrary, is 90.82% (Sestola). For the younger speakers, however, the transition between north and south is slightly more gradual. Especially the speakers in the two most northern locations in Tuscany (Rivoreta and Popiglio) show relatively high amounts of voicing. Further, only one younger speaker (in the southernmost location in Tuscany) has an average voicing percentage of less than 50%; all other younger speakers have an average voicing percentage of more than 50%. Many of the older speakers, on the contrary, show average voicing percentages of less than 50%. It can be concluded that ISV is slowly gaining ground in Tuscany, but it is not yet very pervasive.[12]

[12]Wilcoxon rank sum tests show that the difference between younger and older speakers in average voicing ratios are significant in Monteverdi Marittimo, Popiglio and Rivoreta ($p < 0.05$). For the other speakers the difference is not significant, although it is visible.

3.3.2 Prefix-final fricatives

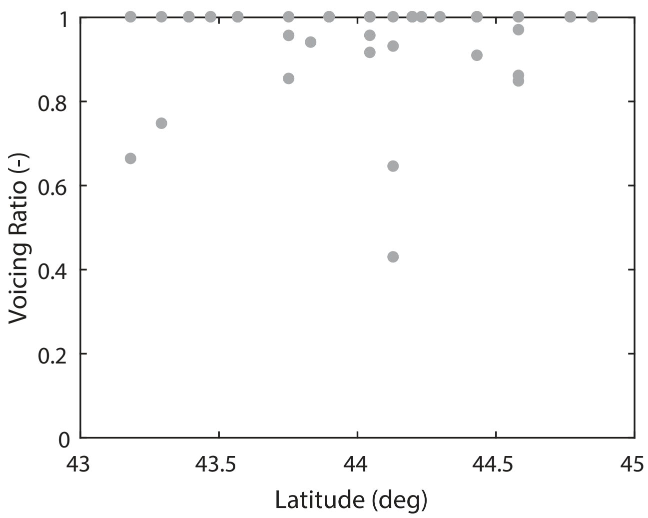


Figure 3.8: Average ratio of voicing in prefix-final fricatives

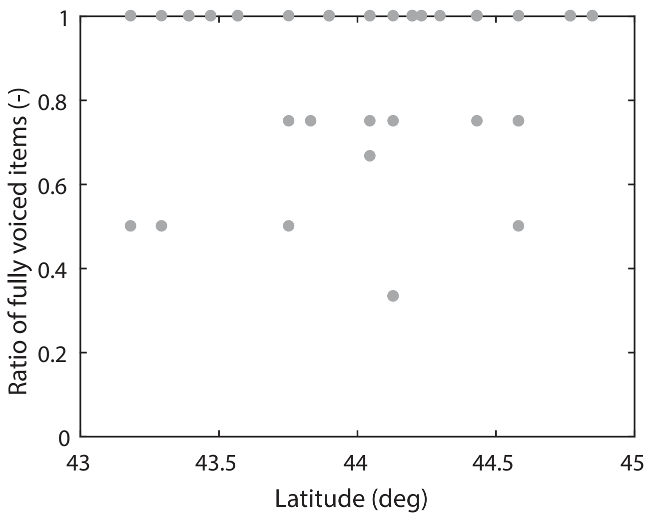


Figure 3.9: Ratio of fully voiced realisations of prefix-final fricatives

Prefix-final fricatives are expected to be realised with full voicing. Most speakers indeed show this pattern (Figures 3.8 and 3.9). Any deviation from this pattern can be epxlained by the difficulty of producing voicing in a fricative.

3.3.3 Stem-initial fricatives (morphologically opaque)

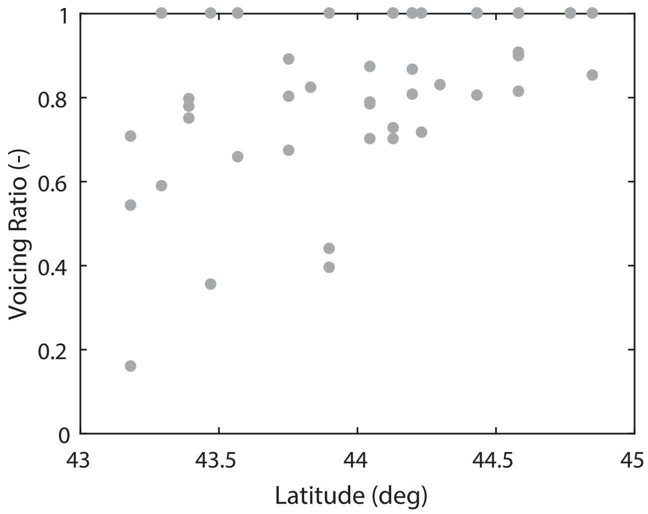


Figure 3.10: Average ratio of voicing in stem-initial fricatives (opaque)

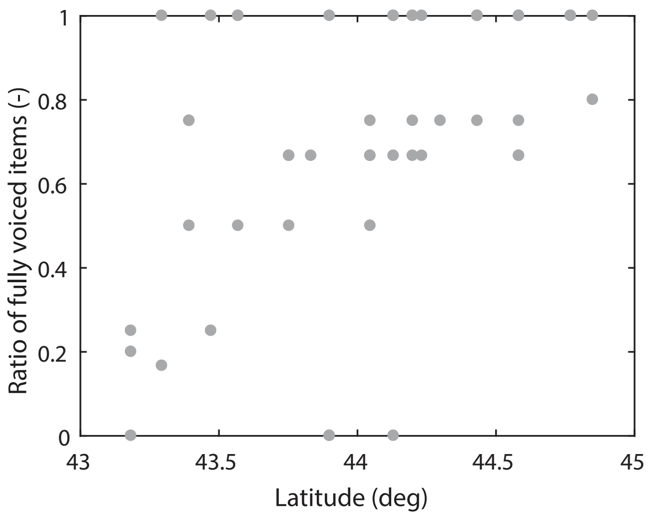


Figure 3.11: Ratio of fully voiced realisations of stem-initial fricatives (opaque)

Fricatives that occur in intervocalic position, but after a morpheme boundary, are not expected to undergo ISV. When the morpheme boundary is no longer interpreted as such, however, it can no longer block the application of ISV, and the fricative is expected to surface as voiced. The pattern for fricatives in this

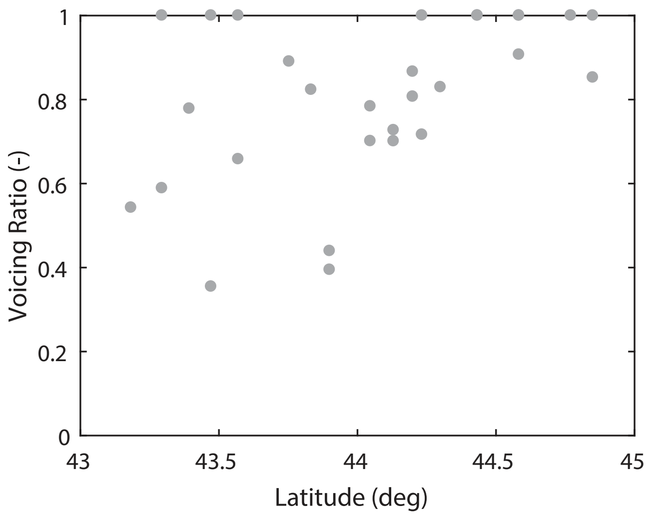


Figure 3.12: Average ratio of voicing in stem-initial fricatives (opaque; old speakers)

category is thus expected to be similar to the pattern for 'truly' intervocalic fricatives. From the graph in Figure 3.10 and 3.11, it can be concluded that for speakers in Emilia-Romagna the |S| is mostly voiced. The absence of voicing in some cases can be explained by the fact that some speakers still analyse the word as morphologically complex, or are at least in doubt of the underlying morphology of the word. Tuscan speakers show slightly more voicing of |S| in this category than of |S| in stem-internal intervocalic position.

Similar to the intervocalic fricatives, younger Tuscan speakers show more voicing of |S| than older Tuscan speakers (see Figures 3.12 to 3.15).

3.3.4 Stem-initial fricatives (morphologically transparent)

The behaviour of stem-initial fricatives at a transparent morpheme boundary is as expected. Only a few items deviate from the pattern, but most fricatives in this context are realised voiceless (Figures 3.16 and 3.17).

3.3.5 Phonological /z/

For most speakers a phonological /z/ is realised with full voicing, but for many speakers voicing is sometimes interrupted (see Figures 3.18 and 3.19). However, because voicing is difficult to maintain during the production of a fricative (Ohala (1983)), this is not unexpected. It can be concluded that there are no unexpected patterns for the voiced fricative. Note, however, that the average amount of voicing gradually decreases from north to south.

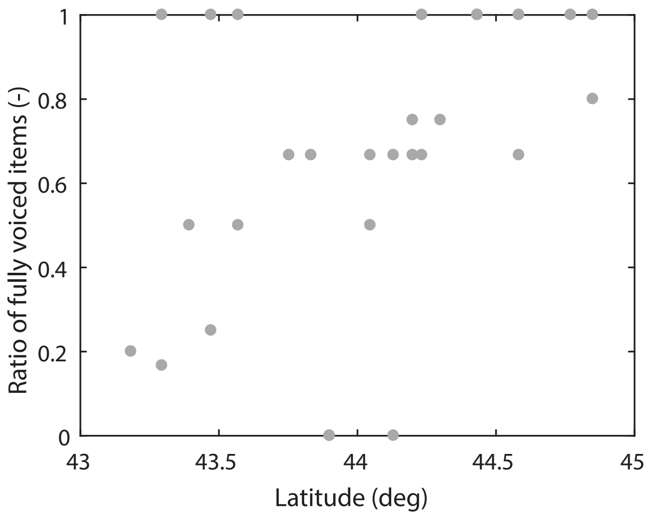


Figure 3.13: Ratio of fully voiced realisations of stem-initial fricatives (opaque; old speakers)

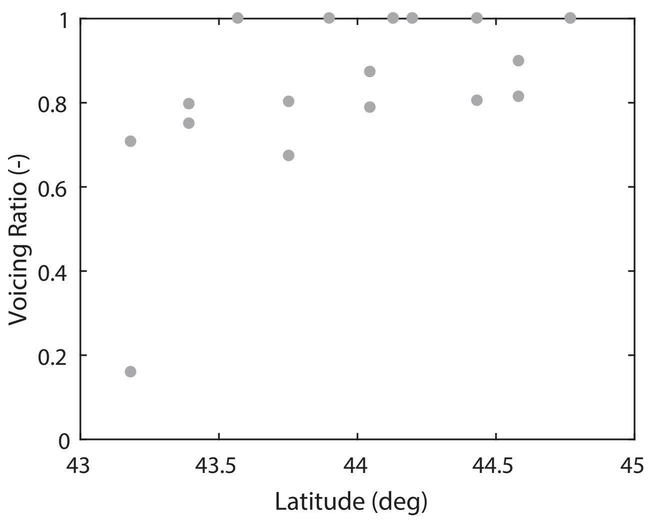


Figure 3.14: Average ratio of voicing in stem-initial fricatives (opaque; young speakers)

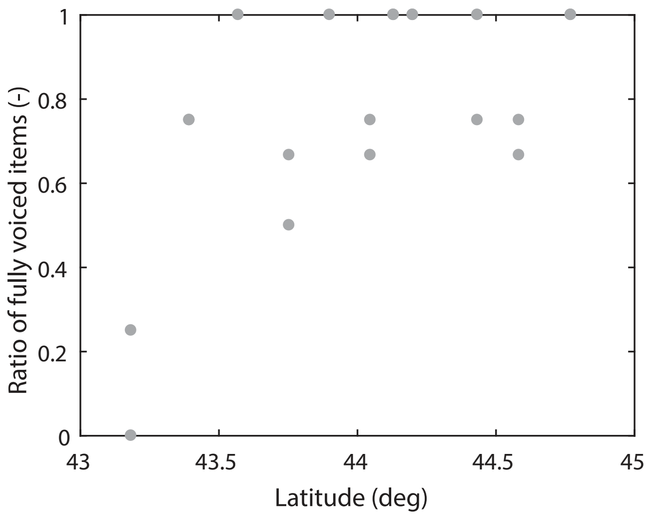


Figure 3.15: Ratio of fully voiced realisations of stem-initial fricatives (opaque; young speakers)

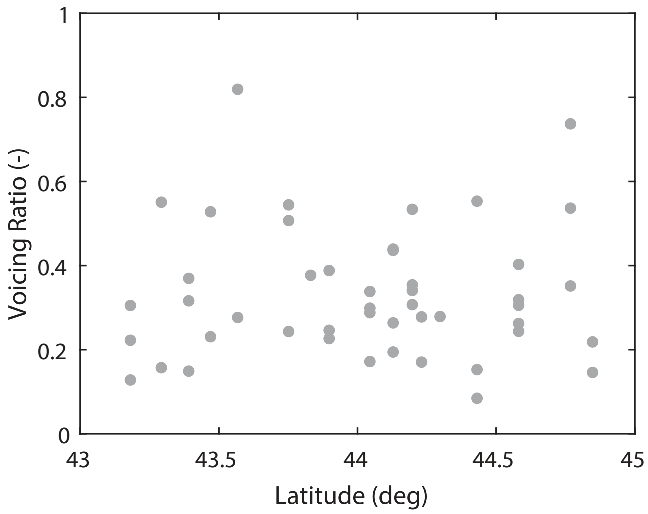


Figure 3.16: Average ratio of voicing in stem-initial fricatives (transparent)

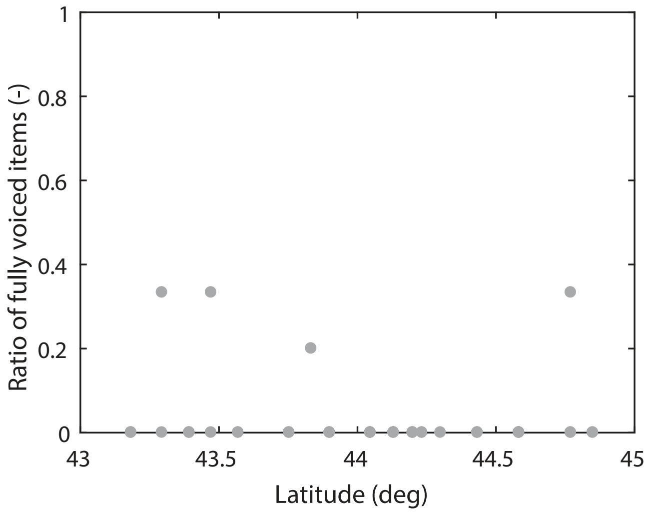


Figure 3.17: Ratio of fully voiced realisations of stem-initial fricatives (transparent)

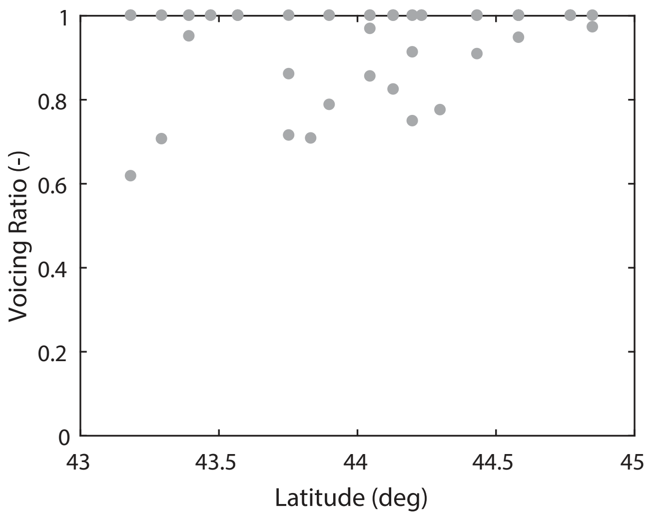


Figure 3.18: Average ratio of voicing in phonological /z/

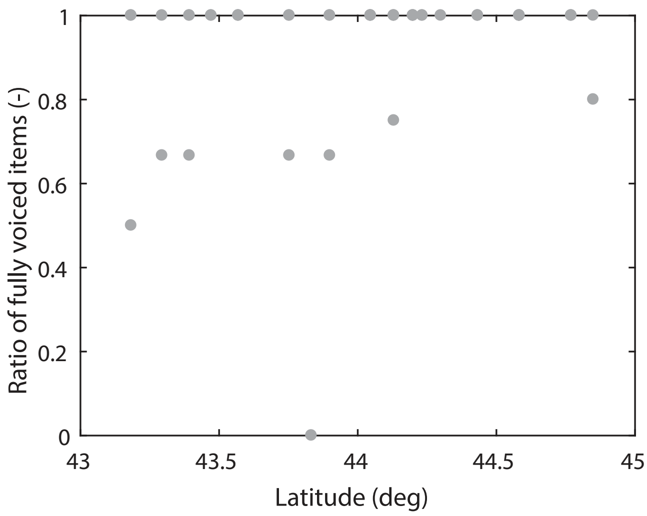


Figure 3.19: Ratio of fully voiced realisations of phonological /z/

3.3.6 Clitics

According to all descriptions of ISV, /s/ at a clitic boundary is not supposed to be voiced. There appears to be some confusion, however, as several speakers (9 out of a total of 49 speakers) realised at least one |S| at a clitic boundary as fully voiced. This is visible in Figures 3.20 and 3.21: for several speakers the average voicing percentage of fricatives at a clitic boundary is 100%, meaning that all fricatives in this position are fully voiced. The confusion seems more apparent if one considers the fact that several speakers showed variation in the voicing of this fricative: some realisations are fully voiced, while others are only partially voiced. Average voicing ratios for these speakers are thus lower than 100%. In Figure 3.21 the confusion is clearly visible for the two speakers who realise 50% of the fricatives with full voicing.

As the voiced realisation occurs across the entire region it does not seem likely that this realisation is indicative of a transitional area, as the results for intervocalic fricatives show that, if there is a transition zone, it is a rather narrow one. Apart from that, it must be mentioned that the number of fricatives at a clitic boundary seems to be too small to draw a conclusion.

3.3.7 Geminates

Figure 3.22 shows that the highest average voicing ratio for geminates is, as expected, lower than 50%. No speaker thus has a consistently voiced realisation of geminates. The percentage of fully voiced realisations (Figure 3.23) shows that there are no fully voiced realisations of geminates: all geminates are realised with partial voicing, i.e. as voiceless. Degeminated fricatives in northern dialects

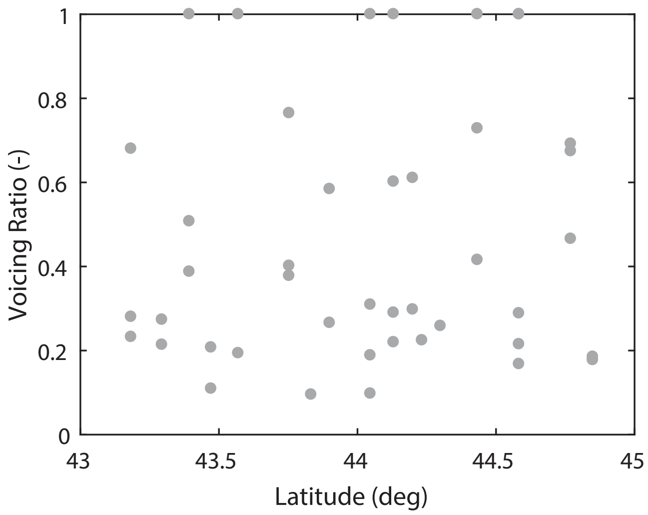


Figure 3.20: Average ratio of voicing in clitics

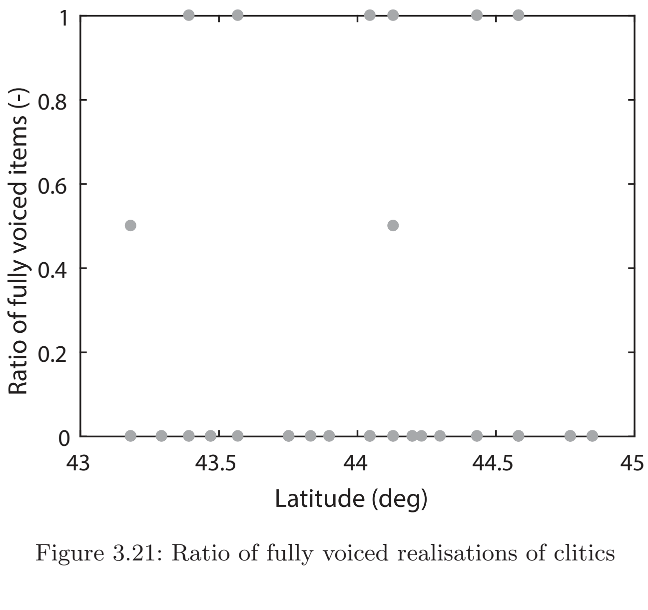


Figure 3.21: Ratio of fully voiced realisations of clitics

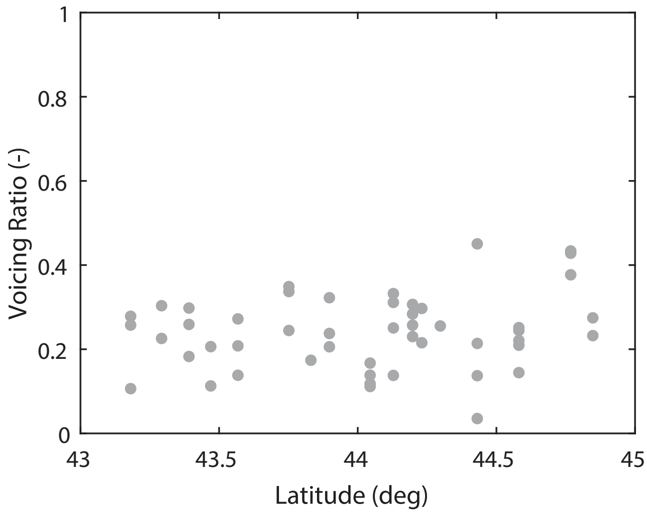


Figure 3.22: Average ratio of voicing in geminates

are thus, as expected, not subject to ISV.

There is, however, an unexpected length difference between 'geminate' and singleton intervocalic fricatives in the northern dialects: the literature predicts that the length contrast, which occurs in intervocalic position, is neutralised, so that the only trace of the contrast is the voicing difference (the degeminated fricatives do not take part in the intervocalic voicing process; Rohlfs (1966)). A linear model with consonant category (geminate vs. singleton intervocalic fricatives) as a predictor of fricative duration, calculated per speaker, shows that for every single speaker in the continuum consonant category is a good predictor of fricative duration. P-values are significant for all speakers ($p \langle 0.05$). The graphs in Figures 3.24 and 3.25 show the average geminate duration per speaker (3.24; latitude plotted on the x-axis and duration plotted on the y-axis) and the average difference in length between geminate and singleton intervocalic fricatives per speaker (3.25; latitude plotted on the x-axis and difference plotted on the y-axis).

These plots show that speakers in both areas have a length difference between geminate and singleton intervocalic fricatives. The smallest difference recorded is approximately 38 ms.

3.3.8 Phonological /s/

The realisation of phonological /s/ is, as expected, consistently voiceless, although two speakers show a small deviation from the pattern. Most fricatives, however, are realised voiceless (Figures 3.26 and 3.27).

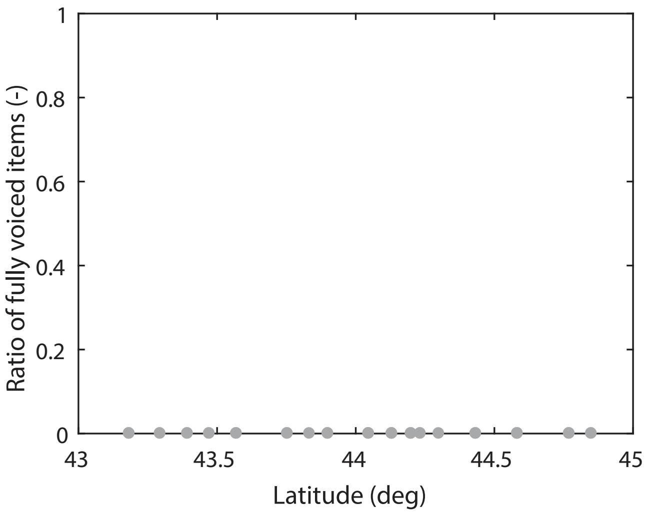


Figure 3.23: Ratio of fully voiced realisations of geminates

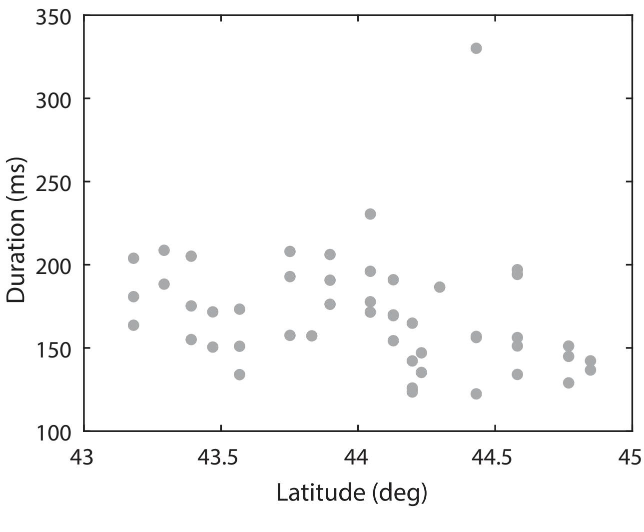


Figure 3.24: Average duration of geminates

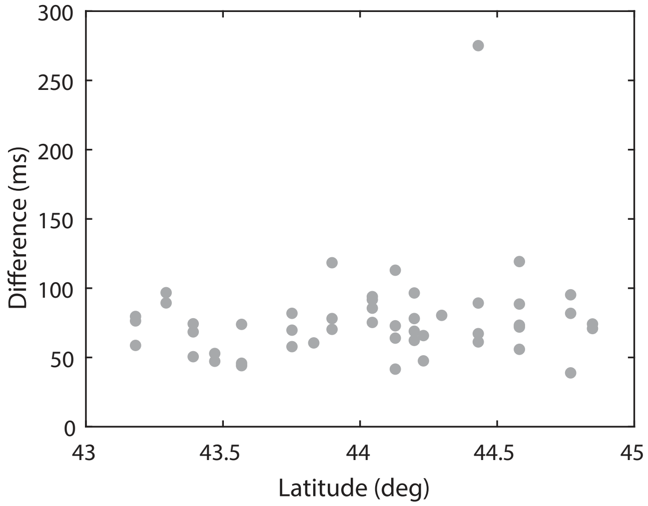


Figure 3.25: Average duration difference between geminate and singleton intervocalic fricatives

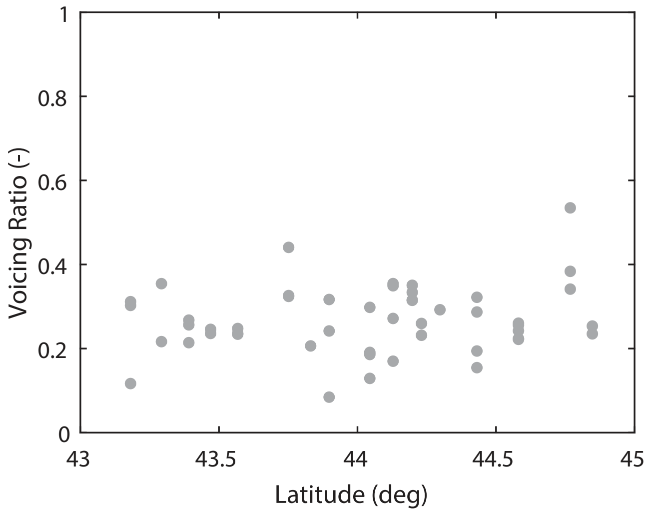


Figure 3.26: Average ratio of voicing in phonological /s/

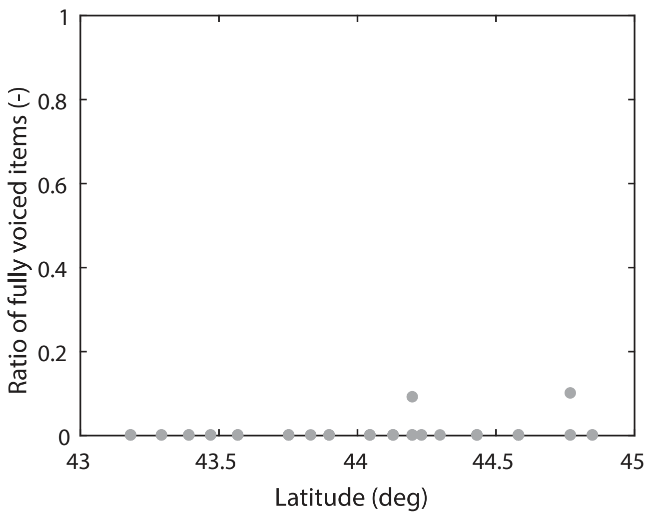


Figure 3.27: Ratio of fully voiced realisations of phonological /s/

3.3.9 Word boundary

As expected, fricatives at a word boundary are realised voiceless (see Figures 3.28 and 3.29). Several speakers deviated from this pattern by realising one or two fricatives at a word boundary as voiced, but as only three speakers showed fully voiced intervocalic fricatives (two speakers realised one fricative as voiced, one realised three as voiced), this cannot be considered a consistent pattern in those varieties.

3.3.10 Word-final fricatives

All word-final fricatives are, as expected, realised voiceless (Figures 3.30 and 3.31). Some speakers show relatively high amounts of voicing, which seems to be caused by the inclusion of words with a truncated final vowel (i.e. words which would fall in the category intervocalic, but are realised with a final fricative). All word-final fricatives, however, are still realised voiceless.

3.3.11 Word-initial fricatives

All word-initial fricatives are, as expected, realised voiceless (Figures 3.32 and 3.33).

3.3.12 Mixing and fudging

Regarding the differences between mixed, fudged and scrambled lects, the Italian transition zone shows a pattern that is very different from the pattern

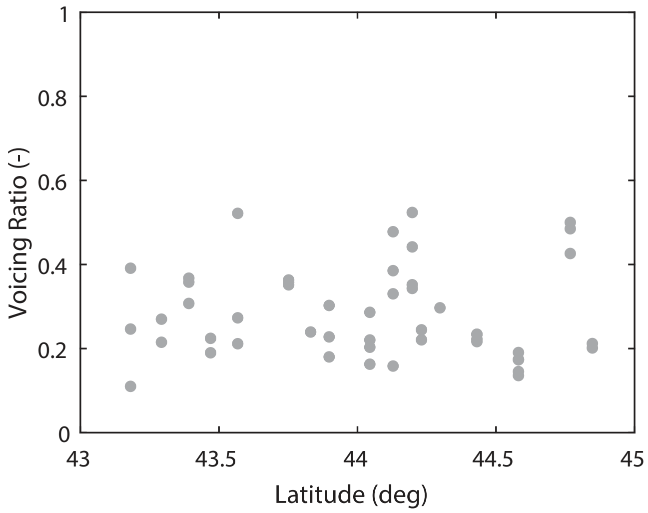


Figure 3.28: Average ratio of voicing in fricatives at a word boundary

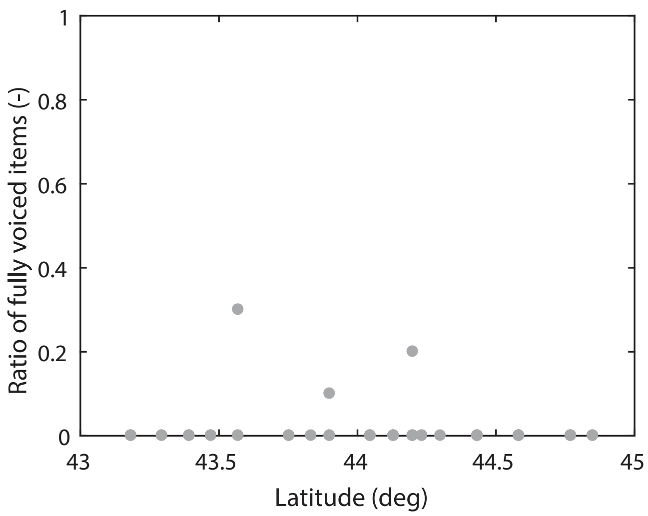


Figure 3.29: Ratio of fully voiced realisations of fricatives at a word boundary

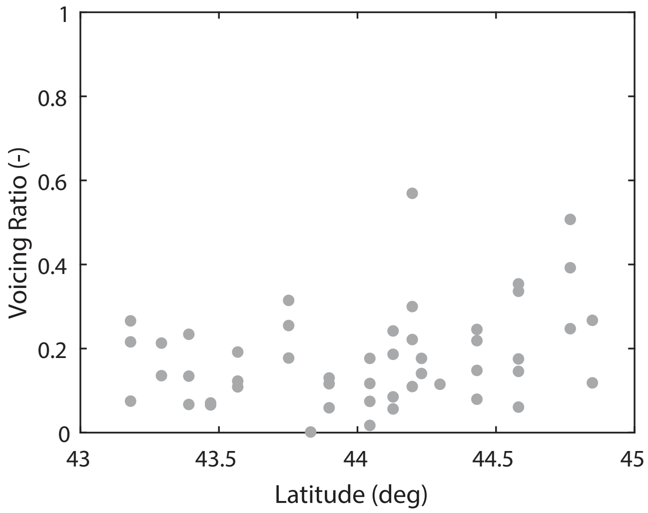


Figure 3.30: Average ratio of voicing in word-final fricatives

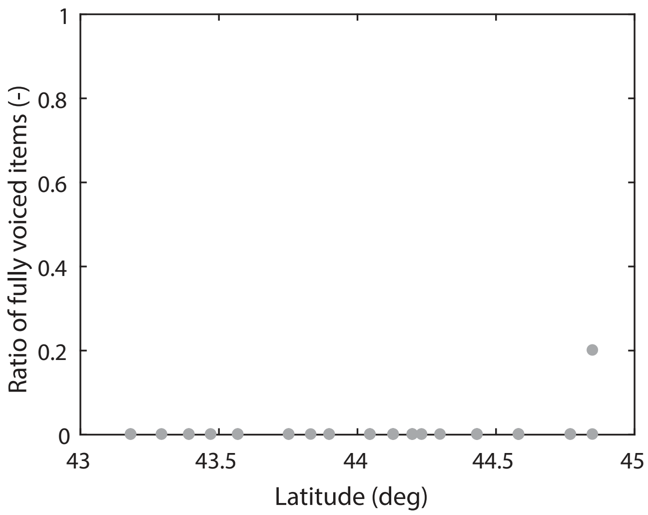


Figure 3.31: Ratio of fully voiced realisations of word-final fricatives

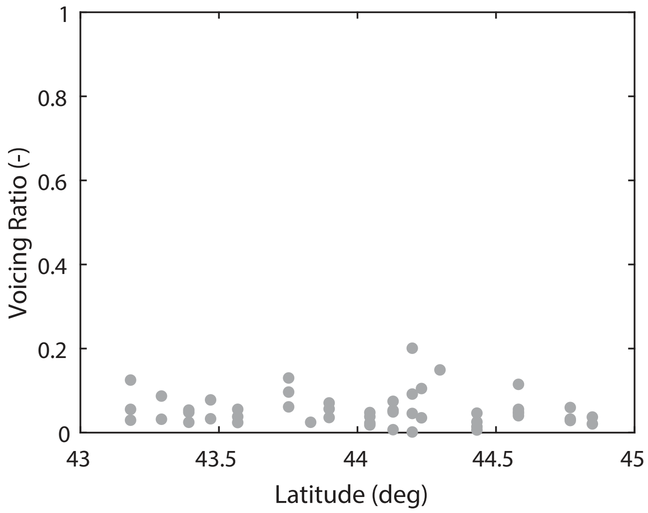


Figure 3.32: Average ratio of voicing in word-initial fricatives

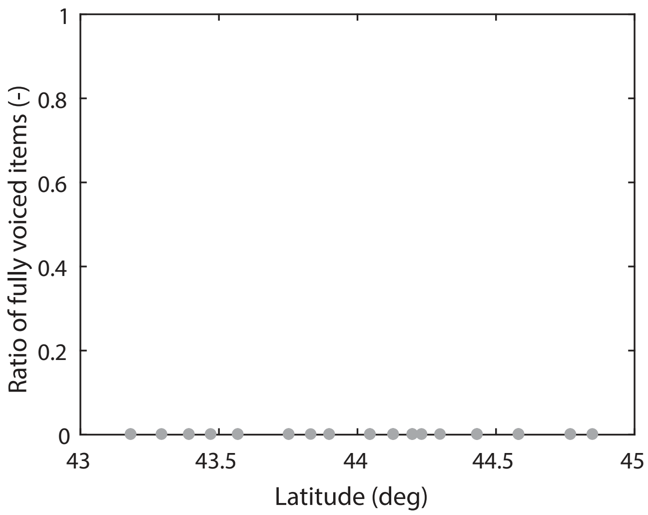


Figure 3.33: Ratio of fully voiced realisations of word-initial fricatives

found in the Dutch-German transition zone. Rather than a transition between two areas with a different realisation for one phoneme, we are looking at a transition between an area with a phonological opposition, and an area with a neutralisation of that contrast. In Tuscan Italian both /s/ and /z/ can occur intervocalically (stem-internal), but in Emilian-Romagnol only /z/ can occur in that position. If the transition zone would have mixed lects, it would have to combine the characteristics of dialects with neutralisation and dialects with a contrast. This means that in some cases there is a contrast, but in other cases there is neutralisation of the contrast. This leads, however, to a situation with an 'irregular neutralisation' of a phonological contrast. In such a situation, there is still a phonological contrast present, as the neutralisation is only partial. This means that mixed lects will have more neutralisation, i.e. more voiced realisations of the intervocalic fricative, than non-neutralising dialects, but less neutralisation than neutralising dialects. The number of items with a fully voiced fricative, as well as the average voicing percentage of intervocalic fricatives, is thus expected to be intermediate between the values for Tuscan and Emilian-Romagnol speakers. Fudged lects in this area would average the realisations of all fricatives. So a fricative that is voiced in both stable areas would simply come out as fully voiced, but a fricative that is voiceless in Tuscan Italian but (predictably) voiced in Emilian-Romagnol would come out as more voiced than the Tuscan fricative but more voiceless than the Emilian-Romagnol fricative. In other words, the percentage of voicing of the fricative would be averaged between the two possible realisations. This would mean that in the case of fudged lects, the average voicing percentage for intervocalic fricatives in the transition zone would be intermediate between the averages found in the neutralising northern dialects and the non-neutralising southern dialects. The percentage of fricatives with a fully voiced realisation, however, would be the same as the percentage of fricatives with a fully voiced realisation in the south: a fricative only counts as fully voiced if at no point in the production of the fricative voicing is interrupted. In the case of a fudged lect the amount of voicing of one fricative will be intermediate between a northern (fully voiced) and southern (partially voiced) realisation, so the fricative will never be 100% voiced, and the percentage of fully voiced fricatives will not be increased.

Looking at the data, it is visible that there is a sudden increase in the percentage of fully voiced fricatives between the two regions. There is no area where an intermediate percentage is found. The same is true for the percentage of voicing of intervocalic fricatives. For both percentages the higher values are found not only in the more northern part of what could possibly be the transitional area, but also in the south of the entire region. It is thus very difficult to argue for the presence of both mixed and fudged lects; rather, it appears that there are no transitional dialects in the case of the Italian data. The border between the two areas is abrupt rather than transitional.

If all data, for young and old speakers combined, are taken into account, a very small transitional area might be visible in the percentage of voicing of intervocalic fricatives. If the data are split into a group of younger speakers and a group of older speakers, it is immediately clear that the older speakers do not show a gradual transition at all. Instead, the change between the two regions is rather abrupt. It is clear that the apparently gradual trend between the two regions is the result of the voicing values of only two speakers in northern Tuscany; one in Rivoreta (the northernmost village) and one in Popiglio (immediately to the south of Rivoreta). The two other younger speakers in these two villages both have lower voicing values. It is thus very difficult to argue for the presence of a gradual transition between the two regions, as the number of speakers that participate in the transition is too small.

3.4 Concluding analysis

Two sets of hypotheses were formulated in Section 3.1.5, one set relating to diachronic change and one set relating to diatopic change. The first diatopic hypothesis, that ISV is consistently present in the speech of speakers from Emilia-Romagna, is confirmed by the data. Several speakers did not realise 100% of the intervocalic fricatives with full voicing, but closer inspection of the data reveals that for most of these speakers only one out of all items is realised without full voicing. This low number of items without full voicing is unlikely to be a sign of absence or decline of ISV. It is more likely to be the result of conflicting phonetic requirements. Ohala (1983, p. 201-2) explains that phonetic voicing requires a low oral pressure, while frication noise requires a high oral pressure: "[F]or the sake of continued voicing the oral pressure should be low, but for the sake of frication the oral pressure should be high, that is, the difference between oral pressure and atmospheric pressure should be high enough to cause high air velocity through the consonantal constriction. Meeting both of these requirements simultaneously may be difficult. To the extent that the segment retains voicing it may be less of a fricative, and if it is a good fricative it runs the risk of being devoiced". It is thus not unexpected that individual speakers realise one or two intervocalic fricatives with partial rather than full voicing. Further, all speakers show an average percentage of voicing of more than 90% for the morpheme-internal fricatives. It can thus be safely concluded that all speakers in Emilia-Romagna actively employ ISV. The percentage of fully voiced fricatives in contexts where ISV is not expected to apply (at a clitic boundary and a word boundary) is too small to be of any significance.

For older speakers from Tuscany ISV is clearly not active. The highest percentage of voicing of the fricative is about 60%, while the average percentage of voiced items is approximately 50% at most. If a speaker would actively apply ISV the percentage of items with a fully voiced realisation would be expected to be above chance level. For older Tuscan speakers it is either at chance level or below it, so ISV cannot be an active process in the language. The voiced intervocalic fricatives are therefore not the result of ISV but instead they must be underlyingly specified as voice, or they might be the result of a phonetic

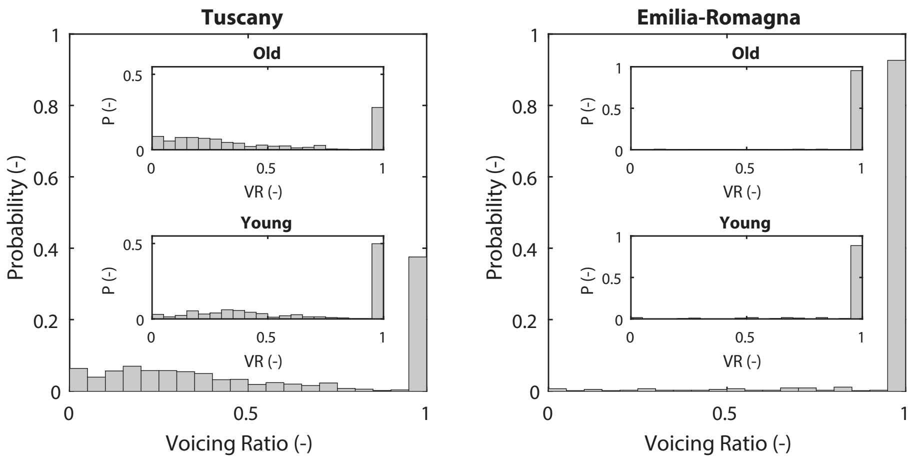


Figure 3.34: Histogram of the probability distribution of voicing ratios, broken down by region and age group

voicing process.

To be able to test several of the hypotheses formulated in Section 3.1.5, the distribution of voicing ratios (i.e. the proportion of the fricative that is voiced) of stem-internal intervocalic fricatives for all speakers is shown in the histograms in Figure 3.34. The larger histograms show the data broken down by region, while the smaller histograms show the data per region further broken down by age group. The plots show voicing ratios on the x-axis, and the probability of different fricatives on the y-axis. The following assumptions can be discussed using these plots: a) there is a difference between older Tuscan and older Emilian-Romagnol speakers; b) there is a difference between older and younger Tuscan speakers; c) there is a difference between younger Tuscan and younger Emilian-Romagnol speakers, but d) there is no difference between older and younger Emilian-Romagnol speakers. In order to examine if these hypotheses in more detail, histograms showing the distributions of voicing ratios are shown in Figure 3.34. The large histograms are broken down by region, while the smaller histograms included in those figures are broken down by region and age. The histograms only include linguistic categories for which a difference between regions and age groups can be expected (stem-internal intervocalic and stem-initial after an opaque morpheme boundary); all contexts where speakers show similar voicing patterns (word boundaries and margins, clitics, geminates, prefix-final fricatives, stem-initial fricatives in transparent words and phonological /s/ and /z/) are excluded from the histograms.

The large histograms, with voicing ratios only broken down for region, show a difference between Tuscan and Emilian-Romagnol speakers. Speakers in both regions show a concentration of voicing ratios of 1. Because in both regions [z] is a possible realisation of the fricative in either of these contexts (in Emilia-

Romagna it is the only expected realisation; in Tuscany there is a phonological contrast between /s/ and /z/), these peaks at voicing ratios of 1 are indeed expected. The large histograms further show a difference in frequencies of partially voiced fricatives. In Emilia-Romagna the probability of encountering a partially voiced fricative is very low, while in Tuscany this probability is clearly higher. A closer look at the histograms further shows that in Emilia-Romagna, the proportion of fully voiced fricatives is much bigger than in Tuscany. As a consequence, in Tuscany the proportion of partially voiced fricatives is much bigger than in Emilia-Romagna. In conclusion, there is a clear effect of region on voicing ratio.

The smaller histograms in the figures show the distribution of voicing ratios in the two regions further broken down by age. In Emilia-Romagna a difference between younger and older speakers is not visible: for both age groups the probability of encountering a fully voiced fricative is extremely high, while the probability of encountering a partially voiced fricative is very small. Younger speakers do not show a tendency towards more partially voiced fricatives than older speakers. The hypothesis predicting that there is no difference between older and younger Emilian-Romagnol speakers is indeed supported by the data, as both generations show all signs of ISV.

In Tuscany some small differences between the two age groups are visible. First, for younger speakers the probability of encountering a fully voiced fricative is higher than for older speakers, while the probability of encountering a partially voiced fricative is higher for older speakers than for younger speakers. Furthermore, the average voicing ratio of partially voiced fricatives appears to be lower for older speakers than for younger speakers. The histograms do indeed seem to support a small difference between the two age groups in Tuscany, as hypothesised.[13]

A difference between the two age groups is only visible for Tuscan, but not for Emilian-Romagnol speakers. It can therefore be concluded that there no is difference between the two age groups in general. Instead, the presence of a generational difference depends on the region speakers are from. In itself, the age of a speaker does not influence its language.

A final comparison that might be worth making is the comparison between speakers of different regions within each age group. For both age groups a difference is immediately apparent: speakers from Emilia-Romagna have relatively many more items with a voicing ratio of 1 than speakers from Tuscany (for both younger and older speakers), while speakers from Tuscany have relatively many more items with partial voicing than speakers from Emilia-Romagna. Within each age group there are thus differences between the two regions.

Although the diatopic data do not show a transition zone, the histograms show that the voicing ratios of young Tuscan speakers are higher than the voicing ratios of old Tuscan speakers, but lower than the voicing ratios of young

[13] A Welch Two Sample t-test shows that the differences in average voicing ratios between younger and older Tuscan speakers is indeed significant: $t = -4.6341$, $df = 611$, $p = 4.3853e{-6}$.

Emilian-Romagnol speakers. If we change our point of view from a diatopic to a diachronic change, younger Tuscan speakers can be argued to be in a transition from a system without ISV to a system with ISV.[14] They show characteristics of both mixed and fudged lects: the average voicing percentage of intervocalic /s/ is higher (so most individual fricatives have a higher percentage of voicing compared to the realisations of older speakers), and the percentage of fully voiced items is higher (meaning that more intervocalic fricatives are realised with full voicing compared to the realisations of older speakers). For both variables, the value is lower than for the speakers from the north. The younger Tuscan speakers might thus be a diachronic transition zone. As they show an increase in both percentage of voicing and percentage of fully voiced items, it can be argued that their linguistic varieties are a combination of both mixed and fudged lects, i.e. scrambled lects.

The question is what the grammatical representation for younger speakers must be. Their voicing averages are slightly higher than voicing averages of the older speakers, and there are two speakers in the two southernmost Tuscan locations that appear to be in a transition from a system without to a system with ISV. The question how this transition should be modelled in the grammar will be addressed in the next chapter.

3.5 Conclusion

Several hypotheses were formulated in Section 3.1.5. Five of them seem to be confirmed by the data. In the Emilia-Romagna, ISV is applied in those contexts where it is expected (first hypothesis) and not in those contexts where it is not expected (second hypothesis). In Tuscany, older speakers do not show clear signs of ISV (hypothesis three), while younger speakers show higher voicing percentages and thus more evidence of ISV (hypothesis four). The fifth hypothesis, however, does not seem to be confirmed: there are very few items with a voicing realisation different from the voicing realisation that is expected. Regarding the diatopic change, it seems that the transition between the two systems is abrupt rather than gradual (first hypothesis). There is no clear evidence of a gradual phonetic or phonological change, at least not for the older generation. The younger speakers do reveal a slightly more gradual trend, but this is mostly the result of the voicing percentages of a few speakers in northern Tuscany. It can therefore be concluded that the transition between the two systems is abrupt rather than gradual.

[14] The endpoint of the young Tuscan speakers is obviously unknown, but their language currently appears to be in a transitional state between the two systems.

CHAPTER 4

Explaining the patterns of variation

The previous two chapters discussed the two transition zones in more detail, including the phonetic data from the transitions themselves and a first analysis of the data. In this chapter a phonological analysis of the data will be presented. Problematic for a thorough analysis of the Dutch-German data is the small amount of data available, especially in the German database. The data do not immediately point to a clear phonological distinction, nor do they immediately exclude other potential distinctions. Therefore, all different contrasts will be considered. Further, they will be considered in the light of insights from literature on the acquisition of phonological contrasts: would it be possible for a child to acquire the hypothesised contrast based on the linguistic input in the transition zone? Section 4.1 will be devoted to an overview of the different theoretical approaches to laryngeal contrasts as discussed in Chapter 2. As these approaches all make different predictions with respect to variation, the data from the Dutch-German continuum will prove to be very useful in the discussion of these different approaches. Section 4.2 will provide an overview of the literature on the acquisition of phonology and, more specifically, on the acquisition of laryngeal contrasts.

Sections 4.3 and 4.4 will be devoted to the actual phonological analyses of the data. As the analyses might not be immediately evident from the data, the advantages and disadvantages of all potential analyses are included for both regions. A thorough discussion of the linguistic characteristics of both regions will be shown to (at least partially) explain the differences in the gradualness of the transitions in the two regions. It must be mentioned here that, because the main goal of this study is to investigate language variation, and the data from the Dutch-German continuum show significantly more variation than the data from the Italian continuum, the larger part of this chapter will be devoted to a discussion of the Dutch-German data.

4.1 Predictions for variation

Laryngeal Realism (LR) argues that languages with a voicing contrast employ a different phonological distinction than languages with an aspiration contrast: [voice] is used to distinguish prevoiced plosives from short-lag VOT plosives, while [spread glottis] is used to distinguish aspirated plosives from short-lag VOT plosives. As the key assumption of the theory is that the difference between voicing and aspiration languages is not merely phonetic but is also represented in phonology, the approach predicts that there is not much variation possible within each category. The feature [spread glottis] is specifically linked to the wide position of the glottis at the time of plosive release, and the feature [voice] is specifically linked to the presence of prevoicing on a plosive. The framework thus seems to predict that variation is possible within the limits of each feature: VOT values of a plosive marked for [voice] may vary as long as they are prevoiced, just like VOT values of a [spread glottis]-marked plosive may vary as long as they are aspirated. The unmarked series has specifically been argued to allow for more variation, but only in the direction of the absent plosive series (Avery and Idsardi (2001)). They argue that the unmarked series in a voicing system is free to be realised with either short-lag or long-lag VOT, as there is no category of [spread glottis] plosives with which they would be in conflict (see also Hamann and Seinhorst (2016), who show that lenis plosives in several varieties of German can be realised with prevoicing). Similarly, unmarked plosives in an aspiration system may be realised with prevoicing as there is no category of [voice] plosives with which they would be in conflict. This implies, however, that unmarked plosives are not free to be realised with prevoicing in a voicing system, as they would then be in conflict with the [voice] series, and that they are not free to be realised with aspiration in an aspiration system as they would then be in conflict with the [spread glottis] series. Also, the marked series in the system (either [voice] or [spread glottis]) should, following Avery and Idsardi's line of reasoning, not be realised with short-lag VOT, as they would then be in conflict with the unmarked series. In theory, however, it should be possible for [voice] plosives to be realised with aspiration, and for [spread glottis] plosives to be realised with prevoicing. Only a link between phonetics and phonology would prevent this from being a possibility. In conclusion, LR thus predicts variation between categories only for the unmarked series, as long as the variation does not cause a conflict with the marked series. The marked series themselves are not expected to show variation outside their 'predicted' values.

What does LR predict for the phonetic realisation of [∅]? Avery and Idsardi (2001) argue that unmarked plosives are relatively free because of the absence of a laryngeal feature, so it can be realised with any VOT as long as that VOT it is not in conflict with another VOT category that is present in the language. So, if there is no laryngeal opposition for plosives, [∅] can be realised with any VOT as no conflict between [∅] and another (marked) plosive series will arise. However, if there is a [voice] series present in the language, [∅] cannot be realised with negative VOT as this would cause a conflict with the [voice] series. The unmarked series can be realised with both short-lag and long-lag VOT in this system, as there is no other plosive series present with which this would create a conflict. Similarly, if a series of [spread glottis] plosives is present besides the unmarked series, the realisation of the unmarked series may not conflict with the realisation of the [spread glottis] series, so it cannot be realised with aspiration. Realisation with prevoicing, on the other hand, does not create a contrast with another plosive series, and is thus possible. Summarising, in Avery and Idsardi's framework the realisation of [∅] depends on the presence of other features in the system. The 'feature' itself (as [∅] represents the absence of a feature it is in fact not right to refer to it as a feature) does not have a direct link with the phonetics. However, the approach does not predict what kind of realisations can be expected if the unmarked series is the only plosive series in the system. It implies that, as there is no other plosive series present, the realisation of [∅] is unconstrained, so any VOT value, from prevoicing to short-lag VOT to aspiration, can occur. In reality, however, it turns out that the most common realisation is a voiceless unaspirated plosive (Maddieson (1984)), implying that the realisation is in fact constrained. Phonetic constraints on the realisation of unmarked plosives are similar to markedness constraints: producing prevoicing or aspiration requires more effort than producing a voiceless unaspirated plosive, so if there is no need to emphasise a phonological contrast, ease of articulation will cause the plosives to be realised with mostly short-lag VOT. Note also that in languages without a laryngeal contrast in plosives, phonetic variation will probably still occur. The prediction is that in these languages the variation can go in both directions (prevoicing and aspiration, instead of only one of them as in languages with a laryngeal contrast), only short-lag VOT is the most frequent realisation because of the minimal effort required to realise it.

Element Theory does not make use of phonological features, but instead uses elements. Contrary to phonological features, elements have an independent phonetic interpretation. The laryngeal elements, |L| and |H|, are linked to respectively slack vocal cords (required for prevoicing) and stiff vocal cords (causing a delay in voice onset) (Kaye et al. (1990)). In other words, the elements are linked to prevoicing and aspiration respectively. This approach does not predict much variation, as the elements come with a phonetic prespecification. This means that |H| can only be posited for phonetically aspirated plosives, and |L| only for phonetically prevoiced plosives. However, it is unclear if the reverse is also true: is a consistently prevoiced plosive necessarily analysed as an |L| plosive, and is a consistently aspirated plosive necessarily analysed as an |H| plosive? This would have to be assumed: if the restriction is this strict in production, it should be equally strict in perception (and language acquisition), otherwise the link could be quite simply lost (i.e. a voicing contrast could then get analysed as a phonological aspiration contrast without being realised as such). Plosives that are incompatible with both |L| and |H| end up unmarked, as elements are monovalent. Essentially, an element approach is thus highly similar to LR.

Element Theory seems to predict that the absence of an element is less free in its phonetic realisation than the absence of a feature in LR, as the elements both require and specify a precise phonetic realisation. In other words, if an element is present, a specific phonetic realisation is required, but if that element is not present the phonetic realisation of that specific element should not be possible. A phonologically unmarked plosive can only be realised according to its baseline, so without prevoicing (which requires an element |L| in the phonological representation) and without aspiration (which requires an element |H| in the phonological representation).

Laryngeal Relativism (Cyran (2011), Cyran (2013)) argues that the phonological features are not directly linked to phonetic values.[1] The approach is similar to LR in the sense that it phonologically distinguishes voicing from aspiration languages, but because the approach assumes that the phonetic realisation of a plosive is not directly linked to its phonological specification (i.e., a plosive marked with the element |L| may be realised as a prevoiced plosive, as in a voicing system, but it may also be realised as a plain voiceless plosive, as a lenis plosive in an aspiration system) it has been called 'Laryngeal Relativism' instead of 'Laryngeal Realism'. Cyran discusses two Polish dialects (Warsaw and Kraków Polish), which show the same voicing patterns on the surface but, according to Cyran, have different underlying phonological representations: Warsaw Polish is phonologically a voicing language, while Kraków Polish is phonologically an aspiration language but its phonetic realisations correspond mostly to the phonetic values found in voicing languages. Cyran uses evidence from a process of intervocalic voicing that can only occur in phonological aspiration languages as evidence for his analysis. In his analysis, the unmarked plosives in the phonological aspiration language are consistently realised with prevoicing, while the |H| plosives (Cyran uses the elements |L| and |H| for [voice] and [spread glottis] respectively) are realised with short-lag VOT. However, the author does not explain why the language changed its underlying voicing system to an aspiration system without changing the phonetics, or why the language changed a phonetic aspiration system to a phonetic voicing system without changing the phonology. Also, he does not explain how speakers learn to acquire and realise the phonological aspiration contrast as a phonetic voicing contrast: how do they know that the unmarked plosive is consistently realised with prevoicing (which cannot possibly be specified by a phonetics-phonology interface as the absence of a feature cannot have a specific meaning), and that

[1] Cyran makes use of phonological elements instead of features, but these elements are not independently pronounceable in his approach. Instead, the elements function in a way very similar to phonological features: they only have a phonetic realisation in combination with other elements.

the [spread glottis] plosive must be realised without aspiration?

Ignoring these questions for now, Laryngeal Relativism predicts more variation than LR.

[±voice] for both voicing and aspiration contrasts requires language-specific implementation rules. [+voice] can be realised with both prevoicing and short-lag VOT, and [−voice] can be realised with both short-lag VOT and aspiration. This means that the approach does not only predict the existence of voicing and aspiration languages, but also languages that contrast prevoiced ([+voice]) with aspirated ([−voice]) plosives. In theory, the exact opposite (aspirated [+voice] plosives and prevoiced [−voice] plosives) should also be possible, but such an analysis will most likely not occur because the reverse analysis is more logical (at least in the phonological literature it will not occur, as the labels [+voice] and [−voice] seem more or less arbitrary. The linguist will most likely choose for the most natural analysis). As the specific phonetic realisation of the plosives is language-specific and needs to be determined by the language-learning child, different children can come to different phonetic implementations in the case of variable input. I.e., in the transition zone the input is variable, so child A can posit a voicing contrast for [±voice], while child B can posit an aspiration contrast, even though they both receive the same input. It does predict, however, that each speaker has a clear distinction between [+voice] and [−voice].

The voicing distinction in the Italian dialect continuum must be represented with a feature [voice]; [spread glottis] cannot be present in this system as there is a clear voicing distinction for fricatives. The only point of discussion is whether the distribution of voiced and voiceless fricatives is predictable, as in varieties with ISV, or not, as in varieties without ISV.

4.2 Acquisition

Are phonological features universal, or do children extract them from the linguistic input? If features are universal, they must not only be present in the linguistic knowledge of all humans, but they must also have phonetic content so that their usage is similar across different languages. If features are not universal, they have to emerge during acquisition, based on the contrasts present in the language a child is learning. This is of course identical to a usage-based approach, even though the latter (Exemplar Theory) does not necessarily require phonological features as we know them (a cloud of linked phonetic realisations can be regarded as a feature). But regardless of the question whether features are innate or not, the potential phonological distinctions in human language have to be present, as has been shown by the different experiments. Children can easily distinguish prevoiced, plain voiceless and aspirated plosives from each other at birth (regardless of the contrast relevant in their language), but they cannot distinguish 40 ms of prevoicing from 70 ms of prevoicing (similar for aspiration). So, one could argue that the question of feature innateness is merely a philosophical one: are all potential human contrasts linked to phonological features already at birth, or are features only linked to the phonological contrasts relevant in the language, and not to the contrasts the child has lost?

Before it is possible to compare different phonological analyses of the transition zones (or rather of any linguistic pattern) on their plausibility it is important to know how children acquire language, as the child should be able to acquire the proposed contrast based on the linguistic input it receives. Further, if the Dutch-German transition zone is indeed characterised by the absence of a contrast, as the data in Chapter 2 seem to suggest, it is likely that this confusion arises in the language-learning child: adult speakers are expected to have a fixed phonological system, and to have lost the flexibility to adapt their phonological system to a great extent (cf. loanword adaptation, accents in L2; see e.g. Scovel (2000), Petrovic (2008), Hartshorne, Tenenbaum and Pinker (2018) for an overview of the critical period literature). This section will first lay out the findings of studies on language acquisition, and subsequently present the relevant acquisitional data from the two different areas.

When presented auditory speech stimuli differing in only one dimension, adult speakers are unable to detect small differences within a phonological category present in their language, but they are very good at detecting the same difference between two phonological categories present in their language (e.g. Miller, Wier, Pastore, Kelly and Dooling (1976), Pisoni (1977), Soli (1983)). So, for example, adult speakers of a voicing language will not detect the difference between a plosive with 100 milliseconds of prevoicing and a plosive with 70 milliseconds of prevoicing (a difference of 30 milliseconds in total), because the two plosives belong to the same phonological category. They are, however, able to detect the difference between a plosive with 20 milliseconds of prevoicing and a plosive with 10 milliseconds of voicing lag (again a difference of 30 milliseconds in total), because the two plosives belong to different categories. This phenomenon is called Categorical Perception (CP): speech sounds are perceived as belonging to one category, and within each category differences are not detected by speakers. So, plosives with (for example) a very short voicing lead will never be perceived as ‘slightly /b/ but also a little bit /p/’, but always as clearly belonging to either one category (which category they will be perceived as belonging to will depend on the speaker’s native language).

Early studies on CP of laryngeal contrasts in young infants found that 1-to-4-month-old (Eimas, Siqueland, Jusczyk and Vigorito (1971)) and 6-to-8-month-old English-speaking children (Eilers, Gavin and Wilson (1979)) were better at noticing the distinction between plain voiceless and voiceless aspirated plosives than they were at noticing the distinction between prevoiced and plain voiceless plosives. Eilers et al. explains these results by assuming that the boundary between plain voiceless and voiceless aspirated plosives is an innate boundary, while the boundary between prevoiced and plain voiceless plosives is acquired during acquisition (nature vs. nurture). These results are surprising for several reasons. First, it is difficult to explain why the prevoiced-plain voiceless contrast is abundantly present in languages of the world, often as the only laryngeal contrast (i.e., in favour of the aspiration contrast which is assumed to be innate). If the aspiration contrast were indeed innate while the voicing contrast were learned, it would be difficult to explain why the voicing contrast is often unaccompanied by the aspiration contrast. Second, the boundary between plain voiceless and voiceless aspirated plosives (which VOT values count as plain voiceless, and which as voiceless aspirated) has never been truly determined; estimations put the border around 30 milliseconds but this has never been fully supported by experimental evidence. The boundary between prevoiced and plain voiceless plosives, however, is by nature clearly defined, as prevoicing requires the vocal cords to start vibrating before the plosive release (negative VOT) and plain voiceless plosives require the vocal cords to start vibrating after the plosive release (positive VOT). If the aspiration contrast were innate and the voicing contrast only acquired, the difference between the two would be expected to go the other way: a clearly defined boundary between plain voiceless and voiceless aspirated plosives, and an unclear boundary between prevoiced and plain voiceless plosives.

Later studies have indeed not confirmed these results, but rather found that both contrasts are equally distinguishable for children; see for example Streeter (1976) for 2-month-old Kikuyu-learning infants, Lasky, Syrdal-Lasky and Klein (1975) for 4-6.5 month-old Spanish-learning infants, Aslin, Pisoni, Hennessy and Perey (1981) for 6-12 month-old English-learning infants. Most studies have thus found that at birth, infants are capable of detecting all phonetic distinctions that can possibly be relevant in human language,[2] but based on linguistic experience they preserve the ability to distinguish only those contrasts that have proven to be relevant in their language and lose the ability to distinguish those contrasts that have not proven to be relevant[3] (cf. Kuhl et al. (1992), Hoff (2009), Abramson and Lisker (1970), who discovered that adult Thai speakers could distinguish both a voicing and an aspiration contrast (Thai has a three-way laryngeal distinction), while adult English speakers could only distinguish an aspiration contrast). This perception of contrast, both by adult and by infant listeners, is not gradient but categorical. When presented with a continuum of sounds (e.g. a continuum of sounds changing from /b/ to /p/, so with gradually increasing VOT values) listeners will classify each sound as either /b/ or /p/ but not as something in between /b/ and /p/ (e.g. Abramson

[2] Note that this is in fact necessary, as bilingual children are not reported to have difficulties with one of the languages they acquire, and adopted children acquire the language of their new home country without problems (not to mention the fact that it is impossible for a fetus to have different neurological developments depending on the country its mother is staying during pregnancy).

[3] An early study by Werker and Tees (1984) proposes that between the age of 6 and 12 months old, infants gradually lose their ability to distinguish speech contrasts irrelevant to their native language. Kuhl, Williams, Lacerda, Stevens and Lindblom (1992), on the contrary, found that at the age of 6 months, children's perceptual abilities have already been altered and adapted to their native language. It is important to mention that the different authors studied different contrasts: while Werker and Tees (1984) studied the perception of place contrasts in plosives, Kuhl et al. (1992) studied the perception of different vowel contrasts.

and Lisker (1970), Miyawaki et al. (1975), Werker and Tees (1984)).[4]

Apart from categorical perception, the language-learning child makes use of the distribution of sounds in the linguistic input to decide which phonetic contrasts are relevant and which are not (Saffran, Aslin and Newport (1996), Maye and Gerken (2000); cf. Maye, Werker and Gerken (2002) for adult speakers). The phonetics of a language will always show variation (no single phoneme is consistently realised with precisely the same phonetic values), so the child needs a way to determine which variation is relevant and which is not. A unimodal distribution of a sound will lead the child to assuming the absence of a phonological contrast, while a bimodal distribution will lead the child to postulating a phonological contrast.[5] In the case of VOT, a child that receives bimodally distributed VOT values (either around prevoicing and plain voiceless values, or around plain voiceless and aspiration values) will postulate the relevant laryngeal contrast, but a child that receives unimodally distributed VOT values will decide that VOT is an irrelevant phonetic dimension in the language it is learning, and will subsequently lose the ability to distinguish different VOT contrasts. In case of a bimodal distribution, the child will preserve its ability to distinguish the different VOT categories around the two modes, but it will lose the ability to distinguish the VOT category that is irrelevant in the language.

The graphs in Figures 4.1 and 4.2 show the statistical distribution of VOT values in, respectively, the western and eastern part of the continuum. The x-axis represents different ranges of VOT values, the y-axis represents the probability of encountering an item within a specific VOT range. The continuum has been divided into three regions based on the graphs in Figures 2.29 and 2.30. The distributions show that in the Dutch-German dialect continuum children in the west and east receive bimodally distributed VOT values in the linguistic input.[6] The child will see this as evidence for a linguistically relevant contrast, and will preserve its ability to distinguish the two series. In the west, however, the child will only preserve its ability to distinguish prevoiced from plain voiceless plosives but lose its ability to distinguish plain voiceless from voiceless aspirated plosives, while in the east the child will only preserve its ability to distinguish plain voiceless from voiceless aspirated plosives but lose its ability

[4] Kuhl and Miller (1975) and Kuhl and Miller (1978) found that categorical perception is not unique to humans, but can also be found in other mammals such as the chinchilla. The authors found that chinchillas could distinguish prevoiced, plain voiceless and aspirated plosives, and hypothesise that the mammalian ear is extremely well-suited to hearing the differences between those categories; hence the differences between these categories are similar across many languages.

[5] A trimodal distribution should lead the child to postulating a three-way contrast, but this is not mentioned by the different authors. However, there is no reason to assume that the presence of a trimodal distribution would be interpreted in a different way by the language-learning child.

[6] The two modes are clearly visible in the west, where there is a mode in the negative area and a mode in the positive area. In the eastern area, the two modes are not as clear as in the west, but still a mode around a VOT of 0.01-0.02 ms and a mode around a VOT of 0.03-0.04 ms are visible.

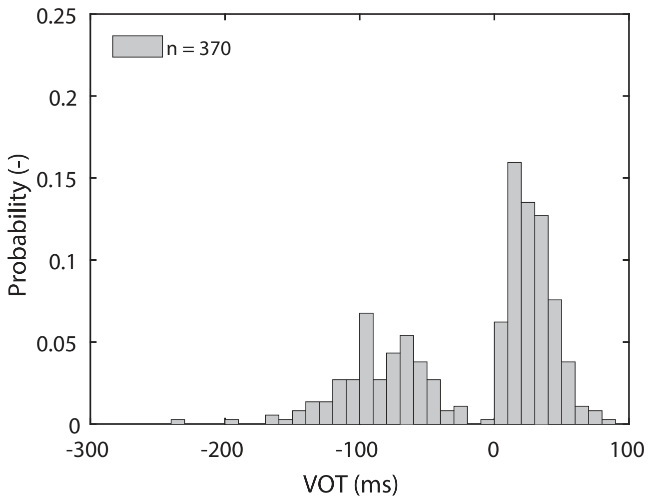


Figure 4.1: Distribution of VOT values in the west

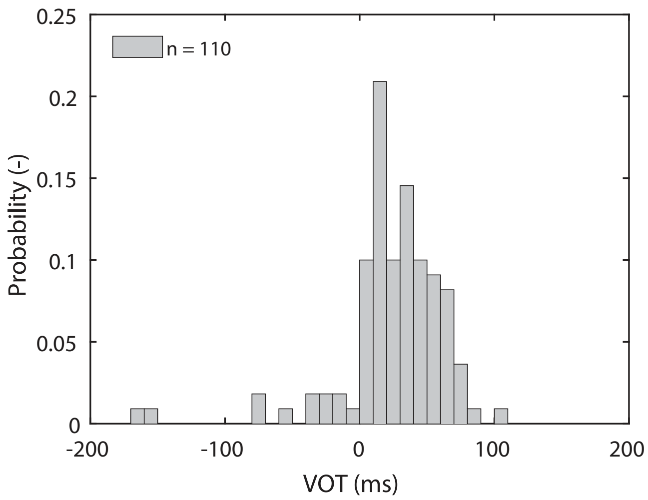


Figure 4.2: Distribution of VOT values in the east

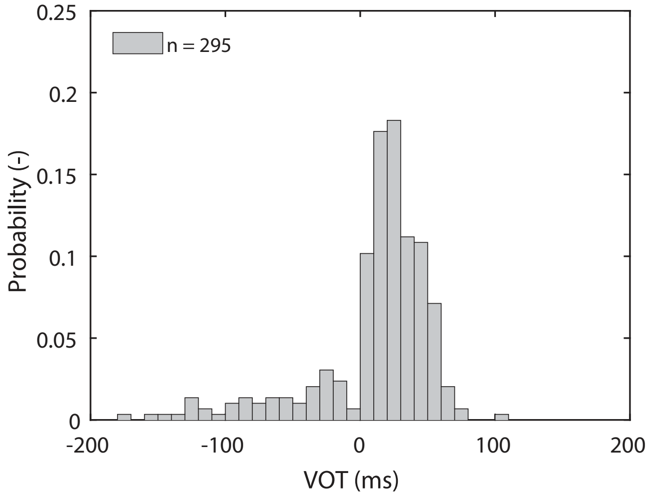


Figure 4.3: Distribution of VOT values in the transitional area

to distinguish prevoiced from plain voiceless plosives.[7]

In the transitional area, however, VOT values do not clearly follow a bimodal distribution, but appear to be scattered around the entire available range. Most values are found in the positive range, but this is influenced by several factors: first, all plosives can be realised with a short-lag VOT, while only the plosives ⟨b⟩ and ⟨d⟩ can be realised with negative VOT values and only the plosives ⟨p⟩, ⟨t⟩ and ⟨k⟩ can be realised with a long-lag VOT. An majority of short-lag VOT values is thus statistically not unexpected. Second, only two lenis plosives (⟨b⟩ and ⟨d⟩) are included in the study, while three fortis plosives (⟨p⟩, ⟨t⟩ and ⟨k⟩) are included. The amount of plosives that can be realised with prevoicing is thus reduced even more by the smaller presence of lenis plosives over all. Third, the range for prevoicing values and aspiration values is bigger than the range for plain voiceless values. This might mean that short-lag VOT values are clustered tightly together around one 'mode', while the values for prevoicing and aspiration value are much more scattered.

In Italy, the relevant distributions concern the values of the voicing ratios. As the entire continuum appears to be divided by a clear boundary, rather than there being a gradual transition, the distributions have been plotted for these two regions (the northern region, containing all and only the Emilian-Romagnol villages, and the southern region, containing all and only the Tuscan villages).

Both graphs show a very clear distribution of voicing proportions, with a large group of fricatives being realised with full voicing (the category 0.99-

[7]Note that this explains why in voicing languages the fortis series may sometimes be realised as a voiceless aspirated instead of a plain voiceless plosive, and why in aspiration languages the lenis plosive may sometimes be realised as a prevoiced instead of plain voiceless plosive, as the child has lost the ability to distinguish the two series.

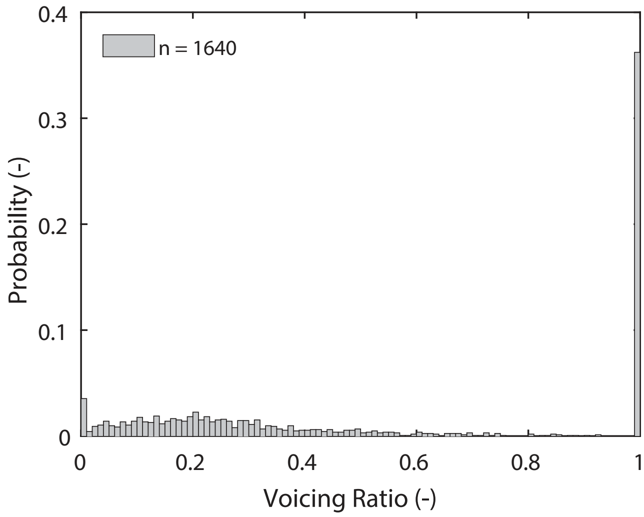


Figure 4.4: Distribution of the proportion of voicing values in the northern region

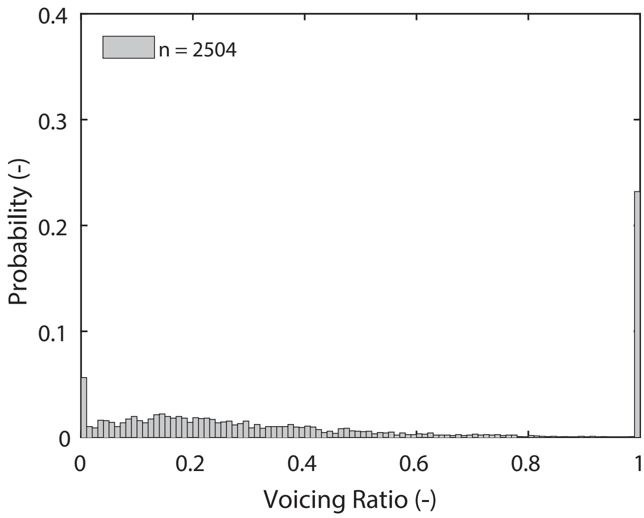


Figure 4.5: Distribution of the proportion of voicing values in the southern region

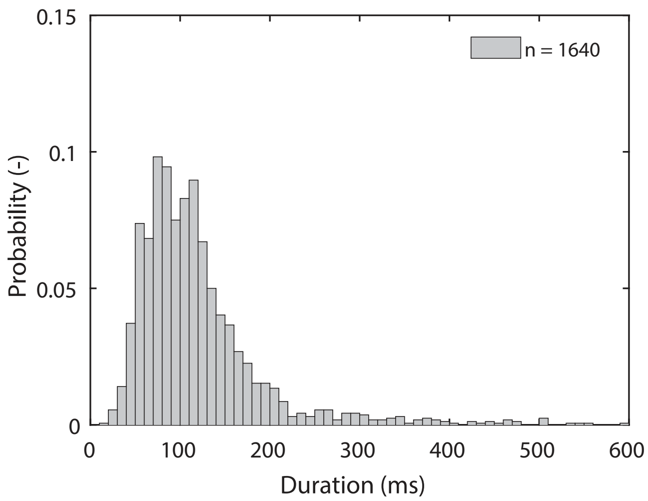


Figure 4.6: Distribution of fricative length in the north

1) and a large group of fricatives realised with partial voicing (the rest of the values). The discrepancy between fricatives in the category 0.99-1 and all the other fricatives is the result of most voiced fricatives being realised with uninterrupted voicing, so that they all fall in the same category, while voiceless fricatives (i.e. the ones with partial voicing) can be realised with any voicing value as long as they are not fully voiced. Although fricatives occur in all categories, there is a bimodal distribution of voicing proportions: the higher proportion values (0.9 and more) occur very infrequently.

As discussed in Chapter 3, both regions employ a length contrast for the alveolar fricative. The linguistic input for both regions does indeed show a bimodal distribution of fricative length, so it is not unlikely that children acquire a phonological length contrast.

4.3 Netherlands-Germany

In this section, the different phonological analyses for the Dutch-German transition zone will be discussed. The different analyses will be compared to the data, the variation that is predicted for the different frameworks and to the literature on acquisition. For the Dutch-German transition zone, the comparison will be between different featural representations, while for the Italian transition zone the comparison will also include different OT constraint rankings. In the Dutch-German transition zone OT rankings are not relevant for the laryngeal distinctions, as these concern the phonetics-phonology interface (although OT rankings are relevant for the assimilation patterns). In this section no distinction between LR and Element Theory will be made, as the two appear to

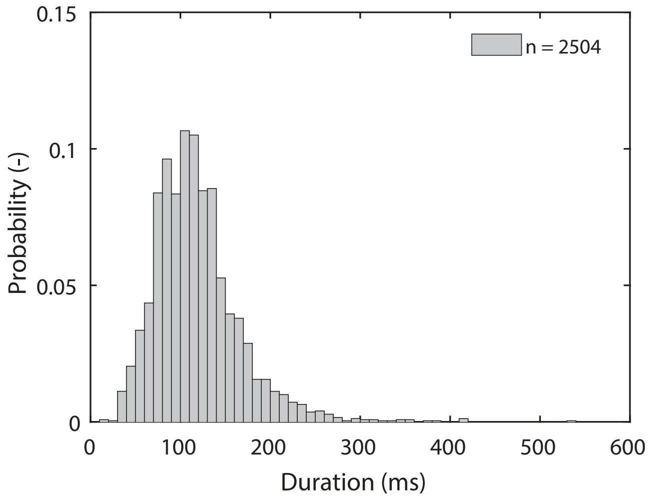


Figure 4.7: Distribution of fricative length in the south

make similar predictions.

4.3.1 [voice - ∅]

If speakers make a phonological distinction between two plosive series, there must be phonetic cues that give them evidence for postulating that particular distinction. If speakers postulate a [voice - ∅] distinction, they must have more evidence for a feature [voice] than for a feature [spread glottis]; or rather, there must be sufficient evidence to postulate [voice] but not sufficient evidence to postulate [spread glottis], as the two features are not mutually exclusive. A phonetic distinction between the two series is expected. If speakers postulate a voicing contrast, it is expected that the lenis series is realised with consistent prevoicing. Some items might be realised without prevoicing as a result of 'phonetic failure', but this number should be relatively low. Looking at the data in the transition zone, it seems unlikely that speakers implement a voicing contrast, as many lenis plosives are realised without prevoicing. The number of lenis plosives with a short-lag VOT is too high to be the result of a phonetic implementation error. Proposing a [voice - ∅] contrast for the transition zone would thus imply that the feature [voice] is not linked to a phonetic implementation. Further, the fortis plosives are unmarked, so they cannot have a prespecified phonetic realisation: the absence of a feature means there is nothing present in the phonology, and 'nothing' obviously cannot be linked to a phonetic realisation. The fortis unmarked plosives are thus expected to receive phonetic implementation based on the phonetic implementation of the [voice] series. Some phonetic variation (short-lag VOT, aspiration) might be expected, but it is likely that a minimal distance requirement would hold. In the transition zone, however, several speakers appear to have reversed VOT values, realising lenis plosives with a higher VOT value than fortis plosives. This requirement would thus be violated here as well, just like the prevoiced realisation of lenis plosives.

If speakers employ a voicing contrast, the assimilation patterns in the transition zone are expected to be similar to the assimilation patterns of other voicing languages. Clusters with a lenis C2 are thus expected to always surface with full voicing, while clusters with a fortis C2 are expected to always surface with partial voicing. If only C1 is lenis (and thus [voice]), assimilation is not expected to apply as [voice] assimilation is regressive, and if both C1 and C2 are fortis (and thus unmarked) phonetic voicing is expected to be blocked by the presence of the feature [voice] in the grammar. The data show a very different pattern, however, as many of the clusters with a fortis C2 are realised with full voicing.

The postulation of a [voice - ∅] contrast, like the postulation of any two-way phonological contrast, requires a bimodal distribution of VOT values. In Section 4.2, it has been shown that in the transition zone VOT values do not follow such a distribution, but rather a unimodal distribution. The likelihood of a child acquiring a two-way phonological contrast is thus very small. Even if one would assume that phonological features are substance-free, and that a [voice - ∅] contrast would in theory not have to be realised as a phonetic voicing contrast, one still needs phonological evidence for the presence of a phonological contrast: if features can be postulated without phonetic and phonological evidence, there are no longer any restrictions and, as a consequence, 'anything goes'. The data in the transition zone do not seem to give any evidence for a [voice - ∅] contrast: the phonetic evidence clearly does not point in that direction, nor does the phonological evidence. Consistent assimilation to the voiced member of the cluster is absent, and more importantly, clusters with a fortis member can surface with full voicing (which would have to be blocked if [voice] were present in the language).

All in all, the data observed in the transition zone do not appear to be very well compatible with a [voice - ∅] analysis.

4.3.2 [∅ - spread glottis]

This representation assumes that language-learners find enough evidence for the presence of a feature [spread glottis], but not enough evidence for a feature [voice]. The two plosive series can be expected to be phonetically distinct. As the fortis series is marked [spread glottis], (relatively) consistent aspiration can be expected for this series. The unmarked lenis series has no phonological feature imposing restrictions on the phonetic realisation, so its phonetic implementation will only be constrained by the feature [spread glottis] (with which it is not supposed to merge), and can thus be expected to show some phonetic variation (not only short-lag VOT but also prevoicing). Looking at the data, the lenis plosives in the transition zone do indeed show phonetic variation between prevoicing and short-lag VOT. However, the fortis plosives are phonetically less stable than expected, as they are not consistently aspirated but rather vary between aspiration and short-lag VOT.

For plosive clusters it would be expected that clusters with two lenis plosives can surface with full voicing. As this has to be a phonetic process instead of a phonological process, it is not necessarily expected to apply consistently. Some clusters might surface with full voicing, while others might surface with only partial voicing. This is indeed what is found in the assimilation data: some of the clusters with a lenis C2 are fully voiced, while others are partially voiced. If C1 is fortis (and hence [spread glottis]), progressive assimilation might be expected to apply, so that the entire cluster is partially voiced (i.e., voicing seizes in the closure phase of the first plosive, and is only resumed after the release of the second). If C2 is fortis, C1 can still have voicing during the closure if it is unmarked, but the closure of C2 will be voiceless so that the entire cluster is only partially voiced. If both plosives are fortis partial voicing for the entire cluster is expected, with voicing seizing during the closure of the first plosive. However, clusters with a fortis C2 in the transition zone show more variation than expected, as a large number of clusters surfaces with full voicing, which is very unexpected if the second plosive is [spread glottis]. It thus appears that the assimilation data in the transition zone are not very well compatible with the [∅ - spread glottis] analysis.

Looking at acquisition, the same argument as put forward in the previous section can be made. Language-learning children can only postulate a [∅ - spread glottis] contrast if the distribution of VOT values in the input is bimodal. In Section 4.2 it was shown that the distribution is, instead, unimodal. It is therefore unlikely that children acquire this two-way contrast based on the input they receive. Summarising, neither the distribution of VOT values nor the phonological patterns in the transition zone point in the direction of this hypothesis.

4.3.3 [voice - spread glottis]

If a speaker postulates the presence of not one laryngeal feature but two, and does not postulate a third, unmarked series in his or her plosive inventory, the speaker must find evidence for the presence of both features. This evidence can be phonetic, phonological or both. A speaker would find phonetic evidence for both features if both prevoicing and aspiration are present in the input; this is indeed the case. If prevoicing is consistently absent from the data there is no evidence to assume the presence of [voice]; if aspiration is consistently absent from the data there is no evidence to assume the presence of [spread glottis]. The representation predicts consistency within both individual series: lenis plosives are consistently prevoiced, and fortis plosives are consistently aspirated. Neither plosive series is likely to show phonetic variation, as there is a phonological feature imposing constraints on the phonetic realisation. The plosives in the transition zone, however, show so much variation in their phonetics that an overlap between the two series is observed. Many lenis plosives are prevoiced, and many fortis plosives are aspirated, but there are just as many lenis plosives with short-lag VOT and just as many fortis plosives with short-lag VOT. These realisations are only compatible with the underlying features [voice] and [spread glottis] if these features do not have a clear link with a phonetic realisation: if the link between phonetic realisation and phonological feature is strong, the transition zone would have to have consistent prevoicing of lenis plosives, and consistent aspiration of fortis plosives. However, if the link is indeed weak, it is not immediately clear how the language-learning child finds evidence for postulating a feature in the first place: if it encounters a plosive with short-lag VOT, it may be underlyingly specified as either [voice] or [spread glottis]. The choice for the feature cannot be based on these realisations, so it will have to be based on the realisations that are consistent with the underlying feature, implying that the link between the feature and the phonetic realisation is in fact quite strong instead of weak.

A speaker can find phonological evidence for both features if both features can be proven to be phonologically active in the language. The predictions for assimilation in plosive clusters, however, are not very clear. In voicing languages, the feature [voice] can spread regressively, while in aspiration languages the feature [spread glottis] can spread progressively. Feature spreading is, however, not attested the other way around (progressive assimilation of [voice] or regressive assimilation of [spread glottis]). It is therefore not clear what will happen in clusters with different laryngeal features. If both C1 and C2 are lenis, the cluster will surface with full voicing as both plosives are marked [voice]. If both C1 and C2 are fortis, the cluster will surface with partial voicing as both plosives are marked [spread glottis]. In clusters with a lenis ([voice]) C1 and a fortis ([spread glottis]) C2, neither feature is expected to be spread, as [voice] is only attested in regressive but not progressive assimilation, and [spread glottis] is only attested in progressive but not regressive assimilation. It is thus expected that the first plosive has voicing during the closure, but that voicing seizes in the closure phase of the second plosive. Voicing of the entire cluster is thus partial. In plosive clusters with a fortis ([spread glottis]) C1 and a lenis ([voice]) C2, both plosives are in a position from which they can spread: [spread glottis] can spread progressively from C1 to C2, and [voice] can spread regressively from C2 to C1. It is unclear which of the two features would be chosen to spread and thus replace the other feature, if spreading would even happen at all. Most likely is that both features have equal phonological status and neither will assimilate to the other. In that case, voicing in the closure phase of the first plosive is expected to be partial, while voicing should be resumed during the closure phase of the second plosive.

The assimilation data in the transition zone do not seem to be very well compatible with the scenarios described above. In plosive clusters with a fortis C2 no full voicing is expected, regardless of the phonological specifications of the first plosive. However, the data show a large amount of fully voiced clusters with a fortis C2.

A bimodal distribution of VOT values in a child's linguistic input is again necessary for the acquisition of this phonological contrast, and either phonetically consistent realisations of both plosive series, or phonological activity of both plosive series is required for the postulation of two features in case of a two-way contrast. As none of these prerequisites is satisfied, the hypothesis can be deemed very unlikely.

4.3.4 Three-way contrast for all speakers

If speakers in the transition zone postulate a three-way contrast,[8] the prediction is that the voiced category contains only plosives that are lenis in both the western voicing system and the eastern aspiration system, that the aspirated category contains only plosives that are fortis in both the western voicing system and the eastern aspiration system, and that the unmarked category combines lenis and fortis plosives (i.e. plosives that would be considered either lenis or fortis in both the western voicing system and the eastern aspiration system). These speakers would thus show some lexical variation compared to the speakers with a two-way contrast. An important first question is how speakers would come to postulate such a contrast system. Most likely, such an analysis can only be postulated if speakers do not allow phonetic variation within a phonological category: if they would allow phonetic variation within a category they would most likely postulate a two-way contrast (either [voice - ∅], [∅ - spread glottis] or [voice - spread glottis]), as the neighbouring dialects all give evidence for a two-way contrast. If speakers do not allow phonetic variation within a phonological category, a plosive in a specific lexical item can only have the feature [voice] in its representation if it is consistently realised with prevoicing. The same is true for the feature [spread glottis]: it can only be postulated in the representation of an individual plosive if that plosive is consistently aspirated. Plosives that are either consistently realised with a short-lag VOT, or whose realisation varies between prevoicing and short-lag VOT or between aspiration and short-lag VOT must be analysed as unmarked. This, in turn, means that speakers are expected to have a stable phonetic realisation as well: if a speaker cannot match variation in the realisation of a plosive with a phonological feature, the presence of a phonological feature must have a consistent phonetic realisation as well. As there are three phonologically distinct plosive series present in the grammar, the three systems are expected to be phonetically distinct. Phonetic overlap between the series is thus not expected. Finally, this hypothesis predicts some stability in the mapping of the two series from the stable areas onto the three series of the transition zone: as lenis plosives are never realised with aspiration (neither in the west nor in the east), they can never be realised with aspiration (and hence be analysed as [spread glottis]) in the transition zone. Similarly for fortis plosives: as they are never realised with prevoicing (neither in the west nor in the east), they can never be realised with

[8] [voice - ∅ - spread glottis].

prevoicing (and hence be analysed as [voice]) in the transition zone.

A three-way contrast is thus only likely to arise if speakers receive sufficient evidence to postulate [voice] for a particular lexical item, sufficient evidence to postulate [spread glottis] for another lexical item, while for yet another lexical item there is not sufficient evidence to postulate either category. In other words, the input across the lexicon must be consistent for such a contrast to arise. This means that speakers must hear a specific set of lexical items being uttered with only prevoicing, which means that they must receive this input from speakers with a voicing system (speakers with an aspiration system might realise lenis plosives with prevoicing but they do not do so consistently, while speakers with a voicing contrast do have consistent prevoicing), while they must hear another set of lexical item being realised with consistent aspiration, which means that they must receive this input from speakers with an aspiration system (speakers with a voicing system might realise fortis plosives with aspiration but they do not do so consistently, while speakers with an aspiration contrast do have consistent aspiration). Speakers are expected to have difficulties in postulating a laryngeal feature only for the set of lexical items for which they receive input from both sides of the continuum (i.e., from both a voicing system and an aspiration system), so that these items end up as unmarked. Which input a single speaker receives is of course unpredictable (depending on social relations, work, education, physical boundaries between locations and many other factors), but in general one can expect to find more [voice] lenis plosives in the west of the transition zone and more unmarked lenis plosives in the east of the transition zone. For fortis plosives, on the other hand, more unmarked plosives can be expected in the west and more [spread glottis] fortis plosives can be expected in the east, simply because of the influence of the systems in the west and east. In the west speakers thus have a relatively large chance to encounter lexical items with a consistently prevoiced lenis plosive, but they have a very small chance to encounter lexical items with consistently aspirated fortis plosives. In the east the chances are reversed. It thus seems unlikely that speakers will encounter a linguistic input that is compatible with a three-way contrast, simply because of the geographical distribution of the two laryngeal systems.

This hypothesis has some clear predictions for assimilation in plosive clusters. Only clusters with a lenis C2 that is marked ([voice]), ánd a C1 that is either marked [voice] or unmarked (either lenis or fortis) can surface with full voicing. If C1 is [spread glottis] there is a clash between features: [spread glottis] can spread progressively, thereby devoicing the entire cluster, but [voice] can spread regressively, thereby voicing the entire cluster. Or, because of the clash, no assimilation takes place in the cluster. However, if C2 is fortis (either [spread glottis] or unmarked), full voicing should not be possible. If C1 is marked [voice], that feature cannot progressively spread to C2 so assimilation is not expected to be possible (even if C2 is unmarked), and [spread glottis] on C2 blocks full phonetic voicing. If C1 is unmarked, phonological assimilation is not likely and phonetic voicing is expected to be blocked by the feature [spread glottis] on C2. Finally, if C1 is [spread glottis] as well, full voicing can only be the result of a phonetic process, but this is likely to be blocked by the two features. If both C1 and C2 are unmarked, phonetic voicing is likely blocked by the presence of the feature [voice] in the system (to 'avoid' a clash). The clusters with a fortis C2 in the transition zone are thus not expected to be realised with full voicing. However, a quick glance at the data shows that a large amount of these clusters in fact does surface with full voicing.

In order for a child to acquire a three-way contrast, finally, the linguistic input needs to contain a trimodal distribution of VOT values. Only in that case will the child assume all three VOT categories are phonologically relevant. The distribution of VOT values in the transition zone has already been shown to be unimodal rather than bimodal or even trimodal. The language-learning child is thus not provided with any evidence for a three-way contrast.

4.3.5 No contrast for all speakers

In a scenario without a laryngeal contrast, speakers postulate the presence of only one plosive series. With respect to the phonetic realisation of the plosives, two possibilities exist. The first is that all plosives are realised with a short-lag VOT. Cross-linguistically this is the most frequent phonetic implementation (Maddieson (1984)). The second option is that there are large amounts of phonetic variation. As there is no laryngeal feature present in the phonological representation of the plosives, there are no restrictions on the plosive's VOT values. Further, there is no other contrasting plosive series which would restrict the specific phonetic realisation of the unmarked series. Based on the transition zone being located between a voicing and an aspiration area, some phonetic variation might be expected, as speakers receive input containing not only plain voiceless plosives (which is the expected realisation for the unmarked plosive series) but also prevoiced and aspirated plosives (cf. Exemplar Theory).

As the phonetic output in the transition zone is highly variable, it is very well possible that speakers fail to postulate a stable phonological contrast for the different plosives. If LR is strict on the link between phonetics and phonology, speakers in the transition zone cannot posit the feature [voice] (or the element |L|) for lenis plosives, as they are not consistently realised with prevoicing. An unmarked lenis plosive is thus the most likely outcome. However, speakers cannot postulate the feature [spread glottis] (or the element |H|) for fortis plosives either, as they are not consistently realised with aspiration. An unmarked fortis plosive is thus the most likely outcome. Speakers in the transition zone would then end up without a laryngeal contrast for plosives, as there is too much phonetic variation present in the realisation of the plosives, and the existence of either feature (or element) cannot be deducted from the input. Speakers thus only have one, laryngeally empty ([∅]) plosive series.

Comparing this analysis to the plosive cluster data, it does not seem to be a very unlikely option. Clusters with both a 'lenis' and a 'fortis' C2[9] can surface

[9] If there is only one plosive series, it is of course impossible to distinguish lenis and fortis

with full voicing phonetically. If speakers would have a phonological voicing distinction, only clusters with a lenis ([voice]) C2 would be able to surface with full voicing, whereas clusters with a fortis ([spread glottis]) C1 would surface as voiceless. If C1 were lenis, and thus marked for [voice], assimilation (leading to full voicing of the cluster) would not be possible, as voicing assimilation is only regressive (from C2 to C1) and not progressive (from C1 to C2). If C1 were fortis, and thus laryngeally unmarked, phonetic voicing of the cluster[10] is prohibited to 'avoid' a merger with clusters with a lenis C2. If speakers would have a phonological aspiration contrast, however, only clusters with two lenis plosives can surface with full voicing. If C2 is fortis, the cluster will surface with partial voicing, as the presence of the feature [spread glottis] on the second consonant will block full voicing. If C1 is fortis, the feature [spread glottis] might spread to C2 causing devoicing of the cluster. If both consonants are fortis, and thus marked [spread glottis], the entire cluster is resistant to voicing (either phonetically or phonologically). Summarising, if the transition zone would be characterised by either a voicing or an aspiration contrast, only clusters with a lenis C2 could surface with full voicing (with the extra requirement of a lenis C1 in the case of an aspiration system). Clusters with a fortis C2 cannot surface with full voicing in either system. In a system without a laryngeal contrast, however, all consonant clusters can surface with full voicing. There is no feature [voice] that prohibits phonetic voicing in unmarked clusters, and there is no feature [spread glottis] that prohibits phonetic voicing in clusters with at least one fortis plosive.

The behaviour of intervocalic plosives further supports this analysis. The lenis bilabial and fortis velar plosive were not present in intervocalic position in the German database, so they cannot be used in the analysis. The other three plosives, however, show that a phonological contrast in the middle area is very unlikely, as fortis plosives in intervocalic position are frequently realised with full voicing. Regardless of the laryngeal contrast, fortis plosives are never expected to be realised with full voicing. If there is no phonological contrast, however, fortis plosives can be realised with either full or partial voicing.[11]

Besides the phonetic data, insights from language acquisition support postulating the absence of a contrast. As was shown in Section 4.2, VOT values in the middle (transitional) area follow a unimodal distribution. In order to be able to postulate a phonological contrast, however, children need to encounter a bimodal[12] distribution of phonetic values. As the input in the transition

plosives. However, the terms are used here to refer to the distinction as made in the two stable areas.

[10] As assimilation is a phonological process, full voicing of a fully unmarked cluster must be phonetic: there is no laryngeal feature that can cause voicing of the entire cluster.

[11] If a language has a process which neutralises a laryngeal contrast in intervocalic position, the presence of fortis plosives with full voicing can be easily explained. Based on only intervocalic plosives it is impossible to distinguish languages without a laryngeal contrast from languages that neutralise the contrast in intervocalic position. However, the data are used here merely to show they support the analysis.

[12] Or trimodal; only a unimodal distribution cannot result in a phonological distinction.

zone contains a unimodal distribution of VOT values, the language-learning child will not find evidence for an underlying laryngeal contrast. The language-learning child will then, obviously, not be able to postulate a laryngeal contrast in the phonology.

The data from the present study can be compared to the data from Caramazza, Yeni-Komshian, Zurif and Carbone (1973) and Caramazza and Yeni-Komshian (1974), who found an overlap in VOT values between lenis and fortis plosives in Canadian French. The authors propose that this overlap can be explained by the influence of Canadian English. They assume Canadian French is changing from a voicing system to an aspiration system.

4.3.6 Interspeaker variation

In a situation with interspeaker variation, different speakers postulate different phonological contrasts. While some speakers postulate a voicing contrast, others may postulate an aspiration contrast, a three-way contrast or no contrast at all. Such a scenario predicts a certain level of stability within individual speakers' realisations of plosives. Lenis plosives are either always [voice] (in a voicing system) or always unmarked (in an aspiration system), and fortis plosives are either always unmarked (in a voicing system) or always marked [spread glottis] (in an aspiration system). The phonetic realisations of these plosives are expected to match their phonological representations: a lenis plosive marked for [voice] is expected to be realised with prevoicing, while a fortis plosive marked for [spread glottis] is expected to be realised with aspiration. In either system the unmarked series may show some phonetic variation. Further, speakers are expected to show consistency in plosive clusters corresponding to the consistency in single plosives. If they have a consistent voicing contrast ([voice - ∅]) they should display consistent voicing assimilation in clusters (a fortis C1 assimilates in voicing to a following lenis C2 so the entire cluster surfaces as voiced, or a lenis C1 does not undergo final devoicing under the influence of a following lenis C2). If they have an aspiration contrast ([∅ - spread glottis]), a consonant cluster is only expected to surface as voiced if both plosives are lenis (in which case phonetic voicing may apply). In all other cases the cluster is expected to surface as voiceless: if C1 is voiceless it is expected to block passive phonetic voicing of a lenis C2, and if C2 is voiceless it is expected to block passive phonetic voicing of a lenis C1.

In the preceding sections the different options considered here have been discussed individually. All hypotheses, except the hypothesis excluding a laryngeal contrast, have been shown to suffer from several different problems. In short, all hypotheses proposing a laryngeal contrast can be ruled out by both the assimilation patterns and the distribution of VOT values in the linguistic input. The proposition that different speakers in the transition zone postulate different contrasts is unlikely, not only because of the problems most individual analyses face, but also because it implies that different speakers would have to receive different types of input. More specifically, different language-learning children would have to find different distributions of VOT values in the linguistic input. Some will have to find a bimodal distribution supporting a voicing contrast, others will have to find a bimodal distribution supporting an aspiration contrast, some will have to find a trimodal distribution supporting a three-way contrast and yet others will have to find a unimodal distribution supporting no contrast. As it is impossible to register all the linguistic input a single child receives (let alone the input several children receive), it is impossible to decide how likely this hypothesis is exactly. It is simply impossible to prove that some children receive different VOT distributions than other children. However, the data show that some speakers have a voicing contrast for one PoA but an aspiration contrast for another PoA, other speakers contrast prevoiced with aspirated plosives for one PoA while they do not show a contrast for another PoA, and yet other speakers have a consistent voicing or a consistent aspiration contrast. Because of this large amount of variability in the data, it is unlikely that some children will receive input compatible with one system, while other children will receive input compatible with another system,[13] if only because of the large amounts of intraspeaker variation. The next section will look at intraspeaker variation in more detail.

4.3.7 Intraspeaker variation

In a situation where there is intraspeaker variation, speakers may postulate different phonological contrasts for different PoA's. Stability within the places of articulation is then expected. The fact that speakers postulate two different phonological representations for the two series would mean that there is too much phonetic variation between the places of articulation for them to be represented by the same contrast. Therefore, phonetic variation must be quite restricted for the phonological features. This means that every speaker must, per PoA, have either a clear voicing contrast or a clear aspiration contrast. In principle, nothing rules out one speaker having a [voice - ∅] contrast for both PoA's, another speaker a [∅ - spread glottis] contrast for both PoA's, another speaker [voice - ∅] for bilabials and [∅ - spread glottis] for alveolars, and yet another speaker [∅ - spread glottis] for bilabials and [voice - ∅] for alveolars. However, the intraspeaker variation does imply that there is not much phonetic variation possible for the phonetic realisation of a phonological feature. This means that speakers with a different representation of the contrast for the same PoA must have had very different input, and that the input per speaker must be consistent. Based on the data observed in the transition zone, this does not seem very likely. Both fortis and lenis plosives for both PoA's can have very different VOT values, depending on the speaker. The input listeners receive will thus vary. If a phonological feature does indeed not allow for much phonetic variation (as predicted by this analysis), then speakers are not likely to posit

[13]It is equally unlikely that children will ignore input compatible with one system but take input compatible with another system into account, as the child has to postulate a phonological grammar based on all input it receives.

such an analysis for the observed data. It must be noted that this analysis does, in fact, not predict much other than stability within each PoA. As the current data only has one item per plosive available, it is impossible to falsify this hypothesis: it can never be proven that there is stability or instability within each PoA.

For plosive clusters the hypothesis makes some new predictions. The possibility of assimilation taking place would depend on the PoA of the plosives if different PoA's can have different phonological representations: if the first plosive would be an unmarked fortis plosive, an unmarked lenis plosive or a lenis plosive marked for [voice], the second plosive can only trigger assimilation if it is marked for [voice]. If the second plosive would belong to the PoA marked with an aspiration contrast, assimilation cannot take place. However, plosive clusters with a fortis C2 could surface as fully voiced if C2 is unmarked, and if C1 is lenis ([voice] or unmarked). In both cases the voicing would be the result of a phonetic process: if C1 is marked for [voice] it cannot spread progressively, so voicing of the entire cluster cannot be the result of a phonological process. If C1 is unmarked, there is no phonological feature present that can cause assimilation, so voicing of the cluster must again be a phonetic process. Because of the small number of items in the study it is difficult to falsify this hypothesis, as it is difficult to determine the specific phonological representation of each PoA. However, the large number of fully voiced clusters with a fortis C2 is striking. As mentioned, these clusters could only surface with full voicing if the fortis C2 is phonologically unmarked (as it would be in a voicing system), and if C1 is either marked [voice] or unmarked. In all cases a fully voiced cluster can only be the result of a phonetic process, so it would have to apply inconsistently. Looking at the realisations of clusters with a fortis C2, a phonetic process is more likely than a phonological process, as voicing indeed applies inconsistently.

If children are to acquire different contrasts for different PoA's, they have to receive different VOT distributions in the linguistic input depending on the PoA of the plosive. An individual child has to receive input consistently compatible with a voicing contrast for alveolars, and input consistently compatible with an aspiration contrast for bilabials (or vice versa).[14] Just like the hypothesis of interspeaker variation, it is impossible to prove which input a child does and does not receive. However, considering the fact that children will receive input from many different speakers in their geographical region, and considering the fact that many different speakers in the transition zone (i.e., the geographical region of the language-learning child of interest here) display different laryngeal contrasts within one PoA, it is unlikely that a child will postulate different laryngeal contrasts for the different PoA's. If a three-way laryngeal contrast or no laryngeal contrast are to be included as well, this hypothesis is rendered even more unlikely.

[14]Note that a prevoicing-aspiration contrast for one PoA, or no laryngeal contrast for another PoA are also possible variations.

4.3.8 Lexical variation

If one assumes lexical variation (i.e., some plosives can be prevoiced, others plain voiceless, yet others aspirated), and the representation of a plosive in one lexical item can differ between individual speakers, then one does in fact not make any predictions at all, other than consistency within one speaker's representations.[15] This hypothesis is not very well testable because of the scarcity of the data, but it can be argued to be very unlikely. If speakers can choose whichever phonological representation for a phoneme in a specific lexical item, one would expect an abundance of variation between different speakers. Where some speakers might choose a prevoiced realisation, others might implement a plain voiceless realisation, and yet others might aspirate the plosive. Such a pattern is not found in the data: plosives that are lenis on either side of the continuum are realised lenis in the transition zone, and plosives that are fortis on either side of the transition zone are realised fortis in the transition zone. Although an abundance of variation is present, there do appear to be limits on it. This seems difficult to match with the hypothesis of lexical variation.

If the three different laryngeal categories are present in an individual speaker's system, it is unclear what the predictions with respect to assimilation are. If an unmarked plosive and a prevoiced plosive are adjacent, assimilation will occur if the prevoiced plosive is the second of the cluster, and if an unmarked plosive and an aspirated plosive are adjacent, assimilation will occur if the aspirated plosive is the first of the cluster. In all other cases (reversed order of the plosive in the two examples just given, two adjacent unmarked plosives, or an adjacent prevoiced and an aspirated plosive in whichever order), no predictions can be made. The fully voiced clusters with a fortis C2, as found in the data, might in theory be possible in the case of lexical variation, if the fortis C2 would in fact be lenis, or if both members of the cluster are unmarked.

If children are to acquire laryngeal contrasts with lexical variation they would have to receive consistent linguistic input with respect to individual lexical items, meaning that the same plosive must always be present in one lexical item. Looking at the data this seems highly unlikely, as different speakers have different realisations of the same lexical item. For a child to receive consistent input with respect to the plosive in an individual item, it would thus have to be in contact only with a specific group of speakers, who all happen to have identical realisations of plosives in individual lexical items. This is already very unlikely for one lexical item (if speakers can show lexical variation between them), but for an entire vocabulary this is even more unlikely.

[15]Note that, if speakers have access to all three laryngeal categories, possible phonetic variation is much smaller than when speakers contrast only two laryngeal categories. In the latter case, one of the laryngeal categories present may show variation not only within its own phonetic 'domain', but also in the domain of the unused laryngeal category. If speakers use all three categories this variation is not expected.

4.3.9 [±voice]

When assuming voicing contrasts in all languages are represented as a [±voice] distinction (not only phonetic voicing contrasts but also phonetic aspiration contrasts), the phonetic implementation of the two features must be language-specific. This means that the language-learning child has to decide which phonetic realisation of [+voice] and [−voice] is most compatible with the linguistic input. Upon receiving consistent input (either a voicing or an aspiration contrast), the child will most likely end up with the same phonetic contrast as found in the input, as there is no evidence for a different phonetic implementation. Contrary to what LR assumes, the child does not have to acquire the relevant phonological contrast based on the phonetic input, as the only possible representation assumes that lenis plosives are specified as [+voice] and fortis plosives as [−voice]. If the child receives highly variable phonetic input, as is the case in the transition zone, it will not be confused about which phonological distinction it has to postulate (this is always [±voice]), but it will only be confused about the phonetic implementation of [+voice] and [−voice]. It is uncertain how the child will handle this variation, but it might be expected that it does not fail to postulate a consistent phonetic implementation. The child must know, to a certain extent, that [+voice] can be realised with either prevoicing or short-lag VOT, and that [−voice] can be realised with either short-lag VOT or aspiration.[16] If a plosive-initial word is realised with prevoicing in one case, and with short-lag VOT in another case, the child will most likely be able to postulate a feature [+voice] for this, as it will be able to postulate [−voice] for plosive-initial words with short-lag VOT in one instance and aspiration in another. As the child is likely to specify a phonetic implementation for both features, and is also likely to keep the two series distinct, a possibility is that the child chooses the most frequent implementation for both plosives. In that case, the child will end up with a contrast (prevoicing vs. short-lag VOT, short-lag VOT vs. aspiration or prevoicing vs. aspiration) if prevoicing for [+voice] is more frequent than short-lag VOT and/or if aspiration for [−voice] is more frequent than short-lag VOT. Only if short-lag VOT is most frequent for both [+voice] and [−voice], the child will not be able to postulate a phonetic contrast between the two series. In that case, as [+voice] and [−voice] are similar, it might be expected that the child no longer feels the need to postulate a phonological distinction between the two series, and ends up with only one phonological plosive series. In this case, children do not necessarily decide on the correct phonetic implementation of plosives: because there is no voicing contrast, exact voicing values are not very relevant. This also means, however, that children are free to implement whatever voicing values they want to, as there will not be a merger with another series. This would be consistent with the patterns found in the transition zone.

[16] There are no languages documented with a prevoiced realisation of [−voice] or an aspirated realisation of [+voice], so the phonetics-phonology interface would have to provide a way to block this.

If the child does find evidence for a distinct phonetic implementation of [±voice] (whether it be a phonetic voicing contrast, a phonetic aspiration contrast, or prevoicing vs. aspiration), it will most likely choose to implement this contrast consistently, so the phonetic variation found in the transition zone is very unexpected. Also, if the child uses frequency to decide what the correct phonetic implementation of the contrast is, it may be expected that different children end up with different implementations: small variations in the number of plosives realised with prevoicing, short-lag VOT or aspiration might lead to a different phonetic characteristic being the most frequent, and thus a different phonetic implementation.

With respect to assimilation, the [±voice] approach does not make clear predictions. Only if speakers end up without a phonological contrast can we predict what will happen in plosive clusters: either they are partially voiced, or they are fully voiced. There is no feature imposing restrictions on phonetic realisation (there is not even a phonetic implementation rule present, maybe, as the absence of a feature cannot be linked to a phonetic implementation). If speakers do employ a [±voice] contrast, however, the predictions are unclear (regardless of the phonetic implementation of the contrast; assimilation is a phonological process, so the phonetic implementation of the contrast is unlikely to have an effect on this). Both [+voice] and [−voice] are present in the feature system of speakers, so they are expected to have identical status in the phonology and thus both should have the ability to be spread in plosive clusters. In clusters with two [+voice] or two [−voice] plosives, it is evident that the cluster will be realised with voicing values corresponding to the implementation of either feature (i.e., fully voiced in the case of two [+voice] plosives, and partially voiced in the case of two [−voice] plosives). However, in clusters with a [+voice] and a [−voice] plosive (in whichever order they might appear), both features have, in theory, the possibility of spreading to the other plosive. There is nothing in the phonology that prevents one feature or the other from spreading or being assimilated to the other feature.[17] In general, and in the transition zone as well, there are four options in total: a) no assimilation occurs; b) assimilation to [+voice] occurs; c) assimilation to [−voice] occurs; d) both plosives change values. The latter option appears to be very unlikely. Assimilation processes affecting both phonemes have not been described in the literature; rather, the process always involves a trigger phoneme and a targeted phoneme. In the final option described here, both phonemes would function as both the trigger and the target phoneme. As this is highly unlikely, the option will not be discussed in any more detail here. The other options described do not appear to fully de-

[17] In traditional phonological voicing representations the same problem exists, as they cannot explain why languages with a phonetic voicing contrast only allow for regressive spreading of [+voice], while languages with a phonetic aspiration contrast only allow for progressive spreading of [−voice]. Progressive spreading of [+voice] and spreading of [−voice] in both directions in voicing languages, and regressive spreading of [−voice] and spreading of [+voice] in both directions in aspiration languages must all be excluded on a language-specific basis, even though this pattern appears to be universal.

scribe the patterns found in the transition zone either. Assimilation to [+voice] would predict that clusters with a lenis C2 always surface as fully voiced, which is evidently not true. If no assimilation would occur, clusters are expected to surface faithful to their underlying representation. This means that neither clusters with a lenis C2, nor clusters with a fortis C2 are expected to surface with full voicing: in clusters with a lenis C2, the first plosive would either be voiceless underlyingly or as a result of final devoicing (so the first part of the cluster should be voiceless), while in clusters with a fortis C2, the second half of the cluster should be voiceless. The assimilation data show a very different pattern, with full voicing for both clusters with a lenis C2 and clusters with a fortis C2. These patterns are not compatible with any of the four possibilities described in this section.[18]

Since the [±voice] contrast is a two-way contrast regardless of its phonetic implementation, a child could postulate this contrast if the linguistic input contains a bimodal distribution of VOT values. Whether VOT values are bimodally distributed around a voicing contrast or around an aspiration contrast, is not relevant for the phonological contrast the child will postulate (as [±voice] can represent both a voicing and an aspiration contrast), but is only relevant for the phonetic implementation of the contrast the child will choose (as the phonetic implementation the child eventually postulates must be in accordance with the linguistic input). However, for the child to postulate a [±voice] contrast, it will have to receive a consistently bimodal distribution of VOT values: VOT values must be bimodally distributed around either a voicing contrast, or around an aspiration contrast. If a child receives input containing VOT values that follow a unimodal or trimodal distribution, it will be impossible to postulate a [±voice] contrast. Looking at the distribution of VOT values in the transition zone, it is unlikely that the individual child will receive a consistently bimodal distribution of VOT values in the linguistic input. Even though it is irrelevant whether the distribution centres around a voicing contrast or around an aspiration contrast, the language-learning child in the transition zone would have to receive input from only one part of the speakers in the transition zone (namely only those with a voicing contrast, or only those with an aspiration contrast). Speakers employing the opposite contrast would have to be (almost) completely absent from the child's input. As both contrasts are present in the entire transition zone, this does not seem very likely.

4.3.10 A voicing contrast for the entire continuum

Cyran (2011) and Cyran (2013) discuss voicing phenomena in two different Polish dialects (Cracow-Poznań Polish and Warsaw Polish). Although on the surface both dialects appear to be voicing dialects, Cracow-Poznań Polish has an extra sandhi process which voices the final obstruent of a word when it is followed by a vowel or sonorant in the following word. As final devoicing is active

[18] The databases used in this study are too small to study the behaviour of clusters with a fortis ([-voice]) C1.

in all varieties of Polish, the word-final obstruent must be either underlyingly voiceless (i.e. laryngeally unmarked), or voiceless because of final devoicing (i.e. laryngeally unmarked). As vowels and sonorant consonants lack a laryngeal feature (voicing is not contrastive for vowels and sonorants, so it is not represented in the phonology) they cannot spread the element |L|[19] to a preceding obstruent; they are unable to trigger voicing assimilation. If Cracow-Poznań Polish is indeed a phonological voicing language, a laryngeally unmarked obstruent (either underlyingly unmarked or unmarked as a result of final devoicing) cannot undergo a phonetic process of intervocalic voicing: the presence of the element |L| in the phonology of the language prohibits the existence of such a process. Cyran proposes to analyse the language as a phonological aspiration language rather than a phonological voicing language: lenis plosives are unmarked, fortis plosives are marked with the element |H|. Such an analysis requires the loss of the link between phonological features and phonetic realisation, as Cracow-Poznań Polish phonetically employs a voicing contrast: in word-initial position the unmarked plosives are realised with prevoicing, and the plosives marked |H| are realised with a short-lag VOT. Cyran therefore refers to his approach as 'Laryngeal Relativism', rather than Laryngeal Realism. The voicing of unmarked obstruents is argued to be the result of passive voicing: a phonetic process which spreads the voicing of a following vowel or sonorant consonant to the unmarked obstruent. Final devoicing must be reanalysed as a process prohibiting the presence of a laryngeal element in word-final position, so that a word-final |H| is deleted. This word-final obstruent cannot be subject to passive voicing as there is no following vowel or sonorant; the same holds for word-final lenis obstruents. In intervocalic position, the lenis (unmarked) obstruent is subject to passive voicing, but as the fortis obstruent has the element |H| in its representation (which does not delete in intervocalic position) passive voicing of fortis obstruents in intervocalic position is blocked. Assimilation from a lenis plosive to a fortis plosive, which can no longer be analysed as spreading of the element |L|, is reanalysed as delinking of the element |H| from the first plosive of the cluster. The entire cluster subsequently undergoes passive voicing. Most importantly, the Cracow-Poznań Polish sandhi voicing can now be analysed in terms of delinking as well: in word-final position the element |H| is delinked, and the obstruent subsequently undergoes passive voicing triggered by the following vowel or sonorant. Note that sandhi voicing and final devoicing are now the result of the same process: delinking of a word-final laryngeal element. The difference between the two is that in the case of final devoicing there is no following vowel or sonorant that can trigger passive voicing, whereas in sandhi voicing there is.

As this approach deals with unexpected voicing patterns, it might be useful to try to apply it to the data from the present study. It could be possible that the entire continuum is represented by a phonological voicing contrast with

[19] Cyran (2011) and Cyran (2013) use elements instead of features. This does not have any effect on the analysis, however.

a different phonetic realisation of the contrast throughout the continuum: in the west the contrast is phonetically realised as a voicing contrast, while in the east the contrast is phonetically realised as an aspiration contrast. The middle area could then be considered to still be in the middle of that phonetic transition (from a phonetic voicing to an aspiration contrast or vice versa) diachronically speaking. As Cyran argues, evidence for such a contrast cannot be found in the phonetics alone; phonological evidence is necessary to prove the existence of such a contrast. In the present study such evidence must be sought in the assimilation patterns, as the behaviour of word-initial plosives cannot be considered evidence for a phonological representation that differs from the surface phonetic patterns (this is, of course, precisely what Cyran argues for). If lenis plosives are marked with the feature [voice], this feature should be able to spread regressively. A preceding plosive in a cluster will always surface with the feature [voice] in that case: either the plosive was already marked with that feature, or [voice] spreads to a laryngeally empty plosive. However, as lenis plosives are realised with a short-lag VOT (even though they are marked with the feature [voice]), these clusters are expected to surface with partial voicing. As Cyran only discusses the example of a phonetic voicing language represented as a phonological aspiration language, but not vice versa, it is unclear whether these clusters are predicted to be able to undergo passive voicing or not. On the one hand, this might be possible because there is no other phonological feature blocking the application of passive voicing.[20] On the other hand, however, the lenis plosives might be expected to be consistently realised with a short-lag VOT as in this approach the relation between [voice] and short-lag VOT might be regulated in the grammar. In the former scenario, the presence of fully voiced clusters with a lenis C2 is not unexpected, but in the latter scenario it cannot be explained. The presence of fully voiced clusters with a lenis C2 in the middle and eastern areas can thus only be explained if the relation between short-lag VOT and [voice] is not present in the grammar.[21]

At any point in the continuum, clusters with an unmarked fortis C2 can only surface with partial voicing. If C1 is lenis (and marked with the feature [voice]), this feature cannot spread progressively: all dialects have an active rule of final devoicing, so the feature [voice] is deleted from C1 so that both plosives in the cluster are unmarked. If C1 is fortis, both plosives are underlyingly unmarked. The unmarked plosives cannot possibly surface with full voicing as they cannot be subject to passive voicing: this process is blocked from applying by the presence of the feature [voice] in the grammar. The abundance of fully voiced clusters with a fortis C2 in the middle area is thus unexplainable. Even if the middle area is still in a diachronic transition between the two systems (a phonetic voicing realisation of the phonological contrast, or a phonetic aspiration realisation of the phonological voicing contrast), the presence of these

[20] Note that the unmarked series is blocked from undergoing passive voicing by the presence of the feature [voice] in the system.

[21] As the western area has not only a phonological but also a phonetic voicing contrast, the presence of fully voiced clusters with a lenis C2 is entirely expected and also predicted.

clusters cannot be explained, as clusters with a fortis C2 cannot surface with full voicing in either system.

The acquisitional data also appear to pose a problem for the analysis. While the west and east show a bimodal distribution of VOT values, the middle area does not, so a child in that region will not find evidence for postulating a two-way phonological contrast in the first place. Further, only the west shows a phonetic and phonological pattern compatible with a [voice - ∅] contrast (a phonetic voicing contrast as well as assimilation to the lenis member of the cluster). The east does not show any patterns compatible with this phonological contrast at all: neither a phonetic voicing distinction nor a phonological process that requires the presence of [voice] can be found. While the child will find evidence for a two-way laryngeal contrast, both phonetic and phonological evidence would lead the child to postulating an aspiration contrast.

Apart from the fact that the input a child receives is not compatible with the proposed analysis, Cyran does not address the question how a child acquires a phonological contrast that does not match its phonetic realisation. Although it does not appear to be impossible, no formalisation of such a process is given. Whether or not it is possible is a question that, for the moment, remains unanswered.

A second, unanswered question is how the change in VOT values come about. Assuming the entire area is represented by a phonological voicing contrast, it is most logical to assume that the entire area used to have a phonetic voicing contrast as well, which subsequently changed to an aspiration contrast for the eastern area. The most likely scenario is one of language contact: under the influence of the phonetics of a different language, speakers from the eastern area started changing the phonetic realisation of the voicing contrast to an aspiration contrast. However, changing the underlying phonological representation is much more difficult. Speakers will try to implement a phonetic aspiration contrast on an underlying voicing contrast, which will cause mismatches between phonetic values and phonological representations, and therefore confusion will arise. Especially in the case of trying to replace a voicing contrast by an aspiration contrast, speakers will encounter difficulties when it comes to the distribution of phonetic aspiration, as aspiration is not always clearly visible in consonant clusters. When an aspirated plosive is followed by a sonorant, aspiration is only visible as devoicing of a part of the sonorant, and when an 'aspirated' plosive is preceded by a fricative, the two consonants share the spread position of the glottis so that the adduction of the vocal cords already commences in the fricative rather than after the release of the plosive closure. As a result, voicing commences shortly after the plosive release (making it sound like a lenis plosive with short-lag VOT), instead of showing a lag (long-lag VOT). It is surprising that the eastern area fully succeeded in implementing this change. However, a second, stronger argument against this hypothesis is the fact that most languages surrounding German and the dialects spoken in Germany are voicing languages: Dutch is a voicing language, the Romance languages are voicing languages, and the Slavic languages are voicing languages. There thus does not appear to be any trigger present for such a change; it should be language-internal instead of contact-driven. However, it is unclear how and why such a change would happen: the language-learning child finds phonetic and phonological evidence for a voicing contrast only.

The hypothesis that laryngeal distinctions in the entire continuum would be represented by a [voice - ∅] contrast has several implications. First, phonological features must be substance-free, i.e. they cannot have a link with phonetic realisations. Apart from finding a bimodal distribution of VOT values in the linguistic input, the language-learning child has to find a reason to assume a phonological voicing contrast for a phonetic aspiration contrast. As this hypothesis assumes the absence of a link between phonetics and phonology, the reason to posit a phonological voicing contrast for a phonetic aspiration contrast cannot be phonetic but must be phonological. The language in question should thus show some pattern that gives reason to assume the presence of a feature [voice] and/or the absence of a feature [spread glottis]. An example can be assimilation to the lenis member of the cluster (even if the language shows an aspiration contrast).[22] If speakers can posit a phonological contrast that does not 'match' the phonetic contrast, the link between phonetics and phonology must be absent. This would then explain why such an apparent mismatch between feature and phonetics can be stable, i.e. why speakers do not change either the phonological representation or the phonetic realisation of the contrast, so that the link between phonetics and phonology is restored.

4.3.11 An aspiration contrast for the entire continuum

The proposal by Cyran (2011) can also be implemented in the other direction: the entire continuum used to be represented by a phonological and phonetic aspiration contrast, but the western area changed the phonetic realisation of the aspiration contrast to a voicing contrast. This implies that speakers in the west have phonological evidence for postulating an aspiration contrast.

In this scenario the trigger for the change of the contrast might be somewhat clearer than in the previous analysis (outlined in Section 4.3.10), as the Dutch varieties have been in contact with different Romance varieties. First, there is language contact between Dutch and French varieties at the Flemish-Walloon border in Belgium. The change between voicing and aspiration languages may have started out in that area, gradually spreading to the north. Second, the Dutch language area has been occupied by both Spain (before and during the Eighty Year's War, which ended in 1648) and France (1795 to 1813, Milis et

[22] One could argue that another phonological process in which the feature [voice] is active, Final Devoicing, could be a relevant example here. However, if a phonological voicing contrast is realised phonetically as an aspiration contrast, word-final lenis plosives, although marked with the feature [voice], are already realised voiceless. If such a plosive is no longer in word-final position because of a following suffix, it can still be realised voiceless (lenis intervocalic plosives in aspiration systems can either be realised with full voicing or with partial voicing). It is thus impossible to decide if a system with a phonetic aspiration contrast has in fact a phonological voicing contrast based on the behaviour of word-final plosives.

al. (2014)), the languages of which both employ a voicing contrast. This latter period, however, may be too short to be of much linguistic influence: in only twenty years time an entire generation would have to acquire a voicing contrast that is only used by the governing class, instead of the aspiration contrast that is used by their families.

Linguistically the analysis appears to be problematic as well. First, the west has a phonetically consistent voicing contrast. In Cyran's approach, lenis plosives can only be passively voiced. Passive voicing must be a phonetic process instead of a phonological process, so it is not expected to apply consistently. Even though the west shows several lenis plosives without prevoicing, the number of prevoiced lenis plosives is too large to be explained by phonetic voicing. The same can be argued for the number of fully voiced plosive clusters with a lenis C2. The full voicing must be the result of passive voicing as well, which makes it a phonetic process. A phonetic process is not expected to apply consistently, but the plosive clusters in the west are all (except for one) fully voiced.[23] The assimilation patterns in the middle area do not fit the expected patterns either, as clusters with a fortis second plosive should not be able to surface with full voicing. In a system that contrasts unmarked plosives to [spread glottis] plosives, only clusters with two unmarked plosives could potentially surface with full voicing. The presence of the feature [spread glottis] on the second plosive is expected to block phonetic voicing or voicing assimilation. The presence of these fully voiced clusters therefore cannot be explained.

Language acquisition poses, again, a problem for this analysis, as children would need a bimodal distribution of VOT values in their linguistic input. Children in the middle area receive a unimodal distribution of VOT values in the input, and cannot be expected to find evidence for a two-way phonological contrast.

The discussions in this section and the previous section show not only that the data from the present study are incompatible with Cyran's approach; in fact they contradict the possibility of such an approach at all. An important assumption in Cyran's approach is that phonological building blocks (be they features, elements or something else yet) are not strictly linked to a phonetic realisation. If this is indeed true for speakers of two Polish dialects, it should be possible for all speakers of any dialect or language. Speakers therefore do not need to encounter a phonetic voicing contrast to postulate a phonological voicing contrast, or a phonetic aspiration contrast to postulate a phonological aspiration contrast: if they find evidence for a phonological voicing contrast in the linguistic input, even though they encounter a phonetic aspiration contrast, they should still postulate a voicing contrast (or vice versa for a phonological aspiration but phonetic voicing contrast). The same should be true for speak-

[23] Note that the presence of fully voiced clusters with a fortis C1, which is marked with the feature [spread glottis], and a lenis C2 can be explained in terms of delinking of the feature [spread glottis] in final positions. This would lead to a cluster with two unmarked plosives which can subsequently be passively voiced. However, for these clusters passive voicing has to be a phonetic process as well, so it is not expected to apply consistently.

ers that do not encounter a true phonetic voicing or a true phonetic aspiration contrast: if the phonology gives evidence for a laryngeal contrast, the speaker should be able to postulate this laryngeal contrast regardless of its phonetic realisation. Based on these assumptions, the patterns found in the transition zone are quite unexpected. First, speakers in the middle area will always encounter a phonetic laryngeal contrast: they will hear speakers in the west producing a voicing contrast and speakers in the east producing an aspiration contrast. The consistent presence of a phonetic contrast in the linguistic input of speakers in the middle area should be enough for them to postulate a phonological laryngeal contrast. The behaviour of plosive clusters in the west, where assimilation to lenis plosives marked for [voice] occurs, can be considered as more evidence of the presence of a phonological contrast. The speaker is, however, very, if not completely free in the phonetic implementation of the contrast. Based on the phonetic input the speaker receives, they may choose for either a phonetic voicing or aspiration contrast, but considering the implementation is very free they may also choose to use, for example, the phonetic values encountered most frequently, or they may use the average phonetic realisation of lenis and fortis plosives. Speakers are thus expected to show a phonetic distinction between lenis and fortis plosives. This is not borne out by the data: speakers in the middle area show a lot of overlap in their phonetic realisation of the plosives. The two series are clearly not distinct from each other. A second point that contradicts the hypothesis is the behaviour of plosive clusters with a fortis second plosive. Note that, whatever the underlying phonology may be, and whatever implementation the speaker may choose, it should be impossible for clusters with a fortis second plosive to surface with full voicing: in a phonological voicing system (whether it is realised with a phonetic voicing or a phonetic aspiration contrast), fortis plosives are unmarked and cannot spread their element or feature to the preceding plosive (note that the presence of a more lenis plosive in the system is expected to block phonetic voicing of the entire cluster, even if both plosives are unmarked); in a phonological aspiration system fortis plosives are marked and are expected to have a fixed phonetic realisation (as this realisation is plain voiceless if the phonological aspiration contrast is phonetically realised as a voicing contrast, and voiceless aspirated if the phonological aspiration contrast is phonetically realised as an aspiration contrast, the presence of the element |H| or feature [spread glottis] must block full voicing of the cluster).

It must be mentioned here that, although Cyran's approach assumes that the link between phonetics and phonology is very weak, it does seem to assume that speakers are free to acquire any phonetic contrast regardless of the phonological contrast only when they are still acquiring language. After the acquisition period, the link between the phonetics and phonology appears to be stable in the approach. In other words, speakers cannot switch between different phonetic realisations of the same phonological contrast.

In conclusion, the fact that the data from the present study cannot be explained using Cyran's approach seems to indicate that the entire approach cannot be true at all. If it were true, speakers in the Germanic transition zone are expected to show a pattern similar to the pattern found in the Polish dialects. The fact that they show a very different pattern implies that they are unable to lose or weaken the link between phonetics and phonology.

4.3.12 [voice - spread glottis] for the entire continuum

The predictions this hypothesis makes are similar to the predictions of the hypothesis [voice - spread glottis] for the transition zone. As both plosive series have a laryngeal feature in their representation, the two series are expected to be phonetically distinguishable. If the phonological features have a strict link with a phonetic realisation, all lenis plosives (in the entire continuum) are expected to be prevoiced, and all fortis plosives (in the entire continuum) are expected to be aspirated. The predictions for assimilation patterns for the entire continuum are the same as the predictions for assimilation patterns as described in Section 4.3.3.

The data in the entire continuum do not seem to be very well compatible with a [voice - spread glottis] analysis. In the west and east the two series are phonetically clearly separated from each other, so a phonological distinction seems likely. In the west, the data do not present any problems for the presence of the feature [voice]; in the east, the data do not present any problems for the presence of the feature [spread glottis]. All other data, however, seem to be a bit problematic. The fortis plosives in the west can be aspirated, but short-lag VOT is a more frequent realisation than aspiration. This is difficult to explain if they are specified by the feature [spread glottis]. The same can be argued for the presence of a feature [voice] on lenis plosives in the east: they can be realised with prevoicing, but short-lag VOT is a more frequent realisation. In the transition zone, both plosive series are realised with either prevoicing or short-lag VOT (lenis plosives) or short-lag VOT or aspiration (fortis plosives). If the link between a feature and a phonetic realisation can be weakened that much (consistent prevoicing or aspiration is not a requirement), it must be explained why this link is weakened for both plosive series in the transition zone, while in the west it is only weakened for fortis plosives, and in the east it is only weakened for lenis plosives. Why is the link between phonological feature and phonetic realisation in the west and east not weakened for both series, but only for one of them? The exact opposite question must be asked as well: why is the link between phonological feature and phonetic realisation weakened for one (west and east) or even both (middle) series?

The predictions for assimilation in this analysis are not entirely clear. The feature [voice] can only spread regressively while the feature [spread glottis] can only spread progressively, so if [voice] is present in the representation of C1 in the cluster and [spread glottis] in the representation of C2 in the cluster, neither feature can spread as the spreading would be in the wrong direction. If both plosives in the cluster have the same feature, assimilation cannot take place. However, if the first plosive in the cluster is represented by the feature [spread glottis] and the second plosive is represented by the feature [voice], both features can spread to the other plosive. It cannot be predicted which feature will 'win the competition', so it is unclear if the data in the transition zone would match the predictions. However, the presence of fully voiced clusters with a fortis C2 cannot be explained: if both plosives are fortis both are marked for [spread glottis], so full voicing should be impossible. If only C2 is fortis while C1 is lenis, the feature [voice] does not appear in a position from which it can spread, so assimilation should still not be possible.

The acquisition data, finally, argue against this hypothesis as well. The language-learning child does not find evidence for a two-way phonological contrast as VOT values follow a unimodal instead of bimodal distribution.

4.3.13 Concluding remarks

In this section, several potential phonological analyses for the transition zone have been discussed. All data, not only the VOT values but also the assimilation patterns and the input for the language-learning child, point in the direction of an area without a phonological distinction: the VOT values cannot be linked to a phonological feature without weakening or even losing the link between phonetics and phonology; the assimilation data cannot be explained if a phonological feature is present as that feature would be expected to either consistently trigger or consistently block assimilation; the distribution of VOT values in the input does not present the language-learning child with any evidence for a two-way or three-way laryngeal contrast.

In the middle of the Dutch-German dialect continuum, the language-learning child finds characteristics of a voicing contrast in one part of the linguistic input, and characteristics of an aspiration contrast in the other part of the linguistic input. However, the child receives all input simultaneously. Because he or she is not yet able to distinguish between the different realisations phonologically, the child cannot distinguish the input compatible with a voicing system from the input compatible with an aspiration system. In an ordinary situation, the VOT values will follow a bimodal distribution, either compatible with a voicing system or with an aspiration system. In the transition zone the situation is different, as the two different systems cannot be distinguished from each other in the input: instead of a prevoiced range, a plain voiceless range and an aspirated range, the different ranges are all merged into one large range of VOT values. The child thus finds a unimodal distribution of VOT values, instead of a bi- or trimodal distribution. As discussed in Section 4.2, a phonological contrast can only be postulated if the linguistic input gives evidence for it. The distribution of VOT values in the middle area will give the child evidence of the absence of any contrast, as there are no clear boundaries between different voicing categories. For the child to acquire a voicing contrast, VOT values would have to follow a bimodal distribution with one mode in the negative region and one in the short-lag positive region, and for the child to acquire an aspiration contrast

VOT values would have to follow a bimodal distribution with one mode in the short-lag positive region and one in the long-lag positive region. There is thus no evidence for the child to postulate the presence of a laryngeal feature. Based on the linguistic input, the language-learning child will thus not be able to postulate a phonological contrast.

The patterns found for laryngeal contrasts in the transition zone resemble characteristics of koiné languages. Britain (2004) lists three characteristics of these languages: a) levelling; b) structurally less complex; c) emergence of interdialect forms. Speakers in the transition zone do not postulate a laryngeal contrast, in clear contrast to speakers on either end of the continuum who postulate either a voicing or an aspiration contrast. As the absence of a contrast implies a simpler structure than the presence of a contrast, and because the absence of a laryngeal contrast cannot be found anywhere else in the continuum, the last two criteria are clearly met by the varieties in the transition zone. Upon closer inspection, it appears that the first criterion is met as well: Britain defines levelling as "[favouring] those variants which are in a majority in the dialect mix, unmarked as opposed to marked, and socially neutral as opposed to those strongly stigmatized as belonging a particular social or geographical grouping (whether standard or non-standard)". The phoneme present in the transition zone, the laryngeally unmarked plosive, is the only plosive that is present in all varieties at both ends of the continuum. It might thus be hypothesised that in the transition zone, koineisation takes place at the level of laryngeal characteristics.

Two questions remain to be answered, however. The first question is why speakers use the entire range of VOT values, instead of realising all (or most) plosives with an articulatorily 'easy' short-lag VOT. The second (related) question is why plosives that are lenis in either system are still mostly realised with either prevoicing or short-lag VOT, and why plosives that are fortis in either system are still mostly realised with either short-lag VOT or aspiration. These questions will be addressed in the next chapter.

One might argue that, based on the lexical distribution of the different plosives, children will eventually be able to postulate a phonological contrast. Indeed, Feldman, Griffiths, Goldwater and Morgan (2013) used computer models to show how children can make use of lexical constraints to discover phonological categories when phonetic values are overlapping. Distributional learning is obviously made much more difficult if two categories show phonetic overlap: statistically, they follow a unimodal instead of a bimodal distribution. It has been shown that children start building a lexicon at the same time they start categorising sounds (Saffran et al. (1996)), and Feldman et al. (2013) propose that children use the distribution of sounds in words when learning how to categorise sounds.

In the transition zone, however, this strategy might not be of much help to the child. While focussing on phonetic overlap in the linguistic input (which causes a problem for the language-learning child), Feldman et al. (2013) assumed the presence of uniform linguistic input. One lexical item will thus always be realised with the same phonemes and the same phonetic characteristics (i.e. within the limits of the variation that is allowed for each phoneme). In the transition zone, however, the language-learning child receives mixed linguistic input. One lexical item will thus not necessarily be realised with the same phoneme and the same phonetic characteristics; on the contrary, the phonological status of the phoneme as well as its phonetic characteristics is exactly what varies in the input and even within one lexical item. The lexical item a phoneme appears in is thus not likely to be helpful in deducing the underlying laryngeal contrast. If a child encounters a plain voiceless plosive in the input, it can only 'know' this plosive is phonologically unmarked, but it cannot decide if it is the unmarked series of a voicing system (and thus a fortis plosive) or the unmarked plosive of an aspiration system (and thus a lenis plosive). Further, the child is likely to encounter the same lexical item with a different plosive (prevoiced or aspirated), and hence cannot use lexical information to decide which contrast is relevant.

Another problem is that for some lexical items, the status of the plosive differs across the entire continuum. An example is the word for 'to do', which is *doen* [dun] (i.e. it has a lenis plosive) in the west, but *tun* [t^hun] (i.e. it has a fortis plosive) in the east. When the child encounters a realisation with either prevoicing or aspiration, the phonological representation of the plosive can easily be deduced, but when the child encounters a realisation with a plain voiceless plosive it is impossible to decide whether this plain voiceless realisation is the fortis series from voicing systems, or the lenis series from aspiration systems. The child will need evidence from other lexical items, but this evidence is, as argued above, not consistent either.

A question that needs to be addressed is why there is so much phonetic variation present in the transition zone: if there is no laryngeal contrast, plosives are expected to be realised with a short-lag VOT (Maddieson (1984)) because this requires the least physical effort (Kager et al. (2007)). The abundant phonetic variation in the transition zone raises an important question: why do speakers show this phonetic variation, rather than having a phonetic realisation consistent with the phonological representation of the plosive contrast? Insights from language acquisition might also help answering this question. The question will be addressed in the next chapter.

In conclusion, the proposal that children do not find enough statistical evidence for a two-way contrast, and no linguistic evidence for a three-way contrast (because one lexeme may be realised with different voicing properties by different speakers), is plausible from different points of view. The distribution of VOT values is such that children cannot postulate a laryngeal contrast, and the phonological patterns in clusters do not support the presence of a feature. The phonetic realisation of word-initial plosives could in theory support the presence of a feature, but only if one assumes that features and their phonetic realisations are not linked to each other. However, all other patterns are evidence for the absence of a linguistic feature.

4.4 Italy

In Italy a clear distinction between a geographical area with ISV and a geographical area without ISV is visible. There is no intermediate area where speakers apply the process inconsistently or where the process applies only in some of the environments. As there is no gradual transition zone between the two systems, a phonological analysis of the patterns found in this area is not very interesting. The distribution of voicing and length, however, is interesting. According to the literature, all non-northern varieties should be characterised by both a voicing and a length contrast. The data from the present study show that this is indeed the case: in the central varieties the distribution of singleton voiced, singleton voiceless and geminate voiceless alveolar fricatives is unpredictable. Both a voicing contrast and a length contrast (which, following Hayes (1989) and S. Davis (2011), is here assumed to be a difference between a segment that projects a mora and a segment that does not) are required to capture the distribution of the three fricatives. The literature also proposes that northern varieties have long degeminated the length contrast. Since geminates do not occur in these varieties, only a voicing contrast should be present in the grammar. Interestingly, the data from the present study show that not only voicing but also length is contrastive for the northern dialects. In this section the representation of the three fricatives in the northern Italian varieties will be discussed: is either one of the contrasts (voicing or length) enough to capture the distribution of the three fricatives, or are both phonological contrasts required? This discussion is followed by a discussion of the abrupt instead of gradual transition: can the grammar explain the absence of a gradual change?

4.4.1 A length contrast

In the south, both a length contrast and a voicing contrast must be present in the language. Although a triplet of the contrast cannot be found, there are several minimal pairs showing that both voicing and length are relevant phonological contrasts. Compare, for example, the nouns *la casa* [kasa] and *la cassa* [kasːa] ('the house' resp. 'the cash'): these only differ in the length of the fricative, showing that length must be contrastive. The words *fuso* [fuso] and *fuso* [fuzo] ('spindle' resp. 'melted') are realised with a voiceless resp. voiced fricative. The distribution of the long and short fricatives and the voiced and voiceless fricatives cannot be predicted from the phonological environment. Therefore, the feature [voice] must be present in the language, and the language must make a distinction between a mora-projecting fricative and a non-projecting fricative.

In the north, the distribution of fricatives is slightly different. Both the short, voiced fricative and the long, voiceless fricative can appear in intervocalic stem-internal position, but a short, voiceless fricative does not appear in that position.[24] There thus is a contrast between *la casa* [kaza] and *la cassa* [kasːa] ('the house' resp. 'the cash'). Contrary to the varieties in the south, there is no contrast between voiced and voiceless short fricatives: the words *fuso* and *fuso* ('melted' and 'spindle') are both pronounced as [fuzo]. The phonology therefore only needs to distinguish a short voiced and a long voiceless fricative. One could propose that both characteristics are phonologically specified, with the language having a distinction between mora-projecting and non-projecting fricatives as well as having the feature [voice] in its inventory. However, as voicing is predictable from length (a short fricative is always voiced and a long fricative always voiceless) or vice versa (a voiced fricative is always short and a voiceless fricative always voiceless), having both characteristics specified in the language is redundant. In this section the scenario in which only length is specified will be explored.

If only length is specified, that means that the varieties have a distinction between a mora-projecting and a non-projecting fricative. The grammar must have a constraint ranking to ensure the link between length and voicing. The link between length and voicing, however, cannot be one to one: although a voiced fricative is always short and a long fricative is always voiceless, a voiceless fricative is not always long and a short fricative is not always voiced. The short voiceless fricative surfaces in a number of contexts, e.g. at word margins, in onset clusters and in intervocalic position with a preceding morpheme boundary. The constraint ranking must therefore not only account for the link between long voiceless and voiced short fricatives, but must also account for the presence of short voiceless fricatives.

An OT-grammar provides two different possibilities for this link. In the first scenario, a fricative can only surface as voiceless if it underlyingly projects a mora. This means that any non-projecting fricative must always be assigned the feature [voice], and that only an underlyingly projecting fricative can surface voiceless. The constraint ranking must then not only assure this link between length and voicing, but must also assure the degemination of the long fricative in all contexts where a short voiceless fricative surfaces: as the fricative surfaces as voiceless, it cannot be underlyingly non-projecting because it would then have to receive a voicing feature. In other words, the constraint ranking must assure that the mora-projecting fricative at word margins, in clusters with a following sonorant or voiceless obstruent, and intervocalically with a preceding morpheme boundary, surfaces as a non-projecting fricative. A problem with this analysis is that the language-learning child would have to find evidence, somewhere in the linguistic input, for a mora-projecting fricative in those environments. As these fricatives are realised as short rather than long regardless of the phonological environment, there is no reason for the child to postulate an underlyingly mora-projecting fricative that surfaces as a non-projecting fricative. Only if there were alternations between short and long

[24] The long, voiced fricative does not occur in any of the varieties studied here, so it will not be considered in this discussion.

fricatives in the contexts mentioned above would the child have evidence for postulating a mora-projecting fricative. As there are no such alternations, the child cannot find evidence for a mora-projecting fricative and will postulate a non-projecting fricative. If, as hypothesised above, the link between length and voicing always applies, the constraint ranking would always assign a feature [voice] to these fricatives. As these fricatives surface voiceless, the constraint ranking would also have to assure that the feature [voice] is deleted after it has been assigned. Besides the fact that only a stratal version of OT, but not classical OT, could capture a subsequent application of two processes, it is difficult to explain why a voiced fricative would devoice in several of the contexts mentioned. Word-finally and in clusters with a following voiceless obstruent devoicing can be explained as a process of Final Devoicing respectively voicing assimilation, but the devoicing of the fricative in a cluster with a sonorant or in word-initial position is difficult to explain.

In short, assuming a strict link between length and voicing cannot explain the voicing patterns found in the varieties. If voicing depends on length, the link must be less strict than assumed previously in this section. The second possibility is a less strict link between length and voicing. First of all, one has to assume that only the fricatives that surface as long are underlyingly projecting. All fricatives that surface as short are underlyingly non-projecting. The grammar must then assign a feature [voice] to underlyingly non-projecting fricatives in some, but not all contexts. The feature is never assigned to an underlyingly projecting fricative. This can be captured with a high-ranked constraint against voiced geminates. The constraint ranking must have a constraint assigning a feature [voice] to a non-projecting fricative, but since not all short fricatives surface voiced, there must be constraints regulating the distribution of voiced fricatives. The constraints must prohibit a voicing feature to be assigned to fricatives in clusters with a voiceless plosive and with a sonorant, at word margins, and intervocalically preceded by a morpheme boundary. As stated, the word-final position can be explained as a process of Final Devoicing, but the other positions cannot be easily explained as a deletion process because the contexts are not commonly known as contexts in which voicing is difficult to maintain. Instead of a deletion process, a licensing process regulating the distribution of a voicing feature is more likely to capture this distribution. The licensing constraint needs to license the presence of the feature [voice] on alveolar fricatives in two contexts: in prosodic word-internal intervocalic position, and in clusters with another segment marked for [voice]. The first licensing constraint ensures the fricative surfaces voiced within a prosodic word, and explains the voiced fricative within a stem (e.g. *la casa*), [z] after an opaque morpheme boundary and [z] before a transparent morpheme boundary. For the latter context one still has to assume that the /s/-final prefix is incorporated into the prosodic word as a result of a high-ranked constraint requiring every syllable to have an onset, and a higher-ranked constraint against epenthesis. When the prefix is incorporated into the prosodic word, the fricative can be assigned the feature [voice]. The second licensing constraint licenses a feature [voice] in clusters with another segment marked for [voice]. This explains why fricatives in clusters with a following voiced plosive surface as voiced, and why fricatives in clusters with a following sonorant or voiceless obstruent surface voiceless. A voiceless obstruent is obviously not marked for [voice], so it cannot license the presence of that same feature on the fricative, and in a monovalent feature framework sonorants are inherently voiced and thus are not marked for the feature [voice].

4.4.2 A voicing contrast

If only voicing is specified in the north while length is not, there must be a phonological contrast between a short voiced and a short voiceless fricative. So, *la casa* [kaza] ('the house') must have a voiced fricative in the underlying representation (/kaza/), while *la cassa* [kasːa] ('the cash') must have a voiceless fricative in the underlying representation (/kasa/). Length is not specified, but must be derived from voicing. In this section it will be assumed that length is a difference in mora projection, i.e. a geminate consonant projects a mora, while a singleton consonant does not. Two options exist in that case: either both the voiced and voiceless fricative project a mora, or neither does.

In the first scenario, where both the voiced and voiceless fricative project a mora, they must do so in all contexts. In other words, they project a mora intervocalically, word-intially and word-finally, in onset clusters, etc. The voiced fricative must then be degeminated, i.e. lose the projected mora, in any context, while the voiceless fricative must be degeminated in any context but intervocalically. This can be modelled by a high-ranked constraint prohibiting voiced mora-projecting obstruents (or mora-projecting fricatives) in general and a constraint prohibiting voiceless mora-projecting alveolar fricatives in non-intervocalic context.

The degemination of the voiced fricative can be expected considering the physiological difficulties of realising a voiced geminate. Combined with the physiological difficulties of realising a voiced fricative, a high-ranked constraint against a mora-projecting voiced fricative is not unexpected. The degemination of the voiceless fricative, except in intervocalic context, is not unexpected either: in obstruent clusters a geminate followed by another obstruent would imply a trimoraic onset, i.e. an onset that is too large. At word margins geminates are also infrequent. At a transparent morpheme boundary degemination could be explained considering the fact that the fricative occurs in prosodic word-initial position. However, although a constraint ranking could probably ensure an output like this, the underlying phonological system would be very rare cross-linguistically. A first question is why a language would only have geminate fricatives but not their singleton counterparts. Especially if one considers the fact that a geminate consonant is a consonant that does project a mora, one would expect a language that has a projecting consonant would also have the non-projecting counterpart of that consonant in its inventory.

A second question raised by this scenario, is why a language would underlyingly only have a mora-projecting voiced fricative, which surfaces as a degeminated fricative in every context in the output. The language-learning child would find zero evidence in the linguistic input for postulating a mora-projecting voiced fricative. The same holds for the voiceless fricative. The language-learning child only finds evidence for a mora-projecting fricative in the stem-internal intervocalic context. In all other contexts, the child can only find evidence for a non-projecting fricative. For both the voiced and voiceless fricatives it is thus extremely unlikely that the child postulates a mora-projecting, i.e. geminate fricative. In conclusion, if only voice is specified it seems rather unlikely that both the voiced and voiceless fricative are underlying geminates.

If all fricatives are underlyingly short, the occurrence of long voiceless fricatives in stem-internal intervocalic position must be accounted for by the constraint ranking. This can be easily done, as previous literature has shown that it is relatively simple to ensure the very same fricative surfaces as a singleton voiced fricative in that context. One could relatively easily change the constraint ranking in such a way that /s/ in intervocalic stem-internal position always projects a mora. However, it is unclear how a mora-projecting fricative in intervocalic position would be a more optimal output than a non-projecting fricative. It must be considered a case of fortition.

Fortition, in Optimality Theory, must be the result of an interplay between faithfulness and markedness constraints (as all processes must be). In this scenario we have assumed that fricatives are underlyingly short, so the faithfulness constraint must ensure a fricative surfaces as non-projecting in the output candidate. If there were only one faithfulness and one markedness constraint regulating the relationship between length in the input and output form, neither constraint can be positional. In that case, ranking the faithfulness constraint over the markedness constraint would ensure that any fricative in the output has to be faithful to the length specifications (mora projection specifications) of the fricative in the input form. Fricatives in the northern Italian varieties would then always have to surface short. If markedness were ranked over faithfulness, the general markedness constraint against non-projecting fricatives would require all fricatives in output forms to surface long. The northern Italian varieties would in that case only have geminate fricatives but no singleton fricatives. Both options are clearly not borne out by the data. It therefore has to be argued that, if fricatives are only marked for voicing and are all underlyingly non-projecting, either the markedness or the faithfulness constraint has both a general and a positional variant (cf. Smith (2008)).

If the faithfulness constraint has both a positional and a general variant, there must be a high-ranked positional faithfulness constraint requiring any prosodic word-internal non-intervocalic fricative in the output to be faithful to its mora-projecting capabilities in the input. This constraint must be ranked above the general markedness constraint requiring any fricative to be a mora-projecting fricative, while the general faithfulness constraint (which would also apply to prosodic word-internal intervocalic fricatives) must be ranked below the markedness constraint. A constraint against voiced geminates in general must be ranked above the three constraints to ensure voiced fricatives never surface as geminates, not even in prosodic word-internal intervocalic position.

If the markedness constraint has both a positional and a general variant, there must be a high-ranked positional markedness constraint requiring any prosodic word-internal intervocalic fricative to project a mora in the output. This constraint must be ranked above the general faithfulness constraint requiring any fricative in the output to have the same mora-projecting capabilities as the fricative in the input. Below the general faithfulness constraint is the general markedness constraint requiring any fricative in the output to have the same mora-projecting capabilities as the fricative in the input. This constraint ranking ensures that only the prosodic word-internal intervocalic fricatives surface as geminates, while any other fricative surfaces true to its input characteristics.

In this section two different options for a length contrast being derived from a voicing contrast have been proposed: an approach invoking positional faithfulness and an approach invoking positional markedness. It is impossible to favour one approach over the other based on their ability of explaining the observed data patterns. However, there is another disadvantage to the positional faithfulness approach. The highest ranked constraint is a constraint against voicing of non-prosodic word-internal non-intervocalic fricatives. A general markedness constraint against non-projecting fricatives is ranked below that constraint, followed by a general faithfulness constraint. This way, only prosodic word-internal intervocalic fricatives are not protected from lengthening by the constraint ranking. However, the positional markedness constraint is a very strange constraint, as the contexts that are affected by it do not share any characteristics with each other, except for the negative characteristic that they do not appear in prosodic word-internal intervocalic position. It seems unlikely, however, that a constraint can refer to a negative characteristic. If the constraint were to refer to positive characteristics, it would have to be split into a number of different positional constraints as the different contexts (or positions) do not share characteristics with each other. It would be difficult to explain why all these contexts have to be excluded from gemination even though they are so dissimilar. In conclusion, a positional markedness constraint requiring prosodic word-internal intervocalic fricatives to be long seems a more likely scenario.

4.4.3 An underlying length or voicing contrast?

In this section two different approaches to the representation of the length and voicing contrast in northern Italian varieties have been discussed. In the first approach voicing values of the alveolar fricative are derived from the underlying length contrast; in the second approach the length contrast is derived from the underlying voicing contrast. Both approaches can easily explain the distribution of voiced and voiceless, short and long fricatives in the varieties. It therefore cannot be concluded which underlying representation a speaker will choose: a speaker will always have to acquire the distribution of voiced and voiceless fricatives and of short and long fricatives, but as long as voicing can be derived from length and length can be derived from voicing, it is impossible for the language learner to choose one of the two approaches based on the data. Only cross-linguistic comparison might offer some insights in the likeliness of either approach.

An advantage of the underlying voicing approach is that the child already has to acquire a voicing contrast for plosives and labiodental fricatives. This would imply a similarity between the alveolar fricatives and the other obstruents. A length contrast, on the other hand, is argued to be absent from the northern varieties: not only alveolar fricatives but all other voiceless obstruents as well are argued to have degeminated. However, since the voiceless alveolar fricative clearly shows a length contrast, it cannot be assumed that the length contrast for the other voiceless obstruents is still absent.[25] A disadvantage of this approach is that cross-linguistically, lengthening of a voiceless intervocalic fricative seems a rare process. Blevins (2004) lists seven different pathways for the evolution of geminates, none of which appears to be able to explain the gemination of singleton fricatives in intervocalic position.[26]

If one assumes that the voicing contrast is derived from an underlying length contrast, an obvious disadvantage is that the alveolar fricatives only actively employ a length contrast while all other obstruents employ a voicing contrast. The distribution of voicing of the short alveolar fricatives, however, can be explained as a phonological licensing process. Cross-linguistically this is much more common than the lengthening of intervocalic fricatives. Based on this comparison, it seems more likely that the northern Italian varieties have an underlying length contrast for alveolar fricatives with voicing being derived from this length contrast.

4.4.4 Phonological process or allophonic distribution?

If we leave the question of the relevant phonological contrast in the northern varieties out of the discussion for a moment, the distribution of voiced and voiceless singleton fricatives poses a second theoretical question that needs answering. In all theoretical approaches that discuss ISV, the child is assumed to acquire a phonological neutralisation rule or a constraint ranking that neutralises a voicing contrast in the input form. If one assumes the child acquires a rule or process that changes an underlying form into a different surface form, or an input form into a different output form, one also has to assume that the child

[25] The present study was not designed to investigate the length contrast in the north. Instead, it was assumed that length was indeed irrelevant in all varieties studied; it was only included in the questionnaire because the central varieties were expected to show a length contrast. The discovery of a length contrast in northern varieties was unexpected. To find out whether or not the north has no length contrast for the other voiceless obstruents another study would have to be designed.

[26] Since the length contrast is assumed to be derived from the voicing contrast, one has to assume that the geminates are underlyingly short. As the short fricatives occur in intervocalic position it is impossible to argue that the geminates are the result of place assimilation in obstruent clusters, although historically this is how the geminates entered the Italian varieties.

acquires a phonological contrast in underlying or input form: children acquiring a northern Italian variety have to acquire underlying or input forms such as /kasa/ and /resistɛnza/ which, because of a rule or constraint ranking, surface as [kaza] and [rezistɛnza]. Although a difference between the underlying or input form and the surface or output form is very common across languages, there appears to be a difference between the Italian situation and the other examples. In most examples of a phonological process there are alternations present in the language: in some morphophonological or syntactic environments the process does apply, while in other environments it does not apply. The underlying segment will thus be visible in the environments where the process does not apply. The Italian case of ISV, however, does not apply in such an environment that alternations between the underlying or input segment and the changed surface or output segment can be found: it only affects prosodic word-internal fricatives. However, a fricative in a lexical item either always appears within the prosodic word, or it never does. There are no examples of fricatives that sometimes appear within the prosodic word and sometimes at a prosodic word boundary: a stem-internal fricative always appears in stem-internal position, as does a fricative preceded by an opaque morpheme boundary.[27] The only exception to this pattern is the prefix-final /s/ in the prefixes *bis* and *dis*: this fricative only surfaces as voiced when it is followed by a vowel-initial stem. When it is followed by a stem with an initial voiceless obstruent, the fricative surfaces as voiceless.

Is the alternation between [s] and [z] in /s/-final prefixes enough for a child to acquire a process of prosodic word-internal intervocalic /s/-voicing? Based on the linguistic input, a child could certainly acquire a process that voices the final fricative of /s/-final prefixes in prosodic word-internal intervocalic position. However, as there is a transparent morpheme boundary present, the child has no evidence whatsoever that voicing has to apply in any prosodic word-internal intervocalic context. Instead, the child will only find evidence for voicing in intervocalic position after a morpheme boundary. Further, the final fricative of this prefix does not only surface as voiced in intervocalic position. It is also realised as voiced when it is followed by a sonorant or a voiced obstruent: in the word *bisnonno* ('grandfather'), the fricative is realised as voiced as well. Only in words like *biscotto* ('biscuit') is the fricative realised as voiceless. Instead of showing evidence for a process of intervocalic voicing, the prefix shows evidence for a process that licenses the presence of the feature [voice]. This is similar to the discussion in Section 4.4.1, where the possibility of an underlying length contrast from which voicing is derived was discussed.

Instead of acquiring a phonological process, the language-learning child might acquire a distribution of voiced and voiceless alveolar fricatives. In the north, [s] and [z] appear in a fully predictable, allophonic distribution. Allophonic distributions are very common cross-linguistically, so children must be

[27] Note that the stem of words with an opaque morpheme boundary never occurs independent of the prefix, so that the fricative indeed always appears within the prosodic word.

able to acquire such a distribution (e.g. White, Peperkamp, Kirk and Morgan (2008)). The situation in the border area between Emilia-Romagna and Tuscany is, however, slightly different, as children do not only receive input from a system with an allophonic distribution, but also from the Tuscan system with an unpredictable, phonological distribution. This means that they will not find a predictable distribution of the two fricatives in the input: some speakers will show this predictable distribution in the input, while other speakers will not. The question is thus whether a child can acquire an allophonic distribution if there is noise present in the linguistic input.

Peperkamp, Le Calvez, Nadal and Dupoux (2006) studied the acquisition of allophonic rules, also including corpora in which noise in the distribution was present. They showed that the child could still successfully acquire a distribution if noise is present. For the current discussion it can thus be assumed that a child acquiring a northern Italian variety, in fact acquires an allophonic distribution of the two sounds rather than a phonological process, the workings of which are completely obscured.

In short, it is very unlikely that a language-learning child will acquire a phonological process that voices prosodic word-internal intervocalic fricatives, because the language does not provide any evidence for it. Instead, the child will probably acquire an allophonic distribution of the fricatives: voiced fricatives can appear only in positions where they are licensed, while the voiceless fricative appears in any other position.

4.4.5 The transition in Optimality Theory

In Chapter 3 it has been suggested that the transition between northern and central Italian varieties might be phonologically gradual, i.e. a gradually increasing number of phonological environments is affected by ISV. Such a pattern is not borne out by the data. In this section it will be shown why the phonologically gradual increase cannot in fact take place. Krämer (2005) proposes the following constraint rankings for Tuscan Italian and Lombardian Italian (where Lombardian Italian is applicable to northern Italian in general):

Tuscan: F-Contiguity, Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)pwd, AlignL ⟩⟩ *[+voice] ⟩⟩ Ident(voice)

Lombardian/northern: Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)pwd, AlignL ⟩⟩ *[+voice] ⟩⟩ F-Contiguity, Ident(voice)

The constraint Dep-IO is a standard anti-insertion constraint; Onset is the standard constraint requiring every syllable to have an onset; AlignL requires that "the left edge of every stem [is aligned] with the left edge of a prosodic word" (Krämer (2005)); *[+voice] is violated by [+voice] segments[28]; and Ident(voice) is violated by segments that change their underlying voic-

[28]This constraint can easily be translated to a privative feature approach by transforming it to *[voice].

ing values.[29] The constraint *(VC̥V)PWD is violated by a voiceless intervocalic consonant within the prosodic word; the constraint F-Contiguity is violated by a change in features of an internal segment (i.e., a segment that is both non-initial and non-final). The difference between privative and bivalent features is irrelevant to the constraint: from [+voice] to [−voice] is just as much a violation of the constraint as from [voice] to [∅]. When F-Contiguity is ranked above *(VC̥V)PWD (as in Tuscan), changing the voicing value of a stem-internal fricative (which is of course also prosodic-word-internal) violates a higher-ranked constraint than a stem-internal voiceless fricative. Stem-internal voiceless fricatives will thus not surface as voiced in these dialects; rather, the fricative will surface as voiceless. A stem-internal voiced fricative does of course not violate F-Contiguity if it is underlyingly voiced.

In dialects where F-Contiguity is ranked below *(VC̥V)PWD, a stem-internal voiceless fricative incurs a worse constraint violation than changing the voicing values of that fricative. In those dialects ISV will apply. Note that ISV is in fact not a rule changing intervocalic voiceless fricatives to voiced fricatives, but rather the result of a constraint against voiceless intervocalic fricatives. Voicing them is the least bad option.

Given Krämer's ranking, the intermediate stages will be the following (as the ranking of *(VC̥V)PWD is the same for both Tuscan and Lombardian/northern, and F-Contiguity is ranked lower in Lombardian/northern than in Tuscan, I will here only consider a reranking of the latter constraint):

- **F-Contiguity**, Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ *[+voice] ⟩⟩ Ident(voice)
- Dep-IO ⟩⟩ **F-Contiguity** ⟩⟩ Onset ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ *[+voice] ⟩⟩ Ident(voice)
- Dep-IO ⟩⟩ Onset ⟩⟩ **F-Contiguity** ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ *[+voice] ⟩⟩ Ident(voice)
- Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ **F-Contiguity** ⟩⟩ *[+voice] ⟩⟩ Ident(voice)
- Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ *[+voice] ⟩⟩ **F-Contiguity** ⟩⟩ Ident(voice)
- Dep-IO ⟩⟩ Onset ⟩⟩ *(VC̥V)PWD, AlignL ⟩⟩ *[+voice] ⟩⟩ **F-Contiguity**, Ident(voice)

Four intermediate stages between the northern and the Tuscan ranking exist. In the first intermediate stage F-Contiguity is ranked below Dep-IO. Because F-Contiguity is still ranked above *(VC̥V)PWD, changing a prosodic-word-internal /s/ to /z/ is worse than not changing it. The reranking of F-Contiguity below Dep-IO does not have an effect either. In the northern grammar, the two constraints are not ranked with respect to each other. This means that neither constraint shows evidence for being ranked above the other,

[29] This constraint can apply in a bivalent feature framework when [+α] changes to [−α] or vice versa, and to a privative framework when a privative feature [α] is inserted or deleted.

so that both rankings (Dep-IO ⟩⟩ F-Contiguity or F-Contiguity ⟩⟩ Dep-IO) would give the same output. Reranking F-Contiguity below Dep-IO therefore does not have any influence on the output the grammar produces.[30]

In a word like 'dis-onesto' (dishonest), the prefix-final /s/ is resyllabified in the onset of the stem. This is represented in the grammar by ranking the constraint Dep-IO over the constraint Onset: insertion of a segment incurs a worse violation than not having an onset at all. However, as a consonant is available in the preceding syllable, it can be resyllabified to the empty onset. The voicing of this fricative is argued to be spontaneous, under the influence of the voicing of the following vowel.

If F-Contiguity is ranked below Dep-IO and Onset, the output produced by the grammar is still the same: ISV cannot apply, nor does it apply in any context where it is not supposed to apply. ISV can only apply when F-Contiguity is ranked below *(VC̥V)PWD (and automatically below AlignL, as this constraint is unranked with respect to *(VC̥V)PWD). In that case it automatically applies to every fricative in stem-internal, intervocalic context, but not to any other fricative. Reranking F-Contiguity below *[+voice] and subsequently below Ident(voice) does not have any consequence, as it does not lead to any other output. This means that, with respect to ISV, a phonologically gradual change between the two systems cannot exist on the surface. Speakers might assume an intermediate constraint ranking, but this does not have any effect on the output: the language either applies ISV perfectly, or not at all.

It is unclear why, according to Krämer (2005), the constraint F-Contiguity has to be ranked above Onset in Tuscan Italian. If it were ranked below Onset, but above *(VC̥V)PWD, ISV still would not apply. However, it is clear that a gradual demotion of F-Contiguity (seen from south to north) does not create a grammar with a phonologically intermediate output. Speakers therefore do not have the option to postulate such a system, as the grammar simply does not provide the option.

It has to be admitted that Krämer's 2005 approach to the voicing of prefix-final /s/ in the northern varieties is somewhat unorthodox. However, even if one assumes a more standard approach to the voicing of prefix-final /s/, it is still possible to explain the absence of a phonologically gradual change between the two regions. Following Nespor and Vogel (1986), most authors assume that the /s/-final prefix has to be incorporated into the prosodic word that is constituted by the following stem, because Italian varieties do not allow prosodic words ending in a consonant. The prefix and stem together form one prosodic word, so that ISV can freely apply in this context. An intervocalic /s/ preceded by

---

[30] Note also that Dep-IO argues against segment insertion: this includes feature insertion as well. In non-initial and non-final position, this constraint therefore (partly) has the same effect as F-Contiguity, which blocks a change in feature values (and thus also blocks the insertion of features associated with segment insertion). The scope of Dep-IO is wider than the scope of F-Contiguity, as it also applies to segments in non-internal positions, but it is narrower as it does not block deletion.

an opaque morpheme boundary will be interpreted by the speaker as occurring in stem-internal position as well: if a speaker does not recognise a morpheme boundary as such, he or she must assume the prefix to be part of the stem. The /s/ in those words then also occurs in prosodic word-internal context, so that, again, ISV can apply freely in this context. In conclusion, all contexts affected by ISV are phonologically identical to each other. A phonologically gradual increase in the number of contexts affected by ISV therefore cannot exist.

Even if one assumes that the voicing of prefix-final /s/ in Tuscany is the result of another process (which one would have to assume, since if it were the result of ISV the other contexts would have to be affected as well), the transition still cannot be gradual. Prefix-final /s/ will always surface voiced, whether this is the result of prosodic word-internal ISV or of another process. A transition between varieties where this fricative is voiced and varieties where it is voiceless therefore does not exist (at least not between the northern and central varieties). The two other contexts still cannot show a phonologically gradual transition because they are the result of the interplay between the same constraints. In conclusion, the nature of the process is such that the phonological transition has to be abrupt.

4.4.6 Other options

In the present study only the length and voicing contrast are relevant. Although several different options have been discussed in the preceding sections, more options exist. An analysis that is missing from the preceding sections, for example, is the analysis that proposes both a length and voicing contrast in the north, but only either a length or voicing contrast in the south. It may be clear from the discussions in the previous sections that a voicing contrast is not required in the north, while it is required in the south. The same is true for a length contrast, which is required not only in the south but also in the north. The remaining hypotheses that exclude either a voicing or length contrast in the south, or a length contrast in the north, will not be discussed here, as these contrasts have proven to be necessary in the representations of these regions.

4.4.7 Concluding remarks

The most likely phonological representation of contrasts in the Italian transition zone postulates a length contrast in the north and south, and a voicing contrast only in the south. The distribution of voiced and voiceless fricatives in the north is guided by a phonological rule or constraint ranking.

Children from both regions indeed find a bimodal distribution of long and short fricatives, enabling them to postulate a length contrast. Children in Tuscany receive a linguistic input with a clear distinction between partially and fully voiced fricatives. They will thus be able to postulate an underlying phonological contrast between voiced and voiceless fricatives. In the north, the data show that children will receive a similar input: partially and fully voiced fricatives are clearly distinguished from each other (which is not unexpected considering the fact that northern varieties contain both voiced and voiceless fricatives; only their distribution is predictable). Why do they, as opposed to children in the south, not postulate a phonological voicing contrast? The most likely explanation is that finding a bimodal distribution is a necessary but not sufficient requirement for postulating a phonological contrast. As the distribution of voiced and voiceless fricatives is predictable in the north, it is not necessary to postulate a phonological contrast, although it would be a possibility. In Tuscany the distribution of voiced and voiceless fricatives is not predictable, so a phonological contrast needs to be postulated to account for the distribution.

A remaining question is why there appears to be no confusion between the two regions: as there is no gradual transition, speakers of a neutralising dialect appear to be able to correctly perceive the speech of a speaker of a non-neutralising dialect, and vice versa. Van der Feest and Johnson (2016) studied a similar situation in the Netherlands, where some dialects have a complete neutralisation of the voicing contrast in fricatives (all fricatives are realised as voiceless), while other dialects maintain the contrast. The authors studied the effect of mixed linguistic input on language acquisition: their study included children who only received input from a neutralising variety of Dutch, as well as children who received input from both a neutralising variety and a non-neutralising variety of Dutch (at least one of the parents spoke the non-neutralising variety, and the child grew up in a region where the neutralising variety is spoken). They found that children who received mixed input could easily adjust to the speaker they hear (i.e., they knew whether the speaker neutralised or maintained the voicing contrast). These children thus regarded the differences between speakers as phonological (some speakers maintain a phonological contrast, others do not), rather than regarding voicing in fricatives as allophonic variation. Children who only received uniform input (i.e., only the neutralising variety) could not readily adjust to the speaker, but merely two minutes of exposure to a speaker who maintained the contrast was enough for them to perform just as well as children who received mixed input. Their study shows that receiving mixed input (containing input from a variety with a contrast and a variety that neutralises that contrast) does not confuse the child. In the Italian transition zone, the linguistic situation is very similar to the situation studied by Van der Feest and Johnson (2016): the Italian children will receive either a uniform linguistic input (at the ends of the continuum), or a mixed linguistic input (in the transitional area). The uniform input contains input only from a neutralising variety (northern dialects) or a non-neutralising variety (Tuscan dialects), while the mixed input contains input from both varieties. Children are thus expected to be able to adjust their speech to the dialect of the speaker, and the mixed input is not expected to lead to any confusion. From an acquisitional point of view, the absence of a gradual transition in the

Italian dialect continuum is thus easily explainable.[31]

Note that the results from the study by Van der Feest and Johnson (2016) cannot be directly transferred to the Dutch-German transition zone, as those patterns do not concern a neutralisation of a contrast but rather a different phonetic realisation of two comparable but not identical phonological contrasts.

4.5 Conclusion

In this chapter the data from both transition zones have been analysed. It has been shown that the confusion in the Dutch-German transition zone must arise during language acquisition, which is supported by the input the language-learning child receives. Based on these data the child does not find any evidence for a phonological feature, and will assume the language only has one plosive series. The absence of variation in the Italian transition zone, on the other hand, is again explained by the input the language-learning child receives and by the phonological representation of the phenomenon, which does not create the possibility of phonological variation (i.e. variation in the application of the process).

[31] Floccia, Delle Luche, Durrant, Butler and Goslin (2012) found that children receiving mixed input (/r/-dropping and /r/-preserving varieties of English) acquire the socially dominant contrast. However, children in their experiment were unable to correctly recognise non-rhotic stimuli, even if (one of) their parents spoke a non-rhotic variety. This would imply that the children were unable to understand (one of) their parents, which seems unlikely. It must be noted that the statistics appear to have been carried out over the entire population of participants, and not, subsequently, for the mixed input children independently. The overall effect might thus be different from what children receiving a mixed input would show.

CHAPTER 5

Modelling variation

5.1 Introduction

In the previous chapters data from two transition zones have been presented. A phonological analysis of the Dutch-German continuum has revealed that the middle area of this continuum is characterised by the absence of a phonological contrast between voiced and voiceless unaspirated, or voiceless unaspirated and voiceless aspirated plosives. Speakers in the middle area do, however, show a large amount of variation in their VOT values. In Chapter 2 it has been mentioned that laryngeally unmarked plosives are commonly realised with a short-lag VOT value. The data from the transition zone thus do not correspond to the predicted patterns, so the question is why speakers in the middle area show as much phonetic variation in their realisations as they do.

The data in the Italian continuum pose a second question. It has been shown that all speakers in Emilia-Romagna consistently apply ISV (or rather have a consistent allophonic distribution of voiced and voiceless fricatives), while older speakers in Tuscany have a phonological opposition between /s/ and /z/. Younger Tuscan speakers, however, follow a pattern that falls somewhere in between the two patterns: they do not have the same phonemic distribution of voiced and voiceless fricatives as the older Tuscan speakers, as they realise a much larger percentage of intervocalic fricatives with full voicing, but they do not follow the same consistent distribution as speakers from Emilia-Romagna either. They can thus be argued to show phonological variation. The main question is how this variation must be modelled in the speakers' grammar.

In this chapter these two question will be addressed, combined with a discussion on the different approaches to variability in phonetics and phonology. Because the data sets are relatively small, they will not solve this discussion, but they might still be able to shed some light on the matter. The chapter will start with a discussion of the phonetics-phonology interface, where variability in phonetics can be modelled. Two different approaches will be discussed. Subsequently several approaches to phonological variability will be discussed.

Throughout this chapter the assumption that features are innate will be maintained. The presence of categorical perception in human speech perception (see Goldstone and Hendrickson (2010) for an overview of the literature on categorical perception), which is present from birth (Eimas et al. (1971)), is linked to this: if infants are able to distinguish different phonological categories at birth this capacity must be innate.

5.2 Implications for the phonetics-phonology interface

The phenomena investigated in the present study are both related to phonetics and phonology, and provide interesting insights in phonetic and phonological variation and the interface between phonetics and phonology. First, the realisation of plosives in the Dutch-German continuum shows that plosives marked with a phonological feature have a restricted phonetic realisation, while unmarked plosives are relatively free in their realisation. The marked plosive series in the west and in the east are both constrained in their realisation ([voice] in the west must be prevoiced, while [spread glottis] in the east must be aspirated), but the unmarked plosives in the entire continuum show large amounts of phonetic variability. If the link between a feature and its phonetic realisation were relatively loose, plosives marked for either [voice] or [spread glottis] would be expected to show phonetic variation such that they would be realised without, respectively, prevoicing or aspiration. The unmarked series shows another interesting pattern: in the regions where there is a phonological contrast, the unmarked plosives can be realised with any VOT value as long as it does not conflict with the marked series. In the region without a phonological contrast the unmarked plosive can be realised with any VOT value. The phonetic realisation of unmarked plosives is thus in principle unrestricted; a restriction can only be imposed by the presence of another laryngeal feature. Further, if the link between a feature and its phonetic realisation were loose, only one laryngeal feature (e.g. [(±)voice]) would be enough to capture the different phonetic realisations of the contrast: [(+)voice] could be realised with prevoicing (initially) and full voicing (intervocalically) or a short-lag VOT (initially) and partial voicing (intervocalically), and [−voice/∅] with a short-lag VOT or aspiration (and partial voicing intervocalically in both cases). In that case, however, the expectation is that in contact areas between a voicing and an aspiration language both lenis and fortis plosives would show some phonetic variability. The fact that in the west of the Dutch-German continuum only the fortis plosives show variability while in the east only the lenis plosives show variability (like many other studies before have shown), shows that such a phonological contrast is not likely. The data show that, instead, LR is a more realistic approach to modelling the contrast.

The data further show that there must be a link between a phonological feature and its phonetic realisation. Phonemes marked for [voice] are consistently realised with prevoicing or as a fully voiced fricative, and phonemes marked for [spread glottis] are consistently realised with aspiration. The phonetics-phonology interface must be able to model this link. Further, it must be able to handle the perception of phonetic categories: in a voicing language, prevoiced plosives or fully voiced fricatives must consistently be perceived as marked for [voice], while in an aspiration language, aspirated plosives must consistently be perceived as marked for [spread glottis] (but see Cyran (2011, 2013) for a different approach).

Assuming there is a link between features and their realisation, and assuming constraints on minimal distance do exist, the interface between phonetics and phonology must be specified. This chapter will discuss two different approaches to the modelling of the phonetics-phonology interface.

5.3 Modelling the phonetics-phonology interface

In this section several different approaches to the phonetics-phonology interface will be discussed. Early approaches to the interface (e.g. *The sound pattern of English* (Chomsky and Halle (1968)), Keating (1984)) assumed the presence of phonetic implementation rules: rules that state how phonemes should be realised phonetically. These rules are learned (i.e. they are not part of Universal Grammar (UG)) and language-specific. Different languages might thus choose for a different phonetic realisation of the same phonological feature. Keating (1984) discusses the realisation of voicing contrasts across different languages. In her approach, all two-way laryngeal contrasts are represented by the [±voice] opposition. The features [+voice] and [−voice] are both associated with a phonetic label: {voiced} or {voiceless} ([+voice]) or {voiceless} or {voiceless aspirated} ([−voice]). The choice for these labels is language-specific. Finally, the realisation of the phonetic labels is guided by phonetic implementation rules. These rules are, as mentioned, language-specific as well, and account for cross-linguistic variation in, for example, VOT values within the phonetic categories (i.e., different languages may have different specific VOT values for the phonetic category voiced, or for the phonetic category voiceless aspirated). However, an approach with language-specific rules for phonetic implementation of a phonological feature is not very powerful. Although some restrictions on the phonetic realisation of a phonological feature could exist (e.g. a plosive marked with the feature [+voice] can never be realised with aspiration, or in Keating's approach a phonetic label (e.g. {voiced}) should have a corresponding phonetic realisation), in theory the approach does not exclude or predict any specific patterns. Variation between languages in the precise realisation of a phonological feature can simply be modelled by using different phonetic implementation rules for the different languages. An example can be found in two different varieties of French: European French and Canadian French. The specific realisation of the voicing contrast differs between these two varieties, the Canadian variety having VOT values that are closer to the VOT values of an aspiration system than the VOT values of the European variety (Caramazza et al. (1973), Caramazza and Yeni-Komshian (1974)). However, this approach is in fact nothing more than a description of the facts: if the realisation of the voicing contrast in the two French varieties would have been different from the actual situation, the solution would simply be to postulate different phonetic implementation rules.

In recent times many different authors have studied the phonetics-phonology interface. As a result, many different approaches to the interface have been proposed, ranging from approaches where phonetics and phonology are strictly separated from one another, to approaches where there is no distinction made between the two. An example of the former approach is Substance Free Phonology (Hale and Reiss (2008), Reiss (2017)), which assumes that phonetics does not have any influence on the phonology of a language. An example of the latter approach is Exemplar Theory. This approach assumes that speakers store the precise phonetics of every utterance they encounter, and create clouds of utterances that are highly similar. Production of such an utterance is based on all utterances that are stored. In this section two different approaches to the interface will be discussed; one generative approache and one usage-based approach.

5.3.1 BiPhon

Bidirectional phonetics and phonology (BiPhon; e.g. Boersma (2009), Hamann (2009), Hamann (2014)) is a model that explicitly specifies the workings of the phonetics-phonology interface, and assumes that both the speaker and the listener make use of the very same grammar. The computational module of the framework is very similar to the computational module of classical OT: an infinite list of output candidates is generated for a specific input. These are compared to each other by a constraint ranking: the optimal output candidate is the candidate that violates the lowest ranked constraint in comparison to the other output candidates. Because the model is bidirectional (modelling both speech production and perception), the constraint ranking must yield the correct forms both for the speaker (who goes from a mental representation to a phonetic output) and the listener (who goes from a phonetic output to a mental representation).

Several differences exist between BiPhon and classical OT. As classical OT only models the phonological processes that change an underlying form into a surface form, only these two forms need to be represented, and only constraints on the output and on the relation between the input and output are required.[1] BiPhon, however, also models the phonology-phonetics interface, so besides the phonological forms (underlying and surface[2]), a phonetic form must also be present, plus a set of constraints to link the phonetic form to the phonological form. Instead of only one phonetic form, two phonetic forms are used in the model: the auditory form,[3] modelling the phonetic form for the listener ("[a] continuous representation of sound; it consists of noises, pitches, spectra, silences, transitions, and durations" (Boersma (2009))), and the articulatory form,[4] modelling the phonetic form for the speaker ("[a] continuous representation of the gestures of the human sound-producing mechanism; it consists of the activities of the relevant muscles of the lungs, tongue, throat, larynx, lips and nose and their coordinations" (Boersma (2009))). Without either one of these forms the model would no longer be bidirectional, as it would no longer represent both the speaker and the listener. Boersma (2009) assumes that only the auditory but not the articulatory form is linked to the surface form, reasoning that this is more economical from the perception point of view as a listener can access the phonological form from the phonetic form "without ever touching the Articulatory Form".[5] Finally, the morpheme "mediates in connecting the phonological underlying form to semantic features in the lexicon" (Boersma (2009)).

The different forms (morpheme, underlying form, surface form, auditory form and articulatory form) are all linked to each other via different types of constraints. The morphemes, first, are linked to the underlying form via lexical constraints. The underlying form is linked to the surface form via faithfulness constraints (similar to classical OT), and the surface forms themselves have to comply with structural constraints. The structural constraints are very similar to the markedness constraints in classical OT, but they have a more restricted area of application: where markedness constraints can refer to both phonological and phonetic characteristics of the surface form, structural constraints can only refer to phonological characteristics. In classical OT, markedness constraints such as *Coda or *ComplexOnset refer to the phonological wellformedness of the surface form, while constraints such as *Retroflex or MinDist refer to the phonetic wellformedness. Hamann (2011) argues that "[i]n such a minimal-functionalist approach it is impossible to define the phonetics-phonology interface, since there is no clear-cut distinction between phonetic and phonological representations, and both are handled in one grammar com-

[1] OT assumes Richness of the Base, stating that there are no restrictions on the input. Constraints on the structure of the input thus do not exist (Prince and Smolensky (n.d.)).

[2] The |Underlying Form| is represented between vertical brackets; the /Surface Form/ between forward slashes.

[3] The [[Auditory Form]] is represented between double square brackets.

[4] The [Articulatory Form] is represented between square brackets.

[5] This means that in speech production a speaker cannot skip the auditory form, but rather must come to the articulatory form via the auditory form.

ponent". BiPhon does clearly distinguish phonological and phonetic representations from each other, and consequently the phonetics-phonology interface is located in the relation between these two forms. The markedness constraints from classical OT that refer to phonetic characteristics thus cannot be constraints on one of the two phonological forms. Rather, these constraints guide the relationship between the phonological surface form and the phonetic auditory form, and are referred to as cue constraints (so-called because certain phonetic characteristics are used as cues for phonological features). This is what constitutes the phonetics-phonology interface. The sensorimotor constraints and articulatory constraints, finally, cover the markedness constraints from classical OT that cannot be considered structural or cue constraints. The sensorimotor constraints guide the relation between an articulatory movement and an auditory characteristic. An example is a constraint stating that "an auditory high F1 (first formant) does not correspond to an articulatory raised jaw" (Boersma (2011, p. 43)). Boersma is not very clear on the rankings of these constraints, stating that they, "*if and when* they have been learned (i.e. *if* the relevant sounds and articulations are used in the language at all, and *when* the learner has finished acquiring their relations), are universal [. . .]" (author's emphasis). This is expected considering the fact that an articulatory gesture must always have the same auditory effect, regardless of the language. However, he then states that "[. . .] perhaps they depend on the speaker". This statement is not further explained,[6] so it remains unclear whether Boersma assumes a universal or speaker-specific ranking of these constraints. What is specified, however, is that the sensorimotor constraints that are of any relevance in a speaker's language have a clear ranking, while "[a]reas in auditory or articulatory space that your language does not use at all will probably lead to poor sensorimotor knowledge (variable constraint ranking) in those areas [. . .]". In the model proposed by a.o. Boersma (2009), all different constraints can interact with each other, as the different forms are computed in parallel rather than serial. This is referred to as 'cross-level parallellism'.

BiPhon thus assumes the presence of five different levels. In the derivation these levels are not only all linked to, but also interact with each other. One could assume a relatively simple derivation, where a speaker would retrieve individual morphemes from the lexicon, which would be transformed into an underlying form via the lexical constraints; the underlying form would subsequently be transformed into a surface form via faithfulness constraints (the surface form itself would also have to comply with the structural constraints); the surface form would then be transformed into an auditory form via the cue constraints; finally, the auditory form would be transformed into an articulatory form via the sensorimotor constraints. In speech perception, the derivation would go the other way around, starting from the auditory form. However, Boersma (2009) proposes a parallel rather than serial derivation. A

[6]It may be argued that there must be a language- or speaker-specific ranking of these constraints, as several acoustic characteristics can be realised in different ways. See for example the work by Sebregts (2015) on the realisation of rhotics in different varieties in Dutch.

speaker would thus not derive the optimal surface form from the underlying form only after the optimal underlying form is derived, but instead the speaker would derive the optimal underlying, surface, auditory and articulatory forms simultaneously. The listener would not derive the optimal surface form from the auditory form, the optimal underlying form from the surface form, etc., but instead would derive the optimal surface form, underlying form and morphemes simultaneously.

An example

In this section the workings of the BiPhon model, and specifically the workings of the cue constraints, will be shown using an example of second language acquisition. The example is copied from Boersma (2009). In the example, Boersma discusses the case of perception of Russian by a native speaker of Japanese. Perception, in the case of BiPhon, is the mapping of the auditory form to the surface form. In the example, the input the Japanese speaker receives is the auditory form [dɾa.ma]. Contrary to Russian, however, Japanese does not allow consonant clusters. The question is thus how the Japanese speaker will perceive the Russian input form.

A possibility is that the Japanese speaker faithfully postulates an underlying consonant cluster in the surface form. Boersma (2009) rules this out as a possibility by proposing a high-ranked constraint */CC/, which prohibits any consonant cluster in the surface form. This constraint must be high-ranked in the speaker's grammar because Japanese does not allow consonant clusters.

A second possibility is that the Japanese speaker does not postulate the plosive [d] in the surface form, but instead realises a simplex onset [ɾ]. This is where the first cue constraint comes in: Boersma proposes that speakers of Japanese have a high-ranked constraint that prohibits linking a [d] in the auditory form to nothing in the surface form.[7]

The third possibility, which is the actually chosen scenario, is the insertion of a vowel in the surface form breaking up the cluster, so that the surface form is tri- instead of bisyllabic. This is ascertained by a low-ranked constraint that prohibits a link between nothing in the auditory form and a vowel in the surface form. The following tableau shows how the ranking of the different constraints ensures a mapping of auditory input [dɾa.ma] to surface form /do.ɾa.ma/.[8]

[7] In perception this constraint thus prohibits deletion of the [d], while in production it prohibits insertion of [d].

[8] Japanese does not allow the sequence /du/ in surface forms, so the tableau must include a high-ranked constraint prohibiting these constraints. The tableau also includes a constraint that prohibits a link between nothing in the auditory input and /u/ in the surface form, merely to show the workings of these constraints.

[dɾa.ma]	*/CC/	*/ / [burst]	*/du/	*/o/ [ ]	*/u/ [ ]
a. /dɾa.ma/	*				
b. /ɾa.ma/		*			
c. /du.ɾa.ma/			*		*
☞ d. /do.ɾa.ma/				*	

The same constraint ranking accounts for perception but also for production of the same form by a Japanese speaker. Note, however, that the auditory output [ɾa.ma] is very unlikely to arise, as an entire syllable would have to be deleted. Further, the constraint that prohibits a link between /u/ in the underlying form and nothing in the auditory form has been replaced in this tableau: as the surface form does not contain /u/, the constraint is irrelevant in production. Instead, a constraint prohibiting a link between /o/ in the input and [u] in the output is added to ascertain the first vowel does not change quality in production.

/do.ɾa.ma/	*/CC/	*/ / [burst]	*/d/ [ ]	*/du/	*/o/ [u]	*/o/ [ ]
a. [dɾa.ma]	*					*
b. [ɾa.ma]			*			*
c. [du.ɾa.ma]					*	
☞ d. [do.ɾa.ma]						

In this example the cue constraints have been proposed as a link between a specific segment in the surface form and another segment in the auditory form. For the sake of the argument, the constraints have been somewhat simplified, as cue constraints need not necessarily refer to segments in the auditory form. Instead, they might also refer to specific phonetic characteristics, such as formant values, duration or VOT values. The latter option will be chosen in the next section, where segments in the surface form are linked to phonetic values in the auditory form.

BiPhon and the two transition zones

The laryngeal systems of the west and east can be easily represented in BiPhon. In the west, a cue constraint requiring lenis ([voice]) plosives to be realised with a negative VOT is required to ensure lenis plosives are realised with prevoicing. Further, some constraint is necessary that links unmarked plosives to any VOT value that is not negative.[9] In the east a cue constraint requiring fortis ([spread glottis]) plosives to be realised with aspiration is necessary, again accompanied by a constraint that requires lenis plosives to be realised without aspiration

[9]Note that a minimal distance constraint would not work in this case, as it is vital that fortis plosives are realised without prevoicing. With only a minimal distance constraint fortis plosives could still be realised with prevoicing, as long as lenis plosives are realised with, for example, 200 ms. prevoicing and fortis plosives with, for example, 50 ms. prevoicing. As fortis plosives are never realised with prevoicing a constraint blocking this from happening is required, rather than a minimal distance constraint.

(lenis plosives can be realised with either prevoicing or short-lag VOT, as long as they are not realised with aspiration).[10]

Note that cue constraints cannot be formulated in terms of a relationship between a phonological feature and a specific phonetic realisation if one assumes privative features: although a feature [voice] (or any other feature, such as [round] for vowels) can be linked to a phonetic realisation, [∅] cannot be linked to a phonetic realisation in principle because it indicates the absence of a feature. Further, it is impossible to make a distinction between [∅] that refers to the absence of a laryngeal feature and another type of [∅], referring e.g. to the absence of a roundedness feature. The approach used in e.g. Hamann (2011), where instead of referring to features the constraints refer to phonemes, is a solution to this problem. This means that in the west of the continuum, /b/ must be linked to prevoicing while /p/ must be linked to positive VOT values, and in the east of the continuum, /p/ must be linked to negative and short-lag VOT values and /p^h/ must be linked to long-lag VOT values. In the middle area the only plosive /p/[11] is not linked to any VOT value.

A total of nine constraints is required to represent all possible realisations of the three different laryngeal categories: every phonological category ([voice], [∅] and [spread glottis]) must be linked to every phonetic realisation (prevoiced, short-lag VOT and aspirated). In a voicing language, the constraints on the realisation of lenis plosives must be ranked as follows[12]:

/b/	*/b/ 30 - ∞ ms	*/b/ 0 - 30 ms	*/b/ -∞ - 0 ms
☞ a. -80 ms			*
b. 10 ms		*	
c. 60 ms	*		

The constraints on the realisation of fortis plosives in a voicing language must be ranked as follows[13]:

/p/	*/p/ -∞ - 0 ms	*/p/ 30 - ∞ ms	*/p/ 0 - 30 ms
a. -∞ - 0 ms	*		
☞ b. 0 - 30 ms			*
☞ c. 30 - ∞ ms		*	

In the eastern part of the continuum the constraints are ranked so that the laryngeal contrast is realised as an aspiration contrast:

[10] Note that, again, a minimal distance constraint would not yield the same results: if fortis plosives would be realised with a VOT of, for example, 200 ms. and lenis plosives would be realised with a VOT of, for example, 50 ms., the minimal distance constraint would most likely be satisfied but both plosive series would surface with aspiration.

[11] The plosive is here represented as /p/ (or /t/ or /k/) to indicate the absence of a laryngeal feature.

[12] - ∞ represents any negative VOT value.

[13] Note that the latter two constraints are unranked with respect to each other: the unmarked fortis series can be realised either with a short-lag VOT or with aspiration.

/p/	*/p/ 30 - ∞ ms	*/p/ 0 - 30 ms	*/p/ -∞ - 0 ms
☞ a. -80 ms			*
☞ b. 10 ms		*	
c. 60 ms	*		

/pʰ/	*/pʰ/ -∞ - 0 ms	*/pʰ/ 0 - 30 ms	*/pʰ/ 30 - ∞ ms
a. -80 ms	*		
b. 10 ms		*	
☞ c. 60 ms			*

In the middle area the unmarked plosives can be realised with any VOT value. The constraint ranking is very simple in this area:

/p/	*/p/ -∞ - 0 ms	*/p/ 0 - 30 ms	*/p/ 30 - ∞ ms
☞ a. -80 ms			
☞ b. 10 ms			
☞ c. 60 ms			

Such a ranking, where the cue constraints are unranked with respect to each other, is unlikely. When a child receives linguistic input, it will always rerank its constraints so that the ranking matches the input. If the child in the transition zone ends up with a set of unranked cue constraints, the input must be such that constraints are either never reranked, or such that constraints are reranked and subsequently reranked back to their old ranking. It is important to show that ending up with such a ranking is indeed possible.

Let us first concentrate on how acquisition in BiPhon works. The assumption is, similar to classical OT, that all constraints are universal and innate. This means that all cue constraints must be universal and innate as well. The task of the language-learning child is to postulate a ranking of the constraints that corresponds to the observed input. The starting point is a set of unranked constraints. Reranking only happens if the input gives evidence for this; if not, the ranking of the constraints will not change.

Let us now assume that a child in the transition zone encounters a speaker with a stable voicing contrast. The child will find evidence for a voicing contrast, postulating a /b/ and a /p/. The cue constraint linking /b/ to prevoicing will then be ranked over the cue constraint linking prevoicing to /p/ (this ranking must be reversed in an aspiration system as it does not have /b/), and the cue constraint linking aspiration to /p/ will be ranked over the cue constraint linking aspiration to /pʰ/ (the voicing system obviously does not have /pʰ/ so this constraint must be ranked low). The cue constraint linking short-lag VOT values to /p/ must be ranked over the cue constraints linking it to /b/ (/b/ can be realised without prevoicing because of articulatory difficulties) and to /pʰ/.

-80 ms	*-∞ - 0 ms /pʰ/	*-∞ - 0 ms /p/	*-∞ - 0 ms /b/
☞ a. /b/			*
b. /p/		*	
c. /pʰ/	*		

10 ms	*0 - 30 ms /pʰ/	*0 - 30 ms /b/	*0 - 30 ms /p/
☞ a. /p/			*
b. /b/		*	
c. /pʰ/	*		

60 ms	*30 - ∞ ms /b/	*30 - ∞ ms /pʰ/	*30 - ∞ ms /p/
☞ a. /p/			*
b. /pʰ/		*	
c. /b/	*		

Following this, the child encounters a speaker with a stable aspiration system. The child will find evidence for an aspiration contrast, postulating a /p/ and a /pʰ/. The cue constraint linking prevoicing to /p/ must now be ranked over the cue constraint linking prevoicing to /b/ (the aspiration system does not have /b/), and the cue constraint linking aspiration to /pʰ/ must now be ranked over the cue constraint linking aspiration to /p/. The cue constraint linking short-lag VOT to /p/ must be ranked over the cue constraints linking it to /b/ or /pʰ/.

60 ms	*30 - ∞ ms /b/	*30 - ∞ ms /p/	*30 - ∞ ms /pʰ/
☞ a. /pʰ/			*
b. /p/		*	
c. /b/	*		

10 ms	*0 - 30 ms /pʰ/	*0 - 30 ms /b/	*0 - 30 ms /p/
☞ a. /p/			*
b. /b/		*	
c. /pʰ/	*		

-80 ms	*-∞ - 0 ms /pʰ/	*-∞ - 0 ms /b/	*-∞ - 0 ms /p/
☞ a. /p/			*
b. /b/		*	
c. /pʰ/	*		

The first thing that it visible is that the cue constraint prohibiting a link between short-lag VOT and /p/ is always dominated by the constraints prohibiting a link between short-lag VOT and /b/ or /pʰ/. The language-learning child thus always has evidence for ranking this constraint the lowest: any short-lag VOT value will be interpreted as /p/. The difficulties arise in the ranking of the cue constraints guiding the phonological interpretation of prevoicing and aspiration. In a voicing system, the cue constraint prohibiting a link between prevoicing and /b/ must be the lowest-ranked constraint, while in an aspiration system this constraint must dominate the constraint linking prevoicing to /p/. When a child encounters the two different systems, it will thus not be able to find consistent evidence for the interpretation of negative VOT values: these values can either be interpreted as /b/ or as /p/. A similar pattern is found for long-lag VOT values. In a voicing system these values are interpreted as /p/ (as /pʰ/ is absent from the system), but in aspiration systems these values are interpreted as /pʰ/. Again, the child will not be able to find evidence for the interpretation of aspiration values: they can be interpreted as either /p/ or /pʰ/. A result of the conflicting input is that the child will have to rerank the constraints every time it encounters a different system in the input.[14] The child can thus never postulate a definitive ranking, and can never postulate a consistent phonological grammar. The only evidence that is consistently present in the grammar is the evidence for the presence of /p/. In conclusion, it is not impossible and even likely that a language-learning child, upon receiving input from two conflicting linguistic systems, can only postulate one laryngeal category.

A question that has not yet been answered is why, if children acquire only /p/ but not /b/ or /pʰ/, their phonetic realisations seem to show not only the presence of phonetic [p] but also the presence of phonetic [b] and [pʰ]. To answer this question we have to take another look at the ranking of the cue constraints linking the different VOT values to /p/. In a voicing system, the link between prevoicing and /p/ is prohibited by ranking that constraint highest, while in an aspiration system, that constraint is ranked lowest (together with the constraint linking short-lag VOT values to /p/). In an aspiration system, the link between aspiration and /p/ is prohibited by ranking that constraint highest, while in a voicing system, that constraint is ranked lowest (together with the constraint linking short-lag VOT values to /p/). Similar to the previous discussion, on why children only postulate one laryngeal category, the answer to this question can be found in conflicting constraint rankings. In a voicing system, the child finds evidence for ranking the constraint that prohibits a link between prevoicing and /p/ above the constraints prohibiting a link between short-lag and long-lag VOT values to /p/, while in an aspiration system, it only finds evidence for ranking the former constraint at the same level as the constraint prohibiting a link between short-lag VOT values and /p/. The child thus does not find definitive evidence for a ranking of the two constraints prohibiting a link between prevoicing and /p/ and between short-lag VOT and /p/, and has to leave these two constraints unranked. In an aspiration system, the child finds evidence for ranking the constraint that prohibits a link between aspiration and /p/ above the constraints prohibiting a link between prevoicing

[14] In a way, this is similar to the "Buttered cat paradox". This thought experiment concerns the question what would happen when a cat, with a buttered slice of toast attached to its back, would fall: as cats always land on their feet, and a slice of toast always falls buttered side down, a falling 'buttered' cat would end up in a paradox where it would keep spinning around and around, without ever hitting the ground.

and short-lag VOT values and /p/, while in a voicing system it only finds evidence for ranking the former constraint at the same level as the constraint prohibiting a link between short-lag VOT values and /p/. Again, the child does not find definitive evidence for a ranking of the two constraints prohibiting a link between aspiration and /p/ and between short-lag VOT and /p/, and has to leave these two constraints unranked. This means that all three constraints, linking /p/ to either prevoicing, short-lag VOT and aspiration, are unranked with respect to each other. In conclusion, the child is very free in the phonetic implementation of the plosive.

Besides the constraints regulating the realisation of plosives in word-initial position, there must be constraints on the realisation of single intervocalic plosives and plosive clusters. Intervocalically the constraint ranking in the west must account for the full voicing of lenis plosives, and in word-final position it must account for the devoicing of lenis plosives. In the east it must allow full voicing of lenis intervocalic plosives (although it should not force them to be fully voiced), while at the same time regulating the final fortition of lenis plosives in word-final position.

In the middle area, the constraints guiding the realisation of intervocalic plosives must be similar to the constraints on the realisation of word-initial plosives. There is only one plosive series (phonologically represented as /p, t, k/ which are 'laryngeally empty'). The cue constraints linking the intervocalic plosives to specific phonetic realisations (with respect to the percentage of voicing during the closure) must be unranked: these plosives can surface either with partial voicing or with full voicing.

The western area has an active rule of voicing assimilation. This is a consistent, phonological process, and must be represented in the ranking of the faithfulness constraints. The cue constraints, at their turn, are responsible for the realisation of these clusters. A cluster consisting of two voiced plosives must be realised with full voicing during the closure, while a cluster with two voiceless plosives must be realised with partial voicing. In the east there is no phonological rule of voicing assimilation, but the cue constraints are evidently still responsible for the realisation of these plosives. The cue constraints can only allow full voicing during the closure if both plosives are lenis; if at least one of the plosives is fortis, the cluster must be realised with partial voicing.

The middle area cannot have a process of voicing assimilation, as there only is one plosive series. The faithfulness constraints are thus irrelevant in this region. The cue constraints must regulate the phonetic realisation of the clusters. As voicing is not contrastive, the cue constraints regulating the percentage of voicing during closure can be assumed to be unranked: an intervocalic plosive can surface either with full or with partial voicing.

In the Italian continuum the absence or presence of ISV is modelled in the relationship between the |Underlying Form| and the /Surface Form/ (faithfulness constraints). The cue constraints are responsible for the phonetic realisation of the fricative. As there is not much variation in this realisation, the constraint ranking simply looks as follows:

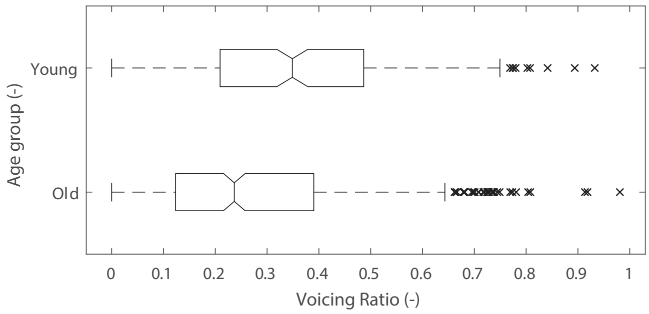


Figure 5.1: Distribution of voicing ratios of partially voiced fricatives, cross-classified by younger and older Tuscan speakers

/kasa/	*/s/[z]	*/s/[s]
☞ a. [kasa]		*
b. [kaza]	*	

/kaza/	*/z/[s]	/z/[z]
☞ a. [kaza]		*
b. [kasa]	*	

The ranking of the four relevant constraints, on the link between the phonetic and phonological form of the intervocalic fricatives, must be similar in the two subregions: the constraints linking /s/ to [s] and /z/ to [z] must be ranked above the constraints linking /s/ to [z] and /z/ to [s]. These constraints cannot explain, however, the differences in voicing percentages between younger and older Tuscan speakers (as shown in Section 3.4 there is a significant difference in voicing percentages between younger and older Tuscan speakers). First, young Tuscan speakers do not show consistent ISV, so their ranking of faithfulness constraints must be the same as the older speakers' ranking of faithfulness constraints. Second, although younger Tuscan speakers realise more intervocalic fricatives with full voicing, the voicing percentages of partially voiced stem-internal intervocalic fricatives is still higher than the voicing percentages of partially voiced stem-internal intervocalic fricatives of older speakers (see Figure 5.1).

A Welch Two Sample t-test shows that the difference in voicing ratios between younger and older speakers is statistically significant.[15] The difference in voicing ratios thus cannot be explained by the higher number of fricatives with full voicing. The difference in voicing ratios between younger and older speakers cannot be explained as a phonological difference either: neither generation has consistent ISV, and the difference in voicing ratios between the generations is still visible if one corrects for the difference in the amount of fully voiced fricatives. The phonetic differences, however, are only visible in stem-internal

[15] $t = -4.6341$, $df = 611$, $p = 4.3853e\text{-}6$.

position and at an opaque morpheme boundary. In other positions the two generations have similar voicing realisations.[16] The cue constraints therefore cannot account for the observed phonetic patterns. Instead, the differences in voicing values can only be explained on lexical grounds (only the faithfulness constraints can refer to the morphophonological context in which a fricative appears and they have been shown not to play a role; the cue constraints cannot refer to the morphophonological context and cannot make a distinction between the fully voiced fricatives in stem-internal contexts and the partially voiced fricatives in the same context).

It is important to note that the precise specification of the cue constraints (whether they link a phoneme to an articulatory form or to a precise phonetic value) is irrelevant. Formulating them in terms of precise phonetic values does not mean that these phonetic values no longer have to be linked to individual lexical items.

5.3.2 Exemplar Theory

Hooper (1976) (but see also Bybee (1999)) found a correlation between two types of morphophonological change (sound change and analogical levelling) and the frequency of words they first apply to: sound change affects frequent items before it affects infrequent items, while analogical levelling affects infrequent items before it affects frequent items. Although her data were only explorative, she argued that if they are correct and if the identified pattern is general, phonological theory should be able to capture these frequency effects. Exemplar Theory is an example of such a model. It is a usage-based theory of phonology and phonetics, which assumes that phonology, or rather grammar in general, is shaped by the way a language is used: linguistic categories are emergent rather than innate, and small changes in language use might change categories or create new categories. It "was first introduced in psychology as a model of perception and categorization" (Pierrehumbert (2001)), and first applied to language perception by Johnson (1996). Roughly, the model assumes that speakers do not acquire a phonological grammar consisting of rules or constraint rankings, but instead store every individual linguistic utterance they have ever encountered. The stored unit contains very specific phonetic details: formant frequencies, formant transitions, pitches, durations, whether the speaker of the utterance was male or female, etc. (Bybee (1999) proposes that speakers in fact store gestural scores, following Browman and Goldstein (1991)). All utterances are stored in a network, where similar utterances form a cloud together (the items are labelled as belonging to the same cloud). This

[16] The differences are never between speakers who realise a phonologically voiceless category with full voicing and speakers who realise those categories with partial voicing, but always between speakers who realise a phonologically voiceless category with partial voicing (one of the speakers has a higher voicing percentage than the other, but for both speakers the fricative is partially voiced) or between speakers where one of the speakers realises a category that might be specified as voiced with full voicing, while the other speakers realises that category with partial voicing.

can be seen as a lexical item: for example, all items characterised by a velar closure, an explosion, a delay in onset of periodicity, a vowel with a relatively high first and second formant and an alveolar closure will belong to the cloud for the lexical item 'cat'.[17] When a new linguistic utterance is encountered, its precise phonetic values are compared to the precise phonetic values of all other stored items. Pierrehumbert (2001) shows, using a clear example, how speakers perceive a specific second formant (F2) value. In her example, a speaker encounters an utterance containing an F2 value of X. This specific F2 value is compared to all other F2 values that are stored within a window relative to the encountered F2 value (the size of the window is not specified, however). All items with an F2 value within the window are activated. If all items belong to the same cloud, the interpretation of the encountered utterance is clear: it will belong to the same cloud as the activated items. However, if items belonging to two (or more) different clouds are activated, interpretation of the encountered utterance will be guided by frequency: it will be perceived as belonging to the same category that most activated items belong to (Lavie (2005)).

Every time a speaker encounters a new realisation of a linguistic item, it is added to the already existing set of stored utterances. A speaker's network of stored units is thus updated every time the speaker encounters a new realisation of a specific linguistic item. Theoretically, this means that speakers need unlimited memory, as in theory there is no limit on the amount of linguistic utterances they can encounter. Even though the human brain has proven to have a very large memory capacity, it seems unlikely that there are no limits on this capacity. Exemplar Theory thus assumes that stored utterances decay: "[m]emories of utterances that we heard yesterday are more vivid than memories from a decade ago" (Pierrehumbert (2001)). Newly stored realisations have a bigger influence on the realisation the speaker calculates than older stored realisations. Johnson (1996) gives a relatively simple model of memory decay: all items are stored in a matrix, with time represented on the x-axis (the y-axis represents different auditory characteristics). The newest items enter the matrix on the right, pushing the older items to the left. Eventually, they will be forced out of the matrix (i.e., they are forgotten by the speaker) (see Pierrehumbert (2001) for more details on memory decay). Another assumption of the framework is that stored items for which a specific phonetic characteristic differs by less than the 'Just Noticeable Difference' (JND) are stored as having the same value for that characteristic. This reduces the number of items that have to be stored.

Although Exemplar Theory started as a model of perception, it can be elaborated to a model of speech production as well. After having selected a cloud (a network of exemplars), the speaker chooses a random exemplar from the cloud to produce. This would mean that any phonetic realisation a speaker produces, is a phonetic realisation he or she has encountered before. However,

[17] Possibly, but this is not discussed in works on Exemplar Theory, such a cloud can also be seen as a phoneme, if it were possible to compare only a part of an encountered utterance to a part of a stored utterance.

the precise realisation will differ somewhat from the stored realisation, as a result of, for example, motor control.[18] By modelling production as the random selection of an exemplar from all stored exemplars, the model automatically accounts for frequency effects: the items that are heard more often are stored more often, consequently have a higher frequency in storage and thus have a bigger chance of being selected in production.

Pierrehumbert (2001) proposes a model with a systematic bias to account for sound changes such as lenition. In her model the bias assumes a specific value (-0.01 in her example[19]) which causes the phonetic realisation a speaker produces to always be different from the phonetic realisation a speaker intended. Even if a speaker intended to already realise a lenited exemplar, the eventual realisation of the exemplar will always be slightly more lenited. Pierrehumbert does not discuss when languages will allow lenition to occur and when they will not (although important, it is not of any relevance to the question of how to model such a process).

Language change can quite easily be modelled in Exemplar Theory. As speakers base their production on every item they have stored, encountering a variant of an item with a small phonetic difference from the majority of stored items will have an effect on production: as the exemplar is added to the set of stored exemplars, this exemplar may be selected as the output candidate in speech production. As more exemplars with that small phonetic difference are added to the set of exemplars, the chances of an exemplar that shows the phonetic change being selected increase.

Exemplar Theory and the two transition zones

Exemplar Theory assumes that speakers store realisations of utterances they encounter. Realisations that are similar enough are stored in the same cloud, thereby forming a network of realisations. This storing of utterances is done very precisely: it is assumed that speakers store phonetic details (e.g. the spectrogram) of every utterance they hear. The precise realisation of an utterance depends on the realisations of the utterances that are stored by the speaker, as they randomly select an utterance from the cloud as the aim for their phonetic realisation. As speakers store all new realisations they encounter, and use these realisations to calculate their own realisation, this realisation might be shifted upon encountering another realisation of the same utterance. All speakers will therefore always show some diachronic variation in their phonetics.

In contact areas speakers are likely to encounter realisations from both different regions, leading to a large amount of variation in their stored realisations. The realisations in the transition zone will be based on the exemplars a speaker

[18]Pierrehumbert (2001) does not discuss "whether the exemplars have a dual acoustic-motor nature, or whether the motor program is computed on the fly in order to match the acoustic goals represented by the exemplar".

[19]This specific example, with a bias of -0.01, causes a lenition. A positive bias value would cause fortition.

has stored. A speaker will select one of the exemplars, whereby the chances of being selected are higher for the more frequent exemplars than for the more infrequent exemplars. The surrounding exemplars all affect the eventual realisation, so an effect of the speech of a speaker's neighbours will be visible as well: if one assumes an initial stage with an abrupt boundary between the voicing and aspiration systems, a speaker will select one realisation for a lenis plosive. As these are prevoiced in voicing languages and either prevoiced or plain voiceless in aspiration languages, the chances of selecting a prevoiced plosive are higher than the chances of selecting a plain voiceless plosive, but over time the same speaker is likely to every now and then select a plain voiceless plosive, so a small trace of the aspiration system might be visible. This way, the realisations of a speaker in the contact area might gradually change, until its realisations are perfectly midway between the voicing values and aspiration values. For fortis plosives the same pattern is predicted.

Some variation in VOT values is possible. It is of course not likely that a speaker will receive precisely equal amounts of input from both sides of the continuum, and the input a speaker receives may vary over time. So, at the time of recording the speaker may have received more input from the voicing side of the continuum than from the aspiration side of the continuum, causing their realisations to shift towards voicing values (or similar for aspiration values in the input). As this variation will depend on the speaker's social contacts, the places they have been, etc., this variation is unpredictable for the linguist. However, what is unlikely is that a speaker in the transition zone will have values perfectly corresponding with a voicing or an aspiration system, as this implies that the speaker is only in contact with speakers from the outer end of the continuum and not with other transitional speakers (like his/her family members, neighbours, colleagues, etc.) that do have linguistic contact with speakers of an aspiration system. Exemplar Theory thus predicts that the transition zone shows a gradual change, accompanied by variation.

A further point is that in the present study the selected items could not always be kept constant, so for different villages different lexical items were used in the study. These different items might all show a different pattern with respect to VOT, depending, for example, on the vowel following the consonant, or depending on the source of individual exemplars the speaker may have stored (and on which they will base their own realisation). Some deviation from the perfectly gradual pattern may thus be expected.

For the Italian transition zone the prediction is quite similar, only the relevant variable is the voicing ratio of the intervocalic fricatives. The lexical items that have a voiced fricative in both regions will have a voiced fricative in the transition zone as well, but for the lexical items that have a voiced fricative in the north and a voiceless fricative in the south, a gradual transition is expected: from south to north, the percentage of voicing of intervocalic fricatives is expected to gradually increase. Some variation is again expected, but it is not expected that speakers in the transition zone have realisations that are perfectly compatible with either one of the systems, as this implies they only have contact with speakers from either one of the stable systems.

If we compare the data in both transition zones to the predictions of Exemplar Theory, we see that both areas show patterns that are not fully compatible with the predictions. The Dutch-German transition appears to be quite compatible with the predictions made by Exemplar Theory, as VOT values for both fortis and lenis plosives (for all PoA's) show a gradual change. However, taking a closer look at the data and the framework, this gradual phonetic increase is not expected if speakers randomly select one exemplar to base their utterance on. Instead, one would expect a gradual increase in the number of items compatible with an aspiration system when moving from west to east (or with a voicing system when moving from east to west). The phonetically intermediate values (with very short prevoicing values or values that could be interpreted as either a short-lag or a long-lag VOT) are not expected to be present, yet the data show an abundance of these values. If speakers in the transition zone would show values that are similar to the values attested in the stable areas, it can be assumed that these values are not predicted by Exemplar Theory. The voicing patterns in consonant clusters with a fortis C2 do not appear to be explainable in Exemplar Theory either. As discussed, these clusters might surface with full voicing, but only in the middle area; in the west and east, these clusters are consistently partially voiced. The full voicing must be an independent innovation of the middle area. Such an innovation is obviously not impossible in Exemplar Theory, but it is unexpected considering the consistent absence of full voicing of these clusters in the west and east.

The data in the Italian continuum are not very well compatible with the framework either, albeit for different reasons. Exemplar Theory predicts that contact situations always lead to (phonetic) variation. Looking at the voicing values of the fricatives in the Italian continuum, a sudden change between the two regions is visible. Especially for the older speakers a clear distinction between the regions is visible, so that it is impossible to locate an area with a gradual change between the two systems; in other words, there is no transitional area between the two different systems. In Exemplar Theory, this would only be possible if there is no contact at all between speakers of the different regions. Although the boundary between the two regions coincides with a political border, viz. the border between Tuscany and Emilia-Romagna, it seems unlikely that speakers have no contact with speakers from the other province. There is no mountain region, river or other geological phenomenon located between the northernmost village in Tuscany and the southernmost village in Emilia-Romagna included in the fieldwork prohibiting contact between speakers from those locations, and it is not too difficult for speakers to move from one region to the other as the roads are easily accessible. Thus, if contact between the two areas is possible, the absence of a gradual change between the two systems is difficult to account for in Exemplar Theory.

It must be noted that, for the younger speakers, a small gradual transition is visible. Not only do they realise more fricatives with full voicing, the average voicing percentage of partially voiced fricatives is higher as well. This corresponds to the predictions made by Exemplar Theory. However, if the behaviour of younger speakers would be explained using only Exemplar Theory, the difference between younger and older Tuscan speakers is highly unexpected: while younger Tuscan speakers' behaviour is fully compatible with Exemplar Theory, the behaviour of older Tuscan speakers, Emilian-Romagnol speakers and Low Saxon speakers is not. To explain these differences without abandoning Exemplar Theory, it would have to be assumed that speakers in the Dutch-German dialect continuum do use exemplars, and thus store every linguistic utterance they have ever encountered. Their own realisation of a lexical item is based on these stored utterances. For speakers in the Italian dialect continuum, however, it will have to be assumed that their mental representation of linguistic representations is very different. While they can be assumed to store every single linguistic utterance they encounter, they cannot be assumed to use these stored representations when deciding on the phonetic realisation of a lexical item. Some other mechanism, which is not influenced by language contact, must be responsible for the phonetic realisation of linguistic items. One could think, for example, of language-specific phonetic implementation rules: when the phonological category of a phoneme is known, the phonetic implementation rule automatically applies to it. This rule may allow for some variation, but the specific phonetic values it allows for do not change under the influence of language contact. One would thus have to assume that the differences observed between the two transition zones are not the result of underlying linguistic differences, but rather of very different linguistic mechanisms. The way speakers acquire language, produce language and maybe even store language would thus have to be very different between speakers of different languages. This has many different implications, not just for linguistics but even for human cognition in general: either the linguistic system is not universal and congenital, but develops at a later age, when the child has discovered his or her native language, or different children are born with different linguistic systems. The first possibility predicts that children are able to acquire at least some language (they can distinguish different languages because they have to 'decide' which cognitive linguistic system to use) without any cognitive linguistic system, which seems very unlikely if not impossible. The second possibility predicts that bilingual children (speaking Italian and Dutch or German) can only be fully fluent in one of the languages, not in both, as they require different cognitive mechanisms, or these children make use of different linguistic systems for the different languages.

In conclusion, it seems very unlikely that speakers in the Dutch-German dialect continuum do calculate phonetic realisations based on stored exemplars, while speakers in the Italian dialect continuum do not use stored exemplars to calculate phonetic realisations.

5.3.3 Lexical variation

A remaining point that needs to be explained is the phonetic variation in the middle of the Dutch-German continuum. In this area, speakers realise plosives with a large amount of variation in VOT values, but the variation corresponds with the status a plosive has in either of the stable areas (i.e., a plosive that is lenis in a voicing or aspiration system is realised with prevoicing or as plain voiceless, and a plosive that is fortis in a voicing or aspiration system is realised as plain voiceless or with aspiration). In Section 5.3.1 it has been shown that the different cue constraints are unranked with respect to each other, so that the plosives can be realised with a broad range of VOT values. The pattern that is found in this variation, however, still needs to be explained. It cannot be modelled as a link between a phoneme and a phonetic realisation, as there is only one laryngeal category present in the middle area, and the phonetic realisation depends on the lexical item it occurs in. The most likely explanation is that speakers' phonetic realisations are influenced by the realisations of speakers in surrounding areas.

As the variation depends on the linguistic context or lexical item the phoneme appears in, the only way this regularity can be modelled is, probably, lexical. VOT values would have to be linked to individual lexical items, rather than to phonemes. As BiPhon is a parallel model of perception and production, in theory it is possible to model this. It would be a large adaptation of the framework, however, as it requires a direct link between entire lexical items and phonetic characteristics instead of a link between individual segments and phonetic characteristics.

In theory, the BiPhon model could be adapted to specify phonetic characteristics of individual lexical items. For every lexical item a cue constraint would have to be formulated, linking the surface form to the auditory form. This means that a very large number of cue constraints is required: any item with a plosive in it needs a cue constraint guiding the phonetic realisation of all its segments. The cue constraints indeed have to specify the relationship between surface and auditory form for all segments: if the entire item is present in the part of the constraint referring to the surface form, it also has to be present in the part of the constraint referring to the auditory form. If only one phonetic characteristic is present in the cue constraint, it is impossible for the speaker to know to which segment the phonetic information must be linked. Further, if this type of cue constraints is required to account for the relationship between surface and auditory form for every item containing a plosive, it is difficult to explain why it should not be required for every item that does not contain a plosive.

If all phonetic characteristics of all lexical items need to be specified in the constraint, the entire principle of constraint ranking and constraint interaction is lost: the constraint prohibiting the link between the surface form and the eventual phonetic realisation must be ranked lowest. This is not very different from simply storing a phonetic form for every phonological form.

A simpler way to model this would be to assume an exemplar-like layer on top of BiPhon, that might influence the specific phonetic realisation of lexical items as long as the values fall within the range specified by the cue constraints. The BiPhon model is not adapted in any way; the cue constraints simply link individual segments to specific phonetic characteristics. In the case of the middle area of the Dutch-German continuum, there are cue constraints linking the plosives /p/, /t/ and /k/ to the three different VOT categories. As the cue constraints are unranked, speakers can freely choose between prevoiced, plain voiceless and voiceless aspirated plosives. The specific phonetic implementation, i.e. within which phonetic category the plosive will be realised, is dependent on the speaker's previous experiences, but not on the constraint ranking as the constraints are unranked. An item with a fortis plosive is realised with a short-lag VOT in the west and a long-lag VOT in the east, so the speaker in the middle area will most likely choose either one of those implementations rather than realising the item with a prevoiced plosive. Lenis plosives are more likely to be realised with prevoicing or a short-lag VOT, because the input speakers in the middle area receive does not contain lenis plosives with long-lag VOT.

5.4 Modelling phonological variation

Phonological processes usually apply across the board. Some processes, however, have been shown to apply variably. An example is the deletion of a word-final /n/ after a schwa in Dutch. Exceptions are the indefinite article *een* 'an' and a stem-final nasal in verbs (e.g. *tekenen* 'to draw' never has deletion of the nasal in the first person singular *ik teken* 'I draw'). Van de Velde and Van Hout (1998) and Van de Velde and Van Hout (2003) show that this process is applied variably: it targets the nasal in the verbal plural, verbal infinitival and nominal plural suffix *-en* more often than a stem-final nasal in nouns such as *molen* 'mill'; it is applied more often in informal speech than in formal speech; it is applied more often by speakers of northern standard Dutch (i.e. standard Dutch as spoken in the Netherlands) than by speakers of southern standard Dutch (i.e. standard Dutch as spoken in Flanders). However, most speakers show variation in the application of the rule: within a specific linguistic context, an individual speaker may choose to either apply or not apply the rule, even though the speech register is the same. A very simple approach to modelling this kind of variation is that speakers might have access to multiple grammars (each grammar corresponding to a different outcome). Within one grammar variation does not occur, but the two (or potentially even more[20]) grammars do differ from each other, thereby causing variation to occur. However, many authors have tried to model phonological variation within the grammar rather than between two different grammars. The approach to modelling variation

[20]To my knowledge there are no studies to the maximum number of grammars a speaker may have access to. Although the number can be relatively large, it seems natural to assume there are limits to this, as the human brain capacity is limited.

differs depending on different approaches to phonology.

In this section one approach in RBP and several approaches in OT will be discussed. Data from the Italian continuum will be used to shed light on the question of modelling variation: as has been shown in Chapter 3, younger speakers in Tuscany realise more fricatives with full voicing than older Tuscan speakers do, but they do not consistently realise all fricatives as voiced. They thus cannot be argued to have acquired the same allophonic distribution as northern speakers, but they cannot be argued to have acquired a stable phonological contrast either. Instead, they must be inconsistently applying a phonological process that voices prosodic word-internal intervocalic fricatives. The realisations of younger Tuscan speakers can thus be very insightful in the discussion on the representation of phonological variation.

5.4.1 Variable rules

In RBP, the phonological surface form is derived from the underlying phonological form via an ordered set of rules (e.g. Chomsky and Halle (1968), Giegerich (1992)). A phonological rule describes a change in phonological representation of a phoneme in a specific context. For example, the rule [-sonorant, +voice] $\rightarrow$ [-sonorant, -voice] / ___σ changes a syllable-final voiced obstruent into a voiceless obstruent (i.e. it enforces Final Devoicing). Modelling variation in a rule-based approach can only be done by making assumptions about the frequency of application of a rule.[21] Cedergren and Sankoff (1974) developed a model in which a variable rule is assigned an application probability, representing the likelihood of the rule to apply. If a rule has an application probability of 0.75, the rule will apply 75% of the time; the other 25% of the time it will not apply. Two problems, however, exist in this approach. First, this approach does not take into account the fact that the application of a rule might depend on the linguistic environment or the speaker. Final /n/-deletion in Dutch, for example, is more likely to occur in polymorphemic words than in monomorphemic words, but applies variably in both contexts. It is also a more frequent process in northern standard Dutch than in southern standard Dutch. Second, the specific application probability that needs to be assigned to a rule can only be based on studying how often a rule is applied by speakers. The application probability cannot be predicted based on characteristics of the process; rather, the process has to be studied to be able to assign it an application probability. This renders the approach merely descriptive instead of predictive.

---

[21] Changing the order in which rules apply would not have the desired effect. By changing the order in which different rules apply, the effect of an individual phonological rule might indeed be smaller if the context in which it applies is changed by a rule that now applies before the variable rule, rather than after. However, the variable rule itself still applies every time its context requirements are met, so that the outcome itself will never be variable.

5.4.2 Variation in Optimality Theory

Variation modelled in the constraint ranking

Several more predictive approaches are possible in OT. A first approach, first introduced by Anttila (1997) (but see also Anttila and Cho (1998), Anttila (2002) and Anttila (2008)), assumes that some OT constraints need not be ranked with respect to each other.[22] The constraint ranking is, in a manner of speaking, incomplete. During the evaluation of the candidates generated by GEN, all candidates must be evaluated with respect to all constraint rankings that are possible by fixing the position of the unranked constraint. If constraint A is unranked with respect to constraint B, two calculations must be made: one for A >> B and one for B >> A. When constraint A is unranked with respect to constraints B and C (which are ranked with respect to each other, e.g. B >> C), three calculations must be made: A >> B >> C, B >> A >> C and B >> C >> A. Based on the number of possible outputs and the frequency of each output, each possible output has a specific likelihood of occurrence. So if for input X candidate Y is the optimal output in two out of three constraint rankings, output Y is found in two-third of all instances, while output Z, which consequently is the optimal candidate in one out of three constraint rankings, can be found in one-third of all instances (Anttila (1997)). This approach is much more predictive than a variable rule approach, as the frequency of occurrence of different forms can be precisely calculated based on the different possible constraint rankings. Besides assuming that speakers have a set of constraints that are unranked with respect to each other, it is of course possible to assume that speakers have access to several grammars, each of which corresponds to a different ranking of the constraints. In the example with constraint A being unranked with respect to constraints B and C (of which B is ranked above C), a speaker would be assumed to have access to three different grammars: one grammar with the ranking A >> B >> C, one grammar with the ranking B >> A >> C and a final grammar with the ranking B >> C >> A. The results would still be expected to be the same if the speaker randomly chooses either one of the grammars.

A quite different approach to OT is stochastic OT (e.g. Boersma (1998), Boersma (1999)). In this model the ranking of constraints is not discrete (A over B or B over A), but continuous. Each constraint is assigned a ranking value. Constraints with a small difference in their ranking value are closer together than constraints with a large difference in ranking value, e.g. if constraint A has a ranking value of 10, constraint B a ranking value of 2 and constraint C a ranking value of 1, constraint A is ranked over both B and C, and constraint B is ranked over C, but the distance between A and B is much bigger than the distance between B and C. This is relevant because a small amount of noise is added to the ranking during the evaluation, which causes the constraint ranking to be potentially slightly different. The final ranking of each

[22] The approaches is called 'stratified OT' by Zuraw (2002).

constraint is calculated by adding a 'rankspreading value' to the ranking value of each constraint, which is multiplied by a variable **z** ("where **z** is a Gaussian random variable with mean 0 and standard deviation 1" (Boersma (1998))). The rankspreading value is similar for every constraint, but the **z** value might differ. As it has a mean of 0 and a standard deviation of 1, **z** might either be a positive or a negative value. If it is positive, it will cause a higher ranking of the constraint during evaluation (as the rankspreading value multiplied by **z** is added to the original ranking value), but if it is negative it will cause a lower ranking of the constraint during evaluation. As the **z** value might differ between different constraints, some constraints will be demoted and other constraints will be promoted during evaluation. If the absolute constraint ranking is not affected by the added noise, the noise will not have an effect on the optimal output. However, adding noise might cause two constraints to switch positions in the absolute ranking. In that case, the optimal output might be different from when no noise is added. The closer two constraints are ranked to each other, the bigger the chance they might switch positions during evaluation.

Variation in EVAL

Another option is provided by Coetzee (2004) (see also Coetzee (2006)), who places the source of variation not in the constraint ranking, but rather in the evaluating function (EVAL). In this framework the EVAL function does not simply make a distinction between the most optimal output candidate and the entire set of non-optimal candidates (the 'losers'). Rather, all possible outputs are compared to each other, leading to a ranking of the best output, the second-best output, third-best, etc. The optimal candidate is ranked first; the second-best candidate is the optimal candidate from the entire set of candidates minus the first-best; the third-best candidate is the optimal candidate from the entire set of candidates minus the first- and second-best; and so forth until all candidates are ranked. In selecting the eventual output, every candidate can in theory be selected. However, the more optimal an output candidate is, the more likely it is to be selected as the eventual output. Coetzee gives an example concerning /t, d/ deletion in Jamaican English (data taken from Patrick (1991)). In this variety of English, a final /t/ or /d/ in a consonant cluster (e.g. [wɛst]) can, but does not have to, be subject to deletion. The sequence 'west bank', for example, sometimes surfaces as 'west bank' (without deletion), but much more often surfaces as 'wes bank' (with deletion). The same is true for the sequence 'west end', which sometimes surfaces as 'west end' (without deletion), but much more often surfaces as 'wes end' (with deletion). For both inputs ('west bank' and 'west end') the output with deletion is optimal, while the output without deletion is the second-best candidate. The deletion candidate will thus be chosen more frequently than the no-deletion candidate.

The story of Jamaican English /t, d/ deletion is, however, slightly more complicated. Deletion is always preferred over no deletion, whether the consonant cluster is followed by a consonant or a vowel, but the preference for deletion is even stronger for clusters followed by a consonant. In other words, the final /t/ or /d/ is deleted more often when it is followed by a consonant than when it is followed by a vowel. Patrick (1991) gives a deletion frequency of 88% for clusters followed by a consonant, and a deletion frequency of 63% for clusters followed by a vowel. Coetzee assumes an extra feature of the EVAL function in order to model this: instead of only comparing output candidates for just one input, it can also compare output candidates across different inputs. In this specific example, it would be possible to compare all outputs for 'west bank' with all outputs for 'west end'. If the constraint banning a consonant +/t, d/ cluster followed by another consonant is ranked higher than the constraint banning a consonant +/t, d/ cluster followed by a vowel, no deletion of the final /t/ or /d/ would lead to a more marked output for a cluster followed by a consonant than for a cluster followed by a vowel. In comparison, 'west end' (without deletion) is more optimal than 'west bank', so 'west end' will be chosen as the eventual output more often than 'west bank'.

A final added feature Coetzee proposes is the presence of a cut-off point. If no such point were available, any output could be selected as the eventual output, even, in theory, a candidate that would violate all constraints in the ranking. Although it is likely that this candidate would be chosen very infrequently, the reality is that it is never chosen because it is ungrammatical. The model should thus be able to prevent this candidate from being chosen. The cut-off point does precisely this. It marks a point in the constraint ranking above which constraints cannot be violated: any output candidate that violates a constraint ranked above the cut-off point can never be chosen as the eventual output; only candidates that violate only constraints ranked below the cut-off point can. In the example given by Coetzee (2004), output candidates with vowel epenthesis between the consonant and the /t/ or /d/ are ungrammatical. This means that the constraint against insertion, Dep, must be ranked higher than the constraints against the cluster being followed by a consonant (*Ct#C, where <t> can be either voiced or voiceless), against the cluster being followed by a vowel (*Ct#V) and the constraint against deletion (Max). If the cut-off point lies between Dep and *Ct#C (with Dep ranked higher than *Ct#C and the other two constraints), both the variants with and without deletion can be selected as the eventual output, but a candidate with epenthesis of a vowel can never be selected.

5.4.3 Phonological variation and the two transition zones

The data from the Dutch-German transition zone are not very useful in a discussion on the different approaches to phonological variation: the variation in VOT values is phonetic, not phonological, and so is the variation in the application of voicing assimilation (although it is phonological in the west, in the middle the process is in fact not assimilation at all but rather phonetic intervocalic voicing of the cluster). The differences in underlying features is of course phonological, but there is no evidence for differences within a speaker or even a region. The Italian region can offer some insights in this discussion. Although there is no geographically gradual transition between the two regions, there is a difference between younger and older speakers in Tuscany. Younger speakers have higher percentages of fully voiced items, meaning that they realise a larger number of lexical items with a voiced fricative. A possibility is that the difference is lexical: the younger speakers have a phonologically voiced fricative in more items than the older speakers. If the difference would be lexical, the distribution of voiced and voiceless fricatives throughout the lexicon is predicted to be consistent. The data show a different pattern, however: there is no clear pattern in which items are realised with a voiced fricative and which with a voiceless fricative. Some items have a voiced realisation for several speakers in the north and south of Tuscany but a voiceless realisation for speakers in the centre of Tuscany or vice versa, for other items the realisation differs between speakers in the same village. A consistent lexical distribution thus cannot be found. It is of course possible that there is interspeaker variation with respect to which fricatives are represented with the feature [voice] and which are not, but based on only one questionnaire it is impossible to test whether an individual speaker's behaviour is consistent or not: if the distribution of voiced and voiceless fricatives would be lexical, consistency on the level of the individual speaker would still be expected. To test this hypothesis the questionnaire would have to be repeated with the same informants. However, a lexical distribution of voiced and voiceless fricatives makes it difficult to explain how individual speakers would come to postulate different underlying voicing features for the fricatives in these words: it could be argued that speakers need consistent input to be able to postulate a voiced fricative in an item. The existence of more items with a voiced fricative in the lexicon of the younger speakers than in the lexicon of the older speakers is difficult to explain based on the lexical inconsistency observed. As the distribution of the two fricatives is unpredictable, this hypothesis can be argued to be unlikely.

A more likely option is that the younger speakers occasionally apply ISV. Their grammars thus must allow variability in phonology. In this section the different approaches to variability (variation in the constraint ranking, variation in EVAL and the presence of multiple grammars) will be compared to the data.

The existence of multiple grammars assumes that variability is not encoded within a grammar but rather in the presence of the different grammars: each of the individual grammars is stable and invariable. The approach does not predict when speakers will use which grammar. This might depend on the extralinguistic context (speakers might for example use grammar A in a more formal situation and grammar B in a more informal situation, or use different grammars depending on the person they are having a conversation with). Some consistency in the use of grammar would be expected, however: if speakers would continuously switch between different grammars it is inherently impossible to make a distinction between the presence of multiple stable grammars and the presence of one variable grammar. For the present study this means that younger speakers, during the interview, would consistently use either the grammar with or the grammar without ISV, but not switch grammars during the interview. The inconsistent behaviour of speakers in Tuscany does not support the assumption of multiple grammars. Speakers often show inconsistent voicing (only one younger Tuscan speakers showed (almost) fully consistent ISV; the other speakers all show variability), implying that variability is encoded within the grammar.

If the variability would be explained by unranked constraints, only *(VC̥V)PWD and F-Contiguity have to be unranked with respect to each other (regardless of how often a speaker chooses to implement ISV). Statistically, a fifty-fifty distribution of voiced and voiceless fricatives would then be expected for all speakers: half of the time the realisation with ISV is chosen as the optimal output, and half of the time the realisation without ISV is chosen as the optimal output. This corresponds quite well with the distribution of voiced and voiceless fricatives in the speech of younger speakers: most speakers realise 40 to 60% of intervocalic fricatives with full voicing.

If the difference were encoded in EVAL rather than in the constraint ranking, the cut-off point for older Tuscan speakers would have to be below F-Contiguity (so that only an output candidate with the same fricative as in the input will be selected) while the cut-off point for younger Tuscan speakers would have to be below *(VC̥V)PWD (so that both output candidates with a changed fricative and output candidates with the same fricative as in the input can be selected). The predictions for the older Tuscan speakers are indeed borne out, but for the younger Tuscan speakers the pattern differs slightly from what is predicted by the theory. As the candidate with the voiceless fricative (assuming the lexical item has an underlyingly voiceless fricative in its representation) only violates the constraint *(VC̥V)PWD but not the higher-ranked constraint F-Contiguity, while the candidate with the voiced fricative (assuming the lexical item has an underlyingly voiceless fricative in its realisation) violates the higher-ranked constraint F-Contiguity, Coetzee (2004) predicts that the candidate with the voiceless realisation will be chosen more frequently. This is not supported by the data, which show that for quite a number of younger speakers the more frequently chosen realisation is in fact the candidate with the voiced fricative.

The data from the Italian continuum are better compatible with the assumption of a variable constraint ranking than with multiple grammars or variation in EVAL. For most younger speakers, the data are compatible with an absolute constraint ranking (within which the order of the constraints F-Contiguity and *(VC̥V)PWD is reversible), but several young Tuscan speakers have high percentages of fully voiced fricatives. This might indicate that the constraint ranking is stochastic rather than absolute, but more data are necessary to verify this assumption.

CHAPTER 6

Maps and Grammar

Geographical patterns in language can roughly be attributed to four different categories. First, geological factors may influence linguistic patterns. For example, a large river or mountain range can be difficult to cross, so language contact between the languages spoken on either side of the river or mountain is inhibited. As the different linguistic varieties will then evolve independently of the other variety, they are likely to diverge from each other. The geological phenomenon then coincides with (or rather, causes) the linguistic border.

A second factor contributing to geographical patterns in language is the socio-political situation. The orientation of speakers on a specific region or social group can influence their linguistic behaviour, and political borders can influence speakers' movement patterns, thereby influencing who they are in contact with and thus influencing their language.

Thirdly, climate has been argued to have an effect on language patterns. An early example is the work by Van Ginneken (1913), who argues that the deletion of intervocalic /d/ and the deletion of /r/ in the cluster /rs/ in the dialect of Zeeland can be explained by the local climate: the inhabitants of Zeeland close off their throat with their tongue to protect the throat from the strong wind.[1] Newer examples are the studies by Everett, Blasi and Roberts

[1] "Evenals de Engelschen en Denen sluiten de Zeeuwen de keel min of meer af met de tong om deze te beschermen tegen den guren wind. (Invloed van klimaat). Met de tong kan aldus niet sterk gearticuleerd worden. Het gevolg hiervan is o.a. dat een *d* tusschen twee klinkers en gelijk in het Engelsch een *r* voor een *s* vaak spoorloos wegvallen." (p. 25) ("Just as the Englishmen and the Danes, the Zealanders more or less close off their throat with their tongues to protect the throat from the wind. (Influence of climate). It is therefore impossible to strongly articulate with the tongue. A consequence is that, a.o., a *d* between two vowels and, just as in English, an *r* before an *s* are deleted.")

(2015), who link the presence of tonal contrasts to warm and humid climates, or Everett (2013) who links the presence of ejective consonants to mountainous areas.

Finally, genetic factors may contribute to linguistic patterns. Dediu and Ladd (2007, 2008) propose a causal correlation between the presence of linguistic tone and the presence of a specific gene.

The present study has as one of its main points of focus the link between geography and language variation. The proposal that climate may influence language patterns is not maintained here, and genetic factors in language have not been the subject of the present study. These two factors are thus not discussed in any further detail. The data from the study are used, however, to see what geographical patterns in language, more specifically transitional patterns, can teach us about language. Although the phenomena in both regions are related to voicing, the two regions show a very different transition between the two systems. In the Dutch-German area the transition between the two systems is both phonetically and phonologically gradual, while in the Italian area the transition is both phonetically and phonologically abrupt. In this chapter the differences between the two regions are discussed in more detail, including an analysis of why the two regions show different linguistic patterns. The linguistic characteristics of the two phenomena can partially explain the differences between these two patterns, but the socio-political situation in the two regions forms another part of the explanation. In this chapter the influence maps and grammar can have on each other, whichever way the direction of influence, is discussed.

6.1 The differences between the two regions

In Chapters 2 and 3 the two transition zones have been shown to follow very different underlying patterns. While in the Dutch-German continuum the transition is both phonetically and phonologically gradual, as the increase in VOT values is gradual and the middle area does not show any evidence for a phonological contrast, the Italian transition is phonologically abrupt. Phonetically the picture is a bit more complicated, as the change is abrupt for older speakers but appears to be a bit more gradual for younger speakers. As both phenomena involve laryngeal characteristics, this might be rather unexpected: both VOT and percentage of voicing (for fricatives) are continuous variables, and could in theory thus be expected to show a gradual increase. Phonologically the Italian region could show a gradual increase in the number of items realised with full voicing. An increase in the number of linguistic contexts where ISV applies has been shown unlikely in Section 4.4.5. The absence of an increase in the number of items realised with full voicing, at least for the older Tuscan speakers, cannot be explained on linguistic grounds. A logical explanation might therefore be the geography of the areas. Often, regions separated by geographical phenomena like rivers or mountain ranges are very different from each other, whereas regions that are not separated by geographical phenomena usually show more similarities between their dialects. Indeed, in the Low Saxon dialect continuum, there is no geographical phenomenon that could interrupt a gradual change. However, it is not very likely that the sudden change between northern and central Italian varieties is caused by only a natural phenomenon. As a geographical phenomenon is a complicating factor in language contact, the phenomenon is expected to be located precisely at the point where a discrete boundary between different language varieties is visible. In the Italian continuum the geographical phenomenon should thus be located precisely at the border between Tuscany and Emilia-Romagna, and is not expected to extend into either of the two provinces, as its effect on the languages is then expected to be visible in a larger area around the geographical border. Although there is a mountain range present in the Italian continuum, the location of the mountain range cannot fully explain the discrete border between the varieties with and the varieties without ISV: several villages where ISV is consistently present (Fanano and Sestola) are located within the mountain range, as well as several villages where ISV is absent (Rivoreta and Popiglio). Language contact should thus not only be complicated between the Emilian-Romagnol varieties on the one hand and the Tuscan varieties on the other hand, but also between the villages that are located within the mountain range, so that the effect of the mountains on the languages should be visible between those locations as well. The presence of a mountain range is thus not enough to explain the differences between the two regions.

Several linguistic factors, however, can explain these differences. The first factor has already been discussed in Chapter 3, and deals with the phonological representation of ISV. In an OT framework, the presence of ISV is represented by a constraint ranking in which a constraint disfavouring voiceless intervocalic alveolar fricatives (*(VC̥V)PWD) is ranked over a constraint against changing features in a segment that is both preceded and followed by another segment (F-Contiguity). As long as the former constraint is ranked over the latter, ISV is applied in all contexts where it is expected to. However, when the ranking is reversed, ISV cannot be applied at all. Phonologically an intermediate state (in which ISV is applied in some of the contexts but not in others) is impossible: as soon as F-Contiguity is ranked above *(VC̥V)PWD ISV cannot be applied at all. The phonologically sudden change is thus explained by the phonological representation of the process.

Related to this point is the question that was addressed in Section 4.4.4, where it has been shown that ISV in the northern varieties is more likely to be synchronically analysed as an allophonic distribution of voiced and voiceless fricatives than as a phonological process that voices a prosodic word-internal intervocalic voiceless fricative. Since children can acquire an allophonic distribution even if noise is present in the signal, children in Emilia-Romagna close to the border with Tuscany will not be confused upon encountering the Tuscan phonological distribution in the linguistic input; they will still acquire an allophonic distribution of the two fricatives.

Second, while the laryngeal contrast in the Dutch-German dialect continuum changes both phonetically and phonologically, in the Italian dialect continuum only the distribution of voiced and voiceless fricatives changes. There are no differences in the phonological representation of the contrast or its phonetic realisation between the two regions. Italian speakers thus do not have to learn which phonological or phonetic contrast is present in their language, while Dutch/German speakers do.

Third, there are differences in the linguistic input speakers receive in the two areas. In the Dutch-German continuum there is an area where VOT values are bimodally distributed and hence do not facilitate the postulation of a phonological contrast. In the Italian continuum, on the other hand, all speakers will receive a bimodal distribution of voicing proportions, facilitating the postulation of a phonological contrast for all speakers. Italian speakers will thus only have to decide whether their dialect has a predictable distribution of the fricatives or not, but as discussed in Section 4.4.7, Van der Feest and Johnson (2016) have shown that children are perfectly capable at keeping apart the different varieties they receive in their input. The children in the Dutch-German continuum are unlikely to be able to apply this strategy to the laryngeal contrast as the data do not involve a neutralisation of the contrast but rather a different phonetic realisation of it.

The most important linguistic difference between the two systems, however, is the fact that the linguistic input from the two systems in the Dutch-German transition zone contains conflicts, while the input from the two Italian systems does not contain conflicts. As a result of the conflicting input in the Dutch-German continuum, the child is unable to postulate a laryngeal contrast, so that it ends up with an intermediate system containing only one laryngeal category; and the child is unable to postulate a constraint ranking for the phonetic realisation of this single laryngeal category, so that it ends up with a broad range of phonetic values it can use for that single category. As children in the Italian continuum do not find any conflicts in the linguistic input, they will not encounter any problems in acquiring their language.

Besides linguistic and geographical factors, political and socio-geographic factors also play a part in explaining the patterns. These will be discussed in the next section.

6.1.1 The influence of social and political geography

A potential non-linguistic difference between the two transition zones is the socio-political situation. The borders in the two regions are both old: the Dutch-German country border can be traced back to the early Middle Ages (De Vrankrijker (1946)) while the border between Tuscany and Emilia-Romagna is traced back to the empire of Augustus (Solari (1932)). The two Italian provinces used to be independent kingdoms, which were united in the kingdom of Italy after the unification of the country in 1861 (Britannica Academic (2017)). Although both regions have since long been separated by a strong political border, the Dutch-German continuum has historically behaved as a linguistic continuum (e.g. Goossens (1984)). This difference may be explained by the fact that the Dutch-German area is a linguistic unity while the Italian continuum is not: in Italy the political border corresponds to a linguistic border (the La Spezia-Rimini line) while the Dutch-German political border does not correspond to a linguistic border.

In Italy the division between the two areas is very clearly related to the political border between the two provinces: to the north of the provincial border ISV is consistently implemented, while to the south it is absent from the language. It is not immediately clear how this would work, as politics typically do not influence language (language is generally not influenced by policies but rather by the speakers themselves). It might be possible, however, that speakers are strongly oriented on their own political region, which influences their language. This assumption is in fact supported by the fact that no Tuscan speaker seemed to speak a clear dialect. Even speakers from Rivoreta, which is 10 kilometres away from Fanano (distance as the crow flies) do not have a clear dialect, while speakers in Fanano spoke a clear dialect. Furthermore, most speakers in Tuscany indicated to speak standard Italian instead of a regional variety of Italian (a believe which is strongly hold by Tuscan speakers in general).

The country border between the Netherlands and Germany seems to have a very different effect on the languages. At first sight the border does not appear to be visible at all, as the data form a clear, uninterrupted continuum: there is no clear difference between dialects to the west and dialects to the east of the border. A closer look, however, reveals that the start of the Dutch-German transition zone coincides with the national border, an effect that might be the result of standardisation. As the choice of standard language obviously depends on the country, a small effect of the border is visible. It is possible that the data from the Netherlands, which were recorded later than the German data, have undergone more standardisation than the German data (cf. Giesbers et al. (2005), Giesbers (2008)). If the German data had been recorded later more standardisation might have been visible, causing the transition between the two laryngeal systems to be more abrupt than it appears to be now. The onset of the transition zone might thus partly be explained by differences in recording dates between the two countries. Although a small effect of the state border, or rather the different standard languages, is visible, it remains to be explained why the influence of the Dutch-German border is so much smaller than the influence of the Italian province border.

The effects of the political borders are not attributable to the political border itself. The border might have an effect on people's movements, however, as it can influence the way people people perceive space (social constructs can shape spatial patterns; cf. Massey (1984)). The political border might, as mentioned above, influence speakers' orientation, so that their movements are oriented mostly within their province or country rather than across the border. This, in turn, influences the extent to which dialect contact occurs.

The presence of the abrupt linguistic boundary in Italy can, at least in part, be an effect of this orientation, as can the onset of the transition zone at the Dutch-German country border.

The difference between the two regions shows that the relationship between maps and grammars is a complicated one. Above all, the relationship is bidirectional: in the Dutch-German region the linguistic patterns found on the map are influenced by the grammar, while in the Italian region the grammar is influenced by the geography. Phonologically, the abrupt change in the Italian continuum can be explained by closely looking at the constraint ranking (the constraint ranking either causes the complete presence of ISV or the complete absence of it),[2] but the absence of a phonetically gradual continuum cannot be explained purely by linguistic factors. Especially for the fricatives it could be relatively easy to realise them with intermediate voicing values (i.e. values that are not clearly voiced but not clearly voiceless either, so that a phonetically gradual transition is visible). This is not the case, however, and cannot be excluded based on language contact as it is both impossible to fully exclude language contact and unlikely that there is no contact between the two regions. The socio-political situation can explain the absence: the socio-political map thus explains the linguistic patterns, or the grammar. In the Dutch-German continuum, on the other hand, a phonetically and phonologically gradual transition is possible. As it is not prohibited by the socio-political situation, the grammar influences the linguistic patterns on the map.

6.2 Conclusion

The data from the present study contribute several points to the discussion on linguistic variation, relating both to intralinguistic and extralinguistic factors. Extralinguistic factors can be divided into geographical and socio-political factors; only the latter factors will be discussed here, as the former have been shown to play merely a minor role.

The relevance of the socio-political situation becomes evident when comparing the two regions with each other. Although the political borders in both regions are very old, the Dutch-German continuum historically behaved as a linguistic continuum, not being separated by influential linguistic borders (Goossens (1984)[3]), while the Italian continuum was separated by the La

[2] Note, however, that it would be perfectly possible for speakers to have access to multiple grammars, one permitting ISV and the other prohibiting it, or to have a grammar with the two constraints relevant for ISV unranked with respect to each other.

[3] Goossens (1984) also mentions the existence of three different languages used in writing, viz. Dutch, high German and low German, the latter being used in the northern part of the region including a small region in the Netherlands. This also suggests the unity of the area. In the 16th century the west Germanic area switched to two writing languages (Dutch and high German) instead of three. In the contact areas, however, Dutch and high German were competing with each other until the 19th century (e.g. Dutch was used in church and at school, while German was the official administrative language).

Spezia-Rimini line. The linguistic border between the Netherlands and Germany only recently started playing a more important role, while the Italian linguistic border is very old (Von Wartburg (1936)). This is visible in the linguistic patterns: the Dutch-German continuum shows a much more gradual trend than the Italian continuum. The linguistic situation is likely to have influenced the social orientation of speakers. This orientation, subsequently, may affect language variation: speakers might adapt their language to the language of other speakers in the region they are oriented on. In both regions the political border is old, and thus unlikely to explain the differences between the regions.

In both regions the changing status of the border is visible. The onset of the transition between voicing and aspiration languages coincides with the border between the Netherlands and Germany (and thus with the border between the standard languages), indicating that the dialects are likely undergoing standardisation. In Tuscany younger speakers show more ISV than older speakers, although they do not apply it consistently. It is possible that, because of the smaller influence of the political border (which was reduced from a country border to a provincial border), and because of globalisation, the orientation of speakers is changing from a local, Tuscan orientation to a broader, Italian orientation. Based on this, it may be concluded that political borders may shape language use: an important border is likely to inhibit contact across this border, while a less important border will make contact across this border easier (all else being equal). An interesting follow-up study would be the nature of the border between voicing and aspiration languages in the Dutch-German dialect continuum nowadays: if the border has indeed such a large influence on language, the prediction is that the linguistic border will now be more abrupt and will coincide with the political border.

This leads us to a second point: language contact is not a sufficient condition for language variation. Language contact cannot be excluded for either region, but its effects are visible in the Dutch-German area and not in the Italian area. Whether or not language contact is a necessary condition for language variation is not discussed here.

Besides the socio-political effects some intralinguistic effects can also be identified, with the distribution of phonetic values being the most important one: as long as the input gives evidence for a phonological contrast, speakers will postulate this phonological contrast.

Several linguistic patterns can be identified in the variation in this study. First, the data unsurprisingly show that a phonological contrast is neither a necessary nor a sufficient requirement for phonetic variation: a phonological contrast is absent from the middle area in the Dutch-German continuum but there is phonetic variation; a phonological contrast is present in the entire Italian continuum but there is no phonetic variation.

Second, phonetic variation may be guided by the lexicon. The VOT values in the middle area of the Dutch-German continuum follow a pattern: plosives that are lenis in both a voicing and an aspiration system are lenis (i.e. realised with a negative or short-lag VOT) in the middle area, and plosives that are fortis in both a voicing and an aspiration system are fortis (i.e. realised with a short-lag or long-lag VOT) in the middle area. As there is no phonological characteristic that can make this distinction, the distinction must be lexical. In Italy the difference in voicing averages between younger and older Tuscan speakers can only be explained on lexical grounds as well: only in specific lexical items (i.e. where the fricative appears in stem-internal position or at an opaque morpheme boundary) can the fricative be realised with higher voicing percentages for younger speakers.

A final point relates to the hypothesis put forward by Thomason and Kaufman (1988), who argue that speakers are not only able to borrow surface structures in linguistic borrowing, but can also borrow deeper, underlying patterns. The data from the present study partly contradict this hypothesis. Speakers in the transition zone between voicing and aspiration languages can mimic either system on the surface (i.e. they can imitate the phonetic realisation of the contrast), but the underlying phonological representation of the contrast is absent from their system. Speakers cannot postulate a phonological contrast based on the linguistic input they receive, but they cannot borrow a phonological contrast from a neighbouring region either.

Comparing this to the Italian region, it appears to be easier to borrow the distribution of a phonological contrast: although older Tuscan speakers do not show any signs of ISV, younger Tuscan speakers have higher average voicing percentages and also realise more intervocalic fricatives with full voicing. They only do so in the contexts where ISV applies, and not in any other context. It thus appears easier to borrow a phonological rule than a phonological contrast.

6.3 Recommendations for future research

The present study has shown some very interesting patterns in two linguistic continua. As this study is relatively new in its approach of transitional phenomena (instead of many linguistic variables, only one is studied in great detail), its conclusions have to be supported by more material. Several follow-up studies can be recommended.

First, the databases used in the Dutch-German continuum are both relatively small. The German database provided only a very small amount of items suitable for the study, and especially the intervocalic context suffered from that. Further, the data from this continuum are rather old. The study would certainly benefit from a second run, in which more speakers per location are interviewed and, most importantly, in which all speakers are presented the same questionnaire and in which many more items per phoneme per context are elicited.

Second, the pattern in the Italian continuum appears to indicate the influence of politics. As the topic of study, intervocalic /s/-voicing, does not provide a context in which a phonologically gradual change is the only possibility, it would be interesting to study the transition of several different linguistic phenomena in the same region. Ideally these phenomena would create a grammatical situation in which an abrupt change is highly unlikely, so that the effects of the political, geographical and geological situation can be tested (an abrupt change, however, can never be excluded as a possibility, as there is always the possibility of the absence of language contact or of speakers' orientation on a specific linguistic region). If these phenomena show a similar, abrupt change between the two provinces, this adds to the evidence of the strong effects of the border. If, on the other hand, some of these phenomena show a gradual change, the linguistic characteristics of ISV will have to be considered in the explanation of the patterns from the present study.

The behaviour of younger speakers in Tuscany might also be further examined. It has been hypothesised that these speakers show evidence of ISV, but the present study does not fully confirm that. As previous research focussed mostly on bigger cities, it could well be possible that this change is limited to younger speakers in these locations, while younger speakers in the smaller villages included in this study are following only later.

A final point to be made concerns the focus point of the present study, which only includes phonetics and phonology. Morphology, syntax and semantics are left out of this study, but would certainly have to be studied to get a complete picture of linguistic patterns in transition zones (see e.g. Barbiers, Van Koppen, Bennis and Corver (2016) for an example).

APPENDIX A

Locations in the Dutch-German dialect continuum

The Netherlands

Arum
Barger-Oosterveld
Bellingwolde
Blokzijl
Coevorden
Eelde
Finsterwolde
Franeker
Giethoorn
Groningen
Grouw
Harkema Opeinde
Harlingen
Havelte
Hollandsche Veld
Koekange
Kuinre
Leeuwarden
Marum
Meppel
Midwolde
Nieuw-Schoonebeek
Oldemarkt
Oude Pekela
Oudega
Roderwolde
Rottevalle
Ruinen
Scheemda
Schoonebeek
Slagharen
Slochteren
Spannum
Steenwijk
Tietjerk
Veendam
Vollenhove
Wagenborgen
Zuid-Sleen
Zwinderen

Germany

- Achternholt
- Apen
- Aschendorf
- Bad Zwischenahn
- Bardenfleth
- Barnstorf
- Barssel
- Blockwinkel
- Bockhorn
- Bremen
- Bunde
- Dedendorf
- Dreeke
- Dreibergen
- Drögennindorf
- Eckfleth
- Elmendorf
- Ermke
- Eversen
- Filsum
- Godensholt
- Gristede
- Hellwege
- Herssum
- Holtorf
- Hude
- Jeddeloh II
- Kayhausen
- Kleibrok
- Lähden
- Lathen
- Lauenbrück
- Leer
- Lohne
- Lüneburg
- Mallinghausen
- Meppen
- Neuenkruge
- Neuenmarhorst
- Ocholt
- Oldenburg
- Osteressen
- Osterholz-Scharmbeck
- Ostertimke
- Otersen
- Ottersberg
- Petersfehn
- Putensen
- Reppenstedt
- Saterland
- Schessinghausen
- Schneverdingen
- Schwanewede
- Schweghaus
- Stederdorf
- Stöcken
- Stukenborg
- Süddorf
- Tange
- Vechta
- Warpe
- Westerscheps
- Wiefelstede
- Winkeldorf

APPENDIX B

Wenker sentences

The English translations of the Wenker sentences are taken from `http://staff-www.uni-marburg.de/~naeser/wenker-e.htm`

1. Im Winter fliegen die trocknen Blätter durch die Luft herum.
   In winter the dry leaves are flying around in the air.
2. Es hört gleich auf zu schneien, dann wird das Wetter wieder besser.
   It will stop snowing soon, then the weather will get better.
3. Thu Kohlen in den Ofen, daß die Milch bald an zu kochen fängt.
   Put coals in the oven (so) that the milk will start boiling soon.
4. Der gute alte Mann ist mit dem Pferde durch's Eis gebrochen und in das kalte Wasser gefallen.
   The good old man has broken through the ice with his horse and has fallen into the cold water.
5. Er ist vor vier oder sechs Wochen gestorben.
   He died four or six weeks ago.
6. Das Feuer war zu stark/heiß, die Kuchen sind ja unten ganz schwarz gebrannt.
   The fire was too hot, the cakes are burned all black underneath.
7. Er ißt die Eier immer ohne Salz und Pfeffer.
   He always eats the eggs without salt and pepper.
8. Die Füße thun mir sehr weh, ich glaube, ich habe sie durchgelaufen.
   My feet hurt so much, I believe I've walked them sore.
9. Ich bin bei der Frau gewesen und habe es ihr gesagt, und sie sagte, sie wollte es auch ihrer Tochter sagen.
   I was with the woman and said it to her, and she said she wanted to say

it to her daughter.

10. Ich will es auch nicht mehr wieder thun!
    I do not want to do it again.
11. Ich schlage Dich gleich mit dem Kochlöffel um die Ohren, Du Affe!
    I'll let you have it about the head with the ladle, you monkey.
12. Wo gehst Du hin? Sollen wir mit Dir gehn?
    Where are you going? Should we go along with you?
13. Es sind schlechte Zeiten.
    These are bad times / the times are bad.
14. Mein liebes Kind, bleib hier unten stehn, die bösen Gänse beißen Dich todt.
    My dear child, keep standing down here, (otherwise) the bad geese will bite you dead.
15. Du hast heute am meisten gelernt und bist artig gewesen, Du darfst früher nach Hause gehn als die Andern.
    You learned the most today and you have been well behaved; you may go home earlier than the others.
16. Du bist noch nicht groß genug, um eine Flasche Wein auszutrinken, Du mußt erst noch ein Ende/etwas wachsen und größer werden.
    You aren't tall enough yet to empty a bottle of wine; you have to grow a bit and get taller first.
17. Geh, sei so gut und sag Deiner Schwester, sie sollte die Kleider für eure Mutter fertig nähen und mit der Bürste rein machen.
    Go, be so good and tell your sister she sould finish sewing the clothes for your mother and clean them with the brush.
18. Hättest Du ihn gekannt! dann wäre es anders gekommen, und es thäte besser um ihn stehen.
    If you had known him it would have been different and he would be better off.
19. Wer hat mir meinen Korb mit Fleisch gestohlen?
    Who stole my basket with meat?
20. Er that so, als hätten sie ihn zum dreschen bestellt; sie haben es aber selbst gethan.
    He acted as if they had ordered him to (come to) thresh.
21. Wem hat er die neue Geschichte erzählt?
    Whom did he tell the new story?
22. Man muß laut schreien, sonst versteht er uns nicht.
    One has to shout loudly, otherwise he doesn't understand us.
23. Wir sind müde und haben Durst.
    We are tired and thirsty.
24. Als wir gestern Abend zurück kamen, da lagen die Andern schon zu Bett und waren fest am schlafen.
    When we returned last night the others were already lying in bed and

were fast asleep.

25. Der Schnee ist diese Nacht bei uns liegen geblieben, aber heute Morgen ist er geschmolzen.
    Last night the snow kept lying on the ground but this morning it melted.
26. Hinter unserm Hause stehen drei schöne Apfelbäumchen mit rothen Aepfelchen.
    Behind our house there are three beautiful apple-trees with small red apples.
27. Könnt ihr nicht noch ein Augenblickchen auf uns warten, dann gehn wir mit euch.
    Can't you wait just one more moment for us? Then we'll go with you.
28. Ihr dürft nicht solche Kindereien treiben!
    You must not do such childish things.
29. Unsere Berge sind nicht sehr hoch, die euren sind viel höher.
    Our mountains aren' t very high; yours are much higher.
30. Wieviel Pfund Wurst und wieviel Brod wollt ihr haben?
    How many pounds of sausage and how much bread do you want?
31. Ich verstehe euch nicht, ihr müßt ein bißchen lauter sprechen.
    I don't understand you; you must speak a bit louder.
32. Habt ihr kein Stückchen weiße Seife für mich auf meinem Tische gefunden?
    Haven't you found a small piece of white soap for me on the table?
33. Sein Bruder will sich zwei schöne neue Häuser in eurem Garten bauen.
    His brother wants to build for himself two beautiful new houses in your garden.
34. Das Wort kam ihm von Herzen!
    That word came from his heart.
35. Das war recht von ihnen!
    That was right of them.
36. Was sitzen da für Vögelchen oben auf dem Mäuerchen?
    What kind of little birds are sitting up there on the little wall?
37. Die Bauern hatten fünf Ochsen und neun Kühe und zwölf Schäfchen vor das Dorf gebracht, die wollten sie verkaufen.
    The farmers had brought five oxen and nine cows and twelve little sheep to the village; they wanted to sell them.
38. Die Leute sind heute alle draußen auf dem Felde und mähen/hauen.
    The people are all out in the field mowing.
39. Geh nur, der braune Hund thut Dir nichts.
    Go ahead, the brown dog won't harm you.
40. Ich bin mit den Leuten da hinten über die Wiese ins Korn gefahren.
    Together with the people I drove over the meadow back there into the grain.

APPENDIX C

Locations in the Italian dialect continuum

Emilia-Romagna

- Novellara
- Correggio
- Casalgrande
- Prignano sulla Secchia
- Lama Mocogno
- Sestola
- Fanano

Toscana

- Rivoreta
- Popiglio
- Pescia
- Chiesina Uzzanese
- Staffoli
- San Pietro Belvedere
- Lajatico
- Montecatini Val di Cecina
- Querceto
- Monteverdi Marittimo

APPENDIX D

Questionnaire

1. La pizza è un piatto molto amato in tutto il mondo.
2. I ragazzi stanno risalendo la montagna.
3. Il coniglio è molto tenero.
4. Ieri ho fuso l'anello di fidanzamento.
5. Il papa è morto; il conclave voterà per un nuovo papa.
6. Nel giardino botanico c'è un cactus.
7. Il cane è al canile.
8. Il figlio di Bernadette e Carlo ha già otto anni.
9. Sono andato/a alla cassa per pagare l'ananas.
10. La domenica le campane risuonano in tutte le case del paese.
11. L'olio motore è unto e bisunto.
12. Non lo voglio fare così.
13. George Clooney è il mio attore preferito.
14. Ho un cattivo presentimento.
15. La torre di Pisa è famosa.

16. Mia madre è francese, ma mio padre è italiano.
17. La pietra è erosa/la pietra è stata erosa dal vento.
18. La casa è grande.
19. La casa ha un soffitto basso.
20. Sai che la madre di Paolo è malata?
21. La nonna di Alberto abita a Firenze.
22. Non ho mangiato niente oggi, ho molta fame.
23. Uomini e donne spesso subiscono trattamenti disuguali.
24. Ho comprato un nuovo cappello per il matrimonio di Emma e Luigi.
25. Lo zio di Anna non ha figli.
26. I ragazzi non vogliono andare dal medico.
27. Non mi ricordo niente di ieri, ho bevuto troppo.
28. Davanti al macellaio c'è un negozio di abbigliamento.
29. Andiamo al cinema, c'è un nuovo film con Monica Bellucci.
30. I dialetti italiani hanno molte peculiarità.
31. Il mio telefono è caduto. Devo comprarne uno nuovo.
32. Per la chiesa, gira a destra e prosegui sempre dritto.
33. Barbara non è una buona cuoca, il cibo è proprio immangiabile.
34. Solamente i bambini possono entrare gratis.
35. La "Primavera" di Botticelli è mozzafiato.
36. Le rose profumano molto.
37. La nuova canzone di Mina non mi piace.
38. Queste cose sono molte belle ma costose.
39. La polizia è arrivata troppo tardi, il ladro era fuggito.
40. Sul cartello c'è scritto "vendesi".
41. Sono molto stanco/a.
42. La bella addormentata si è punta con il fuso.

43. Il mio amico è un po' asociale.
44. L'ossigeno è un gas.
45. Le canzoni sono registrate sulle cassette.
46. Il tuo compleanno è il 1 gennaio.
47. Il cane abbaia forte, è impaurito.
48. Gianni e Maria litigano. Gianni grida a Maria che è infantile.
49. Marco e Maria hanno cotto sei pizze per uno. Le sei di Marco sono deliziose.
50. Mio fratello è piu alto di me.
51. La Resistenza è stata un periodo molto importante per la storia italiana.
52. Prendi un caffè o un tè?
53. Giacomo ha un naso enorme.
54. Ho perso il mio lapis.
55. Il riso è un piatto cinese.
56. Il mio bisnonno si chiamava Massimo.
57. Andrea non è un ragazzo disonesto ed è sempre molto cortese.
58. A luglio vado in Portogallo.
59. La festa era meravigliosa.
60. Marina pensa che una sua amica sia bisessuale.
61. I genitori parlano in dialetto ogni giorno.
62. Il governo ha molti debiti, protrebbe andare in bancarotta.
63. Non mi piacciono molto le banane.
64. L'albergo nuovo è formidabile!
65. Credere che la moda piaccia a tutte le donne è una presupposizione sbagliata.
66. La porta è chiusa.
67. Ieri mi è capitata una terribile disavventura.
68. La polizia ha sbarrato la strada.

1. Finlandese
2. Peruviano
3. Nipponico
4. Brasiliano
5. Lettone
6. Lituano
7. Spagnolo
8. Ungherese
9. Australiano
10. Tedesco
11. Boliviano
12. Francese
13. Norvegese
14. Messicano
15. Giapponese
16. Argentino
17. Olandese
18. Danese
19. Austriaco
20. Polacco
21. Estone
22. Americano
23. Italiano
24. Portoghese
25. Vietnamita
26. Gallese
27. Islandese

28. Svedese
29. Cinese
30. Inglese
31. Russo
32. Neozelandese
33. Irlandese

APPENDIX E

Informed consent

MODULO DI CONSENSO INFORMATO SCRITTO
LEIDEN UNIVERSITY CENTRE FOR LINGUISTICS

TITOLO DELLA RICERCA:

____

COGNOME E NOME DEL SUPERVISORE:

____

COGNOME E NOME DELLO STUDENTE RICERCATORE:

____

Io sottoscritto____

Dichiaro di:

- partecipare volontariamente alla ricerca della studentessa ricercatrice Nina Ouddeken ____;
- aver ricevuto dalla ricercatrice su menzionato tutte le informazioni chiare ed esaurienti sulle finalità e le procedure della ricerca a cui mi è stato chiesto di prendere parte;
- aver letto e compreso il foglio di informazioni che mi è stato consegnato e che conferma quanto mi è stato verbalmente detto;
- aver avuto l’opportunità di porre domande chiarificatrici e di aver avuto risposte soddisfacenti sui particolari dello studio

Pertanto acconsento liberamente alla partecipazione alla ricerca condotta dalla signora Nina Ouddeken.
La firma su questo modulo non verrà ad incidere sui miei diritti legali (del mio tutelato…).
Letto e approvato (scritto a mano)

FIRMA DEL/DELLA PARTECIPANTE ALLA RICERCA SPERIMENTALE

____

LUOGO E DATA

APPENDIX F

Procedure information

Informazioni per partecipanti

Gentile Signore e Signori,

Lei sta participando ad uno studio scientifico sull'uso quotidiano dell'Italiano in Emilia-Romagna e Toscana. La sua partecipazione è volontaria, ma tuttavia abbiamo bisogno del suo consenso. Prima di dare il suo consenso, può trovare maggiori informazioni sulla ricerca in questo modulo. Le preghiamo di leggerlo e di domandare al ricercatore nel caso in cui avesse ulteriori domande.

La ricerca è strutturata in due parti diverse. La prima parte consiste in una conversazione tra due o più partecipanti; nella seconda parte Le viene chiesto di tradurre una lista di frasi. Il ricercatore effettuerà registrazioni audio di entrambe queste parti, in modo da poterle analizzare successivamente. Prima della ricerca il ricercatore Le chiederà di compilare un modulo con qualche domanda relativa alle sue abitudine linguistiche.

Tutti i dati raccolti saranno conservati ed elaborati in forma anonima (il nome è noto solo al ricercatore). I dati raccolti saranno utilizzati (in modo anonimo) nelle presentazioni e pubblicazioni del ricercatore. Le registrazioni audio saranno conservate permanentemente dal Meertens Instituut di Amsterdam (Paesi Bassi), il quale rende disponibile la sua banca dati anche ad altri ricercatori. È importante sapere che tutti i dati sono protetti dalla "Wet Bescherming Persoonsgegevens" (Legge per la Protezione dei Dati Personale) olandese. Questo significa che i suoi dati saranno conservati e utilizzati in forma anonima. Solamente il ricercatore conosce il suo nome, che tuttavia non sarà utilizzato in presentazioni e pubblicazioni.

Se lo desidera, il ricercatore La può informare sullo scopo della ricerca. Il ricercatore La può anche tenere aggiornato/a sul progresso della ricerca. La preghiamo di informare il ricercatore in caso positivo.

Importante
Se lo desidera, può terminare l'intervista in ogni momento per qualsiasi motivo (non è necessario dare giustificazioni al ricercatore e la rinuncia non ha alcuna conseguenza negativa per Lei). In questo caso l'intervista si concluderà immediatamente e il ricercatore non farà ulteriori domande.

Se avesse delle domande sulla stuttura dell'intervista, Le chiediamo di porle prima dell'inizio dell'intervista, ma se nuove domande dovessero sorgere durante l'intervista, può interrompere l'intervista in ogni momento.

APPENDIX G

Personal information

Ulteriori informazioni

Sesso:	M	F
Età:	………………………………………………	
Residenza:	………………………………………………	
Luogo di nascita:	………………………………………………	
È cresciuto/a parlando dialetto:	S	N
Parla dialetto:	S	N
Famigliari	S	N
Amici	S	N
Vicini	S	N
Colleghi	S	N
Parla dialetto quotidianamente:	S	N
Parla altre lingue:	S	N
Lingua	Livello[1]	

1.

2.

3.

4.

5.

[1] Basico, intermedio, avanzato, fluente

References

Abramson, A. S. & Lisker, L. (1965). Voice onset time in stop consonants: Acoustic analysis and synthesis. In D. Commins (Ed.), *Proceedings of the 5th international congress on acoustics.* Liège.

Abramson, A. S. & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the 6th international congress of phonetic sciences, prague, 1967* (pp. 569–573). Prague: Publishing House of the Czechoslovak Academy of Sciences.

Anderson, J. M. & Ewen, C. J. (1987). *Principles of Dependency Phonology.* Cambridge: Cambridge University Press.

Anttila, A. (1997). Deriving variation from grammar: A study of Finnish genitives. In F. Hinskens, R. V. Hout & L. Wetzels (Eds.), *Variation, change and phonological theory* (pp. 35–68). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory*, *20*(1), 1–42.

Anttila, A. (2008, dec). Derived environment effects in colloquial Helsinki Finnish. In K. Hanson & S. Inkelas (Eds.), *The nature of the word: Studies in honor of paul kiparsky.* Cambridge, MA/London: The MIT Press.

Anttila, A. & Cho, Y.-m. Y. (1998). Variation and change in Optimality Theory. *Lingua*, *104*, 31–56.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L. & Perey, A. J. (1981). Discrimination of Voice Onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, *52*, 1135–1145.

Auer, P. & Hinskens, F. (1996). The convergence and divergence of dialects in Europe - New and not so new developments in an old area. *Sociolinguistica*, *10*, 1–30.

Avery, J. P. (1996). *The representation of voicing contrasts* (PhD Thesis). University of Toronto.

Avery, J. P. & Idsardi, W. J. (2001). Laryngeal dimensions, completion and enhancement. In T. A. Hall (Ed.), *Distinctive feature theory* (pp. 41–70). Berlin: Mouton de Gruyter.

Backley, P. (2011). *Element Theory.* Edinburgh: Edinburgh University Press.

Backley, P. & Nasukawa, K. (2010). Consonant-vowel unity in Element Theory. *Phonological Studies*, *13*, 21–28.

Barbiers, S., Bennis, H., De Vogelaer, G., Devos, M. & Van der Ham, M. (2005). *Syntactische Atlas van de Nederlandse Dialecten, Deel I.* Amsterdam: Amsterdam University Press.

Barbiers, S., Van der Auwera, J., Bennis, H., Boef, E., De Vogelaer, G. & Van der Ham, M. (2008). *Syntactische Atlas van de Nederlandse Dialecten, Deel II.* Amsterdam: Amsterdam University Press.

Barbiers, S., Van Koppen, M., Bennis, H. & Corver, N. (2016). MIcrocomparative MOrphosyntactic REsearch (MIMORE): Mapping partial grammars of Flemish, Brabantish and Dutch. *Lingua*, *178*, 5–31. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0024384115002211`

Baroni, M. (1997). *The representation of prefixed forms in the Italian lexicon: Evidence of intervocalic [s] and [z]* (MA Thesis). UCLA.

Baroni, M. (1998). *The phonetic nature of the Northern Italian allophones [s] and [z] in words with variable realization: Electroglottographic and acoustic evidence.*

Beckman, J., Helgason, P., McMurray, B. & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics*, *39*, 39–49.

Beckman, J., Jessen, M. & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs . true voice languages. *Journal of Linguistics*, *49*(2), 259–284.

Benincà, P., Parry, M. & Pescarini, D. (2016, jun). The dialects of northern Italy. In A. Ledgeway & M. Maiden (Eds.), *The oxford guide to the romance languages* (pp. 185–205). Oxford: Oxford University Press. Retrieved from `http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199677108.001.0001/acprof-9780199677108-chapter-13`

Berns, J. (2002). *Taal in stad en land: Zuid-Gelderse dialecten.* Den Haag: Sdu Uitgevers.

Bertinetto, P. M. (1999). Boundary strength and linguistic ecology (mostly exemplified on intervocalic /s/-voicing in Italian). *Folia Linguistica*, *33*, 267–286.

Bertinetto, P. M. & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, *35*(02), 131–151.

Blancquaert, E. & Pée, W. (n.d.). *Reeks Nederlandse dialectatlassen (16 volumes).* Antwerpen: De Sikkel.

Blevins, J. (2004). *Evolutionary Phonology.* Cambridge: Cambridge University

Press.

Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics*, *32*(2), 117–166.

Bloemhoff, H., De Haan, G. J. & Versloot, A. P. (2013). Language varieties in the province of Fryslân. In F. Hinskens & J. Taeldeman (Eds.), *Language and space - dutch* (pp. 721–738). Berlin/Boston: Mouton de Gruyter.

Bloemhoff, H., der Kooi, J. V., Niebaum, H. & Reker, S. (Eds.). (2008). *Handboek Nedersaksische Taal- en Letterkunde.* Assen: Koninklijke Van Gorcum BV.

Bloomfield, L. (1933). *Language.* New York: Henry Holt.

Blumstein, S. E. (1991). The relation between phonetics and phonology. *Phonetica*, *48*, 108–119.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives* (PhD Thesis). Universiteit van Amsterdam.

Boersma, P. (1999). Optimality-theoretic learning in the Praat program. In *Institute of phonetic sciences proceedings 23* (pp. 17–35).

Boersma, P. (2009). Cue constraints and their interactions in phonological perception and production. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 55–110). Berlin: Mouton de Gruyter.

Boersma, P. (2011). A programma for bidirectional phonology and phonetics and their acquisition and evolution. In A. Benz & J. Mattausch (Eds.), *Bidirectional optimality theory* (pp. 33–72). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Boersma, P. & Weenink, D. (2016). *Praat.*

Braun, A. (1996). Zur regionalen Distribution von VOT im Deutschen. In A. Braun (Ed.), *Untersuchungen zur stimme und sprache* (Vol. 96, pp. 19–32). Stuttgart: Frans Steiner Verlag.

Braunschweiler, N. (1997). Integrated cues of voicing and vowel length in German: A production study. *Language and Speech1*, *40*(4), 353–376.

Britain, D. (2004). Geolinguistics - Diffusion of language. In U. Ammon, N. Dittmar, K. Mattheier & P. Trudgill (Eds.), *Sociolinguistics: International handbook of the science of language and society* (pp. 34–48). Berlin: Mouton de Gruyter.

Britannica Academic. (2017). *Risorgimento.* Retrieved from `academic.eb.com.ezproxy.leidenuniv.nl:2048/levels/collegiate/article/Risorgimento/63773`

Browman, C. P. & Goldstein, L. M. (1991). Gestural structures: Distinctiveness, phonological processes, and historical change. In I. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 313–338). Hillsdale, N.J.: Lawrence Erlbaum.

Browman, C. P. & Goldstein, L. M. (1992). Articulatory Phonology: An overview. *Haskins Laboratories Status Report on Speech Research*, *111/112*, 23–42.

Bybee, J. (1999). Usage based phonology. In M. Darnell, E. Moravcsik,

F. Newmeyer, M. Noonan & K. Wheatley (Eds.), *Functionalism and formalism in linguistics, volume i: General papers* (pp. 211–242). Amsterdam: John Benjamins Publishing Company.

Caramazza, A. & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, *2*, 239–245.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B. & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, *54*(2), 421–428.

Cedergren, H. J. & Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language*, *50*(2), 333–355.

Chambers, J. K. & Trudgill, P. (1980). *Dialectology.* Cambridge: Cambridge University Press.

Cho, Y.-m. Y. (1990). *Parameters of consonantal assimilation* (PhD Thesis). Stanford University.

Chomsky, N. (1957). *Syntactic structures.* Berlin/New York: Mouton de Gruyter.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York, Evanston, and London: Harper & Row, Publishers.

Coetzee, A. W. (2004). *What it means to be a loser: Non-optimal candidates in Optimality Theory* (PhD Thesis). University of Massachusetts - Amherst.

Coetzee, A. W. (2006). Variation as accessing non-optimal' candidates. *Phonology*, *23*, 337–385.

Cornips, L. (1998). Syntactic variation, parameters, and social distribution. *Language Variation and Change*, *20*, 1–21.

Cravens, T. D. (2002). *Comparative historical dialectology: Italo-Romance clues to Italo-Romance sound change.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Cyran, E. (2011). Laryngeal realism and laryngeal relativism: Two voicing systems in Polish? *Studies in Polish Linguistics*, *6*, 45–80.

Cyran, E. (2013). *Polish voicing - Between phonology and phonetics.* Lublin: Wydawnictwo KUL.

Daan, J. & Blok, D. P. (1969). *Van Randstad tot Landrand.* Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.

Davis, A. L. & McDavid, R. I. (1950). Northwestern Ohio: A transition area. *Language*, *26*(2), 264–273.

Davis, G. W. (2008, sep). Toward a progression theory of the Old High German Consonant Shift. *Journal of Germanic Linguistics*, *20*(03), 197–241. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S147054270800007X`

Davis, S. (2011). Geminates. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (Eds.), *The blackwell companion to phonology* (pp. 873–897). Oxford: Wiley Blackwell.

De Schutter, G., Van den Berg, B., Goeman, T. & De Jong, T. (2005). *Morphological Atlas of the Dutch Dialects - Volume I.* Amsterdam: Amsterdam

University Press.
De Vrankrijker, A. C. J. (1946). *De grenzen van Nederland.* Amsterdam: Contact.
De Wulf, C., Goossens, J. & Taeldeman, J. (2005). *Phonological Atlas of the Dutch Dialects - Volume IV.* Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde (KANTL).
Dediu, D. & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences*, *104*(26), 10944–10949. Retrieved from `http://www.pnas.org/cgi/doi/10.1073/pnas.0610848104`
Eilers, R. E., Gavin, W. & Wilson, W. R. (1979). Linguistic experience and phonemic perception in infancy: A cross-linguistic study. *Child Development*, *50*, 14–18.
Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.
Ethnologue. (2005). *Ethnologue: Languages of the world.* Retrieved 2017-10-05, from `https://www.ethnologue.com/language/afr`
Everett, C. (2013). Evidence for Direct Geographic Influences on Linguistic Sounds: The Case of Ejectives. *PLoS ONE*, *8*(6), e65275. Retrieved from `http://dx.plos.org/10.1371/journal.pone.0065275`
Everett, C., Blasi, D. E. & Roberts, S. G. (2015). Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences*, *112*(5), 1322–1327. Retrieved from `http://www.pnas.org/lookup/doi/10.1073/pnas.1417413112`
Feldman, N. H., Griffiths, T. L., Goldwater, S. & Morgan, J. L. (2013). A role for the developing lexicon in phonetic category acquisition. *Psychological Review*, *120*(4), 751–778.
Floccia, C., Delle Luche, C., Durrant, S., Butler, J. & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, *124*(1), 95–100. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0010027712000613`
Frau, G. (1984). *Profilo dei dialetti italiani - Friuli* (M. Cortelazzo, Ed.). Pisa: Pacini Editore.
Galli de' Paratesi, N. (1984). *Lingua toscana in bocca ambrosiana. Tendenze verso l'italiano standard: Un'inchiesta sociolinguistica.* Bologna: Il Mulino.
Gally, S. (2015, may). Variation linguistique perçue : quelques localités du Piémont occidental (Italie). Questions de méthodologie en dialectologie perceptuelle. *Corela*, *HS-16*(HS-16). Retrieved from `http://journals.openedition.org/corela/4013`
Giannelli, L. (1976). *Profilo dei dialetti italiani - Toscana* (M. Cortelazzo, Ed.). Pisa: Pacini Editore.
Giegerich, H. J. (1992). *English Phonology - An introduction.* Cambridge:

Cambridge University Press.

Giesbers, C. H. E. (2008). *Dialecten op de grens van twee talen: een dialectologisch en sociolinguïstisch onderzoek in het Kleverlands dialectgebied* (PhD Thesis). Radboud Universiteit Nijmegen.

Giesbers, C. H. E., Van Hout, R. & Van Bezooijen, R. (2005). Dialect op de grens van twee talen. *Taal en Tongval*, *57*, 61–88.

Goebl, H. (2007). Dialektometrische Streifzüge durch da Netz des Sprachatlasses AIS. *Ladinia*, *31*, 187–271.

Goebl, H. (2008). La dialettometrizzazione integrale dell'AIS. Presentazione dei primi risultati. *Revue de linguistique Romane*, *72*, 25–113.

Goeman, T., Van Oostendorp, M., Van Reenen, P., Koornwinder, O., Van den Berg, B. & Van Reenen, A. (2008). *Morphological Atlas of the Dutch Dialects - Volume II.* Amsterdam: Amsterdam University Press.

Goldstein, L. M. & Fowler, C. (2003). Articulatory Phonology: A phonology for public language use. In *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). New York: Mouton de Gruyter.

Goldstone, R. L. & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69–78. Retrieved from `http://doi.wiley.com/10.1002/wcs.26`

Goossens, J. (1965). Die Gliederung des Südniederfränkischen. *Rheinische Vierteljahrsblätter*, *30*, 79–94.

Goossens, J. (1984). Die Herausbildung der deutsch-niederländischen Sprachgrenze. *Mededelingen van de Vereniging voor Limburgse Dialect- en Naamkunde*, *29*.

Goossens, J., Taeldeman, J. & Verleyen, G. (1998). *Phonological Atlas of the Dutch Dialects - Volume I.* Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde (KANTL).

Goossens, J., Taeldeman, J. & Verleyen, G. (2000). *Phonological Atlas of the Dutch Dialects - Volumes II + III.* Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde (KANTL).

Gordon, M. & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*(4), 383–406.

Hale, M. & Reiss, C. (2000). Phonology as cognition. In N. Burton-Roberts, P. Carr & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical studies* (pp. 161–184). Oxford: Oxford University Press.

Hale, M. & Reiss, C. (2008). *The phonological enterprise.* Oxford: Oxford University Press.

Halle, M. & Stevens, K. N. (1971). *A note on laryngeal features* (Tech. Rep.). Cambridge, MA: Massachusetts Institute of Technology.

Hamann, S. (2009). The learner of a perception grammar as a source of sound change. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 111–149). Berlin: Mouton de Gruyter.

Hamann, S. (2011). The phonetics-phonology interface. In N. Kula, B. Botma & K. Nasukawa (Eds.), *Continuum companion to phonology* (pp. 202–

224). London: Continuum.

Hamann, S. (2014). Phonological change. In C. Bowern & B. Evans (Eds.), *Handbook of historical linguistics* (pp. 249–263). London/New York: Routledge.

Hamann, S. & Seinhorst, K. (2016). *Prevoicing in Standard German plosives: Implications for phonological representations?* (Tech. Rep.). Budapest: 13th Old World Conference in Phonology.

Harris, J. & Lindsey, G. (1995). The elements of phonological representation. In J. Durand & F. Katamba (Eds.), *Frontiers of phonology: atoms, structures, derivations* (pp. 34–79). Harlow, Essex: Longman.

Harris, M. & Vincent, N. (1988). *The Romance languages.* London/New York: Routledge.

Hartshorne, J. K., Tenenbaum, J. B. & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, *20*(2), 253–306.

Heeringa, W. J. (2004). *Measuring dialect pronunciation differences using Levenshtein distance* (PhD Thesis). Rijksuniversiteit Groningen.

Heeringa, W. J., Johnson, K. & Gooskens, C. (2009, feb). Measuring Norwegian dialect distances using acoustic features. *Speech Communication*, *51*(2), 167–183. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0167639308001258`

Heeringa, W. J. & Nerbonne, J. A. (2001). Dialect areas and dialect continua. *Language Variation and Change*, *13*, 375–400.

Heeringa, W. J., Nerbonne, J. A., Niebaum, H., Nieuweboer, R. & Kleiweg, P. (2000). Dutch-German contact in and around Bentheim. In D. Gilbers, J. Nerbonne & J. Schaeken (Eds.), *Languages in contact* (pp. 145–156). Amsterdam/Atlanta GA: Rodopi.

Helgason, P. & Ringen, C. (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics*, *36*, 607–628.

Hermans, B. (2013). Phonological features of Limburgian dialects. In F. Hinskens & J. Taeldeman (Eds.), *Language and space - dutch* (pp. 336–355). Berlin/Boston: Mouton de Gruyter.

Hinskens, F. (1998). Dialect levelling: A two-dimensional process. *Folia Linguistica*, *32*(1-2).

Hinskens, F. & Taeldeman, J. (Eds.). (2013). *Language and Space - Dutch.* Berlin/Boston: De Gruyter Mouton.

Hoff, E. (2009). *Language development.* Belmont: Wadsworth.

Honeybone, P. (2002). *German obstruent lenition: Some mutual implications of theoretical and historical phonology* (PhD Thesis). University of Newcastle upon Tyne.

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. M. Christie (Ed.), *Current progress in historical linguistics* (pp. 96–105). Amsterdam: NorthHolland.

Hoppenbrouwers, C. & Hoppenbrouwers, G. (2001). *De indeling van de Nederlandse streektalen*. Assen: Koninklijke Van Gorcum BV.

Hunnicutt, L. & Morris, P. A. (2016). Prevoicing and aspiration in Southern American English. In *Proceedings of the 39th annual penn linguistics conference* (pp. 215–224). Pennsylvania.

Iverson, G. K. & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*(3), 369–396.

Iverson, G. K. & Salmons, J. C. (1999). Glottal spreading bias in Germanic. *Linguistische Berichte*, *178*, 135–151.

Iverson, G. K. & Salmons, J. C. (2003). Laryngeal enhancement in early Germanic. *Phonology*, *20*(1), 43–74. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S0952675703004469`

Jakobson, R., Fant, C. G. M. & Halle, M. (1952). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: The MIT Press.

Jansen, M., Marynissen, A., Van Oostendorp, M., Van Reenen, P., Van Reenen, A. & Stroop, J. (2011). *Dialectatlas van het Nederlands* (N. V. der Sijs, Ed.). Amsterdam: Uitgeverij Bert Bakker.

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: A Laboratory Phonology approach to English, Hungarian, and Dutch* (PhD Thesis). Rijksuniversiteit Groningen.

Jansen, W. (2007). Dutch Regressive Voicing Assimilation as a 'Low Level Phonetic Process'. In J. V. de Weijer & E. J. V. der Torre (Eds.), *Voicing in dutch - (de)voicing - phonology, phonetics, and psycholinguistics* (pp. 125–151). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Jellinghaus, H. F. (1892). *Die Niederländischen Volksmundarten: Nach den Aufzeichnungen der Niederländer*. Norden/Leipzig: Diedr. Soltau's Verlag.

Jessen, M. (1998). *Phonetics and phonology of the tense and lax obstruents in German* (PhD Thesis). Cornell University.

Jessen, M. & Ringen, C. (2002, aug). Laryngeal features in German. *Phonology*, *19*(2), 189–218. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S0952675702004311`

Johnson, K. (1996). Speech perception without speaker normalization. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego: Academic Press.

Kabatek, J., Pusch, C. D., Kortmann, B. & Van der Auwera, J. (2011). The Romance languages. In B. Kortmann & J. V. der Auwera (Eds.), *The languages and linguistics of europe* (pp. 69–96). Berlin: Mouton de Gruyter.

Kager, R., Van der Feest, S. V. H., Fikkert, P., Kerkhoff, A. & Zamuner, T. S. (2007). Representations of [voice]: Evidence from acquisition. In J. V. de Weijer & E. J. V. der Torre (Eds.), *Voicing in dutch - (de)voicing - phonology, phonetics, and psycholinguistics* (pp. 41–80). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kallestinova, E. (2004). Voice and aspiration of stops in Turkish. In G. Dogil (Ed.), *Folia linguistica 38: Special issue on voice* (pp. 117–143). Berlin: Mouton de Gruyter.

Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1985). The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook*, *2*, 305–328.

Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, *7*, 193–231.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–319.

Keating, P. A. (1988). The phonology-phonetics interface. In F. J. Newmeyer (Ed.), *Linguistics: The cambridge survey* (pp. 281–302). Cambridge: Cambridge University Press.

Kehrein, W. & Golston, C. (2004). A prosodic theory of laryngeal contrasts. *Phonology*, *21*(03), 325–357. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S0952675704000302`

Kenstowicz, M. (1996). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand & B. Laks (Eds.), *Current trends in phonology: models and methods* (pp. 363–393). Salford: University of Salford Publications.

Kessler, B. (1995). Computational dialectology in Irish Gaelic. In *Proceedings of the european association for computational linguistics* (pp. 60–67). Dublin.

Kim, C. (1970). A theory of aspiration. *Phonetica*, *21*, 107–116.

Kingston, J. (2007). The phonetics-phonology interface. In P. De Lacy (Ed.), *Handbook of phonology* (pp. 401–434). Cambridge: Cambridge University Press.

Kloeke, G. G. (1927). *De Hollandsche expansie in de zestiende en zeventiende eeuw en haar weerspiegeling in de hedendaagsche Nederlandsche dialecten.* Den Haag: Martinus Nijhoff.

Kohler, K. J. (1984). Phonetic explanation in phonology: The feature fortis/lenis. *Phonetica*, *41*, 150–174.

Köhnlein, B. (2011). *Rule reversal revisited: Synchrony and diachrony of tone and prosodic structure in the Franconian dialect of Arzbach* (PhD Thesis).

Krämer, M. (2001). On obstruent voicing in Breton, German, and Italian. In A. Holmer, J.-O. Svantesson & Å. Viberg (Eds.), *Proceedings of the 18th scandinavian conference of linguistics. travaux de l'institut de linguistique de lund* (pp. 39–55). Lund: Lund University Press.

Krämer, M. (2003a). Intervocalic /s/-voicing, geminates and the Richness of the Base in Veneto Italian. *Rivista di Grammatica Generativa*, *28*, 71–85.

Krämer, M. (2003b). *Variation of s-voicing in two varieties of Italian* (Manuscript). University of Ulster.

Krämer, M. (2005). Contiguity and non-derived environment blocking of s-

voicing in Lombardian and Tuscan Italian. *Probus*, *17*, 227–251.
Krämer, M. (2009). *The phonology of Italian.* Oxford: Oxford University Press.
Kuhl, P. K. & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*, 69–72.
Kuhl, P. K. & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, *63*(3), 905–917. Retrieved from `http://asa.scitation.org/doi/10.1121/1.381770`
Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608. Retrieved from `http://www.sciencemag.org/cgi/doi/10.1126/science.1736364`
Kulikov, V. (2012). *Voicing and voice assimilation in Russian stops* (PhD Thesis). The University of Iowa.
Ladd, D. R., Dediu, D. & Kinsella, A. R. (2008). Languages and genes: Reflections on biolinguistics and the nature-nurture question. *Biolinguistics*, *2*(1), 114–126.
Lai, Y., Huff, C., Sereno, J. & Jongman, A. (2009). The raising effect of aspirated prevocalic consonants on F0 in Taiwanese. In J. Brooke et al. (Eds.), *Proceedings of the 2nd international conference on east asian linguistics.* Vancouver: Simon Fraser University.
Lameli, A., Kehrein, R. & Rabanus, S. (2011). *Language and space - Language mapping.* Berlin/New York: De Gruyter Mouton.
Lasky, R. E., Syrdal-Lasky, A. & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, *20*, 215–225.
Lavie, R.-J. (2005). Exemplar theory in linguistics: A perspective for the cognitive subject. In *Communication to the 11th congress of cognitive linguistics.* Bordeaux.
Lindow, W., Möhn, D., Niebaum, H., Stellmacher, D., Taubken, H. & Wirrer, J. (1998). *Niederdeutsche grammatik.* Leer: Verlag Schuster.
Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.
Loporcaro, M. (1999). Teoria fonologica e ricerca empirica sull'italiano e i suoi dialetti. In P. Benincà, A. Mioni & L. Vanelli (Eds.), *Fonologia e morfologia dell'italiano e dei dialetti d'italia - atti del xxxi congresso della società di linguistica italiana.* Roma: Bulzoni.
Maddieson, I. (1984). *Patterns of sounds.* Cambridge: Cambridge University Press.
Maffei Bellucci, P. (1977). *Profilo dei dialetti italiani - Lunigiana* (M. Cortelazzo, Ed.). Pisa: Pacini Editore.
Maiden, M. & Parry, M. (Eds.). (1997). *The dialects of Italy.* London/New York: Routledge.
Massariello Merzagora, G. (1988). *Profilo dei dialetti italiani - Lombardia*

(M. Cortelazzo, Ed.). Pisa: Pacini Editore.

Massey, D. (1984). Introduction: Geography matters. In D. Massey & J. Allen (Eds.), *Geography matters!* (pp. 1–11). Cambridge: Cambridge University Press.

Matras, Y. & Reershemius, G. (2003). *Low German (East Frisian dialect).* München: Lincom GmbH.

Maye, J. & Gerken, L. (2000). Learning phonemes without minimal pairs. In *Proceedings of the 24th annual boston university conference on language development.*

Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

Milis, L. J. R., Blockmans, W. P., Van Deursen, A. T., Janssens, P., Roegiers, J., Van Sas, N. C. F., ... Lamberts, E. (2014). *Geschiedenis van de Nederlanden* (J. C. H. Blom & E. Lamberts, Eds.). Amsterdam: Prometheus Bert Bakker.

Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J. & Dooling, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. *The Journal of the Acoustical Society of America*, *60*, 410–417.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340.

Montemagni, S., Wieling, M. B., De Jonge, B. & Nerbonne, J. A. (2013). Synchronic patterns of Tuscan phonetic variation and diachronic change. *Literary and Linguistic Computing*, *28*(1), 157–172.

Moosmüller, S. & Ringen, C. (2004). Voice and aspiration in Austrian German plosives. *Folia Linguistica*, *38*(1-2), 43–62. Retrieved from `http://www.degruyter.com/view/j/flin.2004.38.issue-1-2/flin.2004.38.1-2.43/flin.2004.38.1-2.43.xml`

Nerbonne, J. A., Heeringa, W. J., Van den Hout, E., Van der Kooi, P., Otten, S. & Van de Vis, W. (1996). Phonetic distance between Dutch dialects. In G. Durieux, W. Daelemans & S. Gillis (Eds.), *Clin vi: Proceedings of the sixth clin meeting* (pp. 185–202). Antwerpen: UIA Center for Dutch Language and Speech.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology.* Berlin: Mouton de Gruyter.

Nocchi, N. & Filipponio, L. (2011). Lo vuoi co[z]ì o co[s]ì? A sociophonetic study on sibilants in the regional Italian of Livorno (Tuscany). In S. Calamai, C. Celata & L. Ciucci (Eds.), *Proceedings of the workshop "sociophonetics, at the crossroads of speech variation, processing and communication".* Pisa: Edizioni della Normale.

Nocchi, N. & Schmid, S. (2007). Lenition of voiceless fricatives in two varieties of southern Italian. In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 1497–1500).

Saarbrücken.

Noske, R. (2005). A prosodic contrast between Northern and Southern Dutch: A result of a Flemish-French sprachbund. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz & J. Koster (Eds.), *Organizing grammar. linguistic studies in honor of henk van riemsdijk* (pp. 474–482). Berlin: Mouton de Gruyter.

Ögüt, F., Kiliç, M. A., Engin, E. Z. & Midili, R. (2006). Voice Onset Times for Turkish stop consonants. *Speech Communication*, *48*, 1094–1099.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). Berlin: Springer.

Ohala, J. J. (1990). There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics*, *18*, 153–171.

Pape, D. & Jesus, L. M. T. (2015). Stop and fricative devoicing in European Portuguese, Italian and German. *Language and Speech*, *58*(2), 224–246.

Patrick, P. L. (1991, jul). Creoles at the intersection of variable processes: - t,d deletion and past-marking in the Jamaican mesolect. *Language Variation and Change*, *3*(02), 171. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S095439450000051X`

Peperkamp, S. (1995). Prosodic constraints in the derivational morphology of Italian. In G. E. Booij & J. V. Marle (Eds.), *Yearbook of morphology 1994* (pp. 207–244). Dordrecht: Kluwer Academic Publishers.

Peperkamp, S. (1997). *Prosodic words* (PhD Thesis). Universiteit van Amsterdam.

Peperkamp, S., Le Calvez, R., Nadal, J.-P. & Dupoux, E. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, *101*(3), B31–B41. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0010027705002118`

Petrovic, J. (2008). Critical period hypothesis. In J. M. Gonzáles (Ed.), *Encyclopedia of bilingual education* (pp. 194–197). Los Angeles: SAGE Publications.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–158). Amsterdam: John Benjamins Publishing Company.

Pinget, A.-F. (2015a). *The actuation of sound change* (PhD Thesis). Universiteit Utrecht.

Pinget, A.-F. (2015b). *Personal communication.*

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *The Journal of the Acoustical Society of America*, *61*(5), 1352–1361.

Prehn, M. (2011). *Vowel quantity and the fortis-lenis distinction in North Low Saxon* (PhD Thesis). University of Amsterdam.

Prince, A. & Smolensky, P. (n.d.). *Optimality Theory - Constraint interaction in generative grammar.*

Reiss, C. (2017). Substance Free Phonology. In S. Hannahs & A. Bosch (Eds.), *Handbook of phonological theory.* London/New York: Routledge.

Revithiadou, A., Van Oostendorp, M., Nikolou, K. & Tiliopoulou, M.-A. (2006). Vowel harmony in contact induced systems: The case of Asia Minor dialects of Greek. In M. Janse, B. D. Joseph & A. Ralli (Eds.), *Proceedings of the 2nd international conference of modern greek dialects and linguistic theory* (pp. 350–365). Patra: University of Patras.

Rietveld, A. C. M. & Van Heuven, V. J. J. P. (1997). *Algemene fonetiek.* Bussum: Uitgeverij Coutinho.

Rivas, M. V. (2006). *Does the perception of fricatives correspond to their production? The case of Italian vs. Dutch* (MA Thesis). Universiteit van Amsterdam.

Rogers, D. & D'Arcangeli, L. (2004, jan). Italian. *Journal of the International Phonetic Association*, *34* (1), 117–121. Retrieved from `http://www.journals.cambridge.org/abstract{\_}S0025100304001628`

Rohlfs, G. (1966). *Grammatica storica della lingua italiana e dei suoi dialetti - Fonetica.* Torino: Piccola Biblioteca Einaudi.

Romanello, M. T. (2002). The perception of urban varieties: Preliminary studies from the south of Italy. In D. Long & D. R. Preston (Eds.), *Handbook of perceptual dialectology - volume 2* (pp. 331–350). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Romero, J. & Riera, M. (2015). *The phonetics-phonology interface.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Roos, N. G. J. (2009). *The weak past tense in Dutch and Low German* (PhD Thesis). Radboud Universiteit Nijmegen.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996, dec). Statistical Learning by 8-Month-Old Infants. *Science*, *274* (5294), 1926–1928. Retrieved from `http://www.sciencemag.org/cgi/doi/10.1126/science.274.5294.1926`

Savoia, L. (1997). The geographical distribution of the dialects. In M. Maiden & M. Parry (Eds.), *The dialects of italy* (pp. 225–234). Abingdon: Routledge.

Scalise, S. (1986). *Generative morphology.* Dordrecht: Foris Publications.

Scherrer, Y. & Stoeckle, P. (2016, jan). A quantitative approach to Swiss German Dialectometric analyses and comparisons of linguistic levels. *Dialectologia et Geolinguistica*, *24* (1), 92–125. Retrieved from `https://www.degruyter.com/view/j/dialect.2016.24.issue-1/dialect-2016-0006/dialect-2016-0006.xml`

Schmeller, J. A. (1821). *Die Mundarten Bayerns grammatisch dargestellt.* München: Karl Thienemann.

Schoemans, M. & Van Oostendorp, M. (2004). Stemhebbende eindfricatieven in het Twents. In *Tabu* (pp. 1–13).

Scholtmeijer, H. (2003). *Taal in stad en land - Utrechts, Veluws en Flevolands.* Den Haag: Sdu Uitgevers.

Scholtmeijer, H. (2011). *Veluws handwoordenboek.* Almere: Uitgeverij Van de

Berg.

Schrijver, P. (2011). The High German consonant shift and language contact. In R. V. Pareren, H. P. Houtzagers & C. Hasselblatt (Eds.), *Language contact in times of globalization* (pp. 217–249). Amsterdam: Brill Academic Publishers.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, *20*, 213–223.

Sebregts, K. D. C. J. (2015). *The sociophonetics and phonology of Dutch r* (PhD Thesis). Universiteit Utrecht.

Séguy, J. (1973). La dialectométrie dans l'Atlas linguistique de la Gascogne. *Revue de linguistique Romane*, *37*, 1–24.

Slis, I. H. & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction I. *Language and Speech*, *12*(2), 80–102.

Sloos, M. & Van Oostendorp, M. (2010, jan). The relationship between phonological and geographical distance. Umlaut on the diminutive in Dutch dialects. *Taal en tongval*, *62*(2), 204–234. Retrieved from `http://openurl.ingenta.com/content/xref?genre=article{\&}issn=0039-8691{\&}volume=62{\&}issue=2{\&}spage=204`

Smith, J. L. (2008). Markedness, faithfulness, positions, and contexts: Lenition and fortition in Optimality Theory. In J. B. de Carvalho, T. Scheer & P. Ségéral (Eds.), *Lenition and fortition* (pp. 519–560). Berlin/New York: Mouton de Gruyter.

Solari, A. (1932). L'unità storica dell'Aemilia. *Athenaeum*, *10*, 40–46.

Soli, S. D. (1983). The role of spectral cues in discrimination of voice onset time differences. *The Journal of the Acoustical Society of America*, *73*(6), 2150–2165.

Spruit, M. R. (2008). *Quantitative perspectives on syntactic variation in Dutch dialects* (PhD Thesis). Rijksuniversiteit Groningen.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, *259*(5538), 39–41. Retrieved from `http://www.nature.com/doifinder/10.1038/259039a0`

Te Winkel, J. (1901). *Geschiedenis der Nederlandsche taal (translation of Geschichte der niederländischen sprache, by F.C. Wieder)*. Culemborg: Blom & Olivierse.

Ten Kate, L. H. (1723). *Aenleiding tot de kennisse van het verhevene deel der Nederduitsche sprake - Eerste deel*. Amsterdam: Rudolph en Gerard Wetstein.

Thomason, S. G. & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley/Los Angeles/London: University of California Press.

Tosi, A. (2004). The language situation in Italy. *Current Issues in Language Planning*, *5*(3), 247–335. Retrieved from `http://www.tandfonline.com/doi/abs/10.1080/14664200408668259`

Van Alphen, P. M. (2007). Prevoicing in Dutch initial plosives - Production, perception and word recognition. In J. Van de Weijer & E. J. Van der Torre (Eds.), *Voicing in dutch - (de)voicing - phonology, phonetics, and psycholinguistics* (pp. 99–124). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Alphen, P. M. & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of prevoicing. *Journal of Phonetics*, *32*(4), 455–491. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0095447004000324`

Van Bree, C. (1987). *Historische grammatica van het Nederlands.* Dordrecht: Foris Publications.

Van de Velde, H. & Van Hout, R. (1998). Dangerous aggregations: A case study of Dutch (n) deletion. In C. Paradis, D. Vincent, D. Deshaies & M. Laforest (Eds.), *Papers in sociolinguistics.* Québec: Éditions Nota Bene.

Van de Velde, H. & Van Hout, R. (2003). De deletie van de slot-n. *Nederlandse Taalkunde*, *8*(2), 93–114.

Van der Feest, S. V. H. & Johnson, E. K. (2016). Input-driven differences in toddlers' perception of a disappearing phonological contrast. *Language Acquisition*, *23*(2), 89–111. Retrieved from `http://www.tandfonline.com/doi/full/10.1080/10489223.2015.1047096`

Van der Hulst, H. G. (2006). Dependency phonology. In K. Brown (Ed.), *The encyclopedia of language and linguistics* (pp. 451–458). Oxford: Elsevier.

Van Ginneken, J. (1913). *Handboek der Nederlandsche taal - Deel I.* 's-Hertogenbosch: L.C.G. Malmberg.

Van Oostendorp, M. (1999). Italian s-voicing and the structure of the phonological word. In S. Hannahs & M. Davenport (Eds.), *Issues in phonological structure* (pp. 197–214). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Oostendorp, M. (2007). Exceptions to Final Devoicing. In J. V. de Weijer & E. J. V. der Torre (Eds.), *Voicing in dutch - (de)voicing - phonology, phonetics, and psycholinguistics* (pp. 81–98). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Oostendorp, M. (2012). *Ungrammaticality in language variation and acquisition.*

Van Son, R. J. J. H. & Pols, L. C. W. (1996). An acoustic profile of consonant reduction. In *Proceedings of the fourth international conference on spoken language* (pp. 1529–1532).

Vietti, A., Alber, B. & Vogt, B. (2018). Initial laryngeal neutralisation in Tyrolean. *Phonology*, *35*(1), 79–114.

Vignuzzi, U. (2010). *Aree linguistiche.* Retrieved from `http://www.treccani.it/enciclopedia/aree-linguistiche{\_}{\%}28Enciclopedia-dell{\%}27Italiano{\%}29/`

Von Wartburg, W. (1936). Die Ausgliederung der romanischen Sprachräume. *Zeitschrift für romanische Philologie*, *56*(1), 1–48.

Weijnen, A. A. (1966). *Nederlandse dialectkunde.* Assen: Koninklijke Van

Gorcum BV.

Weinreich, M. (1959). History of the Yiddish language: The problems and their implications. *Proceedings of the American Philosophical Society*, *103*(4), 563–570.

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Wetzels, W. L. & Mascaró, J. (2001). The typology of voicing and devoicing. *Language*, *77*(2), 207–244.

White, K. S., Peperkamp, S., Kirk, C. & Morgan, J. L. (2008). Rapid acquisition of phonological alternations by infants. *Cognition*, *107*(1), 238–265. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0010027707003009`

Wieling, M. B. (2007). *Comparison of Dutch dialects* (Unpublished doctoral dissertation). Rijksuniversiteit Groningen.

Wieling, M. B., Heeringa, W. J. & Nerbonne, J. A. (2007). An aggregate analysis of pronunciation in the Goeman-Taeldeman-Van Reenen-Project data. *Taal en Tongval*, *59*.

Winkler, J. (1874). *Algemeen Nederduitsch en Friesch dialecticon.* Den Haag: Martinus Nijhoff.

Wrede, F., Mitzka, W. & Martin, B. (n.d.). *Deutscher Sprachatlas. Auf Grund des Sprachatlas des deutschen Reichs von Georg Wenker begonnen von Ferdinand Wrede, fortgesetzt von Walther Mitzka und Bernhard Martin.* Marburg: N. G. Elwert Verlag.

Yip, M. (2002). *Tone.* Cambridge: Cambridge University Press.

Zamboni, A. (1974). *Profilo dei dialetti italiani - Veneto* (M. Cortelazzo, Ed.). Pisa: Pacini Editore.

Zuraw, K. (2002). Optimality Theory in linguistics. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (2nd editio ed., pp. 819–822). Cambridge, MA/London: The MIT Press.

Samenvatting in het Nederlands

Introductie

De studie van taalvariatie is een belangrijk onderdeel van de taalwetenschap. De belangrijkste aanname in de generatieve taalkunde (Chomsky (1957)) is dat alle mensen beschikken over dezelfde 'language faculty', waarmee alle verschillende talen gegenereerd moeten kunnen worden. De studie van taalvariatie speelt hierbij een belangrijke rol, aangezien ze inzicht geeft in de structuren die de language faculty moet kunnen genereren, en omdat ze laat zien dat diezelfde language faculty in staat moet zijn om variatie te genereren. In dit proefschrift wordt een nieuw licht op taalvariatie geschenen door variatie te bestuderen vanuit een geografisch perspectief. De centrale vraag hierbij is wat geografische patronen in taalvariatie ons kunnen leren over deze variatie. De focus ligt op regio's met veel taalcontact, in het bijzonder transitiezones, omdat in deze regio's naar alle waarschijnlijkheid veel taalvariatie gevonden kan worden.

Transitiezones kunnen worden gedefinieerd als gebieden waar taalsysteem A langzaam overgaat in taalsysteem B (Chambers and Trudgill (1980)). Omdat de taalsystemen in deze regio's geleidelijk veranderen, is er naar alle waarschijnlijkheid sprake van taalcontact en taalvariatie. Deze gebieden kunnen dus veel inzicht leveren in taalvariatie. Een zoektocht naar transitiezones in de dialectologische literatuur levert echter direct een probleem op. Hoewel het begrip transitiezone heel makkelijk te definiëren lijkt te zijn (een langzame overgang van taalsysteem A naar taalsysteem B, of in de woorden van Daan and Blok (1969) een 'terrassenlandschap'), laat een literatuurstudie zien dat het begrip niet op die manier gebruikt wordt. De geografische factor lijkt een minimale rol te spelen in hoe het begrip gebruikt wordt: dialectologen classificeren dialecten als transitiedialecten of overgangsdialecten als het dialect niet bij één van de grotere dialectgroepen ingedeeld kan worden. Het gaat daarbij meestal wel om classificatie van dialecten in een geografisch beperkte regio, maar de classificatieproblemen zijn niet per sé terug te leiden op de geografie (i.e., het daadwerkelijk transitionele karakter van het dialect). Een vergelijking van verschillende dialectologische studies van het Nederlandse taalgebied laat zien dat verschillende dialectologen op verschillende dialectclassificaties uitkomen. Een belangrijke oorzaak daarvan ligt in het feit dat de classificaties vaak op verschillende taalkundige verschijnselen gebaseerd zijn: als iedere dialectoloog een eigen set isoglossen hanteert, zullen alle classificaties er anders uitzien. In deze dissertatie is er daarom gekozen om niet te kijken naar de verschillende taalkundige eigenschappen van één transitiedialect, maar naar één taalkundige eigenschap van dialecten in één geografische regio.

De focus ligt in dit proefschrift op voicingverschijnselen. Een eerste reden is dat er in de literatuur veel geschreven is over voicingcontrasten. Veel verschillen tussen taalvariëteiten zijn uitgebreid beschreven. Hierdoor is het goed mogelijk om regio's te definiëren waartussen mogelijk variatie te vinden is. Daarnaast zijn verschillen tussen voicingcontrasten niet alleen fonetisch, maar ook fonologisch van aard, waardoor voicingcontrasten het mogelijk maken niet alleen fonetische maar ook fonologische variatie te bestuderen.

Voicingcontrasten

Het eerste verschijnsel dat centraal staat in dit proefschrift is de overgang tussen voicingtalen en aspiratietalen in het Nedersaksisch taalgebied. Dit taalgebied strekt zich uit van Drenthe en Overijssel in het noorden van Nederland, naar Mecklenburg-Vorpommern en Brandenburg in het noorden van Duitsland. Taalkundig gezien vormen de Nedersaksische dialecten een eenheid, met als één van de belangrijkste verschillen de realisatie van stemhebbende en stemloze klanken. Door de taalkundige eenheid vormt deze regio een ideaal gebied om de overgang tussen voicingtalen en aspiratietalen te bestuderen.

Voicingtalen en aspiratietalen verschillen zowel fonetisch als fonologisch van elkaar. Fonetisch is het belangrijkste onderscheid de Voice Onset Time (VOT) in woord-initiële positie, en fonologisch is het belangrijkste onderscheid het actieve fonologische feature. In voicingtalen vinden we een contrast tussen lenis plosieven waarbij stembandtrilling inzet vóór de explosie van de plosief (negatieve VOT, ofwel prevoicing) en fortis plosieven waarbij stembandtrilling kort na de explosie van de plosief inzet (kort-positieve VOT). Hierbij is de lenis plosief fonologisch gemarkeerd met het kenmerk [voice], terwijl de fortis plosief fonologisch ongemarkeerd ([∅]) is. In aspiratietalen vinden we een contrast tussen lenis plosieven waarbij stembandtrilling kort na de explosie van de plosief (kort-positieve VOT) inzet en fortis plosieven waarbij stembandtrilling relatief lang na de explosie van de plosief (lang-positieve VOT, ofwel aspiratie) inzet. Hierbij is juist de lenis plosief fonologisch ongemarkeerd ([∅]), terwijl de fortis plosief fonologisch gemarkeerd is met het kenmerk [spread glottis].

Naast deze verschillen speelt de realisatie van intervocalische plosieven een rol, alsmede de realisatie van stemhebbendheid in clusters met twee plosieven. In voicingtalen worden lenis intervocalische plosieven altijd gerealiseerd met volledige stemhebbendheid tijdens de sluitingsfase van de plosief, terwijl fortis plosieven altijd met gedeeltelijke stemhebbendheid worden gerealiseerd. In aspiratietalen worden fortis intervocalische plosieven altijd met gedeeltelijke stemhebbendheid gerealiseerd, terwijl lenis plosieven met ofwel volledige ofwel gedeeltelijke stemhebbendheid worden gerealiseerd. Daarnaast speelt het fonologische kenmerk dat in de taal aanwezig is een rol in clusters met twee plosieven. In voicingtalen, waar het kenmerk [voice] aanwezig is, spreidt het kenmerk [voice] van de tweede plosief in het cluster naar de eerste plosief in datzelfde cluster. In deze talen is stemassimilatie dus actief. In aspiratietalen, waar het kenmerk [spread glottis] aanwezig is, is er geen spreiding van dit kenmerk in een cluster van twee plosieven. De aanwezigheid van het kenmkerk in het cluster blokkeert echter wel intervocalische voicing van dat cluster.

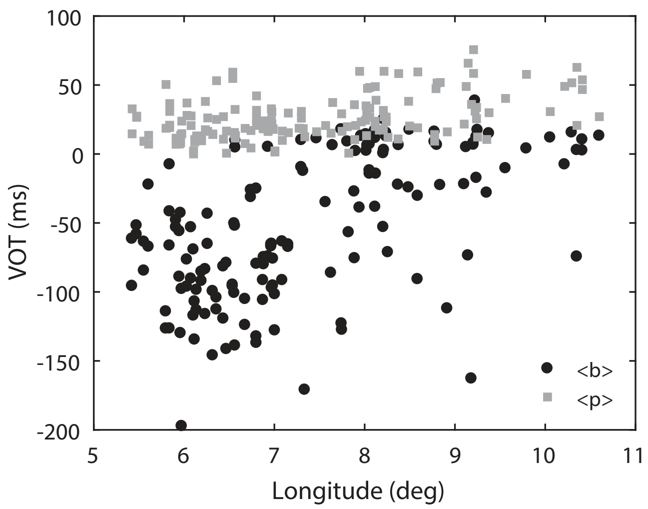
Het westen van het Nedersaksisch taalgebied wordt gekenmerkt door een voicingsysteem, terwijl het oosten van het gebied wordt gekenmerkt door een aspiratiesysteem. De vraag is juist wat er gebeurt in het middengebied, waar de twee systemen met elkaar in contact zijn. Om deze vraag te beantwoorden zijn plosiefrealisaties van sprekers uit het hele continuüm geanalyseerd. De data voor de analyse komen uit twee verschillende databases: aan de Nederlandse kant van de landsgrens komen de data uit de database van het Goeman-Taeldeman-Van Reenen-Project (GTRP), aan de Duitse kant van de grens zijn opnames van de Wenkersätze gebruikt.

In woordinitiële positie is de VOT van de plosieven <b, d, p, t, k> gemeten. Omdat de velaire lenis plosief niet in alle dialecten gesproken in Nederland voorkomt, is deze uit de analyse gelaten. Wanneer de gemeten VOT-waardes van zowel lenis als fortis plosieven (per plaats van articulatie) in één grafiek worden uitgezet tegen de lengtegraden van de locaties waar de opnames gemaakt zijn, wordt een interessant patroon zichtbaar: in het westen en oosten laten sprekers een duidelijk onderscheid zien tussen lenis en fortis plosieven, maar in het middengebied is er geen duidelijk onderscheid tussen lenis en fortis te zien (Figuur 1 en 2).

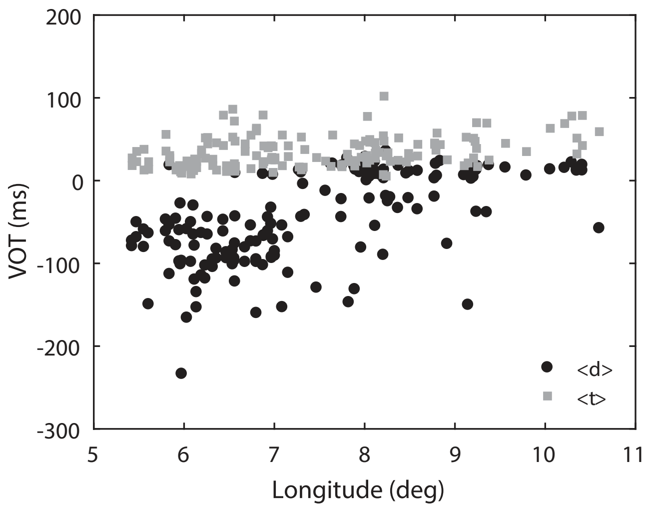
Sprekers in het westen laten een consequent contrast zien tussen plosieven met een negatieve VOT en plosieven met een kort-positieve VOT, overeenkomstig met een voicingsysteem. In het oosten laten sprekers een consequent contrast zien tussen plosieven met een kort-positieve en plosieven met een lang-positieve VOT, overeenkomstig met een aspiratiesysteem. De fonetische waardes van sprekers in het middengebied beslaan wel een groot bereik, maar een consequent contrast is niet zichtbaar.

In clusters met twee plosieven is een vergelijkbaar patroon zichtbaar. In het westen laten sprekers consequent assimilatie naar de stemhebbende plosief zien (in clusters met een lenis tweede plosief), terwijl sprekers in het oosten geen assimilatie laten zien (in clusters met een fortis tweede plosief) (Figuur 3 en 4).

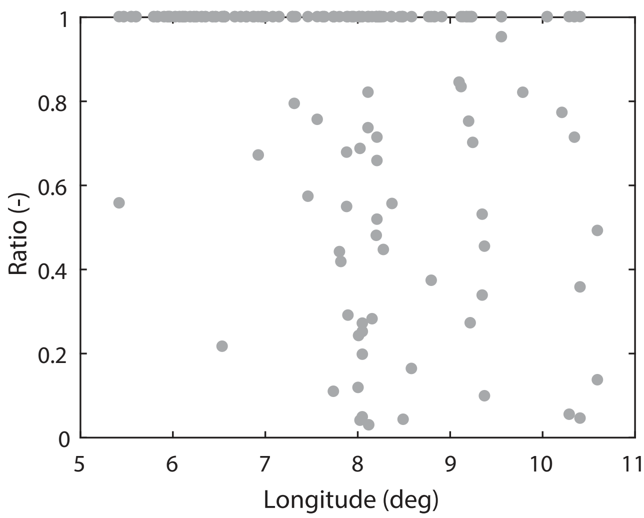
Hoewel sprekers in het middengebied veel fonetische variatie laten zien, moet geconcludeerd worden dat zij fonologisch geen onderscheid maken tussen lenis en fortis plosieven. Fonetisch gezien maken sprekers geen consequent onderscheid tussen lenis en fortis plosieven. Daarnaast laten plosiefclusters geen



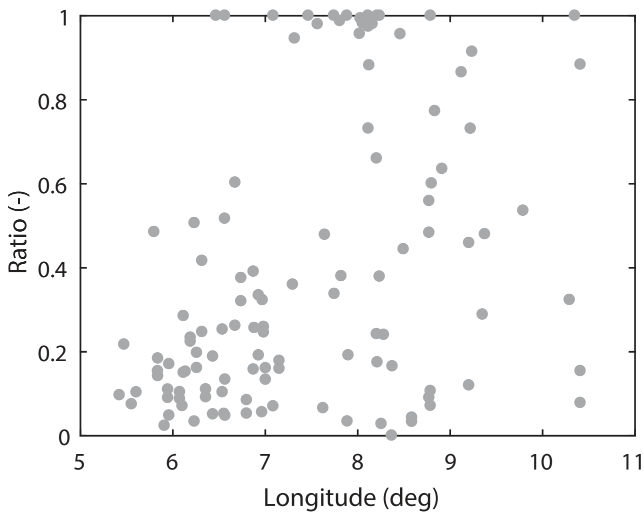
Figuur 1: Overlap in VOT-waardes voor bilabiale plosieven



Figuur 2: Overlap in VOT-waardes voor alveolaire plosieven



Figuur 3: Assimilatie in clusters met een lenic C2



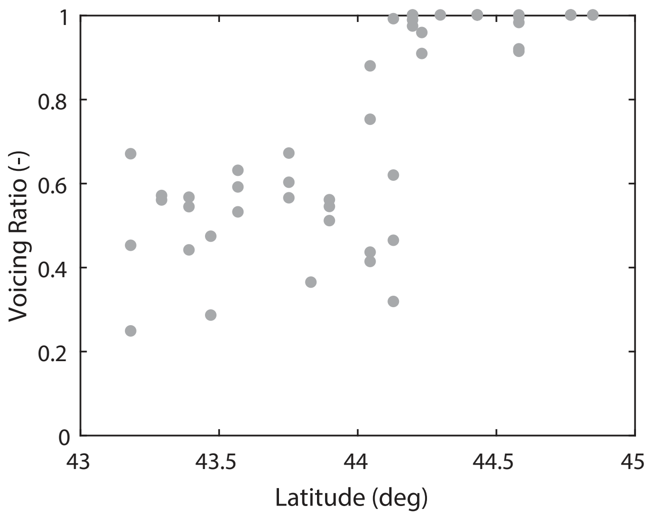
Figuur 4: Assimilatie in clusters met een fortis C2

bewijs zien voor een fonologisch feature. Liever gezegd, clusters laten bewijs zien voor de afwezigheid van een fonologisch feature. Als er een feature [voice] aanwezig is zouden de clusters met lenis plosieven op de tweede plek consequent volledig stemhebbend moeten zijn; als er een feature [spread glottis] aanwezig is zouden de clusters met fortis plosieven consequent slechts deels stemhebbend moeten zijn. In werkelijkheid zijn beide types clusters echter af en toe volledig stemhebbend, en af en toe deels stemhebbend. Assimilatie wordt dus niet afgedwongen ([voice] is niet aanwezig), maar fonetische intervocalische voicing wordt ook niet geblokkeerd ([spread glottis] is niet aanwezig). Kortom, het middengebied tussen de Nedersaksische voicingtalen en aspiratietalen wordt niet gekenmerkt door veel fonologische variatie, maar juist door de afwezigheid van een fonologisch contrast.

Hoe verklaren we de afwezigheid van een fonologisch contrast als beide aangrenzende regio's wel een contrast hebben? Taalverwerving speelt hier een belangrijke rol: een taallerend kind wordt geconfronteerd met conflicten in de taalinput. Voor het verwerven van een tweeledig fonologisch contrast is een bimodale verdeling van fonetische waardes in de input nodig, voor het verwerven van een drieledig contrast is een trimodale verdeling van fonetische waardes in de input nodig. De input in het Nedersaksisch taalgebied laat een unimodale verdeling zien, en geeft dus geen mogelijkheid om een fonologisch contrast te verwerven. Daarnaast is er ambiguïteit van het fonologisch systeem: een plosief met een kort-positieve VOT is altijd ongemarkeerd, maar deze plosief kan ofwel de ongemarkeerde fortis plosief in een voicingtaal zijn, ofwel de ongemarkeerde lenis plosief in een aspiratietaal. Voor het taallerend kind is er niet genoeg evidentie voor ofwel een voicingsysteem ofwel een aspiratiesysteem.

Intervocalische /s/-voicing

In het tweede deel van dit proefschrift staat intervocalische /s/-voicing in Italiaanse dialecten centraal. In dialecten die ten noorden van de La Spezia-Riminilijn worden gesproken, wordt een staminterne /s/ gerealiseerd als [z] wanneer deze zich in intervocalische positie bevindt. Ten zuiden van de La Spezia-Riminilijn is dit fenomeen afwezig: in deze dialecten wordt een /s/ in diezelfde context als [s] gerealiseerd. De staminterne context waarin /s/-voicing optreedt in de noordelijke dialecten, kan worden opgedeeld in verschillende morfologische contexten: monomorfemische woorden (zoals *la casa* 'het huis'), complexe maar morfologisch opake woorden (zoals *la resistenza* 'het verzet') en woorden met een prefix waar de /s/ de laatste consonant van het prefix is (zoals *disonesto* 'oneerlijk'). In de noordelijke dialecten wordt de intervocalische /s/ in al deze woorden gerealiseerd als [z], terwijl in de niet-noordelijke dialecten alleen de fricatief in de laatste context als [z] wordt gerealiseerd; in de andere twee contexten wordt ze gerealiseerd als [s]. In het contactgebied tussen deze twee regio's is het te verwachten dat sprekers inconsequente /s/-voicing laten zien, waarbij in sommige staminterne contexten wel /s/-voicing optreedt en in

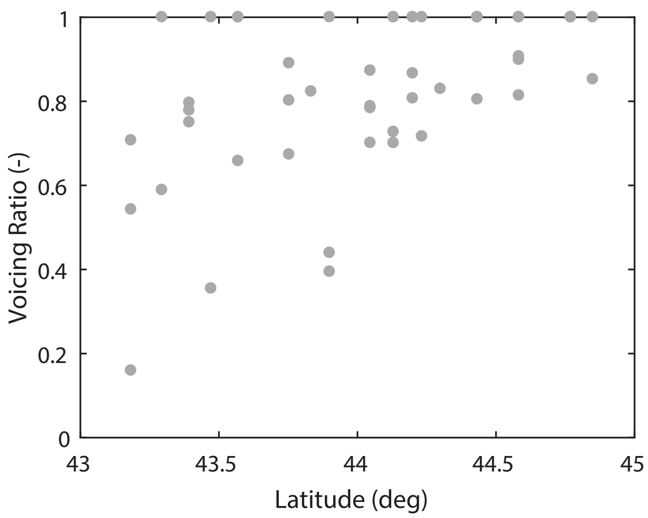


Figuur 5: Gemiddelde voicingratio in intervocalische fricatieven

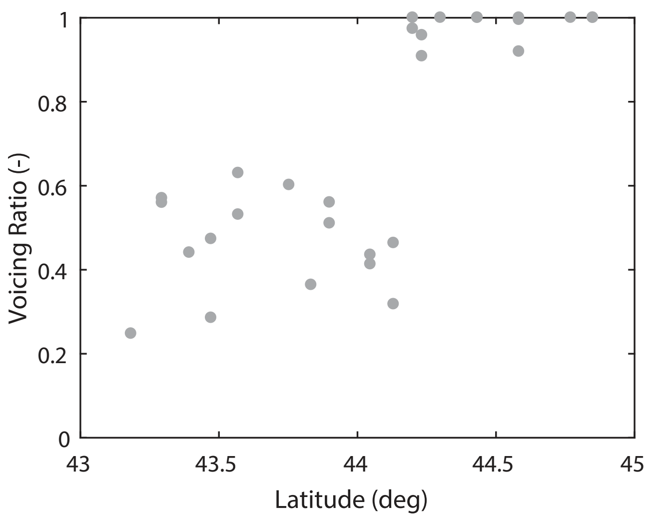
andere staminterne contexten niet. Om dit te onderzoeken zijn de realisaties van fricatieven in deze context, van fricatieven in contexten waar in beide regio's een [s] verwacht wordt, en contexten waar in beide regio's een [z] verwacht wordt, geanalyseerd. De data hiervoor zijn verzameld tijdens een veldwerksessie in de provincies Emilia-Romagna (ten noorden van de La Spezia-Riminilijn) en Toscane (ten zuiden van deze lijn) in september en oktober 2015. Voor iedere fricatief is het percentage van voicing berekend, door de duur van de voicing te delen door de duur van de fricatief. In de contexten waar altijd een [s] of [z] verwacht wordt laten de twee regio's geen verschillen zien. Wanneer we echter kijken naar fricatieven op een opaque morfeemgrens, en vooral naar fricatieven in staminterne contexten, wordt een duidelijk verschil tussen de twee regio's zichtbaar:

In de grafieken in figuren 5 en 6 zijn de voicingratio's op de y-as geplot, tegenover de breedtegraden van iedere opnamelocatie op de x-as. In plaats van een geleidelijke overgang tussen de twee systemen met een duidelijk tussengebied, vinden we hier een abrupte overgang tussen systemen met en systemen zonder intervocalische /s/-voicing. Sprekers in Emilia-Romagna realiseren de fricatief in staminterne context vrijwel consequent als [z] (wanneer ze een fonologisch onderscheid tussen [s] en [z] in intervocalische context zouden hebben, zouden de voicingpercentages veel lager liggen), terwijl de lagere voicingpercentages van sprekers in Toscane laat zien dat zij wel een fonologisch contrast tussen de twee fricatieven hebben.

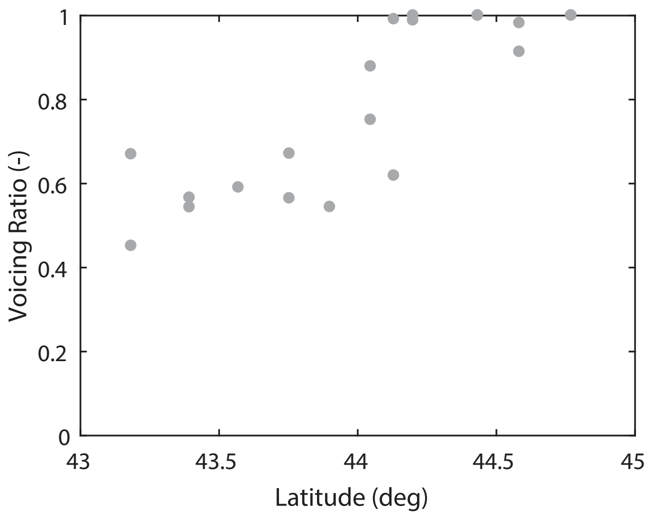
Hoewel de grafieken hierboven een abrupte overgang laten zien, wordt het patroon ietwat complexer wanneer we de realisaties van oudere (Figuur 7) en jongere (Figuur 8) sprekers apart bekijken.



Figuur 6: Gemiddelde voicingratio in staminitiële fricatieven (opaak)



Figuur 7: Gemiddelde voicingratio in intervocalische fricatieven (oudere sprekers)



Figuur 8: Gemiddelde voicingratio in intervocalische fricatieven (jongere sprekers)

De realisaties van oudere sprekers laten een volledig abrupte overgang tussen de twee systemen zien, terwijl deze abrupte overgang veel minder duidelijk is bij de jongere sprekers. Hoe dichter in de buurt van Emilia-Romagna jongere Toscaanse sprekers wonen, hoe hoger hun gemiddelde voicingpercentages zijn. Ook zijn hun voicingpercentages hoger dan de voicingpercentages van oudere Toscaanse sprekers.

De interface tussen fonetiek en fonologie

In geen van de twee bestudeerde gebieden vinden we fonologische variatie. In de Italiaanse regio is er geen tussengebied tussen de twee regio's aanwezig, en in de Nedersaksische regio wordt het tussengebied gekenmerkt door de afwezigheid van een fonologisch contrast. Toch laten sprekers in beide regio's veel fonetische variatie zien. In het Nedersaksische middengebied zien we dat sprekers geen fonologisch contrast hebben tussen lenis en fortis plosieven, maar fonetisch wel alle verschillende realisaties kunnen gebruiken die voorkomen bij voicingcontrasten en aspiratiecontrasten. In het Italiaanse middengebied zien we dat voor noch de jongere noch de oudere sprekers een fonologisch middengebied bestaat (i.e. alle sprekers in Emilia-Romagna hebben consequente intervocalische /s/-voicing, en alle sprekers in Toscane hebben een consequent fonologisch contrast tussen /s/ en /z/), maar de fonetische realisaties van jongere Toscaanse sprekers liggen veel dichterbij de fonetische realisaties van sprekers uit Emilia-Romagna dan de fonetische realisaties van oudere Toscaanse sprekers daarbij in de buurt liggen. Fonologisch gebruiken jongere en oudere Toscaanse sprekers echter hetzelfde systeem. Als we aannemen dat fonemen direct gelinkt zijn aan een fonetische realisatie is het onmogelijk om de gevonden patronen in de twee regio's te verklaren. In plaats daarvan moeten we aannemen dat fonologie en fonetiek tot op zekere hoogte aan elkaar gelinkt zijn.

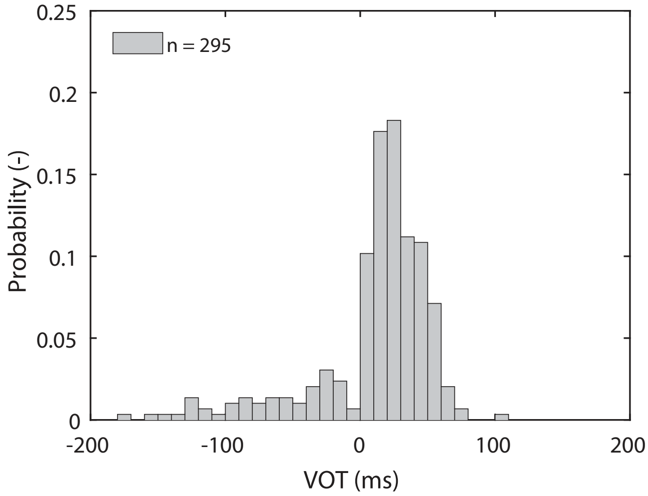
Een framework waarin de relatief losse link tussen fonemen en hun fonetische realisaties gerepresenteerd kan worden is BiPhon (Bidirectional Phonetics and Phonology, zie bijvoorbeeld Boersma (2009)). In dit theoretisch model zijn fonemen en fonetische waarden aan elkaar gerelateerd op dezelfde manier als onderliggende en oppervlaktevorm in Optimality Theory (OT) aan elkaar gerelateerd zijn. Door constraints op verschillende manieren te ranken, kunnen fonemen aan verschillende fonetische waardes gelinkt worden. Zo kunnen de variatie in het middengebied van het Nedersaksisch dialectcontinuüm, en de fonetische verschillen tussen jongere en oudere Toscaanse sprekers, verklaard worden.

Maps and grammar

Een groot verschil tussen de twee regio's is de aard van de overgang tussen de systemen. Waar de overgang in het Nedersaksisch gebied geleidelijk gaat, is deze in het Italiaanse gebied zeer abrupt. Hoe is dit verschil te verklaren? Een eerste factor die hierin een rol speelt is het taalsysteem zelf. Om een tweeledig fonologisch (of allofonisch) contrast te verwerven hebben kinderen een bimodale distributie van fonetische waardes in de input nodig. Kinderen in het westen en oosten van het Nedersaksisch continuüm vinden inderdaad zo'n bimodale distributie van VOT-waardes, en zullen zowel een lenis als een fortis plosief in hun foneeminventaris aannemen. In het middengebied vinden kinderen echter een heel andere distributie van VOT-waardes (Figuur 9).

In het middengebied laat de taalkundige input een unimodale distributie van VOT-waardes zien. Een taallerend kind vindt in deze input geen evidentie voor een fonologisch contrast, en zal in de foneeminventaris voor stemhebbendheid dus slechts één categorie aannemen. Daarnaast vindt het taallerend kind conflicten in de input: de status van fonologisch ongemarkeerde plosieven (die altijd met kort-positieve VOT gerealiseerd worden) verschilt tussen de twee systemen. In een voicingsysteem is de ongemarkeerde plosief fortis en heeft ze een lenis tegenhanger (met negatieve VOT), terwijl de ongemarkeerde plosief in een aspiratiesysteem lenis is en een fortis tegenhanger (met aspiratie) heeft. Ongemarkeerde plosieven in de input moeten dus af en toe geïnterpreteerd worden als fortis plosieven, en af en toe als lenis plosieven. Het taallerend kind vindt daardoor iedere keer evidentie voor verschillende systemen, die niet compatibel met elkaar zijn. Uiteindelijk is er daardoor onvoldoende evidentie in de taalkundige input om een consequent contrast tussen lenis en fortis plosieven aan te nemen, waardoor het taallerend kind uiteindelijk slechts één categorie plosieven aan kan nemen.

Kinderen in het gehele Italiaanse continuüm vinden een bimodale distributie van voicingpercentages, en zullen altijd zowel een [s] als een [z] in hun



Figuur 9: Distributie van VOT-waardes in de transitiezone

foneeminventaris aannemen. Het contrast tussen deze twee fricatieven is in de beide regio's immers hetzelfde, waardoor er geen conflicterende input in het contactgebied is. Het verschil tussen de regio's ligt in de distributie van de /s/ en de /z/: in de noordelijke dialecten is deze distributie allofonisch en dus voorspelbaar, terwijl de distributie in de niet-noordelijke dialecten fonologisch en dus onvoorspelbaar is. De constraint ranking die ten grondslag ligt aan de allofonische distributie van /s/ en /z/ in de noordelijke dialecten, kan niet zo gemuteerd worden dat /s/-voicing in slechts een deel van de contexten met een staminterne /s/ optreedt. Met andere woorden, een spreker kan alleen een constraint ranking hebben die met de noordelijke dialecten correspondeert, of een constraint ranking die met de niet-noordelijke dialecten correspondeert. Ook hier is dus, taalkundig gezien, geen tussenvorm tussen de twee systemen mogelijk.

Naast taalkundige factoren spelen geografische en politieke factoren mogelijk een rol. Ten eerste wordt het Italiaanse continuüm in tweeën verdeeld door een bergketen, waarvan de ligging redelijk overeenkomt met de locatie van de taalkundige grens. Dit kan echter niet de enige verklaring van de verschillen zijn, aangezien de bergketen contact tussen alle verschillende bergdorpen zou moeten bemoeilijken. Gezien de abrupte grens lijkt het er echter op dat contact alleen bemoeilijkt wordt tussen het zuidelijkste dorp in Emilia-Romagna en het noordelijkste dorp in Toscane, hoewel verscheidene dorpen in beide provincies zich ook in de bergen bevinden. De politieke grens tussen Emilia-Romagna en Toscane zou echter wel een grotere rol kunnen spelen. De taalkundige grens valt precies samen met de politieke grens tussen de twee provincies. Hoewel het Nedersaksisch continuüm ook in tweeën gedeeld wordt door een politieke grens, lijkt deze grens historisch gezien weinig invloed te hebben gehad op het taalkundig continuüm. In het Italiaanse gebied lijkt de invloed van de politieke grens veel sterker te zijn. De politieke oriëntatie van sprekers kan mogelijk een grote rol spelen in de abrupte overgang in het Italiaanse taalgebied.

In dit proefschrift is door middel van een studie van geografische variatiepatronen in taal een nieuw licht geworpen op taalvariatie. Een vergelijking van twee verschillende processen in twee verschillende regio's laat zien dat variatie niet alleen afhangt van extralinguïstische factoren, maar ook van de taalkundige factoren zelf. Daarnaast geeft de variatie die in de twee gebieden gevonden wordt meer inzicht in de interface tussen fonetiek en fonologie.

Curriculum Vitae

Nina Ouddeken was born on August 10, 1990 in Leiderdorp, the Netherlands. After completing her secondary education at het Gymnasium Felisenum (Velsen-Zuid) in 2008, she started studying general linguistics at Leiden University. During her bachelor, she had her first encounter with the fields of phonology and dialectology.

In 2011, she started studying the research master Linguistics at Leiden University. During this time she started specialising further in phonology and dialectology. In 2012, she did an internship at the Meertens Instituut under the supervision of Björn Köhnlein and Marc van Oostendorp. For this internship she worked for the project Taalportaal (`www.taalportaal.org`), where her interest in doing research was awakened. In 2013, she spent five months at the University of Edinburgh as part of a student exchange program. In the summer of 2013, she graduated cum laude, with a thesis entitled *Metaphony in Italian. An analysis in Element Theory.*

In September 2013, she started working as a PhD-student on the project Maps and Grammar at the Meertens Instituut. In September and October 2015, she spent six weeks in Italy to collect the data used in this study. This dissertation is the result of her research.