Incomplete neutralization and maintenance of phonological contrasts in varieties of Standard German

Felicitas Kleber

München 2011

Incomplete neutralization and maintenance of phonological contrasts in varieties of Standard German

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität
München

vorgelegt von
Felicitas Kleber

aus
München

München, den 22.03.2011

Referent: Prof. Dr. Jonathan Harrington

Korreferent: PD Dr. phil. habil. Philip Hoole

Tag der mündlichen Prüfung: 29.06.2011

Meiner Mutter

Acknowledgments

It’s a pleasure to thank all those people who have contributed to this thesis in one way or another.

Foremost, I would like to express my gratitude to my supervisor Jonathan Harrington for his advice and support over the past years. He drew my interest to sound change and challenged me to explore dialect change in German. His critical reading was invaluable. Additionally, he provided for numerous resources that allowed me to carry out my research.

I thank Anke Werani for her encouragement, advice, and interesting discussions about dialect and identity. Thanks are also due to Phil Hoole for his support; his door was always open.

I am grateful to my colleagues at the IPS for valuable comments during the institutes’ weekly colloquium, NOSH. In particular, I want to thank Ulrich Reubold and Tina John for being such great workmates in the DFG projects, for technical help with anything, discussion of the data, and for proofreading of the chapters of this thesis.

Many thanks are also due to Christoph Draxler who helped with carrying out the online perception experiments and with recordings using SpeechRecorder. A number of people were at some point involved in running the experiments and the pre-analysis of the data. My thanks (in alphabetical order) to Michaela Adlwart, Aysen Demir, Laura Folk, Benjamin Hamer, Regina Kaiser, Ann-Kathrin Killguss, Katharina Liebich, Veronika Neumeyer, Julia Orlitzky, Theresa Thoma, Dominique Vornwald, and Lisa Wälischmiller for recruitment and recording of subjects, data segmentation, and formant correction. I am also indebted to the 42 speakers and 19 listeners who participated in the experiments described in this thesis.

I want to thank all members of the PhD & PostDoc colloquium(s) for valuable comments and fruitful discussions. Uwe’s comments during the preparation of our seminars were very much appreciated. I’m grateful to Lasse, Susi, and Raphael for entertaining lunch and coffee breaks on a daily basis and to Katalin for office chats that always lasted longer than anticipated.

Special thanks go to Klemmy (from Kletty) for proofreading and sharing her native-dialect knowledge of Bavarian with me. Many thanks also to Annegret and Jessica for last minute proofreading and Anna for participating in thousand and one experiments.

Großer Dank gebührt meiner Familie; ohne die Unterstützung und das in mich gesetzte Vertrauen meiner Eltern, Dorothee und Gerd Kleber, und meines Bruders Christoph, hätte ich es nie so weit geschafft. Ihre Liebe und natürlich ihre Bereitschaft, jedes Experiment mitzumachen, andere Versuchspersonen zu rekrutieren und Literatur aus sächsischen Bibliotheken zu besorgen, bedeuten mir sehr viel. Jutta und Dietrich Ristow haben mich vor fast fünf Jahren in Neubiberg herzlich aufgenommen und mir einen unbeschwerten Start in München bereitet; dafür bin ich sehr dankbar.

Matthias danke ich für seine fortwährende Unterstützung und natürlich für die vielen Ablenkungen, die immer zur rechten Zeit kamen; für das alles und noch viel mehr.

Contents

List of Figures

List of Tables

Introduction

Chapter I

1.1 Contrast and neutralization

The production and perception of phonological contrasts is one major aspect of speech communication: contrasts help to decode and differentiate words and their various meanings. However, phonological contrasts can be neutralized depending on the phonetic context in which they occur or on the variety of a language. Neutralization is a phonological process in which the contrast between two different underlying representations is lost at the surface realization. One phonological contrast that is neutralized in a number of languages (e.g. Bulgarian, Catalan, Czech, Dutch, German, Polish, Russian, Turkish) is the voicing opposition. Voiced and voiceless obstruents in postvocalic position are frequently differentiated by means of various acoustic parameters like the vowel to consonant duration ratio or voice onset time. In word-medial, intervocalic or syllable-initial position, obstruents differ in voicing, but in domain-final[1] position, this contrast is neutralized with a bias in favour of the voiceless component, i.e. voiced obstruents are subject to the phonological process of fortition. The bias and the loss of the voicing contrast in this position are traditionally termed final devoicing and can be generalized as in the following phonological rule:

[+obstruent] → [-voice]/_db (where db stands for domain-boundary)

The classic example for demonstrating word final devoicing in German are the two words *Rad* ('wheel') and *Rat* ('advice') which differ in the underlying morphophonemic representation, but are homophonous on the surface structure: the voicing realization of the final stop in *Rad*

[1] The domains in which the voicing neutralization rule applies are still disputable, particularly, in the context of incomplete neutralization. For example, the morpheme boundary has been proposed as one possible domain (Wurzel, 1970; Kloeke, 1982; Iverson & Salmons, 2006, 2007; Jessen & Ringen, 2002). However, there is a broad consensus that the syllable boundary is the prosodically lowest domain where devoicing occurs, and that the rule is applied in word or utterance final position.

alternates depending on the morphological form (e.g. singular /raːt/ vs. plural /rɛːdɐ/) whereas the stop in *Rat* is always voiceless (singular /raːt/ vs. plural /rɛːtə/). In a number of Central and Upper German regional varieties the voicing contrast is also neutralized in intervocalic or even domain-initial position but in the opposite direction, i.e. with a bias towards the voiced (or more precisely lenis)[2] category. This process is called lenition and examples are *Pass* ('passport') and *Bass* ('bass') as well as *baten* ('requested') and *baden* ('to bath'), which are neutralized as [bas] and [baːdən][3], respectively, in various German dialects.

Lexical phonology (Kiparsky, 1982, 1985; Mohanan, 1986; Booij & Rubach, 1987) differentiates between phonological rules that are applied (and thus stored) in the lexicon (hence lexical rules) and those that are not (hence post-lexical rules). Neutralization can be derived either by post-lexical or lexical rules. An example for the latter case would be the neutralization of /d/ and /s/ when the affix *–ive* is added to a verb (e.g. *expand – expansive* or *include – inclusive*). That is, this particular contrast is only neutralized in a certain morphological environment and only in verbs but not nouns (e.g. *gerund – gerundive*), which means that lexical rules have access to grammatical categories (e.g. verbs vs. nouns). These kinds of neutralizations often involve a replacement of phonemes rather than an allophonic variation (Spencer, 1996) and thus may not appear to be neutralizations at first sight. Post-lexical rules, on the other hand are applied automatically (e.g. in loan words such as *job,* which is realized as [dʒɔp] in German; that is, the final stop is devoiced) and outside the lexicon (Spencer, 1996; Gussenhoven & Jacobs, 1998). In lexical phonology, final devoicing and dialectal lenition are both considered post-lexical processes since there are no exceptions (e.g. irrespective of the grammatical category) and native speakers' intuition regarding the surface realization are often incorrect, e.g. native speakers regard segments such as the last consonant in *Rad* and *Rat* as

[2] The phonological voicing contrast is also described as a lenis vs. fortis contrast. In particular, the terms lenis and fortis are used to describe the intervocalic voicing contrast in German and regional varieties of German. For a discussion on the advantages of a [±fortis] feature see Kohler (1984). In this thesis, we differentiate between underlying voiced versus voiceless stops in final position, because this concept is predominantly used in the phonetic literature on final devoicing. However, when we refer to intervocalic stops we use the fortis/lenis dichotomy commensurate with the literature on German dialects. Additional information on the use of terminology is given in the respective chapters.

[3] Dialect realizations are denoted by square brackets, but are only broadly transcribed, i.e. irrelevant realization subtleties were ignored and not indicated by diacritics.

different although they are phonetically identical (or at least very similar, Gussenhoven & Jacobs, 1998: 122).

Within phonological theory, phonological processes in general have long been considered to cause a structural, i.e. categorical change at the surface realization and German final devoicing has been (probably just as) long regarded as a prototypical example for the phonological neutralization process of the binary [±voice] feature. Most dialectologists probably also consider the voicing contrast in intervocalic position to be completely neutralized in those dialects in which the so-called Central German lenition rule applies. In the past three decades, however, experimental phonetic studies have presented increasing evidence that the domain-final voicing contrast is not completely neutralized in Standard German but partially maintained (e.g. Port, Mitleb & O'Dell, 1981; Port & O'Dell, 1985; Port & Crawford, 1989; Charles-Luce, 1985, 1993).[4] Nevertheless, contrasts may not only be incompletely neutralized, they may also be incompletely maintained. This dissertation investigates whether the domain-final voicing contrast is (in)completely neutralized in the perception of Standard German and whether the intervocalic voicing opposition is (in)completely maintained in the speech of speakers from two regional varieties of German. Furthermore, we examine whether the degree of contrast maintenance is conditioned by linguistic factors (such as phonological frequency) and extra-linguistic factors (such as age and regional background). Extra-linguistic factors are included in this investigation because we are interested whether the degree of contrast maintenance in regional varieties is also an indicator for a sound or language change in progress.

In this first chapter we will give an overview of the literature on incomplete neutralization (1.2) as well as short introductions to the linguistic factors (1.3), the two regional varieties of German that are investigated (1.4) and to sound change research (1.5). The two latter paragraphs are important to understand why the extra-linguistic factors regional background and age were included. Literature that is relevant to the particular hypotheses tested in each of the three empirical studies of this thesis will be given in the introduction paragraphs

[4] Lexical phonology may under certain circumstances account for partial neutralization (Clark & Yallop, 1990: 350).

of the respective chapters. In 1.6 we will outline the structure of the thesis and present the empirical basis for the acoustic analyses in this dissertation.

1.2 Incomplete neutralization

The completeness of the neutralization process was called into question when Port and colleagues found word-final underlying voiced versus voiceless obstruents in minimal pairs to differ on various acoustic parameters that are used to cue the voicing contrast (Port et al., 1981; O'Dell & Port, 1983; Port & O'Dell, 1985): derived voiceless (i.e. underlying voiced) obstruents were acoustically – particularly with respect to the duration of the preceding vowel – somewhere in between voiced and voiceless obstruents (although they were closer to the voiceless obstruents). Port and colleagues called the process yielding this "semi-voicing-contrast" in word-final position "incomplete neutralization". Since then, evidence for incomplete neutralization has been presented for a number of final-devoicing languages such as Catalan (Dinnsen & Charles-Luce, 1984) and Polish (Slowiaczek & Dinnsen, 1985), but not for all. There are, for example, contradictory findings for Dutch, which is also a final-devoicing language: Warner, Jongman, Sereno, and Kemps (2004), Warner, Good, Jongman, and Sereno (2006) and Ernestus and Baayen (2006) found incomplete neutralization in their data, whereas Jongman, Sereno, Raaijmakers, and Lahiri (1992) and Baumann (1995) reported complete neutralization. The extent to which a contrast is neutralized was shown to depend on semantic information, context and phonetic environment (Charles-Luce, 1985, 1993; see below 1.3) as well as dialectal background (Piroth & Janker, 2004). These studies have been taken as evidence that neutralization is not a categorical but instead a gradual process.

Two of the first and persisting main points of criticism have been orthographic influence and speaking style causing the differences in the results to be an experimental artefact. Fourakis and Iverson (1984) purported that incomplete neutralization comes about because the data are obtained in an artificial setting which causes speakers to hyperarticulate, in particular when these languages maintain the underlying voicing difference in orthography. Various studies have shown that the degree of neutralization depends on speaking style (Port & Crawford, 1989) and the experimental setting: in laboratory speech, the contrast was found to be incompletely neutralized by speakers who had shown no evidence that the voicing feature in domain-final position was incompletely neutralized as in previous recordings of (semi-) spontaneous speech (Fourakis & Iverson, 1984 for German; Jassem & Richter, 1989 for Polish). However, since languages like Turkish and Catalan, in which the morphophonemic voicing feature is not represented on the orthographical level, show both complete (e.g. Turkish, cf. Kopkalli, 1993) and incomplete neutralization (Catalan, cf. Dinnsen & Charles-Luce, 1984), the explanation of orthographic interference is not sufficient. Warner et al. (2004, 2006) found small yet persistent voicing differences in words that differed only in the orthography but not in underlying voicing differences and argued in favour of orthography causing incomplete neutralization effects. Ernestus und Baayen (2006) also presented evidence for the spelling-induced incomplete neutralization in Dutch but argued against this being an experimental artefact, since their speaker also produced differences when spelling was kept identical for obstruents differing in the voicing feature. They interpreted this finding as incomplete neutralization being relevant for the morphological process of past-tense formation in Dutch. All these studies have provided extensive evidence that phonological neutralization in production is not categorical but gradual and that the degree of voicing distinction is influenced by factors such as speaking style.

Subsequent perception tests revealed that listeners discriminated words with underlying voiced and underlying voiceless stops better than chance (Port & O'Dell, 1985; Port & Crawford, 1989 for German; Slowiaczek & Szymanska, 1989 for Polish). However, the authors argued that the discrimination ability served no communicative function because word identification performance was poorer than expected and concluded that the contrast was "perceptually neutralized" (Slowiaczek & Szymanska, 1989: 211). Nevertheless, listeners can make use of the acoustic cues in order to identify the intended minimal word pairs when forced, and one interpretation of the findings was that the discrimination performance is rather a function of the acoustic information provided in the speech signal than the native listeners' knowledge of phonological rules. The difficulty with these perception studies is that they used naturally produced stimuli which they obtained from the production experiments described above. In other words, the stimuli contained all sorts of acoustic cues that could be used perceptually by listeners and which they did not control for, i.e. no conclusions can be drawn as to which acoustic cues listeners actually exploit for word identification[5]. Thus, the first aim of this thesis is to investigate systematically the final voicing contrast in perception and to test for linguistic effects such as phonological frequency and the potential for resyllabification in order to better understand the implication of incomplete neutralization.

1.3 Linguistic factors

As was mentioned above, linguistic factors such as phonetic environment (Charles-Luce, 1985) or semantic information (Charles-Luce, 1993) affect the degree of incomplete neutralization. Charles-Luce (1985) showed for German that the position of the target word in a clause and the phonetic environment influenced whether the voicing contrast was completely neutralized: only clause-final obstruents differed significantly, either in vowel duration (only for fricatives) or in voicing into closure (only for stops that are followed by a clause starting with a vowel). Piroth and Jancker (2004) found incomplete neutralization in utterance-final position (and only in the production of Southern German speakers) but neither in syllable- nor word-final position.

We extend the research to other potential internal factors such as phonological frequency of certain vowel plus stop combinations or the potential for resyllabification. The latter factor may influence the degree of contrast maintenance insofar as a syllable-final obstruent in a disyllabic word may be resyllabified with the second syllable leading to the originally final consonant being in onset position in which the voicing contrast is maintained in Standard German. Giegerich (1992), for instance, proposes that "domain-final consonants are

[5] Port & Crawford (1989: 272) reported on results as revealed by discriminant analysis, which suggested that the differences found for obstruents differing in underlying voicing were "distributed across most of the spectro-temporal variables [vowel duration, stop duration, burst intensity and burst duration] in the word".

constituents defined by a generalised version of the Maximum Onset Principle […] and that they become onsets wherever possible and codas only by default" (p. 135). For example, /dl/ is an illegal onset cluster in Standard German and thus /d/ in *Adler* (/aːtlɐ/, 'eagle') cannot be resyllabified with the onset of the second syllable. Consequently, /d/ occurs in syllable-final position and is realized as [t] in *Adler*.[6] On the other hand, /d/ in *Adria* (/aːdria/, 'Adriatic Sea') may be resyllabified with the second syllable's onset because /dr/ is a legal onset cluster in German and thus realized as [d].[7] The former factor is important, since recent studies have shown that the frequency with which words or combinations of sounds occur in a language's lexicon significantly affect speech perception and production: for example, phonotactically frequent combinations are more accurately identified than rare combinations (Vitevitch & Luce, 1999) which may also be reanalyzed perceptually as more frequent combinations (Hay, Pierrehumbert, & Beckmann, 2003; Pitt & McQueen, 1998); and vowels in frequent words with fewer lexical competitors tend to be produced with a more centralized vowel quality than rare words with more lexical competitors (Wright, 2003). In usage-based models of speech perception (and to some extent also speech production, cf. Pierrehumbert, 2001) such as the exemplar theory (Pierrehumbert, 2001, 2002, 2003), phonological categories are defined by the density distribution of all instances of stored exemplars in an acoustic-perceptual space. Thus, phonological frequency forms an intrinsic part of these models. The exemplar theory was found to model gradient differences between categories more easily than, for example, phonological theories within a generative framework (e.g. Chomsky & Halle, 1968).

The effect of syllable position is examined in the first study on the perception of final devoicing (chapter II). Whether phonological frequency affects the degree of neutralization is tested in both the first study as well as the second investigation of the intervocalic voicing contrast in two regional varieties of German (chapter III).

Final devoicing is the prototypical example for the phonological process of phonemic contrast neutralization; research on whether it is incompletely neutralized has been extensively

[6] Note that *Adler* is transcribed as /aːdlɐ/ in the German pronunciation dictionary *Duden* (Mangold, 1990).
[7] Examples are taken from Giegerich (1992).

conducted. Of course, there are other neutralization processes, such as the neutralization of the vowel length contrast in Dutch (e.g. Booij, 1981; van der Hulst, 1985) or intervocalic lenition in German varieties, that might also be only partially neutralized. The second aim of this dissertation is to explore whether dialectal lenition is similarly prone to incomplete neutralization. The next paragraph provides an overview of the dialects of German and the varieties analyzed in the present thesis.

1.4 The regional varieties of German

Traditionally, there is a three-way division into Low, Central (or Middle)[8] and Upper German dialects based upon where and to what extent the Second or (Old) High German sound shift (A.D. 600-800) has taken place (for a comprehensive overview, see Schmidt, 2000 and Barbour & Stevenson, 1990). Low German dialects, which are spoken in the northern part of Germany, were not affected by any of these sound shifts and are not part of the so-called High German dialect area. Central and Upper German varieties can be grouped together as High German[9] dialects because the ancestors of these modern dialects were influenced by the High German sound shift. However, the scope of the sound shift was different for the medieval High German dialects. Upper German dialects were almost completely affected by the Second German sound shift: for example, both /p/ and /t/ changed diachronically to the affricates /pf, ts/ in medial position. On the other hand, not all sound changes have occurred in Central German dialects: e.g. in medial position /t/ became /ts/, but /p/ was not shifted to /pf/. Thus, Old Saxon *settian* ('to set') became *setzan* during the Old High German period in the East Central German area, but Old Saxon *aphul* ('apple') remained *aphul* and is still realized as [abəl] in present-day Upper Saxon. The Central German dialects can be further subdivided into the West and East

[8] Since the term Middle German is also used to refer to the historical period between the 11th and 15th century in which for example Middle High German was spoken, we only use the term Central German, because it is unambiguously used to indicate the geographical area in which these dialects are spoken.

[9] The term High German also refers to Standard German as it is transcribed in the German pronunciation dictionary *Duden* (Mangold, 1990). This is because Standard German has emerged from the High German and not the Low German dialects. E.g. Standard German shows all historic changes that are connected to the High German sound shift.

Central German dialects (*'Westmittel-* and *Ostmitteldeutsch'*). The Upper German dialect area comprises three major dialect areas: East Franconian (*'Ostfränkisch'*), Alemanic (*'Alemanisch'*) and Bavarian-Austrian (*'Mittelbairisch*'). The dialect divisions are shown in Figure 1.1.

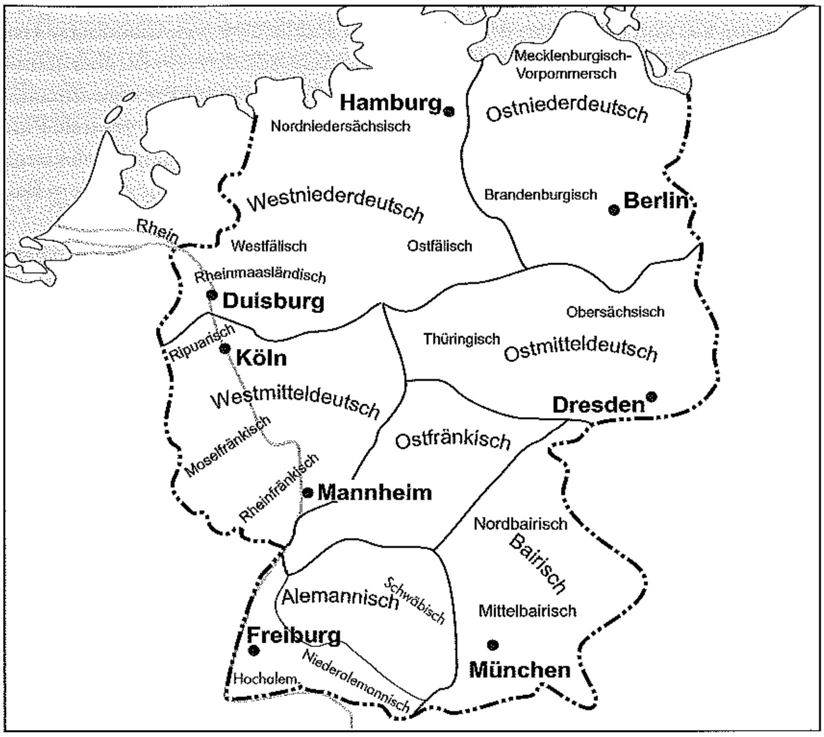


Figure 1.1: Conventional divisions of German dialects. (Source: Gilles, 2005)

German has many small local dialects which were spoken by all social groups within a geographical area (as opposed to English, in which social factors play a greater role, e.g. Barbour & Stevenson, 1990: 135-137). There is a vast amount of literature on these (small) German dialects but almost entirely from the perspective of German Linguistics, in which the phonological descriptions of dialects are based on auditory analyses and focus on historical developments within and across German varieties; by contrast, there are only very few instrumental phonetic analyses (e.g. the phonology of Central Bavarian by Bannert, 1976). These linguistic studies describe broad dialects spoken by old, immobile speakers from rural areas. Research has only recently been concerned with regional varieties spoken by young speakers of larger geographical areas and cities that are highly influenced by the standard variety (e.g., Schlobinski, 1987; Lameli, 2004; but again only on the auditory-phonological level).

The two regional varieties that are investigated in this dissertation are Upper Saxon and Central Bavarian. Both dialects show neutralization patterns of Standard German phonological contrasts and are High German dialects. At the same time the two dialects are very different (because Saxon and Bavarian belong to the Central and Upper German dialects, respectively) with regard to their phonological systems and their prestige. In the following we introduce the dialects, present the most important dialect features that have been described for both the core dialects as well as the regional standard variants, and illustrate the extent to which dialect vs. vernacular is produced in these regions.

Most phonological differences described in the following paragraphs are systemic according to Wells' (1982) classification of phonological differences between regional accents (i.e. varieties in our terminology) because there is not a one-to-one correspondence between the varieties' phoneme inventories. However, it is important to note that not all realizations of the same Standard German phoneme are affected. This is because the dialectal pronunciation depends on the underlying Middle High German[10] (henceforth MHG) realization of a sound. For example, the diphthong /aʊ/ in Standard German *Haus* (/haʊs/, 'house') is derived from the MHG monophthong /uː/ (MHG *hûs*)[11], but /aʊ/ in Standard German *Baum* (/baʊm/, 'tree') is derived from the MHG diphthong /ou/ (MHG *boum*). Thus, the difference between Standard German and, for example, Upper Saxon is systemic since all /aʊ/-diphthongs that can be traced back to MHG /ou/ are produced as [oː][12] in Upper Saxon. However, we cannot generalize that Standard German /aʊ/ is always pronounced as [oː] or [oʊ] in Saxon, because only MHG /uː/ has changed in Upper Saxon to /aʊ/ but not MHG /ou/, and thus *Baum* and *Haus* are realized as [boːm] and [haʊs], respectively. Therefore, most dialect descriptions (e.g. Bergmann, 1990; Wiesinger, 1990) include the MHG pronunciation as a reference in order to make clear which

[10] The term Middle High German refers to the German language spoken between 1050-1350.
[11] MHG transcriptions are taken from Schmidt (2000).
[12] The diacritic ː is used after all vowel phonemes that are considered phonologically long in German in order to enhance the contrast to the phonologically short vowel phonemes.

sounds differ between the Standard and the regional pronunciation. Nevertheless, we only compare dialectal and Standard pronunciations without referring to the MHG sounds, because the aim of the dialect overviews below is to illustrate the extent to which the dialects investigated in this thesis differ from Standard German, especially with respect to the features analyzed and transcriptions provided in chapters III and IV (see below 1.5).

1.4.1 Upper Saxon (Obersächsisch)

Upper Saxon is spoken in most parts of the German state of Saxony and partly in the neighbouring states of Thuringia, Saxony-Anhalt and Brandenburg. Upper Saxon is one of the East Central German dialects. The dialect area is rather heterogeneous considering the numerous isoglosses running from West to East (Bergman, 1990). Among the smaller Upper Saxon dialects are varieties spoken at the geographical peripheries of Saxony such as the Vogtland and the Erzgebirge region in the Southwest, Lusatia in the East as well as parts of the Osterland and the Elbe-Elster area in the North (see Bergmann, 1990). Each of those varieties are again influenced by their ambient dialect region; for example, the dialect spoken in Vogtland contains quite a number of Franconian (i.e. Upper German) dialect features and the variety of the Elbe-Elster area exhibits numerous Low-German characteristics. Lusatia has to be considered exceptional in that Sorbian, a Slavic language, was spoken by a majority of the population until the beginning of the 20th century and parts of the population are bilingual up to the present day: that is, there is a strong Slavic influence. The core Upper Saxon dialect region can be located on a north south axis from Leipzig to Dresden. In this thesis, we will investigate the regional variety spoken in the Dresden area.

The description of the regional varieties spoken in East Central Germany is considered a research desideratum (Dingeldein, 1997). The first and most comprehensive study is that of Bergmann and Becker (1969). The underlying data mostly comes from transcribed interviews, i.e., we are dealing with auditory analyses of spontaneous speech. The Upper Saxon variety is very similar to Standard German, but contains a small number of nonstandard features (cf. 1.4.3

below). The main dialect features are (cf. Bergmann & Becker, 1969; Bergmann, 1990; Auer, Barden, & Großkopf, 1993; Barden & Großkopf, 1998; Rues, Redecker, Koch, Wallraff, & Simpson, 2007):

- The diphthongs /aɪ/ and /aʊ/ are realized as [eː] and [oː], respectively. Thus, Standard German *Bein* (/baɪn/, *'leg'*) and *Baum* (/baʊm/, 'tree') become [beːn] and [boːm].
- Derounding of the front rounded vowels /yː, ʏ, ø, œ/. Standard German *Hüte* (/hyːtə/, 'hats'), *Hütte* (/hʏtə/, 'hut'), *Höhle* (/hølə/, 'cave'), and *Hölle* (/hœlə/, 'hell') are realized as [hiːdə], [hɪdə], [helə], and [hɛlə], respectively.
- Tendency towards centralization of all vowels except /ɛː/ and /ɛ/.
- More open realization of /ɛː/ and /ɛ/.
- Lenition of the voiceless stops in word-medial, i.e. intervocalic, and initial, i.e. prevocalic, position. E.g. *passen* (/pasən/, 'to fit'), *Tasse* (/tasə/, 'cup'), and *Kasse* (/kasə/, 'cash point') are realized as [basən], [dasə] and [gasə].
- Spirantization of the voiced labial stop [b] to the voiced labiodental fricative [v] and the voiced velar plosive [g] to the voiceless velar fricative [ç] after front vowels or [x] after back vowel. For example, Standard German *leben* (/leːbən/, 'to live'), *legen* (/leːgən/, 'to put'), and *lagen* (/laːgən /, 'lay') may be realized as [leːvən], [leːçən], and [laːxən] in Upper Saxon.
- The affricate /pf/ is realized as [b] in medial (e.g. Standard German *Apfel,* /apfəl/, 'apple', Upper Saxon [abəl]) or [p] in final position (e.g. Standard German *Strumpf,* /ʃtrʊmpf/, 'sock', Upper Saxon [ʃdrʊmp]) and as [f] in initial position (e.g. Standard German *Pferd,* /pfeːɐt/, 'horse', Upper Saxon [feːɐt]).
- The contrast between [ç] and [ʃ] is often neutralized towards the alveolo-palatalized fricative as, for example, in Standard German *technisch* (/tɛçnɪʃ/ 'technical') which is realized as [dɛɕnɪɕ] (Auer et al., 1993; Rues et al., 2007)

1.4.2 Central Bavarian (Mittelbairisch)

Central Bavarian is an Upper German dialect and, more specifically, belongs to the Austro-Bavarian dialects (cf. Figure 1.1). It is spoken in Upper and Lower Bavaria, in the South of the Upper Palatinate, in Upper and Lower Austria, in Vienna and the Flachgau, which is the northernmost administrative district of the state of Salzburg. Central Bavarian may be subdivided into a Western and an Eastern part with Munich and Vienna, respectively, as the centres. Standard German and Central Bavarian differ in numerous dialect features (as opposed to Upper Saxon). In what follows only a selection of important dialect features of the Eastern part of Central Bavarian are presented. For a comprehensive overview of the Central Bavarian dialect features see Wiesinger (1990). The main differences are (unless otherwise indicated all examples and transcriptions[13] are taken from Wiesinger, 1990):

- Diphthongization of tense /iː, yː, uː/: e.g. Standard German *lieb* (/liːp/, 'dear'), *müde* (/myːdə/, 'tired'), and *gut* (/guːt/, 'good') are realized as [liɐb], [miɐd], and [guɐd][14].
- Derounding of the front rounded vowels /yː, ʏ, øː, œ/. Standard German *Zügel* (/tsyːgəl/, 'rein'), *Schüssel* (/ʃʏsəl/, 'dish'), *mögen* (/møːgən/, 'to like'), *Böcke* (/bœkə/, 'billy-goats') are realized as [dziːgə], [ʒisl], [meː(g)ŋ], and [bek], respectively.
- The diphthong /aʊ/ may be realized as [aː] before labials, e.g. Standard German *Baum* (/baʊm/, 'tree') is realized as [baːm].
- /a, aː/ are retracted and raised to [ɔ, ɔː] or [ɒ] but only in native German and old loanwords. For example, *Platte* (/platə/, 'plate'), *Hase* (/haːzə/, 'hare'), and *Mantel* (/mantəl/, 'coat') are pronounced as [blɔtn], [hɔːz], and [mɒntl], respectively.
- More recent loanwords that were borrowed after the 16th century contain open, front [a, aː], e.g. *Lack* (/lak/, 'varnish') or *brav* (/braːf/, 'well-behaved') are realized as [lak] and [braːv][15].

[13] We do not necessarily agree with the transcriptions, in particular with respect to voicing in initial and final position. However, we quote these transcriptions as we have no empiric evidence for realizations differing from those transcribed in Wiesinger (1990).
[14] Stops after long vowels may be realized as lenis in syllable-final position in Central Bavarian.

- Standard German /ɛː/ is lowered to front [aː] (possibly because of the retraction and raising of <a>), e.g. Standard German *Käse* (/kɛːzə/, 'cheese') and *Gläschen* (/glɛːsçən/, 'small glass') are realized as [gaːz] and [glaːsl].

- Standard German lax, mid-open /ɛ/ is realized as [e, ɛ, a] depending on the underlying MHG realization. Thus, Standard German *Bett* (/bɛt/, 'bed'), *beten* (/beːtən/, 'to pray') and *Blättlein* (/blɛtlaɪn/, 'small leaf') may be realized as [bet] (MHG *bette*), [bɛtn] (MHG *bëten*), and [blatɐl] (MHG *blätelîn*).

- Elision of <e> in unstressed final syllables, e.g. *Hase* (/haːzə/, 'hare') and *müde* (/myːdə/, 'tired') are [hɔːz] and [miɐd], respectively.

- Vowel length is considered to be allophonic: vowels are long before lenis stops and short before voiceless stops. In Central Bavarian, domain-final stops may be realized as lenis stops when the preceding vowel is long.

- Lenition of initial voiceless obstruents /p, t/ to [b, d], e.g. *packen* (/pakən/, 'to pack'; König & Renn, 2005) and *Tag* (/taːk/, 'day') are realized as [bɔkən] and [dɔː(x)].

- Spirantization of the voiced labial stop /b/ in intervocalic position and before /l/, e.g. Standard German *Kübel* (/kyːbəl/, 'bucket') becomes [ghiːwe].

- Standard German /s/ becomes [ʒ, ʃ] after /r/, e.g. *Durst* (/dʊɐst/, 'thirst') is realized as /dʊɐʃt/.

- Laterals are vocalized after vowels, e.g. *Tal* (/taːl/, 'valley') is realized as [dɔɪ].

- Word-final nasals are deleted and the preceding vowels are nasalized, e.g. *Mann* (/man/, 'man') becomes [mɒ̃].

[15] The IPA symbol /a/ is generally used for the transcription of Standard German /a, aː/ which is phonetically a central vowel (cf. Harrington, Hoole, Kleber, & Reubold, 2011). Central Bavarian /a, aː/, however, is indeed a phonetically front vowel.

Both varieties exhibit dialect features in which a contrast is neutralized that is phonemic in Standard German: the vowel length contrast in Bavarian and the voicing distinction in Upper Saxon. One important factor that has to be taken into account in dialect research (especially in German) is the use of a given dialect, which we will deal with in the next paragraph.

1.4.3 The sociolinguistic situation in Saxony and Bavaria

Sociolinguistically, we have to differentiate between various levels of language use located along a continuum from the basic dialect to a regional colloquial variety to the standard language (Veith, 2005). Whether we tend to use a more dialectal or standard pronunciation depends on the communicative situation (e.g. in the family or at the office; see Wiesinger, 1990: 443-447 for an overview of the distinction between various sociolinguistic levels of regional varieties). The basic dialect is spoken in a very small area and the communicative range is restricted. The standard language, on the other hand, has a large communicative range, as most speakers, irrespective of the regional background, understand the Standard language, yet only very few speakers produce it. Most speakers speak a colloquial variety along the continuum. The number of speakers who (and the degree to which they) speak a basic dialect or standard language has changed in accordance with the development of the standard language. Since Standard German is the language of the school and media, it is superseding dialects more and more (Besch, 1983). Nowadays, a decreasing number of language users speak a basic dialect and an increasing number speak a standard variety. However, Standard German pronunciation is almost always influenced by the regional background of a speaker. The dialectal areas in Germany also differ as to the extent to which a dialect or a regional variety is spoken. In Northern Germany, on the one hand, basic dialects are rarely spoken, whereas in the South, basic dialects are still frequently spoken.

In Saxon, no clear division between dialect and vernacular speech can be made. For almost 400 years since the 15th century, Saxon has conditioned the development of the standard or literary German language in many ways (e.g. Luther translated the bible into German as it was spoken in Dresden-Meißen). As a consequence, Saxon vernacular emerged quite early (in the 16th century) from the literary language. Nowadays, the Upper Saxon dialect is hardly ever spoken, but there is instead a colloquial variety that is very similar to Standard German (Bergmann, 1990). Consequently, Saxon speakers do not switch between dialect and Standard German depending on the communicative situation, but almost always speak in their colloquial variety.[16] Nevertheless, this colloquial variety is far from being an entity but ranges from strongly regionally-coloured colloquial to regionally-coloured standard speech from which Saxon speakers choose depending on the communicative situation. Among the German dialects, Upper Saxon is nowadays regarded a low prestige variety with negative connotations even by Saxon speakers (Barden & Großkopf, 1998; Bergmann, 1990). Since German reunification, Saxon has in particular been identified as the main variety of the former German Democratic Republic, which has negative connotations for many first language German speakers. So it is possible that Saxon speakers avoid the use of dialect features, especially if they are being recorded or come into contact with speakers of another variety with a higher prestige. Furthermore, younger speakers may have lost more dialect features than older speakers of Saxon due to the low prestige, higher speaker mobility after the reunification and an increasing influence of the Standard language (Barden & Großkopf, 1998). Recent studies have provided evidence that speakers accommodate to another variety not only over a long term (Trudgill, 1986; Kerswill, 1986; Barden & Großkopf, 1998) but adapt even after very short exposure to another variety (Giles, 1973; Evans & Iverson, 2007; Delvaux & Soquet, 2007).

The use of the primary dialect in Bavaria is more widespread than in Saxony. That is, the Bavarian primary dialect, Central Bavarian, is still spoken especially in the rural areas; and speakers produce either the dialect or a regional variety of Standard German depending on the communicative situation (Wiesinger, 1990). One reason for this is, that the development of a colloquial speech has emerged later (in the 19th century) than in Saxon (Besch, 1983). Those

[16] However, because Standard German speakers with a Saxon accent are easily recognizable, native speakers of German would probably refer to present day Saxon as a true dialect such as Bavarian.

Bavarian speakers who primarily speak dialect may switch to a regional variety of Standard German in a recording situation.

There are a substantial number of studies on dialect change and levelling in German (Auer, 1988; Auer, Barden & Großkopf, 1993, 1997). Dialect levelling is defined as the "reduction or attrition of marked variants" (Trudgill, 1986) which, following Kerswill (2003), results from speech accommodation (Giles & Powesland, 1997), a "social psychological mechanism" (Kerswill, 2003: 223) that includes, for example, the prestige of a dialect. Speech accommodation can occur on both a horizontal and/or a vertical level: in the former one dialect influences another dialect and the latter describes the influence of a standard variety on a dialect. As can be seen from the overviews of the two dialects above, dialects are mostly described in terms of categorical divergences from the standard language and dialect levelling has often been measured by counting standard divergent features over the course of the years (e.g. Wagener, 2002). For example, Lameli (2004), in his apparent time study of Rhine Franconian speakers "intend[ing] to use standard language" (p. 253), presented evidence for a decreasing number of dialect features, and Barden and Großkopf (1998) found that there was a greater loss of dialect features in younger than older Saxon speakers who had lived in another German dialect area for a longer period of time.

A third aim of this thesis is to test whether speakers from the two varieties not only incompletely neutralize phonemic contrasts but also whether younger speakers neutralize the contrast to a lesser extent than older speakers because of dialect levelling. In spite of the new information from the studies on dialect change, there are – in contrast to English – not enough acoustic or empirically based studies of dialect levelling and sound change in German. In the next section, we report on theories and experimental phonetics in sound change research.

1.5 Sound change

Historical linguistics is a branch of linguistics concerned with systematic phonological, morphological, and syntactic changes within and between languages over time. One major aim of modern historical linguistics is to compare and identify genetically related languages (thus belonging to one language family), to reveal earlier stages of a language (which are frequently unattested) within a language family by means of comparing languages, and to reconstruct so-called proto-languages which are considered to be the precursors of modern languages (Rankin, 2003). The objects of study within historical linguistics were most often earlier stages of a language that survived only in written documents and were no longer spoken, and the language changes described were completed, e.g. the Great Vowel Shift in England (15^{th}-17^{th} century, see e.g. Lass, 1997). Since the 1960s, Labov and colleagues (Weinreich, Labov, & Herzog, 1968; Labov, Yaeger, & Steiner, 1972) have expanded the study of language and sound change to include changes that are still in progress and to analyses of internal or external factors (Labov, 1994, 2001) that have an influence on language and sound changes. The use of experimental techniques in sociophonetic investigations has provided a new way of examining a variety of diachronic changes and especially vowel changes that are in progress in various English varieties: e.g. the Northern Cities Shift (Labov et al., 1972; Labov, 1991, 1994; Labov, Ash, & Boberg, 2006) and the Southern Shift (Labov et al., 1972; Labov, 1991, 1994; Labov & Ash, 1997; Fridland, 1998, 2000) in American English, as well as the vowel shifts in Australian (Cox, 1999) and New Zealand English (Langstrof, 2006; Maclagan & Hay, 2007). These ongoing sound changes can be examined either in longitudinal studies, in which data from two points in time by the same speaker are analyzed (Harrington, Palethorpe & Watson, 2000a, b, Harrington, 2007), or in apparent time studies, in which the speech data from two age groups at the same time are compared. However, only very few investigations of current sound changes in progress using experimental phonetic methods have looked at non-vocalic changes (e.g. Maclagan, Watson, Harlow, King, & Keegan, 2009).

External factors for linguistic change are language or dialect contact due to, for example, conquests (e.g. the Norman conquest in 1066 and its effect on the development of the English language, Baugh & Cable, 2002), economic reasons (e.g. the influence of Low German on the Scandinavian languages during the Hanseatic era in the Late Middle Ages) or speaker mobility (e.g. Clopper & Pisoni, 2006). So-called 'extra-linguistic' factors have been considered to be independent of external factors only fairly recently: they include "sociopolitical and economic" (Farrar & Jones, 2002: 1) as well as "social-psychological factors, especially identities and attitudes" (Torgensen & Kerswill, 2004).[17] Thus the prestige of a community and their language may also be motivation for language change. Other linguistic changes, such as vowel shifts, splits and mergers, are motivated by the existing language system itself and thus called internal or even "natural" (Torgensen & Kerswill, 2004). Labov (1994: 117) observes three principles for internally motivated chain shifts based on evidence from the literature and, following Martinet (1955), gives a functional explanation for a chain shift: "chain shifts reflect the functional economy of the vowel system: vowels move together to avoid merger and preserve their capacity to distinguish words" (p. 117). Watson, Maclagan, and Harrington (2000) presented evidence for various changes in the New Zealand vowel system that were led by /ɪ/ and argued that "/i/ changed in quality not only because of crowding in the front vowel space, but also it would be less likely misperceived as an unaccented vowel" (p. 51). Whether the trigger for sound change is internally motivated or externally driven is a complex issue (Croft, 2000; McMahon, 1994; Torgersen & Kerswill, 2004).

There are phonetic models that provide phonetic explanations for both vocalic and consonantal sound changes (without actually investigating ongoing sound changes) in order to answer the crucial question: why do sound systems that have been stable for quite a while begin to change? Browman and Goldstein (1988, 1990, 1991), for example, present evidence for overlapping articulatory gestures that result in a masking of one of the segments in the acoustic signal that is then perceptually undetectable. These forms of perceptual elision may lead to sound change. Ohala (1974, 1981, 1983, 1989, 1990, 1993, 2003) has developed a theory of sound change in which the synchronic variability in speech production can be perceptually ambiguous. Speech production is highly variable because of the coarticulation of adjacent sounds. First-language speakers have learned to factor this context-dependent variability out

[17] This definition of 'extra-linguistic factors' is not to be confused with our broader use of the term 'extra-linguistic factors' as described in 1.1 (which refers to age and dialect group).

which is why sound change occurs rarely. However, children, second language learners, or even adults may fail to attribute certain acoustic characteristics to the context but instead assign the context-dependent features to the modified sound. That means that misperception is a potential trigger for sound change, but not all perception errors lead to sound changes if they do not spread throughout the community. And even if a listener fails to correct his/her perceptual errors and proceeds to modify his/her pronunciation norm, this mini sound change does not necessarily spread and turn into a maxi sound change. It may instead remain an idiosyncratic pronunciation of the speaker. Thus a mis-perception of coarticulatory relationships only rarely results in a mini sound change that then spreads throughout the community. The transition from a mini to a maxi sound change is completed once a pronunciation form that had previously been characterized by certain acoustic characteristics only in certain contextual environments has become phonologized (Hyman, 1976), i.e. the pronunciation form is now found in all environments. According to Ohala, the communicative 'error' is on the side of the listener – be it in the form of hypocorrection as in assimilative sound changes or in the form of hypercorrection as in dissimilative sound changes (Ohala, 2003). Furthermore, Ohala claims that sound change is non-teleological: since the origin of sound change is an error, then by definition it cannot be planned. Lindblom, Guion, Hura, Moon, & Willerman (1995), on the other hand, argue that sound change is adaptive and that its origin lies in the variation between hypo- and hyperarticulated speech (cf. Lindblom, 1990). In particular, potential sound changes are evaluated according to their social cost: so it is in this sense that sound change in the model of Lindblom et al. (1995) is teleological. These two theories contrast strongly, but they may explain different kinds of sound change. While Ohala's model seems to account particularly for internally driven sound changes, the theory by Lindblom et al. (1995) explains sound changes that are motivated by external or extra-linguistic factors.

Ohala's approach does not aim to describe sound changes but instead to analyse the conditions under which sound changes that occur in many unrelated languages take place. Analyses of the perception and production of coarticulation and assimilation form the empirical basis of his analysis. Harrington, Kleber and Reubold (2008) have extended Ohala's model of sound change to the fronting of /u/, which is a well-established sound change in English. They found that younger speakers not only produced very fronted /u/-variants (in terms of a very high second formant) irrespective of the context (i.e. both in fronting and in backing context), but also perceived more synthetic stimuli with a higher second formant as /u/ than older listeners. More importantly, while there was a context-dependent difference in the placement of the category boundary for old listeners, younger subjects had given up compensating perceptually for the effect of coarticulation (cf. Mann & Repp, 1980). In one analysis of a sound change in progress in Swedish, Janson (1983, 1986) considered how the perception and production of speech were related in a sound change in progress. For example, Janson and Schulman (1983) found that speakers produced differences between short /e/ and /ɛ/, although these two contrasts have merged in perception. In order to test whether sound changes come about because of a waning of perceptual compensation for coarticulation, Kleber, Harrington and Reubold (in press) investigated lax /ʊ/-fronting in Southern Standard British English, which is a more recent sound change (Hawkins & Midgley, 2005). The results showed that the coarticulatory influence of the surrounding consonants on /ʊ/ was smaller in perception than in production for younger than for older subjects. This finding was explained by a model of sound change in which perceptual compensation decreases before coarticulatory relationships begin to change in speech production.

Although there are numerous experimental studies on sound changes in progress in various varieties of English, there are by contrast scarcely any (e.g. Wagener, 2002; Geiger & Salmons, 2006; Blevins & Grawunder, 2009) descriptions, let alone experimental studies on ongoing sound changes in other languages. The studies described in the second part of this thesis are a first attempt to investigate a potential sound change in progress in German that may come about because of (in)complete neutralization or maintenance of phonological contrasts in the course of dialect levelling (see 1.4.3 above). Neutralization is a well-known common sound change that has been frequently reported as a historic change in dialectology. Depending on the position in the word or even phrase, the contrast is either lenited towards the voiced category, or strengthened towards the voiceless category. The latter case is found in many languages phrase-finally, presumably because "lack of articulatory effort results in poor voicing in final position" (Pierrehumbert, 2001: 152). In this position, voiceless obstruents are also more frequent and thus unmarked (cf. Greenberg, Ferguson & Moravscik, 1978) in the world's languages compared to voiced obstruents. Consequently, voiced stops belong to the marked category, which is more unstable and tends to be misperceived as the unmarked counterpart (Pierrehumbert, 2001: 152). On the other hand, in spontaneous speech intervocalic voiceless obstruents are frequently lenited "under conditions such as faster or more relaxed speech [… and …] in the course of time [lenition] may be completely dissociated from these contextual factors and become a petrified alternation" (Kohler, 1984: 157, see also his example of the change from Latin *vita* ('life') to Spanish *vida*). That is, a diachronic sound change has emerged from the synchronic lenition process. Incomplete neutralization of an opposition is an even more promising candidate for a language change in progress, because the gradual change away from (as in phonemic splits) or towards another phonemic category (as in mergers) makes both the phonemic contrast and also the neutralized or merged form phonologically unstable. Two potential outcomes regarding the intervocalic ((in)completely neutralized) voicing contrast are conceivable depending on the following factors:

- **Internal factors**: Different phonetic contexts often lead to fortition and lenition and thus to the neutralization of a contrast. That is, internal factors would lead to decreasing contrast maintenance in young vs. old speakers.
- **External and extra-linguistic factors**: Neutralization of the voicing distinction is partly stigmatized in the varieties of German (see above 1.3). The potentially increasing contact with other (high prestige) varieties and Standard German may have caused younger speakers of a low prestige, neutralizing variety to maintain the voicing contrast to a greater extent which would amount to a split in this variety.

These factors lead to different predictions regarding the degree of contrast maintenance. If internal factors were the driving force, we should expect more neutralization in younger age groups of non-neutralizing varieties, but if external factors were more important, we would predict less neutralization in younger speakers of neutralizing varieties. Since the latter prediction is supported by various studies (Wagener, 2002; Lameli, 2004; Barden & Großkopf, 1998), we will therefore formulate the hypotheses to be tested in the investigations based on these assumptions. However, this thesis is not concerned with resolving the ongoing theoretical discussion of whether internal factors are the primary motivation for sound change (e.g. Lass, 1997) or whether all factors contribute interdependently to sound change (e.g. Torgensen & Kerswill, 2004).

1.6 This thesis

1.6.1 The structure of this thesis

The aims of this thesis are twofold: to explore incomplete neutralization of phonetic contrasts (1) in the perception of Standard German and (2) in the production of two regional varieties of German. This thesis consists of two parts. In the first part (chapter II), we address domain-final voicing neutralization. While the studies on incomplete neutralization described above have shown repeatedly that speakers make use of fine phonetic detail that may be perceptible to the listener, there are no investigations of whether the domain-final voicing contrast is perceptually (in)completely neutralized, i.e. which cues listeners actually relied on when they discriminated stimuli better than chance. The second part of this thesis (chapters III and IV) looks at the intervocalic voicing and vowel length neutralization in two regional varieties of German and with respect to contrast maintenance at potential differences between age groups. That is, phonemic contrasts that were considered to be neutralized in regional varieties of German are possibly maintained to a greater extent in the productions of younger compared to older speakers due to the increasing influence of the standard language.

The investigation in chapter II is concerned with the perception of incomplete neutralization of the syllable-final voicing contrast in Standard German. The studies on incomplete neutralization described above were mostly concerned with the acoustics and used natural stimuli (i.e. they could not control for any particular cues) in the subsequent perception experiments. First of all, we aimed to test whether the fortis/lenis contrast as signalled by the most relevant cue (vowel to vowel+consonant duration ratio, Kohler, 1977), is perceptually incompletely neutralized, i.e. whether subjects' response curves show a gradual change from a more voiced to a voiceless category. Furthermore, in two separate experiments we tested whether (1) the phonological frequency with which a vowel plus stop occurs in the German lexicon and (2) the possibility of resyllabifying the final stop with the onset of the following syllable affected the extent to which the phonemic contrast is neutralized in perception.

In chapter III the focus is on the intervocalic lenis/fortis opposition, which is neutralized in many varieties of German but maintained in Standard German. The contrast is supposed to be neutralized in Saxon, but not in Bavarian. In the latter variety, the voicing contrast is maintained by means of the vowel length of the preceding vowel relative to the stop duration (Bannert, 1976). We used acoustic techniques to investigate the extent to which young and old Saxon and Bavarian speakers neutralize the voicing contrast after tense vowels in an alveolar (*Boden* vs. *boten*) and labial (*Ober* vs. *Oper*) context. There is a frequent co-occurrence of tense vowels and voiced labial stops as well as lax vowels and voiceless labial stops. The underlying voicing of an alveolar stop following a tense vowel, however, is not predictable since the distribution is balanced. One hypothesis was that neutralization was more likely in the labial context because of the bias towards the voiced stop in the lexicon. Another aim was to investigate whether young speakers neutralize the contrast to a lesser extent than old speakers based on the assumption that younger speakers tend towards a more standard-like pronunciation. That is, in this thesis we assume that dialect levelling in the two regional varieties originates from independent levelling towards the standard (vertical accommodation) and not in terms of accommodation between Saxon and Bavarian speakers.

The co-dependency of vowel length and stop voicing was further explored in chapter IV. In order to test whether the vowel length contrast is neutralized before voiceless stops or maintained to a greater extent by an additional variation of the stop's underlying voicing (in terms of duration), we analyzed acoustically various minimal pairs that differ in the tenseness of the lexically accented vowel such as *bieten* vs. *bitten* and compared them to minimal pairs containing sonorants such as *Sohne* vs. *Sonne.* In this analysis, age was again included as a predictor of the extent of contrast maintenance. In chapter V we will briefly summarize the findings from chapters II-IV and give an outlook on further research. The empirical basis for the analyses in chapters III and IV is described in the next paragraph.

1.6.2 The speech corpus

1.6.2.1 Speech materials

The speech corpus was designed in order to allow for investigations of a large range of research questions concerning various types of regional differences (both in terms of the classificatory differences laid out in Well, 1982, as well as the German sound system) in any dialect of German. The materials comprised 46 different trochaic Standard German words (see Appendix A, Table A.1). 31 words contained all 15 monophthongs that are regarded as phonemes in German and each test word belonged to at least one minimal pair that contrasts with respect to one of the four possible vocalic features: tongue height, tongue position, lip rounding and tenseness. Nine words contained one of the six oral stop phonemes in either initial or medial position. Again the test words belonged to sets of near minimal pairs which contrast in phonological stop voicing and place of articulation. The corpus was completed by six additional filler words. Out of the 46 words, 14 test words were analyzed in either chapter III or chapter IV. All other words (minimal pair and filler words) served as distracters, but some of the unused minimal pair words will be tested in other analyses. Although the corpus contained many minimal pairs, 6 additional filler words were considered sufficient since the minimal pair words themselves contained all vowel phonemes of German except for the three diphthongs (and there was also a large degree of variation in the syllable onsets and offsets).

1.6.2.2 The speakers

22 Saxon speakers from Dresden and 20 Bavarian speakers from Munich and Altötting were recorded and grouped into two age groups: speakers older than 50 were assigned to the Old group and subjects younger than 50 were assigned to the Young group (cf. the procedure in Harrington et al., 2008). There were 13 old (five female) and nine young (six female) Saxon speakers as well as eight old (5 female) and twelve young (eight female) Bavarian speakers. All speakers were born and raised in one of these two dialect areas; they speak Standard German with perceivable regional influences. None of the speakers reported any speech, reading or eyesight disorders.

1.6.2.3 Recordings and pre-analysis

The 42 speakers were recorded using the SpeechRecorder software (Draxler & Jänsch, 2004), a Laptop computer (Toshiba Tecra), and a Sennheiser USB headset in quiet rooms at the subjects' houses. The software is frequently used at the IPS, Munich, and has proven to be suitable for recordings outside the studio. Each of the 46 words was repeated ten times and presented in isolation and in randomized order on the computer screen. The subjects were asked to read the word as quickly as possible as there was a preset time limit of 2 seconds for the presentation of each word. One recording session took about 30 minutes.

The acoustic data was automatically segmented with the Munich Automatic Segmentation System (MAUS, Schiel 2004) and segment boundaries were manually corrected in Praat (Boersma & Weenink, 2008) when necessary. All data files were then converted into Emu format and all further analyses (which will be described in detail in the relevant chapters) were carried out in the Emu speech database system (Harrington, 2010) and the programming language/environment R.

1.6.2.4 Why laboratory speech?

Most research on German dialects was based upon transcribed interviews which were considered to be most adequate for studying sociolinguistic aspects of dialect use (e.g. Wagener, 2002). It is indeed very likely that subjects will not produce many of the regional dialect features described in 1.4 in laboratory speech. However, regional varieties of Standard German are likely to differ in fine phonetic detail (cf. studies on English regional varieties, e.g. Docherty & Foulkes, 1999, 2000). It was therefore important to use minimal pair words of a similar structure in order to be able to measure precisely potentially small phonetic difference between the varieties. A large number of stimuli were presented in quick succession so that speakers would be less inclined to hyperarticulate and thus not neutralize the contrast (Fourakis & Iverson, 1984; Port & Crawford, 1989).

Incomplete neutralization in the perception of final devoicing in Standard German*

Chapter II

Abstract

We investigated the perceptibility of stop voicing in a domain-final neutralizing context in German that according to various phonological models is completely neutralized in favour of the voiceless category but that according to various empirical studies is distinguishable phonetically. A primary aim was to determine whether acoustic cues that were available for the stop voicing distinction were perceptible in a neutralizing context. A secondary aim was to assess whether voicing perception was influenced by phonotactic frequency and the potential for resyllabification. Nineteen listeners of a Standard German speaking variety made forced-choice judgments to synthetic stimuli spanning a voiced-voiceless continuum containing domain-final alveolar and velar stops in various neutralizing contexts that differed in terms of phonotactic probability and the potential for resyllabification. Our results showed that voicing information could be distinguished but that the perceptibility of this distinction also depended on statistical properties of phoneme sequences and whether a domain-final stop could potentially be perceptually resyllabified as domain-initial. Our general conclusion is that a categorical neutralization model is insufficient to account for stop voicing perception in German in a domain-final context: instead, voicing perceptibility in these contexts depends on an interaction between acoustic information and phonological knowledge which emerges as a generalization across the lexicon.

* A version of this chapter was published in the *Journal of Phonetics*, 38 (Kleber, John, & Harrington, 2010).

2.1 Introduction

Recent studies have provided increasing evidence that fine phonetic detail, at segmental and prosodic levels is an integral part of speech communication in both production and perception (see e.g. a detailed description by Hawkins, 2003). For example, vowels in frequently occurring words with fewer lexical neighbours tend to be produced with a more centralised quality as opposed to rare words with more lexical neighbours (Wright, 2003) and onset /l/ is longer before voiced versus voiceless coda stops in /liːd/ versus /liːt/ (Hawkins & Nguyen, 2004). These subtle phonetic differences are perceptible and used for word recognition (Davis, Marslen-Wilson & Gaskell, 2002; Hawkins & Nguyen, 2003; Manuel, 1995). The type of fine phonetic detail which is the main concern of the present study involves incomplete neutralization of the voicing contrast in oral stops which has been demonstrated experimentally in Dutch (Warner et al., 2004), German (Port & O'Dell, 1985), and Catalan (Charles-Luce & Dinnsen, 1987). In a number of languages such as Polish, Dutch, Catalan and German, obstruents contrast in voicing when they occur in a prosodically domain-initial or medial position[1], whereas in domain-final position this contrast is neutralized with a bias in favour of the voiceless component. The classic example of neutralization in German arises from word-final devoicing which causes *Rad* ('wheel') and *Rat* ('advice'), which differ underlyingly in the voicing status of the final consonant, to become surface homophonous. The process of final devoicing that results in this type of neutralization is often generalized phonologically by a rule such as

[+obstruent] → [-voice]/_D (where D stands for domain-boundary)

However, implicit in this type of rule is that neutralization is complete, i.e. that neutralized and underlyingly voiceless forms are indistinguishable from each other.

The issue as to whether German final devoicing really is incomplete remains controversial and it has received considerable attention in the last 20-30 years. For German,

[1] This phonological contrast in German is also described in terms of lenis versus fortis. For a better comparability with other languages we use throughout this chapter the terms voiced and voiceless. See Kohler (1984) for a discussion of the lenis/fortis dichotomy.

acoustic analyses have revealed small but significant longer vowel durations preceding word-final underlying voiced obstruents than their voiceless counterparts (O'Dell & Port, 1983, Port et al., 1981; Port & O'Dell, 1985). That is, in these studies derived voiceless obstruents were shown to be acoustically intermediate between voiced and voiceless obstruents in domain-initial or medial position – though they were far closer to voiceless than voiced stops. Other languages for which there is evidence in favour of incomplete neutralization include Catalan (Dinnsen & Charles-Luce, 1984), Polish (Slowiaczek & Dinnsen, 1985), and Dutch (Warner et al., 2004; Ernestus & Baayen, 2006), the latter being a final-devoicing language for which previous studies had shown complete neutralization (Jongman et al., 1992; Baumann 1995). The durational differences between stop categories in a neutralising context varied depending on the language, but were always pervasive. Among all the acoustic correlates measured in these studies (e.g. voicing into and during closure), vowel duration stood out as the most important cue for preserving the voicing contrast to some extent, even in Polish (Slowiaczek & Dinnsen, 1985) where vowel duration is not the primary cue for differentiating voiced and voiceless obstruents word-medially (Keating, 1979). Most of these have also shown individual speaker and dialect dependent significances (see especially Piroth & Jancker, 2004 for German) while in others the extent to which the voicing contrast was neutralized was shown to be dependent on sentence position, semantic information and phonetic environment (Charles-Luce 1985, 1993; Slowiaczek & Dinnsen, 1985), pragmatics (Port & Crawford, 1989) as well as morphology (Ernestus & Baayen, 2006). Various studies showed that orthography and speaking style affect the degree of incomplete neutralization (sometimes resulting in hyperarticulation) and some authors concluded that incomplete neutralization was an experimental artefact of orthography (e.g., Fourakis & Iverson, 1984; Jassem & Richter, 1989; Warner et al., 2006; but see Port & Crawford, 1989 and Ernestus & Baayen, 2006 for a different interpretation).

One way of resolving these contradictory arguments is to test whether acoustic incomplete neutralizations are perceptible and if so to assess whether listeners make use of the different cues for distinguishing between words that differ underlyingly minimally in stop voicing. Some of these issues have already been addressed by Port and colleagues (Port & O'Dell, 1985; Port & Crawford, 1989) for German and by Slowiaczek and Szymanska (1989) for Polish. In these three studies, it was found that the words of minimal pairs differing in underlying voicing were discriminated better than chance in neutralising contexts and it was concluded that listeners can make use of the acoustic cues in order to identify minimal word pairs in a forced choice task[2] but argued against the idea that this discriminability had a functional role. Since, however, these experiments were based on natural speech materials that contained various acoustic confounds, no clear conclusion can be drawn as to which acoustic cues listeners actually exploit for word identification and discrimination in these neutralizing contexts. In Slowiaczek and Szymanska's (1989), the same Polish material was also presented to English listeners whose performance was similar to that of the Polish listeners; they also found that both groups of listeners showed a perceptual bias towards the voiceless variants. According to Slowiaczek and Szymanska (1989), the combination of poorer-than-expected (though better-than-chance) identification and the perceptual bias in favour of the voiceless obstruent suggest that the acoustic cues are not reliably used in regular communication to differentiate between members of a minimal pair (and are hence not primary cues) and are "perceptually neutralized" (p. 211). Moreover, since the data from speaker groups of both languages were quite similar, it was concluded that the performance must have been a function of the acoustic information in the speech signal rather than being mediated by native listeners' knowledge of Polish phonological rules. By contrast, Broersma (2005) found that language background did influence listeners' responses. When a choice of cues was available, listeners preferred to use familiar or primary cues (e.g. cues they know from their language or from other syllable position), but in the absence of such a choice listeners were able to exploit less familiar cues to improve their disambiguation performance (Warner et al., 2004; Broersma, 2005). Listeners are very well able to discriminate voicing even in positions where there usually is no

[2] Unfortunately, Port and Crawford (1989) tested only tokens from the dictation and the word list reading conditions, i.e. they tested words that were recorded in situations in which hyperarticulation may come about due to the communicative tasks.

contrast, as long as this contrast exists in their phonological system (cf. also the Perceptual Assimilation Model, Best, 1994).

One of the factors that has not been considered so far that may influence incomplete neutralization is phonological frequency. Indeed, since listeners' perception of fine phonetic detail to disambiguate domain-final obstruent voicing has been shown to be language-dependent (Broersma, 2005), it seems quite possible that the extent of these subtle acoustic differences may also be conditioned by language-dependent factors such as phonological frequency and effects of syllable position. As far as phonological frequency effects are concerned, various studies have shown how perception is influenced by phonotactic co-occurrences, word frequency and neighbourhood density. For example, Hay et al. (2003) found that listeners often tend to misperceive statistically infrequent (e.g., /np/) as statistically frequent (e.g., /mp/) nasal-obstruent clusters. Nonsense words containing phoneme sequences that occur frequently in real words are better memorized (Frisch, Large, & Pisoni, 2000) and repeated faster (Vitevitch & Luce, 1999; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997) as opposed to non-words containing rare sequences. Compatibly, Pitt and McQueen (1998), found evidence for a perceptual bias in listeners' responses to acoustically ambiguous consonants depending on the probability with which the consonant occurs in a consonant sequence. Following exemplar theory, all these examples of 'misperception' are very likely to come about because of a perceptual adjustment depending on lexically based expectations of the listeners.

Such frequency related predictions may also carry over to final devoicing in German in which the degree of neutralization could depend on properties of statistical frequency in the lexicon and more specifically on phonotactic frequency. In German, there is a phonological distinction between tense and lax vowels (e.g. /biːtən/, 'to offer' vs. /bɪtən/, 'to beg') and intervocalically also between voiced and voiceless obstruents (e.g. /miːdən/, 'to avoid' vs. /miːtən/, 'to rent'). However, the combination of vowel tensity and obstruent voicing is anything but equally distributed in the German lexicon. For example, lax vowels almost always precede voiceless stops. There are only a handful of words containing a lax vowel plus voiced stop sequence and most of these are loan words (e.g. /ɛbə/, 'tide'; /klɛvə/, 'clever'; /meːʃʊgə/, 'crazy') which, according to Féry (2003), are not part of the "truly core native German vocabulary" (p. 150). At the same time, underlyingly voiced labial or velar obstruents are almost always preceded by tense vowels. Thus surface /iːp/ is almost always a reflex of underlying /iːb/ (/liːbə/, 'love'; /ʃiːbən/, 'to push') and, with the exception of one or two infrequent words (e.g., /piːksən/, 'to prickle'), surface /iːk/ indexes underlying /iːg/ (/fliːgə/, 'fly'; /kriːgən/, 'to get'; /liːgən/, 'to lie'). The same predictable relationships hold for other tense vowels. Only alveolar stops that are preceded by tense vowels show a balanced distribution of both underlying voicing categories (e.g. /miːdən/, 'to avoid'; /miːtən/, 'to rent'; /baːdən/, 'to take a bath'; /baːtən/, 'they asked for'; /boːdən/, 'floor'; /boːtən/, 'carriers').

Previous studies have controlled for *phonetic* vowel length in combination with the following obstruents (Broersma, 2005; Warner et al., 2004) but only as part of the investigation of vowel duration as a cue for voicing perception. By contrast, our aim in the present study is to test whether these types of imbalances in combinations of *phonological* vowel length and voicing influence perception. Thus we would predict that listeners are better able to perceive a domain-final /t, d/ contrast after a tense vowel than after a lax vowel, because the voicing contrast is only frequent in the lexicon for tense vowel plus alveolar stop sequences (e.g. /laɪdən/, 'to suffer' vs. /laɪtən/, 'to lead'). When on the other hand the contrast is infrequent, then voicing judgements should be guided by lexical statistics: thus alveolars are more likely to be perceived as voiceless following lax vowels since combinations of lax vowels and voiced stops are almost non-occurring in the lexicon.

The second conditioning factor that we will consider here is the *degree* of neutralization which may vary depending on syllable position and the extent to which the sequence containing a syllable final but pre-consonantal obstruent is resyllabifiable. The starting point for this research question comes from two phonological approaches to the analysis of final devoicing: the *licensing-by-prosody* account (Itô, 1986, 1989; Goldsmith, 1990; Rubach 1990; Lombardi, 1999; Beckman, 1997), in which final prosodic position is the main determiner of neutralization (see also e.g. Brockhaus, 1995; Hall, 1992; Vennemann, 1972; Wiese, 1996) and the *licensing-by-cue* approach (Steriade, 1997, 1999, 2000) in which acoustic properties of the potentially neutralising context are primary. Based on the analysis of neutralization in different languages, Steriade predicts that phonological contrasts are neutralized first in environments in which the perceptual differentiation can only be maintained by additional articulatory effort. Irrespective of the differences between these models, an important question that is relevant for both is whether acoustic cues to a phonological contrast are perceptually *masked* depending on either the phonetic or prosodic context. This issue was touched upon by Cutler (2002) who reasoned that the phonological generalisation that vowel length is neutralized utterance-finally in Japanese may in itself contribute to listeners' inability to hear the contrast, even when there is acoustic evidence for its distinction (Kubozono, 2002).

As far as the present study is concerned, the basis for this kind of perceptual masking is as follows. It is well known that the perceived differences between allophones of a phoneme are much less than those between different phonemes and this is consistent with findings from the child language acquisition literature (Werker, 1995; Werker & Tess, 1998) showing that the perceptual discriminability between allophones of the same phoneme diminishes in the first year of acquisition, presumably because children learn to focus on acoustic cues that are important for distinguishing between phonemes and words and also to ignore (or pay less perceptual attention to) those that are not. It is possible that this kind of perceptual masking operates not only allophonically but also between phonemes that are in a neutralising context. Thus, perhaps listeners filter out perceptually any acoustic cues that might be present to the voicing contrast in domain-final position, because they interpret this as a neutralizing context in which there is usually no surface distinction according to their phonological knowledge (however, see Ernestus & Baayen, 2006 for some results in Dutch that are not consistent with this position). If the phonological grammar does exert a top-down influence on the acoustic signal in this way, then we might expect acoustic cues to be more perceptible in a context in which a domain-*final* stop has the potential to be interpreted as domain-*initial* (as a result of which it is no longer in a neutralising context). More specifically, consider that in a /vowel-stop-l/ context, the stop is necessarily domain-final if it is alveolar because in German, as in English, initial */dl/ and */tl/ are excluded. Therefore, any potential cues to the voicing contrast might be perceptually strongly masked precisely because the alveolar stop is necessarily domain-final and therefore in a neutralising context. But this might not be so if the stop is a velar because /kl, gl/ are legal onset clusters in German (and English). Thus, because velar stops could be interpreted by the listener to be domain-initial, the perceptual masking which is predicted to filter out acoustic information in the domain-final neutralising context would not apply in this case. Consequently, the probability of hearing any distinguishing voicing cues should be greater for the velar than the alveolar context according to this perceptual masking hypothesis.

In this paper, we describe two experiments designed to test whether the domain final voicing contrast in German obstruents is incompletely neutralized in perception and whether the degree of the incompleteness depends on phonotactics and statistical co-occurrences of phonemes in the lexicon. The first hypothesis was the starting point for all other hypotheses and therefore tested in both experiments; it can be summarised as follows.

(H1) The voicing contrast is incompletely neutralized in the perception of German domain-final obstruents.

Hypotheses H2 – H4 that gave rise to the experiments were all motivated by the relative frequency with which various patterns of segment sequences occur in the lexicon. Experiment 1 was conducted to investigate syllable internal sequences. In this experiment, hypotheses H2 and H3 were tested; they can be summarised as follows:

(H2) Listeners show a perceptual bias towards the more frequently occurring voiceless stop when preceded by lax vowels as opposed to tense vowels (e.g. /vɪd/ is predicted to be perceived as /vɪt/ because /ɪt/ co-occurs frequently and /ɪd/ does not).

(H3) The voicing distinction is perceptually less neutralized when there are analogous frequent contrasts in the lexicon. Thus since /iːd/ vs. /iːt/ is lexically frequent but /iːg/ vs. /iːk/ is not, then listeners should be better able to distinguish /d, t/ in the former sequence in a neutralising context (such as in a domain-final position, e.g. before an obstruent in which the stop is necessarily domain-final).

The central research question of Experiment 2 was whether resyllabification affects the degree of incomplete neutralization: this is formulated in hypothesis H4.

(H4) Perception of the voicing contrast is more likely in a consonant cluster that can be resyllabified with the onset consonant of the following syllable (e.g. there is more neutralization of the voicing contrast in a /stop#l/ sequence for alveolars than for velars).

2.2 Experiment 1: effect of probabilistic co-occurrences of phoneme sequences on incomplete neutralization

2.2.1 Speech materials

We created four continua one each consisting of resynthesized stimuli of four minimal pair disyllabic compounds: *Widdlinn – Wittlinn* (henceforth /$V_{lax}C_{alv}$/), *Bigglinn – Bicklinn* (henceforth /$V_{lax}C_{vel}$/), *Niedlinn – Nietlinn* (henceforth /$V_{tns}C_{alv}$/), and *Mieglinn – Mieklinn* (henceforth /$V_{tns}C_{vel}$/). These compounds were hypothetical German town names. The reason for choosing different onset consonants for the compounds was as a reminder to the listener that s/he would be perceiving a tense vowel for example in ‘nie[d/t]’ but a lax vowel in ‘wi[d/t]’. If we had used the same onset consonant for both, then listeners might have confused the tense and lax continua given that decreasing vowel duration, which is one of the variables manipulated here, is also a positive cue for a lax as opposed to a tense vowel. We have no reason to expect the different [v] vs. [n] onset consonants to affect voicing judgments.

Analogously we chose different onset consonants for the minimal pairs differing in place of articulation of the syllable-final stop.

In order to create the continua in the compounds, a male speaker produced (together with the test words from Experiment 2 reported below) the trochaic words *Widden* (/vɪdən/), *Witten* (/vɪtən/), *Biggen* (/bɪgən/), *Bicken* (/bɪkən/), *Nieden* (/niːdən/), *Nieten* (/niːtən/), *Miegen* (/miːkən/), and *Mieken* (/miːkən/). Each of these words was repeated in isolation 10 times. We then chose one token of each of the 10 repetitions of the four voiced tokens and spliced out the first syllable at the stop release thus leaving /vɪd-/, /bɪg-/, /niːd-/, and /miːg-/. Our choice of this syllable was based on two criteria: firstly, the second syllable in the original trochaic word had to be produced as a syllabic /n/ with an elided schwa; and secondly, the durations of the /iː/, /ɪ/, /d/, and /g/ should be closest to the mean duration of these segments across all ten tokens. The second syllable /lɪn/ was spliced out of either a production of *Britlinn* or a production of *Ricklinn* (again both context words were German pseudo town names). We cut out these second syllables at the onset of periodicity of /l/ and, depending on the syllable-final stop's place of articulation, appended the *-linn* taken from *Britlinn* to /vɪd-/ and /niːd-/ and the *-linn* from *Ricklinn* to /bɪg-/ and /miːg-/ to create the spliced compounds with the voiced stop /vɪdlɪn/, /bɪglɪn/, /niːdlɪn/, and /miːglɪn/. The syllable final stops in the target words and the context word matched in place of articulation in order to avoid a disruption of the acoustic cues at the splice point (in front of the lateral). These spliced blends served as the endpoint stimuli at the voiced ends of the continua.

The tokens towards the voiceless end of the continuum, i.e., towards /vɪtlɪn/, /bɪklɪn/, /niːtlɪn/, and /miːklɪn/ were derived from these voiced endpoints by reducing the V:C duration ratio (ratio of vowel duration to vowel plus closure duration, henceforth V:C), where V = /iː/ or /ɪ/ and C = the following alveolar or velar stop closure. V:C has been found to be the most powerful acoustic cue for disambiguating voiced from voiceless stops in a semi-intervocalic context (Kohler, 1979). Before we applied this shortening, we had to determine V:C at the voiceless *endpoint* (i.e. at the most extreme /vɪtlɪn/, /bɪklɪn/, /niːtlɪn/, and /miːklɪn/ tokens). This was calculated by weighting the V:C averaged across all voiceless tokens by the VC duration in the selected voiced context. More specifically where $Rhyme_{voiced}$ is the duration of the voiced rhyme (e.g., duration of the selected /ɪd/, /ɪg/, /iːd/ or /iːg/ token) and $Ratio_{voiceless.m}$ is the mean V:C duration ratio in voiceless consonants (mean duration of V:C in /ɪt/, /ɪk/, /iːt/ or /iːk/), then the duration of the vowel preceding the voiceless stop, $Vowel_{voiceless}$, that was used for the voiceless endpoint in the synthesis continuum was calculated from:

(1) $$Vowel_{voiceless} = Rhyme_{voiced} \ x \ Ratio_{voiceless.m}$$

The calculation by means of equation (1) ensured that the total duration of each stimulus item within a continuum remained constant.

Table 2.1: Segment durations and V:C duration ratios for each stimulus of each continuum.

VC-sequence	Segment	Stimulus number						
		1[a]	2	3	4	5	6	7
/ɪd/ - /ɪt/	V	90	83	77	70	63	56	50
	C	122	129	135	142	149	156	162
	VC	212	212	212	212	212	212	212
	V:C duration ratio	0.42	0.39	0.36	0.33	0.30	0.27	0.23
/ɪg/ - /ɪk/	V	75	68	62	55	48	41	34
	C	96	103	109	116	123	130	137
	VC	171	171	171	171	171	171	171
	V:C duration ratio	0.44	0.40	0.36	0.32	0.28	0.24	0.20
/iːd/ - /iːt/	V	192	178	165	151	137	124	110
	C	73	87	100	114	128	141	155
	VC	265	265	265	265	265	265	265
	V:C duration ratio	0.72	0.67	0.62	0.57	0.52	0.47	0.42
/iːg/ - /iːk/	V	192	172	153	133	114	94	75
	C	71	91	110	130	149	169	188
	VC	263	263	263	263	263	263	263
	V:C duration ratio	0.73	0.66	0.58	0.51	0.43	0.36	0.28

We then derived four seven-step continua between these voiced and voiceless endpoints. To calculate the step size for the stimuli of each continuum, we divided the vowel duration difference between these endpoint stimuli by six. The step size was 7 ms for both the /$V_{lax}C_{alv}$/- and the /$V_{lax}C_{vel}$/- continuum, 14 ms for the /$V_{tns}C_{alv}$/-continuum, and 20 ms for the /$V_{tns}C_{vel}$/-

[a] Selected voiced token.

continuum. In order to compare, analyze and evaluate the perception results for the various continua that differ both in phonological vowel quantity as well as in their segmental structure, it was necessary to use proportionally equal distances instead of absolute step sizes. The vowel durations of the selected voiced tokens were then progressively shortened and the stop closure durations were progressively lengthened by the calculated step sizes, so that the VC duration remained constant (see Table 2.1 for further details).

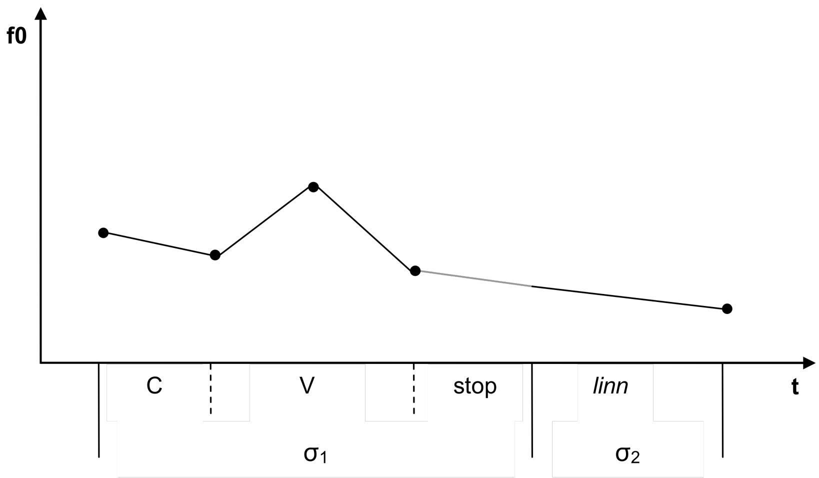


Figure 2.1: Schematic representation of the f0-contour manipulation.

The f0-contour was stylized such that there was a rise towards the midpoint of the accented syllable and a fall from there linearly over the rest of the test blend with all stimuli having the same pitch heights for the five f0 target points (Figure 2.1). For all stimuli, any evidence of voicing during the closure was also removed by high-pass filtering. Listening tests with 10 subjects showed that the endpoints could be unambiguously distinguished in all continua.

All manipulations were done by means of the manipulation function and then resynthesized with the “overlap and add” function in Praat (version 5.0.27, Boersma & Weenink, 2008).

2.2.2 Participants

The 168 stimuli (6 repetitions x 4 continua x 7 steps) were made available in an online forced choice identification experiment (together with the stimuli of Experiment 2 reported below as well as stimuli from another experiment which we will not report on here). 19 native speakers of Northern Standard German, all of them students at Kiel University, participated in the online experiment at the Institute of Phonetics and Digital Speech Processing, University of Kiel. Participants were paid a small amount for participation. None of the subjects reported any hearing, eye-sight, or reading problems.

2.2.3 Experimental procedures

The subjects performed a two-alternative forced choice task. All stimuli (including all other stimuli, which served as distracters in this experiment) were presented to the listeners over headphones in one session. Upon presentation of an auditory stimulus, the subject saw an orthographic representation corresponding to the minimal pair distinction. For example, upon being presented auditorily with one of the tokens from the /niː[d/t]lɪn/ continuum, the subject saw *Niedlinn* or *Nietlinn* on the screen and had to judge which of these was more similar to the perceived stimulus. The experiment was self-paced, i.e. the next item was only presented after the participant had made a decision. On average, the entire experiment took about one hour. The order of the stimuli was random for each participant to avoid any presentation effects. The responses were saved to a server located at the Phonetics Institute in Munich.

Two different types of statistical analyses were carried out in the programming language/environment R. We first analyzed our data by means of logistic regressions, i.e. we calculated the log of the ratio of the voiceless responses to the corresponding voiced responses for each stimulus in each continuum. This logit was the dependent variable and V:C Ratio (stimulus 1 to 7, in which V:C ranged from maximal to minimal) was the independent variable. Second, repeated measures generalized linear mixed models (GLMM – see Baayen, 2008) were fitted to predict incomplete neutralization as a function of the decreasing V:C duration ratio and to determine whether the degree of incompleteness depended on the lexical frequency of the voicing contrast which was in this case dependent upon vowel tensity and place of articulation. The 'voiceless' responses served as our dependent variable. As predictors (independent variables) we entered Tensity (lax vs. tense), Place (alveolar vs. velar) and V:C Ratio (stimulus 1 to 7). Subject was entered as a random effect factor.

2.2.4 Expectations and predictions

To answer the question whether the postvocalic domain final voicing contrast is completely or only incompletely neutralized, we shall first make some assumption regarding different forms of psychometric curves representing the listener's judgments.

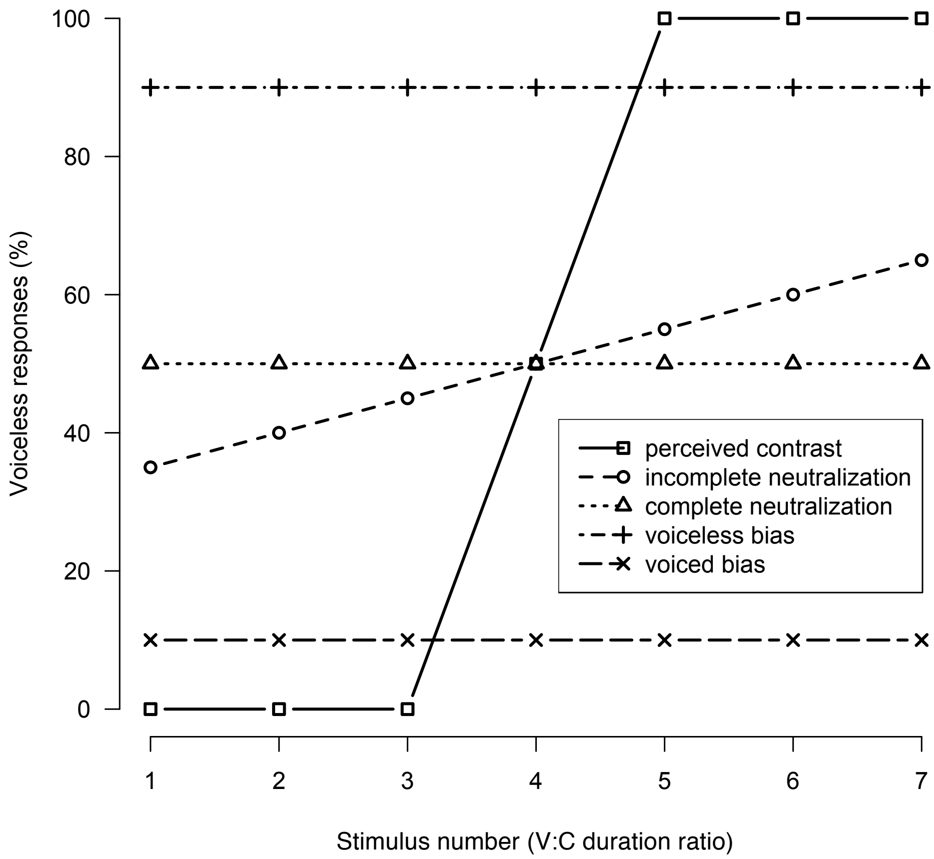


Figure 2.2: Schematic representations of five possible psychometric curves representing the percentage of 'voiceless' responses as a function of decreasing V:C duration ratio.

Figure 2.2 is a schematic outline of five possibilities: (1) no neutralization of the contrast, showing a steep rise between the endpoints which signals the presence of two categories, (2) incomplete neutralization of the contrast with no bias towards either category, showing a slightly rising slope around the 50% cross-over point, (3) complete neutralization with no bias towards either category, (4) complete neutralization of the contrast with a bias towards the voiceless category in which responses are above the 50% cross-over point, (5) as (4), but in which the bias is towards the voiced category.

2.2.5 Results and discussion

Figure 2.3 (a) gives the proportion of voiceless responses as a function of decreasing V:C duration ratio for the four continua. In order to test whether there were discernible trends in the proportional responses along the continua, logistic regression lines were calculated (Figure 2.3 (b)). All continua showed slightly, but gradually increasing identification functions, which indicated that the voicing contrast remained to some extent perceptible though there were no abrupt changes which would point to the perception of two distinct categories. The increases in the voiceless responses along the continua from left to right followed /$V_{tns}C_{alv}$/ ($\chi^2(1) = 28.2$, $p < 0.001$), and /$V_{tns}C_{alv}$/ ($\chi^2(1) = 61.8$, $p < 0.001$). The responses to most stimuli of all continua were in the voiced range, i.e. the percentage was below the 50% boundary. As Figures 3 (a, b) show, listeners labelled more stimuli as voiceless when the vowel was lax (36.6%) than tense (10.9%), which suggests that the lexical frequency of the tensity plus stop-voicing combination influenced voicing judgments: this is consistent with H2. Listeners also labelled more stimuli as voiceless when the place of articulation was velar as opposed to alveolar, which is not compatible with H3.

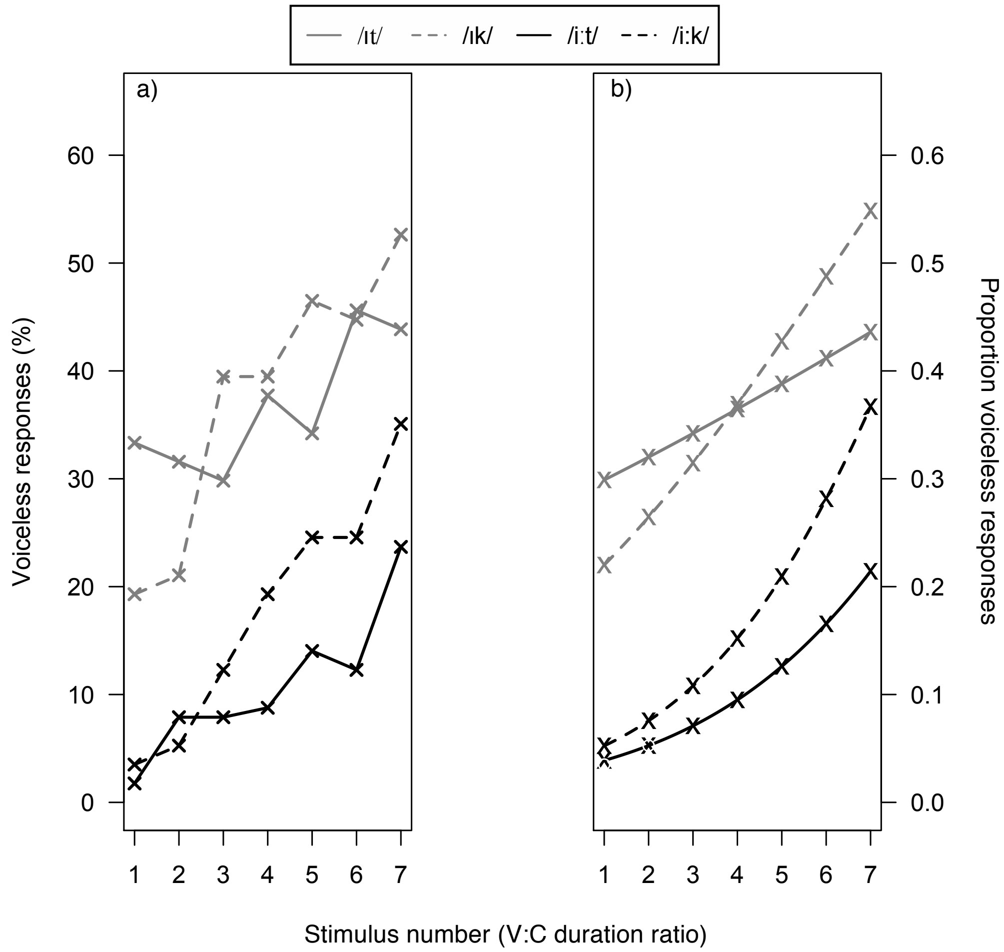


Figure 2.3: Proportion of 'voiceless' responses as a function of decreasing V:C duration ratio (stimulus number) to the four continua differing in vowel tensity and place of articulation (a) and the corresponding regression curves (b): lax (grey), tense (black), alveolar (solid), and velar (dashed).

The results of the GLMM in which the voiceless responses to all four continua were included as the dependent variable showed significant main effects for the independent variables Tensity (z = -11.8, $p < 0.001$) and V:C Ratio (z = 3.0, $p < 0.01$) but not for Place. The significant main effect for V:C Ratio is compatible with H1 that neutralization is incomplete. The significant main effect for Tensity is compatible with H2 that judgements of voicing are biased by the frequency distributions in the lexicon of vowel tensity plus stop voicing combinations. The non-significant effect for Place, however, indicates that there is no difference in the degree of perceptual neutralization of the voicing distinction when there are analogous frequent contrasts in the lexicon, which is prima facie not compatible with H3.

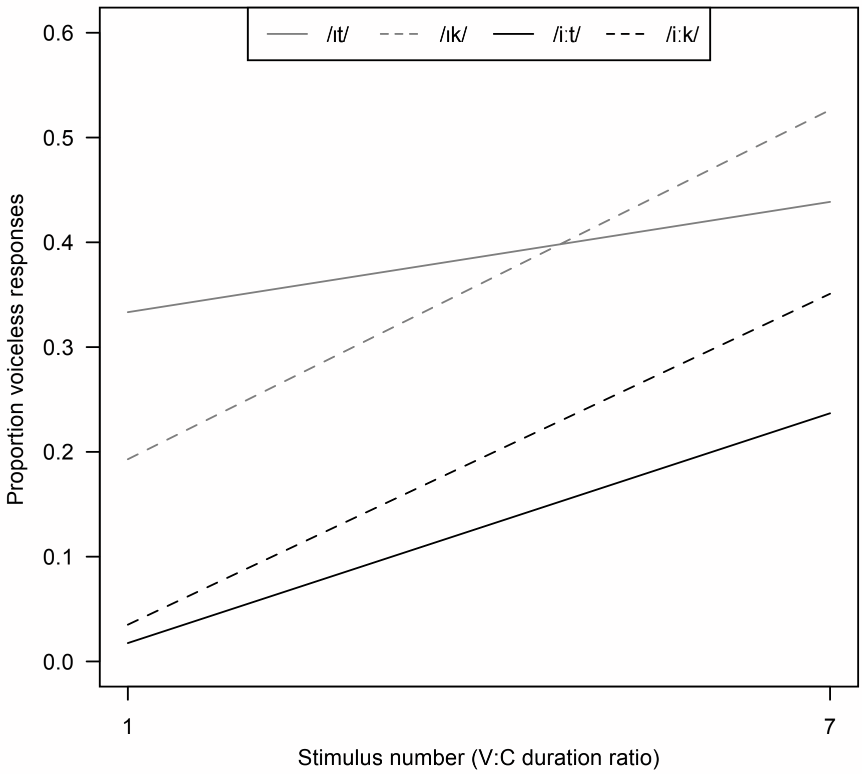


Figure 2.4: Linear interpolations between the proportional 'voiceless' responses to stimuli 1 and 7 of the four continua differing in vowel tensity and place of articulation: lax (grey), tense (black), alveolar (solid), and velar (dashed).

In order to assess further the validity of H3, we compared the effect of changing V:C on lax vs. tense vowels and alveolar vs. velar stops. According to H3, there should be less neutralization and therefore a steeper rise in the regression curve for tokens with a voicing contrast that is lexically frequent. Recall from the introduction that the post-vocalic voicing contrast is lexically frequent in the /$V_{tns}C_{alv}$/-continuum but not in the other three continua. Figure 2.4, in which the difference between the first and the last stimuli of each of the continua are linearly interpolated, shows that the rise of the /$V_{tns}C_{alv}$/-tokens' identification function is steeper than that of the /$V_{lax}C_{alv}$/-tokens' function but slighter than that of the velar tokens' functions.

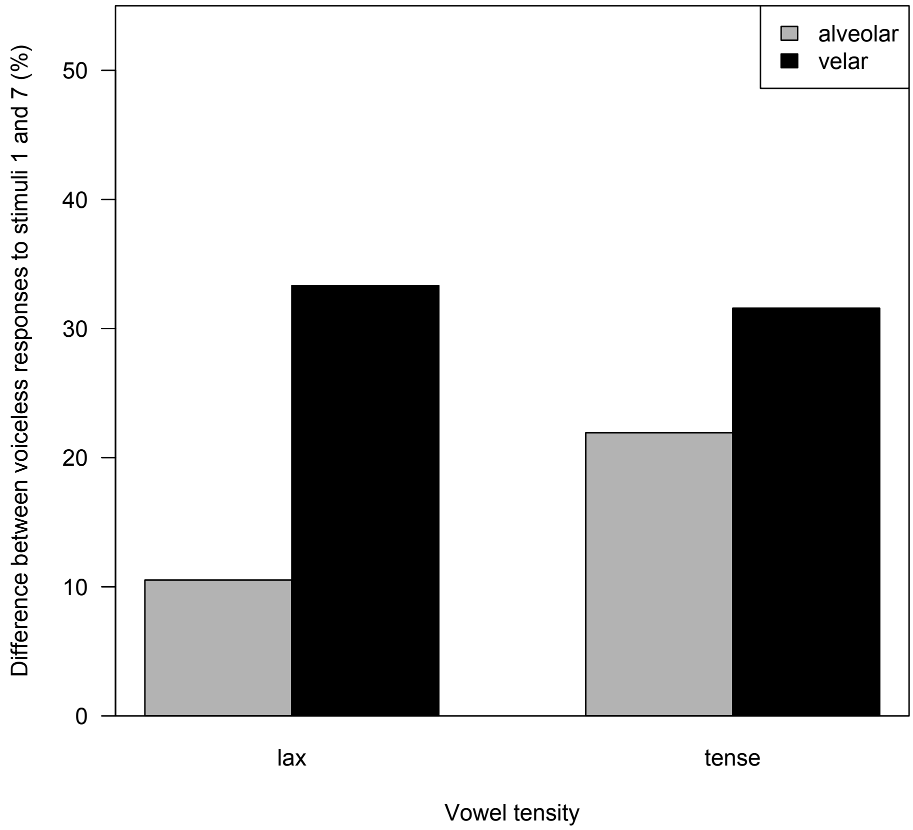


Figure 2.5: Proportional differences of 'voiceless' responses between stimuli 1 and 7 of the four continua.

Figure 2.5 illustrates that the difference in the proportion of voiceless responses between stimulus 1 and stimulus 7 is, on the one hand, greater for /$V_{tns}C_{alv}$/ (21.9%) than for /$V_{lax}C_{alv}$/ (10.5%), but, on the other hand, greater for /$V_{tns}C_{vel}$/ (31.6%) than for /$V_{tns}C_{alv}$/ and greatest for /$V_{lax}C_{vel}$/ (33.3%). The four identification functions diverge significantly and this is consistent with the significant interaction effects V:C Ratio x Tensity ($z = 2.7$, $p < 0.01$) and V:C Ratio x Place ($z = 3.2$, $p < 0.01$). This result means that V:C had a significantly different impact on the four continua depending on vowel tensity and place of articulation. To test for the significance of the rise of the identification functions – defined by the difference between the voiceless responses to stimuli 1 and 7 – we ran *post-hoc* analyses separately for the lax, tense, alveolar and velar continua. There was a significant interaction between V:C Ratio x Place for the lax continua ($z = 3.0$, $p < 0.01$), but no such significant interaction for the tense continua. This significant interaction confirmed that the difference between stimuli 1 and 7 in the proportion of voiceless responses was greater for velar than for alveolar tokens in the lax vowel context, but not in the tense vowel context. The non-significant interaction in the tense continua is not compatible with H3 that there is less perceptual neutralization for contexts in which the voicing distinction is lexically frequent. There was also a significant interaction between V:C Ratio and Tensity for the alveolar context ($z = 3.0$, $p < 0.01$), but no such significant interaction for the velar context. That means, that the difference between the two endpoint stimuli in the proportion of voiceless responses was significantly greater for tense vowels than for lax vowels but only in the alveolar context. This is compatible with H3, because tense vowel plus alveolar sequences frequently co-occur with both underlying voiced and voiceless stops while lax vowels are almost always followed by underlying voiceless stops. Since, on the other hand, velar stops are either underlying voiced when preceded by tense vowels or underlying voiceless when preceded by lax vowels (i.e. there is more or less complementary distribution of voicing in velar stops depending on vowel tensity), then velars should pattern differ from alveolars: indeed, as our results showed, for velars, in contrast to alveolars, the difference in voicing responses between Stimuli 1 and 7 was about the same in lax vs. tense vowel contexts.

The *post-hoc* analyses partly support H3 that the same acoustic cues to the voicing distinction are less effective in contexts in which the voicing contrast is lexically infrequent (i.e., in contexts in which only a handful of lexical items are distinguished by post-vocalic voicing). This prediction however, is only applicable with respect to vowel tensity. In this case, listeners seem to collapse the voicing distinction in favour of the more likely category, when the voicing contrast is lexically infrequent, whereas the acoustic cues to the voicing distinction are much more effective for contexts in which the voicing distinction is lexically frequent.

Our results for place of articulation were, however, inconsistent with H3. According to H3, there should be a sharper discrimination between the endpoint stimuli if the voicing contrast occurs in a context which is lexically frequent, i.e. listeners should have been better able to hear the voicing distinction in a tense vowel plus alveolar context (analogously to frequent contrasts such as /laɪtən/ ‘to lead’ vs. /laɪdən/ ‘to suffer’) than in a tense vowel plus velar context (in which the velar is almost always underlyingly voiced) but this is not what we found. One possible explanation for this result could be the potential for resyllabification in the velar continua. In both velar continua, the syllable-final stop is resyllabifiable with the sonorant onset consonant of the second syllable and can also be interpreted perceptually as syllable initial, while in the alveolar sequences the stop is not resyllabifiable and can therefore be only interpreted as syllable final. In the second experiment, we, therefore, sought to shed more light on this possible explanation for our non-significant result by investigating whether the potential for resyllabifying the syllable final stop influences the extent of perceived incomplete neutralization. To do this, we re-tested H1 and H4. A by-product of this experimental analysis is also a test of the *licensing-by-cue* versus *licensing-by-prosody* hypotheses of neutralization, as outlined in the introduction.

2.3 Experiment 2: effect of resyllabification on incomplete neutralization

2.3.1 Speech materials

In Experiment 2, we reused the *Niedlinn-Nietlinn* (henceforth /C_{alv}-l/) and the *Mieglinn-Mieklinn* (henceforth /C_{vel}-ʃt/) continua from Experiment 1 as well as two newly created continua each formed from resynthesized stimuli of two minimal pairs that were once again based on hypothetical German town names: *Niedstein – Nietstein* (henceforth /C_{alv}-ʃt/) and *Miegstein – Miekstein* (henceforth /C_{vel}-ʃt/). As in Experiment 1, the blends were derived by combining the first syllable of a trochaic target word with the second syllable of two other context words. The target syllable of the test compounds were derived from the intervocalic productions of the two trochaic target words that were selected for Experiment 1: *Nieden* (/niːdən/) and *Miegen* (/miːgən/). The second syllable of the compound was, as before dependent on the syllable-final stop's place of articulation, either the suffix *–stein* taken from *Wirtstein* (/vɪrtʃtaɪn/) or the suffix *–stein* taken from *Birkstein* (/bɪrkʃtaɪn/). The velar place of articulation was chosen because, as described earlier, it is potentially resyllabifiable in a *–linn* context (since /gl/ and /kl/ are legal onset clusters in German) but not in a *–stein* context (because /kʃt/ or /ɡʃt/ are illegal onset cluster in Standard German). The method for generating the stimuli was the same as in Experiment 1. The vowel and stop closure durations as well as the step sizes of the decreasing V:C duration ratios were the same as for the stimuli of Experiment 1 (cf. Table 2.1). The alveolar place of articulation was also chosen to test the *licensing-by-cue* hypothesis. In Standard German both /tl, dl/ and /tʃt, dʃt/ do not occur in syllable-initial position. Therefore, according to the *licensing-by-prosody* hypothesis, /d-t/ is necessarily domain-final and therefore categorically neutralized preceding both /l/ and /ʃt/. On the other hand, according to the *licensing-by-cue* hypothesis, the /d-t/ distinction should be more perceptible preceding the sonorant /l/ than preceding the obstruent /ʃt/ cluster since, according to this theory, sonorants but not obstruents provide a favourable context for the perceptibility of the voicing contrast.

2.3.2 Participants and experimental procedures

The experimental and analysis procedure was the same as in Experiment 1 and was run with the same subjects. Again, we first analyzed our data by means of logistic regressions, in which the logit of the voiceless responses was the dependent variable and the V:C Ratio was the independent variable. Second, repeated measures generalized linear mixed models (GLMM) were fitted to predict incomplete neutralization as a function of the decreasing V:C duration ratio and to determine whether the degree of incompleteness also depended on the potential for resyllabification. The 'voiceless' responses served as our dependent variable. As predictors (independent variables), we entered Place (alveolar vs. velar), Manner (sonorant vs. obstruent), and V:C Ratio (stimulus 1 to 7). Subject was entered as a random effect factor.

2.3.3 Results and discussion

Figure 2.6 (a) gives the proportion of voiceless responses as a function of a decreasing V:C duration ratio (i.e. stimulus number) for the four continua. In order to test whether there was a discernible trend in the proportional responses along the continuum, the corresponding logistic regression lines were calculated, as shown in Figure 2.6 (b). These lines show that the identification functions for three of the four continua increase slightly, but gradually. Although these perceptual changes were not categorical, the increases in the voiceless responses along the continuum from left to right followed significant trends in /C_{vel}-l/ ($\chi^2(1) = 61.8$, $p < 0.001$), /C_{vel}-ʃt/ ($\chi^2(1) = 22.4$, $p < 0.001$), and /C_{alv}-l/ ($\chi^2(1) = 28.249$, $p < 0.001$), but not in /C_{alv}-ʃt/. The responses to /C_{vel}-ʃt/ were predominantly voiceless, but predominantly voiced to the other three continua. The results also show that listeners were much more likely to perceive a voiceless stop preceding an obstruent than a sonorant. From Figures 6 (a, b) it is also clear that velars were more likely to be perceived as voiceless than alveolars.

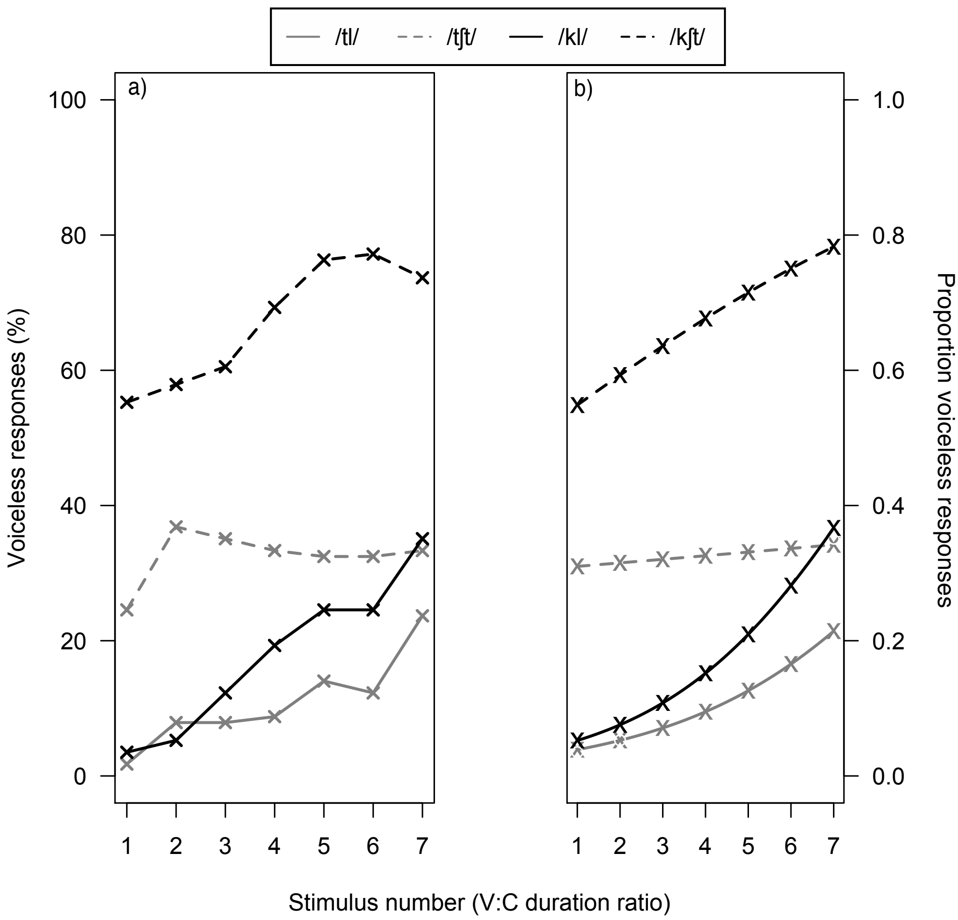


Figure 2.6: Proportion of 'voiceless' responses as a function of decreasing V:C duration ratio (stimulus number) to the four continua differing in the offset's place and the onset's manner of articulation (a) and the corresponding regression curves (b): alveolar (grey), velar (black), sonorant (solid), and obstruent (dashed).

The results of the overall GLMM in which the voiceless responses to all four continua were included as the dependent variable showed significant main effects for Place ($z = 15.0$, $p < 0.001$) as well as for Manner ($z = -10.4$, $p < 0.001$), but none for V:C Ratio. There were also significant interaction effects for Place x Manner ($z = -5.5$, $p < 0.001$), Place x V:C Ratio ($z = 3.3$, $p < 0.001$), and Manner x V:C Ratio ($z = 3.6$, $p < 0.001$). The non-significant main effect for V:C Ratio together with the significant interaction effects for V:C Ratio x Place and V:C Ratio x Manner support the idea that the perceptibility of the voicing contrast was influenced by both place and manner of articulation of the two consonants.

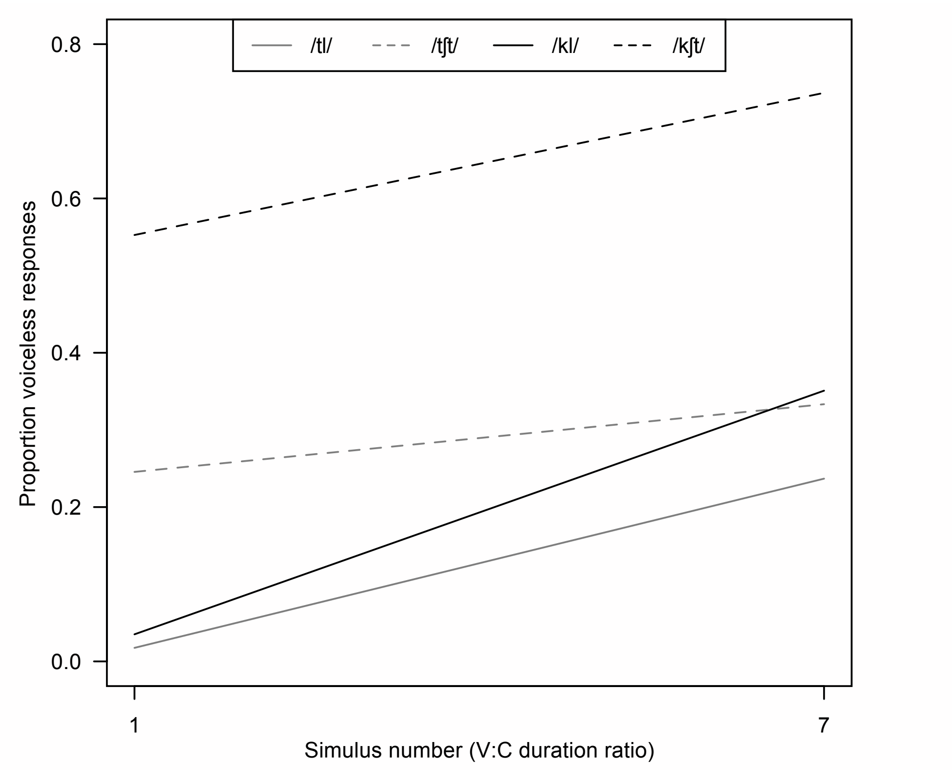


Figure 2.7: Linear interpolations between the proportional 'voiceless' responses to stimuli 1 and 7 of the four continua differing in the offset's place and the onset's manner of articulation: alveolar (grey), velar (black), sonorant (solid), and obstruent (dashed).

In order to assess the validity of H4, we compared the different effects that the acoustic cue V:C duration ratio had upon the disambiguation of the four continua's stimuli. According to H4, there should be less neutralization for resyllabifiable /C_{vel}-l/ than for the other three non-resyllabifiable continua. As can be seen in Figure 2.7, in which the difference between the first and the last stimuli of each of the four continua is linearly interpolated, the rise for the /C_{vel}-l/ identification function was steeper than for those of the other continua. Figure 2.8 also shows that the difference in the proportion of voiceless responses between stimulus 1 and stimulus 7 was greater for resyllabifiable /C_{vel}-l/ than for the other non-resyllabifiable continua, but the difference was also greater for stops preceding sonorants than obstruents.

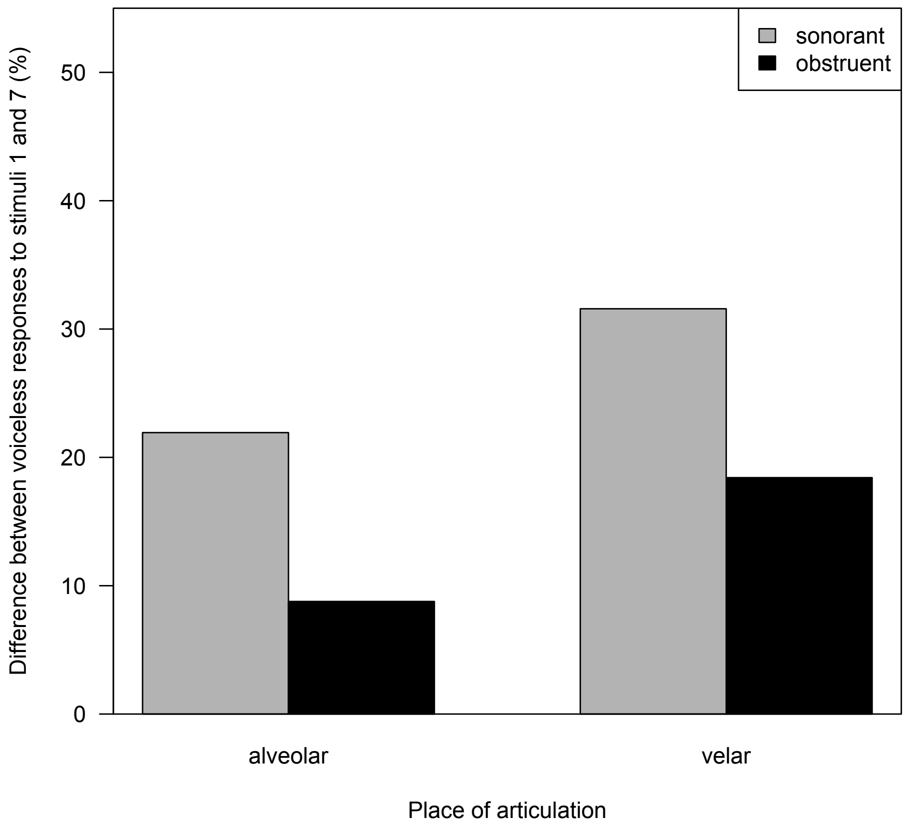


Figure 2.8: Proportional differences of ‘voiceless’ responses between stimuli 1 and 7 of the four continua.

In order to test whether identification functions differed significantly with respect to place and manner of articulation, we ran *post-hoc* analyses separately for the velar, alveolar, sonorant, and obstruent continua. There was a significant main effect for V:C Ratio (z = 3.3, p < 0.001) and also for Manner (z = -9.0, p < 0.001) and a significant interaction effect for V:C Ratio x Manner in the velar continua (z = 3.1, p < 0.01), which means that the /C_{vel}-l/ and

/C_{vel}-ʃt/ slopes differed from each other significantly. More specifically, the *post-hoc* analysis showed that the difference between stimuli 1 and 7 in the proportion of stimuli judged to be voiceless was greater in the sonorant than in the obstruent context. There was also a significant main effect for Manner (z = -4.5, $p < 0.001$) and a significant interaction effect for V:C Ratio x Manner (z = 3.1, $p < 0.01$) but no significant main effect for V:C Ratio in the *post-hoc* analysis for the alveolar continua. Therefore, the divergence between the sonorant and the obstruent series in the proportion of stimuli 1 versus 7 that were judged to be voiceless was about the same for both places of articulation.

On the other hand, there was a significant main effect of V:C Ratio in the sonorant continua (z = 4.1, $p < 0.001$), but neither a significant main effect for Place, nor a significant interaction effect for V:C Ratio x Place. This means that /C_{vel}-l/- and /C_{alv}-l/ did not differ significantly in their identification functions. Neither did /C_{vel}-ʃt/ differ from /C_{alv}-ʃt/ with respect to V:C Ratio: That is, there was neither a significant main effect for V:C Ratio nor a significant interaction for V:C Ratio x Place. However, there was an overall significant main effect for Place (z = 8.4, $p < 0.001$). Thus, there was a greater probability of perceiving the voicing contrast in a sonorant context than in an obstruent context irrespective of the final stops' place of articulation. But the potential for resyllabification also played a role in the perceptibility of the voicing contrast: resyllabifiable clusters with a sonorant onset consonant showed the least degree of perceptual neutralization whereas non-resyllabifiable clusters with an obstruent in syllable onset showed the highest degree of neutralization.

The results provide some support for H4 that the voicing contrast is less perceptible given the same acoustic cues in non-resyllabifiable than in resyllabifiable clusters. The *post-hoc* tests showed that listeners exploit the acoustic cue V:C duration ratio to a greater extent for the differentiation of voiced from voiceless stops when the resyllabifiable cluster contained a sonorant onset consonant as opposed to an obstruent onset consonant.

2.4 General discussion

Our aim has been to establish whether the fine phonetic detail in the neutralizing context of German final obstruents, for which evidence has been presented in a substantial number of production studies (e.g. Charles-Luce, 1985; Port & O'Dell, 1985; Port & Crawford, 1989) is perceptible and moreover whether these subtle acoustic differences are conditioned by factors such as phonological frequency and effects of syllable-position.

Our first hypothesis (H1) was that listeners perceive fine phonetic differences in the speech signal – in our experiment fine phonetic differences in the V:C duration ratio – but that its power to differentiate voiced from voiceless obstruents in neutralising contexts is substantially diminished compared with non-neutralising contexts. Our results were consistent with this hypothesis. Although listeners were able to perceive differences between voiced and voiceless stops in a neutralizing context when acoustic cues were available for their distinction, listeners' judgements between the stimulus endpoints shifted continuously rather than categorically. These results are consistent with the idea that there is incomplete perceptual neutralization to the stop voicing contrast, or at least that the fine phonetic details of the incompletely neutralized contrast are perceptible in this kind of neutralising context, but it is far from clear whether the cues are sufficiently powerful to distinguish unambiguously between stops. Previous research (e.g. Port & O'Dell, 1985) on the perception of word-final German obstruents using natural speech showed that listeners discriminate between derived and underlying voiceless obstruents better than chance. Our results are consistent with findings from other final-devoicing languages in which the voicing contrast was shown to be perceived gradually (Warner et al., 2004) or even categorically (Broersma, 2005) when obstruents were in a neutralising context. Our results are also consistent with those for Dutch listeners who were shown to 'borrow' the intervocalic duration cue to obstruent voicing word-finally (Warner et al., 2004). Taken together, all these results show that acoustic information in a neutralising context can be used for disambiguating voicing. There is further support from our results that incomplete neutralization is not an artefact of the experimental design (Fourakis & Iverson, 1984).

According to H2, there should be a greater probability of identifying stops as voiceless after lax than after tense vowels (i.e., *Widdlinn* and *Bigglinn* in our experiment should be perceived more often as *Wittlinn* and *Bicklinn* respectively), since lax vowels are so rarely preceded by underlying voiced obstruents in German. Our results confirmed that there was a greater probability of listeners identifying more stops as voiceless when they occurred after lax vowels. More specifically, for the same V:C duration ratio, listeners identified a greater proportion of syllable-final consonants as voiceless after lax than after tense vowels irrespective of the syllable-final stop's place of articulation. This finding supports the idea that incomplete neutralization also depends on the frequency with which a vowel and following consonant co-occur and that the phonological [± tense] feature is not syntagmatically independent of a following [±voice] in syllable-final vowel plus consonant sequences in German. That is, lexical frequencies of phoneme sequences strongly affect listeners' phoneme identification and categorization. In our materials, the lax vowel was a phonological cue that evoked the listener's expectation of a following voiced stop, i.e. listeners made predictions about the voicing category of the following stop based on their knowledge of phonotactic constraints. Our results are in line with findings that listeners adjust to high frequent categories (Hay et al., 2003; Pitt & McQueen, 1998) and therefore perceive more tokens as containing a frequent cluster (as opposed to an infrequent one).

As far as tense vowels and the stops' place of articulation are concerned, H3 predicted that the voicing contrast should be more perceptible for alveolar stops than for velar stops given that stop-voicing occurs frequently in the former (e.g., /laɪdən/ 'to suffer', /laɪtən/ 'to lead') but not the latter context. Velar stops after tense vowels should be perceived predominantly as voiced because of the paucity of underlying tense vowel plus voiceless velar sequences (i.e., sequences like /piːksən/, 'to prickle' with an underlying /iːk/ are very rare in German). Our results were not compatible with this position (H3) and showed instead that stop voicing was facilitated in a velar compared with an alveolar context. However, compatibly with H3 we did find that the stop voicing contrast was more perceptible in alveolars when they were preceded by tense than by lax vowels. This presumably comes about because the voicing contrast in alveolar stops preceded by lax vowels is infrequent, i.e. the stop is almost always a reflex of underlying /t/, whereas, the voicing contrast in alveolar stops preceded by tense vowels is frequent (e.g. /miːdən/, 'to avoid' vs. /miːtən/, 'to rent'). In the latter case, listeners are not biased by the lexical voicing distribution and therefore cannot make predictions about the following stop's voicing status. Instead, the voicing disambiguation has to be based on the acoustic information that is available. After lax vowels, the perceptibility of the acoustic cue is constrained due to the perceptual bias towards voiceless stops. Within this interpretation, the listener's meta-linguistic knowledge facilitates or restricts the perceptibility of a strong acoustic cue at least in certain environments (in this case in *Niedlinn – Nietlinn* and *Widdlinn – Wittlinn* respectively) and this together with H2 partially supports the claim made in a number of statistically based phonological models (Bybee, 2001, 2004; Pierrehumbert, 2001; Coleman, 2003) that probability has to be taken into account when modelling phonological neutralization.

The starting point for hypothesis 4 was to test whether resyllabification could explain our finding (which ran counter to H3) that stop voicing was facilitated in a velar compared with an alveolar context. H4 also allowed us to adjudicate between the two diverging phonological approaches to modelling final devoicing – *licensing by cue* (Steriade, 1997, 1999, 2000) and *licensing by prosody* (e.g. Itô, 1986, 1989; Goldsmith, 1990; Rubach, 1990). According to the latter, final devoicing is controlled by prosodic position, whereas according to the former, segmental context is decisive. More specifically, according to the *licensing-by-prosody* hypothesis, there should be a perceptual advantage for the voicing distinction in a velar as opposed to an alveolar context. This is because, since velars but not alveolars can form legal onset clusters with /l/ in German (e.g., /gl/ vs. /kl/, /glɔtsən/ 'to stare' vs. /klɔtsən/, 'to slog away'), listeners should be able to identify the voicing status of a velar stop in /V_1-stop-l-V_2/ in which the stop can potentially be interpreted as part of an onset cluster of the second /l-V_2/

syllable and is therefore not subject to domain-final neutralization. By contrast, since velars and alveolars precede the same sonorant context, then according to the *licensing-by-cue* hypothesis, voicing distinctions should be equally favourable in both contexts. In addition, voicing distinctions should be more perceptible preceding the sonorant /l/ context than preceding the /ʃt/ context because cues for the voicing distinction are less likely to be obscured preceding sonorants than obstruents.

We found some support for the idea that the voicing distinction is more perceptible in resyllabifiable velar plus lateral sequences as opposed to the non-resyllabifiable alveolar plus lateral sequences which is compatible with the prediction from *licensing-by-prosody*. We also found that voicing distinctions were in general more perceptible for both alveolars and velars preceding sonorants than fricatives. This lends some support to the *licensing-by-cue* hypothesis, since sonorants provide a more stable context than fricatives for the voicing distinction to be realised. It runs counter to the *licensing-by-prosody* hypothesis, however, which predicts that the perceptibility of the voicing distinction in alveolars preceding either sonorants or fricatives should be the same, since in both cases the alveolar is necessarily domain-final (because it is not resyllabifiable).

In general, our results provide some evidence that listener's knowledge of phonotactic constraints may perceptually mask lexical redundant acoustic cues (Cutler, 2002). Perceptual masking is the presumed process by which listeners ignore cues in the signal if they are irrelevant phonologically. We propose that perceptual masking varies with the extent to which neutralization is complete. For example, there was less neutralization in the resyllabifiable /kl/ cluster than in the non-resyllabifiable /kʃt/ cluster which suggests that listeners take advantage of the cues as long as they are consistent with phonological distributions.

A comparison of the results of both experiments shows predominant /g/ responses after tense vowels only preceding the sonorant -*linn*, but not preceding the obstruent -*stein* suffix. If listeners' responses had been guided by phonotactic probability, then they should have perceived /g/ preceding both -*linn* and -*stein*. The fact that they perceived /g/ before -*linn* and

/k/ before -*stein* might instead suggest that their perceptions were guided by regressive voicing assimilation and not by phonotactic probability: that is, the voiced sonorant context induced (regressively) primarily /g/ perceptions in the preceding stop whereas the voiceless obstruent context elicited mostly preceding /k/ perceptions. But on the other hand, if right context is the only factor in determining voicing perception of the preceding stop, then listeners should have perceived predominantly /t/ when the tense-vowel plus alveolar sequence preceded voiceless obstruents, but this is not what we found: they instead perceived /d/. Therefore, while phonotactic probability (H3) cannot entirely explain why tense vowel and following velar sequences should be perceived as /g/ before -*linn*, but as /k/ before -*stein*, neither in view of the results from alveolars can this be explained by regressive assimilation alone.

The different perceptual responses to the continua in this study – which depend to some extent also on probabilistic co-occurrences of V+C sequences, the phonetic environment, and the potential for resyllabification – cannot be easily modelled in a generative framework. Phonological rules in generative theories are either applied or not. For that reason these rules are less able to give expression to gradient changes in phoneme perception (unless they include a vast number of constraining rules). In probabilistic phonological theories such as exemplar or episodic models of speech perception (Pierrehumbert, 2001, 2002, 2003), a phonological category such as voicing is defined by the density distribution in an acoustic-perceptual space. This distribution is continuously expanded by new remembered exemplars, which are compared with previous stored exemplars and added to the neighbourhood of the most similar exemplars within that space. In such a usage based model, the density distribution for lexically frequent sequences (e.g. /bɪtən/, ‘to beg’) is very likely to be high and enlarged across the perceptual space and very low and narrowed for rare sequences (e.g. /ɛbə/, ‘tide’/), i.e. language dependent lexical frequency distributions are stored in the listener’s mental lexicon. As a consequence listeners re-interpret rare as frequently occurring sequences (e.g. Pitt & McQueen, 1998; Hay et al., 2003). On the other hand, in phoneme sequences that show a balanced distribution across the lexicon (e.g. /miːdən/, ‘to avoid’ vs. /miːtən/, ‘to rent’), a misinterpretation of a phonological feature is less likely since similar density distribution are to be expected. These predictions from exemplar theory are compatible with our findings for Experiment 1. Similarly, an exemplar framework predicts that the voicing contrast in velar stop + /l/ sequences should be readily perceptible analogous to the frequently occurring /kl/ vs. /gl/ contrast in German; and it also predicts that the voicing contrast should be much less perceptible before obstruents because there are no instances of a voicing contrast in this context in the lexicon. These predictions are compatible with our findings.[3]

Moreover, an exemplar model also explains the finding from Experiment 2 that listeners exploit the acoustic cue V:C duration ratio to distinguish voiced from voiceless velar stops in the non-resyllabifiable /C_{vel}-ʃt/ context. Despite the obstruent context and the non-resyllabifiable cluster, the discrimination performance of the stimuli of this continuum was rather good. One partial reason for this fairly good discrimination could be that forms such as [gʃtoːsn] ‘pushed’ and [gʃtoːln] ‘stolen’ in which the schwa from the first syllable is elided (the citation forms have initial /ɡə/ in both cases) are quite common in spontaneous and colloquial speech in German: consequently, native German listeners are likely to have been exposed a good deal to onset clusters such as /ɡʃt/. According to exemplar theories, remembered exemplars for these alternative pronunciations with /ɡʃt/ might be stored as alternative pronunciation forms in the mental lexicon. Furthermore, the discrimination performance for the obstruent and non-resyllabifiable /C_{alv}-ʃt/ continuum was very poor. Besides the unfavourable phonetic context and the missing potential for resyllabification – which were no such strong impediments to the perception of the voicing contrast in the /C_{vel}-ʃt/ cluster – the /tʃt/-cluster is, in contrast to the corresponding form with initial velars, a more or less non-occurring onset cluster in German even in spontaneous and colloquial speech. Therefore, we expect that German listeners cannot draw on stored /tʃt/ or /dʃt/ exemplars when discriminating voicing in this cluster. In the case of /ɡʃt/, however, German listeners will only have been exposed to devoiced [ɡ̊ʃ] (because of the

[3] For an overview of the problems arising from different models dealing with the phenomenon of incomplete neutralization (including a derivationalist and an exemplar based account) as well as a presentation of a non-linear dynamic framework to model incomplete neutralization (which includes context dependency as a predictor) see Gafos (2006).

high frequency of occurrence of the past tense marker /gə/ with the reduced or elided schwa) but scarcely to /kʃ/ onset clusters (since /kə/ is far less frequent than initial /gə/), so that we are not dealing with a potentially frequent voicing contrast in velar stop plus /ʃ/ sequences, which would, according to H3, have yielded a greater voice contrast perceptibility.

To conclude, we have found that the extent of perceptual neutralization is influenced by phonological frequency. Furthermore, the potential for resyllabification of the final stop enhances the perceptibility of the voicing contrast in domain-final obstruents. The phonetic environment is another important factor that affects the perceptibility. Final devoicing very likely cannot only be characterized in terms of either prosodic position or phonetic context. Instead prosodic position, phonological and phonotactic frequency as well as phonetic context have to be included to model adequately the neutralization of the final voicing contrast in German. But there may well be other contextual factors (which have not been addressed in the present paper) that affect the degree of incomplete neutralization such as, for example, placement of morphological boundaries in relation to the obstruent (e.g. Iverson & Salmons, 2007). Evidence has been provided that incomplete neutralization is an instance of fine phonetic detail that is perceptible when listeners are forced to distinguish underlying voiced from voiceless stops in a neutralizing context. Listeners exploit acoustic cues but only as long as these are consistent with phonotactic generalisations. Our findings so far show that incomplete neutralization plays a “substantial role in language processing” (Ernestus & Baayen, 2006: 27), but whether it is morphologically functional in German as has been shown for Dutch, where “it appears to be a subphonemic cue to past-tense formation” (Ernestus & Baayen, 2006: 27), still needs to be addressed. That is, do German listeners use fine phonetic information in potentially neutralizing environments, such as in the German imperative /heːp/ (‘lift’), in order to infer the correct obstruents in the corresponding German infinitive /heːbən/)? Or do German listeners reconstruct the correct forms based on the predictability from the phonological context (i.e., in this case underlying voiced labial stops after tense vowels) which was shown to be a contributing factor in the syllable-final voicing contrast in the present study?

Incomplete maintenance of the intervocalic voicing contrast in regional varieties of German

Chapter III

Abstract

The main purpose of this chapter was to investigate the extent to which Saxon and Bavarian speakers neutralize the post-vocalic stop voicing contrast in an alveolar (*boten* vs. *Boden*) and labial context (*Oper* vs. *Ober*). Another aim was to investigate whether the extent of neutralization was affected by age. Finally, the purpose of including the two contexts was to test whether neutralization was more likely in a labial context than in an alveolar context commensurate with the different statistical co-occurrence of tense vowels with stop voicing in these contexts. More specifically, we predicted that the extent to which the voicing contrast is maintained is greater for alveolar vs. labial stops, commensurate with a phonologically frequent voicing contrast in alveolar but not labial stops after tense vowels (although there is no such phonological frequency distribution in the broad dialects of Saxon and Bavarian). The second prediction was that Saxon speakers would neutralise the voicing contrast in Standard German to a greater degree than Bavarian speakers, commensurate with their complete neutralization of all intervocalic stops towards the lenis component in broad Saxon but contrast maintenance in broad Bavarian (in terms of varying vowel:stop duration ratio). The results showed that the voicing contrast was neutralized to a greater extent in labial than in alveolar stops except for young Saxon speakers. This supports the idea that the probabilistic co-occurrence of vowel quantity and stop voicing affects the strength with which a voicing contrast is produced. Secondly, Saxon speakers neutralized the intervocalic voicing contrast to a greater extent than Bavarian speakers. There were no significant main effects for Age.

3.1 Introduction

Standard German distinguishes between two sets of stops, /b d g/ and /p t k/, which are frequently referred to as voiced vs. voiceless stops. Yet, the question as to which distinctive feature(s) are most adequate to describe this contrast has been disputed for a very long time, e.g. voiced/voiceless, fortis/lenis, tense/lax, *tenues*/*mediae*, and long/short (for an overview see Braun, 1988). Voicing, which is manifested by a low frequency component in the acoustic signal, may distinguish /b, d, g/ from /p, t, k/: however, vocal fold vibration is also very often absent during the production of stops, because of physiological reasons. In the present study, we use the fortis/lenis distinction (i.e., we refer to underlying voiced and voiceless stops as lenis and fortis stops, respectively) since force and timing of articulation (duration of closing, closure, and release) and phonation (voicing during closure, aspiration) aspects are inherent in this feature (see Kohler 1984), and it is also consistent with the terminology in the literature on German dialects.

A number of parameters such as the duration and intensity of bursts, aspiration, voice onset time (VOT), fundamental frequency (f0) and formant transitions as well as the vowel to stop duration ratio (henceforth, V:C duration ratio) are important acoustic cues that signal the difference between fortis and lenis stops (not only in German, but also in many other languages that have such a contrast, e.g. English). For example, /b, d, g/ have weaker burst and shorter VOT durations than /p, t, k/. Many of the cues are directly related to each other: a longer closure duration is likely to lead to greater supra-glottal air pressure behind the closure which may then result in a longer aspiration phase and a burst with higher intensity. The acoustic parameters differ especially in their perceptual significance, i.e. there is a hierarchical ordering of acoustic cues for the differentiation of lenis from fortis stops. The duration of aspiration is a very pronounced acoustic correlate, which listeners use to perceive the phonological contrast (Jessen, 1998). The presence of a cue depends on syllable position and phonetic environment. E.g. in word-initial position a long vs. short aspiration phase indicates fortis and lenis stops respectively while the V:C duration ratio cue is of course absent and thus cannot be used. In disyllabic trochaic words that end in *-en* and in which the stops occur in intervocalic position, stops are very often released by means of nasal instead of oral plosion, as a result of which the aspiration cue is lost. The 'trading relations' (Jessen, 2001: 264) between the acoustic cues enable the listener to use the second most important cue that distinguishes fortis and lenis stops. According to Kohler (1979), in stops with a nasal release the V:C duration ratio is the most important cue, followed by formant transition and voicing; in stops with an oral release, however, aspiration supersedes the V:C duration ratio as the most relevant acoustic cue. Evidence for the importance of this acoustic cue has been provided in other studies (Kohler, 1982; Mansell, 1979; Jessen, 1998). In an acoustic analysis of tense vowels followed by lenis or fortis stops, Kohler (1977) found that the durations of the vowel (V) plus stop (C) sequences were about the same, irrespective of the stop's voicing category, but that the vowels were relatively shorter before fortis stops which had longer closure duration and vice versa for sequences containing lenis stops. VC sequences differed slightly with respect to the stop's place of articulation (Kohler & Künzel, 1978) with labial stops having longer closure and shorter vowel durations than alveolar stops, but the variation in intrinsic duration did not alter the findings regarding the overall VC duration or the V:C duration ratio. This place-dependent closure duration, which was also found for other languages (e.g. for English, Stathopoulos & Weismer, 1983; Byrd, 1993), may be due to the fact that the cavity behind a labial closure is greater than behind an alveolar closure, and thus, a longer closure duration is necessary to reach an intraoral air pressure that is high enough to cause an explosion (cf. Maddieson, 1997).

3.1.1 Syllable position and the fortis/lenis contrast in Standard German

This phonological voicing contrast, however, is not always maintained in Standard German, but depends on syllable position (e.g. Vennemann, 1972, 1982; Hall, 1992, 2000; Wiese, 1996) and phonetic environment (Steriade, 2000). In syllable initial and word medial, intervocalic position the contrast is preserved, but in syllable final position the voicing contrast is phonologically neutralized towards the fortis category. That is words like *Bad* ('bath') and *bat* (1 sg.

'requested') or *Rat* ('advice') and *Rad* ('wheel') are homophones on the phonetic surface, even though they differ in the underlying voicing status of the final stops. Phonetically the voicing contrast is retained in the following (near) minimal pairs where the stop occurs inter-vocalically in the syllable onset: *baden* (/baːdən/, 'to take a bath') vs. *baten* (/baːtən/, 1 pl. 'requested') and *radeln* (/raːdəln/, 'cycling') vs. *raten* (/raːtən/, 'to advice').

The articulatory, and thus acoustic and perceptual strength of consonants depends in many languages on the syllable position and phonetic environment. Evidence has been presented for domain-initial articulatory (Fougeron & Keating, 1997 on American English; Jun, 1993 and Cho, 1998 on Korean; Keating, Cho, Fougeron, & Hsu, 2003 on French, Korean and Taiwanese) as well as prosodic strengthening (Cho & McQueen, 2005 on Dutch; Kuzla, Cho, & Ernestus, 2007 on German) and domain-final devoicing is a phonological process that operates in Standard German as well as in many other languages such as German, Catalan, Dutch, Russian, etc. On the other hand, intervocalic stops frequently become lenis as the contact between the active and passive articulator is weakened (Kohler, 1984). This intervocalic lenition occurs both synchronically, e.g. in connected speech (especially in fast, colloquial speech), and diachronically as numerous historical sound changes demonstrate (e.g. /b/ in Latin *caballum* (*horse*) has become /v/ in Italian *cavallo*). Historical sound changes are also the reason for many differences between German dialects and Standard German regarding the fortis/lenis contrast.

3.1.2 Fortis/lenis distinction in German dialects

The structural effects of syllable position and phonetic environment can also be observed in many German dialects[1] which differ from Standard German with respect to the positions in which the voicing contrast can be neutralized. In the so-called Central High German lenition (CGL) the fortis stops /p, t, k/ become /b, d, g/ (cf. 1.4). The term refers to the area in Upper and Central German where the neutralizing process takes place. Depending on the regional variety,

[1] There is a fine differentiation between various dialect levels from primary dialect to the standard language in the sociolinguistic literature. We differentiate between primary dialect, regional variety (colloquial speech) and standard language.

lenition occurs in different word positions. For example, in some low Alemannic areas the voicing contrast is neutralized only domain-initially, but word medially it is maintained. In other areas such as East Franconian, the contrast is neutralized in all positions, i.e. *Garten* ('garden') and *Karten* ('cards') are both realized as /gaːrdən/.

The German dialects Upper Saxon and Central Bavarian[2] both belong to the area where the CGL lenition applies. In both dialects, all initial stops are regarded as lenis, since they are unaspirated though voiceless (e.g. Rues et al., 2007 for Upper Saxon and Wiesinger, 1990 for Central Bavarian), i.e. both underlying lenis and fortis stops are realized as lenis stops in initial position, e.g. *dir* ('your') and *Tier* ('animal') are realized as [di:r]. However, in hyper-corrected or hyper-articulated speech, lenis plosives may be initially strengthened and realized as fortis, e.g. the female first name /grɛːtə/ may become [kre:te]. That is, the initial voicing opposition seems to be collapsed in any case, but whether it is neutralized towards the lenis or fortis component may depend on speaking style and situation.

In medial position, fortis plosives again are realized as lenis plosives in both dialects. However, there is one major difference between Upper Saxon and Central Bavarian concerning the postvocalic stops. Whereas in Upper Saxon fortis stops are neutralized towards lenis in all contexts, there is a correlation between vowel quantity and the voicing of the subsequent obstruent in Central Bavarian: lenis obstruents are preceded by tense or long vowels while fortis consonants are always preceded by lax or short vowels[3].[4] That is, one of the most important acoustic cues that distinguish fortis from lenis stops has become phonologized resulting in a syntagmatic relation between quantity and voicing. As a result of the so-called Central Bavarian Lenition, the phonological status of vowel quantity is at issue in Central Bavarian, which we

[2] When we refer to the primary dialects and their phonological systems we use the correct dialect names Upper Saxon and Central Bavarian. The simple forms Saxon and Bavarian refer to the regional varieties as spoken by our test groups.
[3] We use the tense/lax dichotomy to denote a phonological vowel length contrast without addressing differences in vowel quality. The terms short and long are used to describe fine phonetic differences within one phonological category.
[4] The so-called Central Bavarian lenition applies also word-finally causing the preceding vowel to be long.

will investigate in chapter IV. Table 3.1 summarizes the different realizations of stops in different domain positions as they are typically presented in the literature on German dialects.

Table 3.1: Traditional summary of neutralization in Upper Saxon and Central Bavarian.

Position	Upper Saxon	Central Bavarian
initial	p t k > b d g	p t k > b d g
medial	p t k > b d g	After tense vowel: p t k > b d g After lax vowel: p t k > p t k

3.1.3 Statistical co-occurrence of vowel tensity and stop voicing: asymmetries in the German lexicon

Similarly to Central Bavarian, the productivity of vowel stop combinations is restricted with respect to phonological vowel quantity and the voicing of the following consonant in Standard German (Braunschweiler, 1997: 357). Lax vowels are almost always followed by fortis stops, irrespective of the place of articulation, for example *Lippe* (/lɪpə/, 'lip'), *Lappen* (/lapən/, 'cloth'), or *Bock* (/bɔk/, 'billy goat'). Combinations of lax vowels and lenis stops are very rare and occur mostly in loan words from Low German such as *Robbe* (/rɔbə/, 'seal') and *Ebbe* (/ɛbə/, 'tide').[5] On the other hand, there is a bias for lenis labial and velar stops to be practically exclusively preceded by tense vowels, e.g. *lieben* (/liːbə/, 'love'), *laben* (/laːbən/, 'to feast on sth.'), or *Bogen* (/boːgən/, 'bow'). Lax vowels preceding lenis labial or alveolar stops by contrast are very infrequent (/piːksən/, 'to prickle').[6] In contrast to Central Bavarian, this is not a phonological regularity in Standard German, but an unequal distribution which can be observed for many, but not all combinations in the lexicon. After tense vowels, both lenis and fortis alveolar stops occur frequently as the examples given in 3.1.1 show. That is, in Standard German, vowel quantity and stop voicing have to be considered phonemic since stop voicing

[5] The double consonant in a number of Southern German place and family names such as for example *Eggersdorf, Glanegg* or *Ruggenbauer* is a spelling reflection of the historical geminates which were shifted from /g/ to /k/ in Bavarian during the Second High German sound shift (Wiesinger, 1990: 458; Schmidt, 2000).

[6] On the other hand, some of the words containing these rare sequences are quite frequent, e.g. *Opa* (/oːpa/, 'grandpa').

cannot be predicted by the preceding vowel and it is not necessarily the trigger for a phonological vowel quantity (or vice versa) as it is in Central Bavarian. However, for many combinations, one or the other category is statistically more likely. To summarize, vowel length and obstruent voicing always correlate in Bavarian. A similar correlation holds true in Standard German, but with one exception: in tense vowel plus alveolar stop combinations fortis and lenis consonants are equally distributed.

Recent studies have shown increasing evidence that neighbourhood density (i.e. the amount of words that are phonologically similar) and word frequency are important factors that have to be included in models of speech productions and speech perception. For example, Wright (2003) has shown that vowels were produced with greater centralization when they occurred in frequent words with a low number of lexical neighbours as opposed to rare words with a high number of lexical neighbours. The starting point for Wright's study was a pilot investigation by Goldinger and Van Summers (1989) who found an increased VOT in words from dense neighbourhoods as opposed to words from sparse neighbourhoods. Munson (2007) found both that neighbourhood density and word frequency had a marked influence on the vowel space dispersion and considered the word frequency effect to be important during the process of lexical access, but considered neighbourhood density to play a greater role after a certain lexical item has been activated, thereby having an immediate effect on speech production. Semantic predictability was shown to affect vowel duration and dialect-dependent changes regarding the vowel space: General American speakers who spoke a Southern variety showed an overall vowel space reduction in a high predictable semantic context, while subjects who spoke a Northern variety revealed a more extreme Northern cities vowel shift (Clopper & Pierrehumbert, 2008). All those studies share the assumption that speech is reduced to an extent to which it is still intelligible, with the degree of reduction depending on the context: a high predictable context leads to hypoarticulated speech with an increase in reduction, a low predictable context to hyperarticulated and more clear speech (Lindblom, 1990).

Word frequency effects are often investigated in speech perception studies and more specifically with word recognition studies. Neighbourhood density was added as a factor since word frequency only does not predict intelligibility very well. Luce and Pisoni (1998) provided evidence that high frequent words or low frequent words with a sparse neighbourhood density were better identified than infrequent words especially when they are from a dense neighbourhood. Listeners are inclined to misperceive statistically infrequent with frequent sequences (Hay et al., 2003) and show a perceptual bias towards the more likely consonant when asked to respond to ambiguous consonant stimuli in consonant clusters (Pitt & McQueen, 1998). In chapter II it was shown that the probabilistic co-occurrence of vowel tensity and stop voicing affects the voicing perception of domain-final stops, i.e. in a position in which the contrast is neutralized: compatibly with the distribution in the lexicon, stops following lax vowels were perceived more often as fortis and stops after tense vowels were more frequently perceived as lenis. However, the results did not clearly support the hypothesis in which we predicted that listeners perform better in the discrimination of lenis and fortis stops when the voicing is equally distributed in the lexicon (as in the case of tense vowels plus alveolar stops): on the one hand, the underlying voicing of alveolar stops was better identified in a tense vowel than in a lax vowel context, but on the other hand, the underlying voicing contrast was perceived to a greater extent in velar as opposed to alveolar stops following tense vowels, although there is a paucity of underlying tense vowel plus fortis velar sequences in the German lexicon.

Thus, one aim in this chapter is to investigate whether the statistical co-occurrence of vowel and stop voicing affects the extent to which the voicing contrast is maintained in production, although the phonological frequency distribution in those dialects differs from Standard German (cf. 1.2 above). In the present study (as in the last chapter II) we only address the factor phonological frequency, which comprises both word frequency and neighbourhood density.

3.1.4 Incomplete neutralization

The voicing contrast is often considered a dichotomy and thus the neutralization of this contrast is treated as a categorical change, irrespective of whether it occurs dialect-dependent intervocalically or word-finally because of a phonological rule. For example, the phonological process of domain-final voicing neutralization is often generalized in a phonological rule such as

[+obstruent] → [-voice] / ___db (where db stands for domain boundary)

This rule, at the same time, suggests that neutralization is complete and results in an absolute phonemic category change. However, acoustic analyses have repeatedly provided evidence that the phonological process of neutralization is gradual in many languages, i.e. speakers make subtle differences in their productions of different voicing categories in neutralizing positions. Among those languages were Catalan (Dinnsen & Charles-Luce, 1984; Charles-Luce & Dinnsen, 1987), Polish (Slowiaczek & Dinnsen, 1985), and German (Port & O'Dell, 1985). Port and colleagues, for example, found in the acoustic analysis of read isolated German words significant longer vowel durations and more voicing into and during closure when vowels preceded word-final underlying lenis obstruents as opposed to morpho-phonemically fortis obstruents (Port et al., 1981; O'Dell & Port, 1983; Port & O'Dell, 1985), i.e. the values of the acoustic cues for derived fortis obstruents such as vowel duration were somewhere between the values for lenis versus fortis obstruents in domain-initial or medial position – though the measurements were closer to the fortis cues. Vowel duration was shown in all studies, independently of the language, to be the most prominent parameter that is used to distinguish lenis and fortis stops. The degree to which the voicing contrast is maintained also depends on the phonetic environment, sentence position, and semantic information (Charles-Luce, 1985, 1993; Slowiaczek & Dinnsen, 1985). More recently Piroth and Jancker (2004) found (in)complete neutralization to be not only speaker- but also dialect-dependent: Only Southern German speakers maintained differences in the duration of the obstruents, though only in utterance-final position, while speakers from the Central German area neutralized the voicing contrast in all domain-final position. Obstruents in all other positions (syllable-final, word-final, and utterance-final) where completely neutralized by all speakers. The differences in the phonological systems of the dialects spoken in the three areas from which the subjects originated may well have contributed to these findings. Given that the Central Bavarian dialect allows domain-final voicing (see above footnote 4 in this chapter and cf. 1.4.2), this may have been the reason why the South German speakers only partially neutralized the stop voicing contrast. That is, the dialect feature has carried over to some extent in the standard variety leaving its marks in form of fine phonetic detail. Some researchers have argued that incomplete neutralization only occurs in laboratory speech where speakers may hyperarticulate, and indeed, orthography (Fourakis & Iverson, 1984; Jassem & Richter, 1989; Warner et al., 2004) as well as speaking style and situation (Port & Crawford, 1989) have been found to affect whether neutralization is complete or incomplete. Testing the perceptibility is an important step into resolving these contradictory arguments and it was shown in the last chapter that linguistic factors such as phonological frequency and the potential for re-syllabification affect the extent of incomplete neutralization in speech perception (see above 3.1.3). However, in the present study these inconsistent interpretations are not relevant.

Similarly, dialect dependent neutralization of intervocalic obstruents may also be incomplete (especially when speakers of those dialects produce a variety of Standard German). To our knowledge, there are no acoustic analyses of fine phonetic differences in voicing neutralization in intervocalic, word-medial stops between dialect groups in German. The study by Piroth and Janker (2004) on word-final devoicing may be taken as a starting point indicating that a speaker's regional background affects the degree of incomplete neutralization. Piroth and Janker (2004) provide a very detailed acoustic analysis of a large set of word paradigms. However, the basis for the analysis are the recordings only of six speakers, two each per dialect region, and the dialect classification was rather coarse. For example, the South German subjects were from two different dialect areas, East Franconian and Central Bavarian, and Central

Bavarian Lenition occurs only in the latter but not in the former dialect (see Rowley, 1990a for East Franconian and Wiesinger, 1990 for Central Bavarian). Recent studies provide evidence that speakers of Standard German with regional backgrounds that are part of the Central High German Lenition area show neutralization patterns word-medially in Standard German. Saxon speakers, for example neutralize the contrast both in production and perception (John, 2004). The degree of neutralization, however, seems to depend also on age: while Old East Franconian speakers do neutralize the contrast in perception and in production, young East Franconian speakers distinguish the contrast again both perceptually and acoustically (Müller, 2010).

3.1.5 Dialect levelling

Another aim of this study is to establish whether younger speakers of a variety tend to produce a dialect feature to a lesser degree based on the more general assumption of dialect levelling (Trudgill, 1986; Auer, Hinskens, & Kerswill, 2005; see above 1.4.3 and 1.5) as well as the findings for German dialects, that the number of dialect features has decreased in the speech of younger generations compared to older generations (e.g. Lameli, 2004; Wagener, 2002). More specifically, Barden and Großkopf (1998) found that among Saxon speakers who moved to other dialect areas, younger speakers lost or gave up dialect features more quickly than did older Saxon speakers. Recent studies on sound change have shown how important it is to incorporate fine phonetic detail (Harrington et al., 2008) within models of sound change such as the one proposed by Ohala (1993). Acoustic analyses of sound changes in different English varieties have shown that age and/or dialect groups differ at an earlier stage only in fine phonetic details, with phonetic differences becoming more marked at later stages (Hawkins & Midgley, 2005 on Standard Southern British English; Gordon et al., 2004 on New Zealand English), which at some point become perceptible and may even turn into new variants or phonemes (cf. Wells, 1982). For example, evidence has been provided for diachronic Southern Standard British English /u/-fronting within a single speaker (Harrington et al., 2000a, b; Harrington 2007) and between age groups within a single speech community (Hawkins & Midgley, 2005; Harrington et al., 2008). Presuming that dialect features are less present in young than in old speakers for German varieties, we should expect a general difference in the neutralization degree between young and old Saxon speakers because of the intervocalic lenition in the broad Upper Saxon dialect. Furthermore, we should also expect differences between young and old speakers in general, i.e. irrespective of the regional background, because the statistical co-occurrence of vowel tensity and stop voicing is asymmetric in Standard German with respect to the place of articulation but symmetric in both broad dialects (stops are lenited in all contexts in Saxon and only lenis stops occur after long vowels in Bavarian irrespective of the place of articulation). However, in this chapter we do not make predictions concerning which of the two dialects show a greater potential for dialect levelling.

3.1.6 Hypotheses tested in this study

To summarize, there is a voicing contrast in intervocalic, word medial position in Standard German, although there is a place of articulation dependent asymmetry in the German lexicon regarding the phonological frequency with which tense vowels and lenis vs. fortis stops co-occur: the voicing contrast is equally distributed in alveolar stops but there is a bias for lenis labial stops to follow tense vowels. In Upper Saxon, the intervocalic voicing opposition is neutralized in all contexts. In Central Bavarian, by contrast, both lenis and fortis stops can occur in intervocalic position but they are complementarily distributed since vowel length and stop voicing correlate.

The aim of the paper is to investigate the extent to which Saxon and Bavarian speakers maintain the intervocalic stop voicing contrast in an alveolar (*boten* vs. *Boden*) and labial context (*Oper* vs. *Ober*). The main research questions are whether phonological frequency affects the degree to which the voicing contrast is maintained and whether Saxon and Bavarian speakers differ with respect to the neutralization degree since the dialects show different neutralization patterns. Another aim was to investigate whether the extent of neutralization was affected by age. The hypotheses tested in this experiment can be summarized as follows:

(H1) The voicing contrast is maintained or only incompletely neutralized in the speech of both Saxon and Bavarian speakers when they produce a variety of Standard German.

(H2) There is a bias towards more voicing neutralization in the production of labial than in alveolar stops, since in Standard German there is a phonological frequency bias for lenis labial stops after tense vowels, while lenis and fortis alveolar stops are equally distributed after tense vowels.

(H3) The difference between the (incomplete) maintenance of the voicing contrast in labial vs. alveolar words is greater for young than for old speakers, irrespective of the dialectal background based on the general assumption of dialect levelling in the direction of Standard German.

Despite our prediction in hypothesis H1, that both speaker groups maintain the contrast, we predict that Saxon speakers tend to neutralize the contrast to a greater extent than Bavarians (although incomplete), because Upper Saxon is traditionally considered to be a dialect in which the voicing opposition is neutralized whereas in Central Bavarian the contrast between lenis and fortis stops is maintained by means of the preceding vowel length. Thus, we formulate hypothesis H4 as follows:

(H4) There is a greater tendency for voicing neutralization in lenis vs. fortis stops following tense vowels in the production of Saxon compared to Bavarian speakers.

(H5) There is more voicing neutralization in the productions of old vs. young Saxon speakers because of dialect levelling in the direction of Standard German.

3.2 Method

3.2.1 Materials and speakers

To test hypotheses H1 – H5, all minimal pairs that contained tense vowels but differed in the underlying voicing of the word medial stop were chosen from the corpus described in 1.6.2 for analysis: *boten* (/boːtən/, 'offered' 1 pl.) vs. *Boden* (/boːdən /, 'floor') and *Oper* (/oːpɐ /, 'opera') vs. *Ober* (/oːbɐ/, 'waiter'). It is important to test the hypotheses on Standard German words that differ only in the voicing contrast but not in phonemic vowel length in order to compare the neutralization degree between the two varieties and to examine whether phonetic vowel duration relative to the duration of the vowel plus stop sequence is an acoustic correlate of stop voicing in these two varieties. There were only two minimal pairs in the corpus that fulfilled these criteria (cf. Appendix A, Table A.1).

One young Saxon speaker was excluded from the analysis because of too many missing *Boden* tokens. That is, the data of 8 young (6 female) and 13 old (5 female) Saxon speakers as well as 12 young (8 female) and 8 old (5 female) Bavarian speakers were analyzed in the present study. A total of 1624 tokens were included in the analysis. Three *boten*, six *Boden*, two *Ober* and five *Oper* tokens were excluded because they were misread or produced too late. The distribution of the 1624 tokens by dialect and age is shown in Table 3.2.

Table 3.2: Number of tokens for each word analyzed in this experiment for young and old, Saxon and Bavarian speakers shown separately by stop.

Stop	Saxon		Bavarian	
	Young	Old	Young	Old
/t/	80	127	120	80
/d/	80	126	119	79
/p/	79	130	116	80
/b/	78	131	120	79

3.2.2 Segmentation and data analysis

The data was first automatically segmented with MAUS (Schiel, 2004) and then segment boundaries were manually checked and corrected in Praat (Boersma & Weenink, 2008) when necessary. The test words were segmented according to the following criteria: the beginning and end of the vowel was marked by the second formant's on- and offset; the stop's beginning corresponds to the vowel's right boundary and its end is set to the beginning of voicing either of the following schwa or nasal where stops were released nasally. That means burst, aspiration phase and VOT were included in the duration of the stop.

Word, vowel, and stop durations were extracted in Emu/R and normalized for word or VC duration by dividing the duration of the VC sequence by (i.e. normalized for) word duration (henceforth VC:Word duration ratio) and the vowel duration by the VC duration (henceforth V:C duration ratio). Only the various duration parameters were chosen for analysis since the V:C duration ratio parameter was shown to be one of the most robust acoustic cues to signal the fortis/lenis distinction (Kohler, 1977), especially in words with an unaccented *–en* ending in which stops are often released by nasal plosion (see 3.3.5 below for a distribution of nasally vs. orally released vowels). To quantify the degree of neutralization, we calculated the mean difference between the V:C duration ratio in all lenis and the V:C duration ratio in all fortis tokens for each minimal pair and speaker resulting in 82 values (two minimal pairs x 41 speakers).

Two types of statistical analyses were carried out in R. Four mixed models were fitted to predict whether lenis and fortis stops differ with respect to absolute (1) word, (2) vowel, and (3) stop duration as well as (4) the VC:Word duration ratio. One of the four duration values was always the dependent variable. Voicing (two levels: underlying fortis vs. underlying lenis), Age (two levels: old vs. young), and Region (two levels: Saxon vs. Bavarian) were entered as fixed factors, Speaker and POA (place of articulation) as random effect factors. Since the determination of the degrees of freedom in the denominator is problematic, we set df to be 60 and chose an alpha level of 0.01 in all mixed models (for a detailed explanation see Reubold, Harrington & Kleber, 2010: 641). One reason for running mixed models on these duration measures was that we wanted to factor out POA because the two minimal pairs differ with respect to word duration and VC:Word duration ratio irrespectively of any intrinsic differences (Kohler & Künzel, 1978) or the predictions made in hypothesis H2.

In order to test hypotheses H2 – H5 we calculated two repeated measures MANOVAs with independent variables POA (two levels: alveolar vs. labial), Age (two levels: young vs. old), and Region (two levels: Saxon vs. Bavarian). The dependent variable was either the mean V:C duration ratio (averaged over the 10 repetitions per speaker) or the mean V:C duration ratio difference between fortis and lenis tokens. Repeated measures MANOVAs were chosen because POA was one of the predictors of the V:C duration ratio in lenis vs. fortis stops and in the case of the mean V:C duration ratio difference the number of analyzed tokens was too small (82) for a mixed model. In case of significant interactions post hoc tests with Bonferroni correction were carried out.

3.2.3 Expectations and predictions

The question whether a phonological contrast is preserved, neutralized or even incompletely neutralized in speech production is not easy to answer, considering that the V:C duration ratio will never be 1 for lenis (because then the stop would have been elided) or 0 for fortis stops (because then the vowel would have been elided) and therefore the mean V:C duration ratio difference will neither never be 1, which would correspond to no neutralization at all, or zero, which would signal 100% neutralization. On the other hand not all values between 0 and 1 signal incomplete neutralization. In this experiment, a contrast is preserved when the two components of the voicing contrast differ significantly at the 1% level. Our measure of the neutralization degree, the mean V:C duration ratio difference, allows us to compare speaker groups with respect to the tendency with which a certain group neutralizes the voicing contrast. The predictions from hypotheses H1 – H5 are schematically outlined in Figure 3.1: We expect more neutralization in sequences with labial vs. alveolar stops (H2) and for Saxon vs. Bavarian speakers (H4) as well as a greater neutralization degree difference between alveolar and labial stops for young vs. old speakers (H3). Furthermore, we expect more neutralization in the productions of old vs. young Saxons (H5).

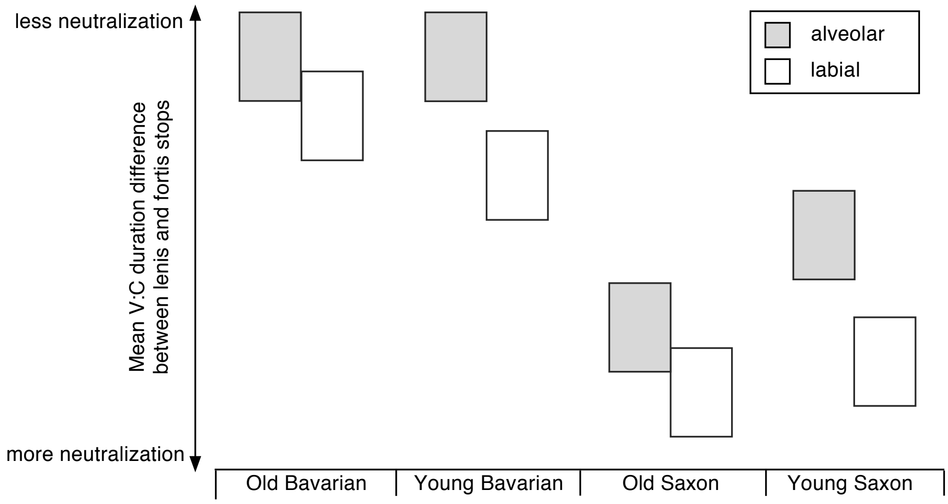


Figure 3.1: Schematic representation of hypothetical mean V:C duration ratio differences between lenis and fortis stops for alveolar (grey) and labial (white) plosives separately for old and young, Bavarian and Saxon speakers.

3.3 Results and discussion

3.3.1 Word duration

In both regional varieties, the word duration was longer for sequences containing a fortis than a lenis stop. Figure 3.2 shows that words with fortis stops have longer durations. The great variation in word duration for all speaker groups comes about because of the different word length of *boten*/*Boden* vs. *Oper*/*Ober*. A Mixed Model with dependent variable word duration showed a significant main effect for Voicing ($F[1,60] = 210.0$, $p < 0.001$) and no other significant effects. The significant main effect indicates that word duration differs only with respect to the underlying voicing of the stop in both varieties, but the word duration difference between lenis and fortis tokens is about the same for old and young, Bavarian and Saxon speakers.

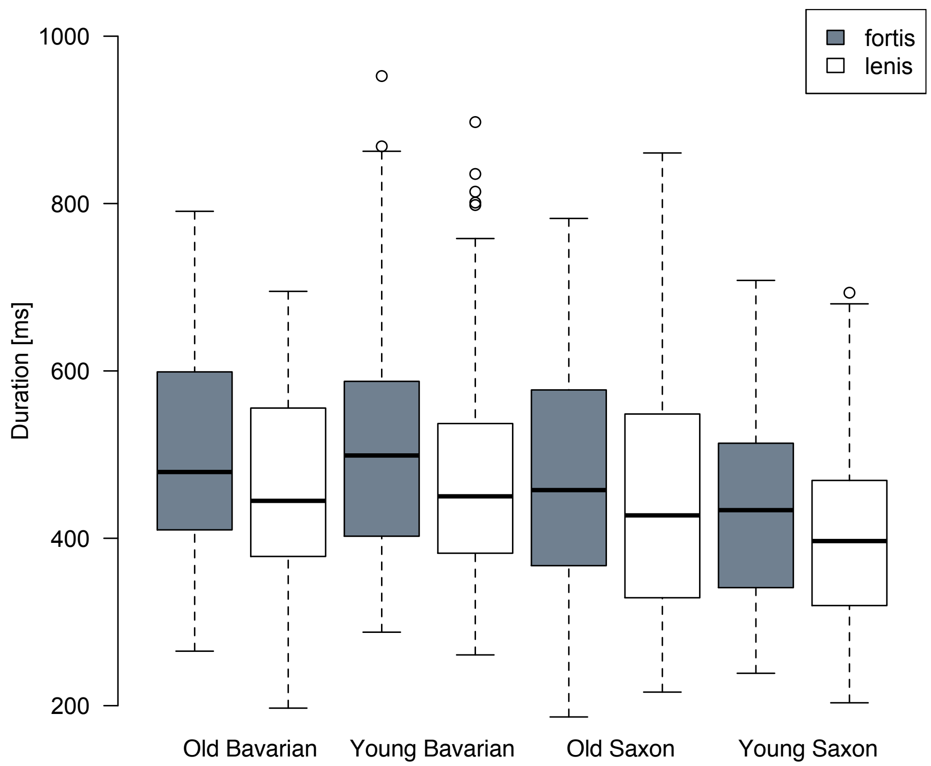


Figure 3.2: Absolute word durations for words containing fortis (grey) and lenis (white) stops separately for old and young, Bavarian and Saxon speakers.

3.3.2 VC:Word duration

For both dialect groups longer VC:Word duration ratios were found for fortis than lenis stops. A Mixed Model with dependent variable VC:Word duration ratio revealed a significant main effect for Voicing ($F[1,60] = 479.8$, $p < 0.001$) and a significant interaction for Voicing x Region ($F[1,60] = 29.6$, $p < 0.001$). The significant main effect indicates that in both regional varieties the VC:Word duration ratios differ with respect to the underlying voicing of the stop age group. As can be seen in Figure 3.3, VC sequences with lenis stops had significantly shorter VC:Word duration ratios than VC sequences with fortis stops. To what extent lenis and fortis stops differed depends on the variety of the speakers as the significant interaction effect between Voicing x Region suggests: the VC:Word-duration difference was greater for Bavarian speakers than for Saxon speakers (cf. Figure 3.3), which indicates that Saxon speakers tended to neutralize the voicing contrast more than Bavarian speakers.

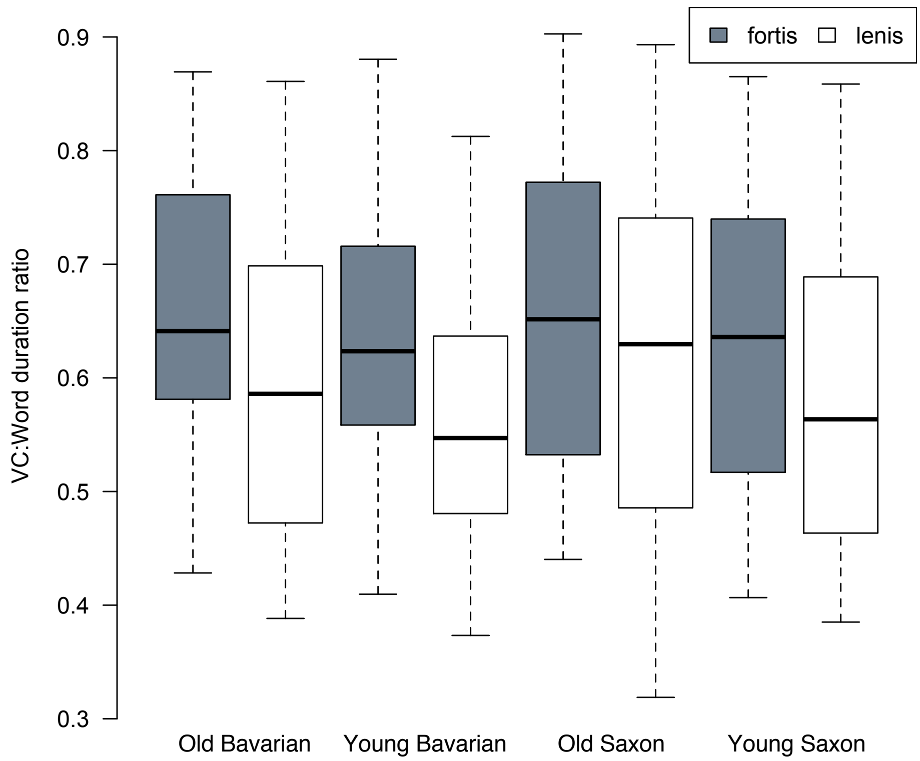


Figure 3.3: Proportional vowel plus stop durations for fortis (grey) and lenis (white) stops across both minimal pairs separately for old and young, Bavarian and Saxon speakers.

3.3.3 Vowel and stop duration

Because there is a difference in word and relative VC duration, we also analysed the absolute vowel and stop duration in order to test which of the two sounds, vowel or stop, is more prone to variation. As can also be seen from Figure 3.4, there is a greater difference between lenis and fortis sequences in stop duration as opposed to vowel duration. While sequences with fortis stops have similar vowel and stop durations, the stop duration is considerably shortened in lenis sequences and simultaneously the vowel duration only slightly lengthened in contrast to vowels preceding fortis stops. For the dependent variable vowel duration, there was a significant main effect for Voicing ($F[1,60]=582.0$, $p < 0.001$) as well as a significant interaction between Voicing x Region ($F[1,60]=24.3$, $p < 0.01$). For the dependent variable stop duration, there was again a significant main effect for Voicing ($F[1,60]=4507.3$, $p < 0.001$) and a significant interaction between Voicing x Region (F[1,60]=199.4, p < 0.001). That means that both Saxon and Bavarian speakers have different vowel and consonant durations depending on the underlying voicing of the stops. But, although region significantly affects vowel duration depending on the underlying stop voicing, stop duration seems to be more strongly affected by underlying stop voicing and regional background.

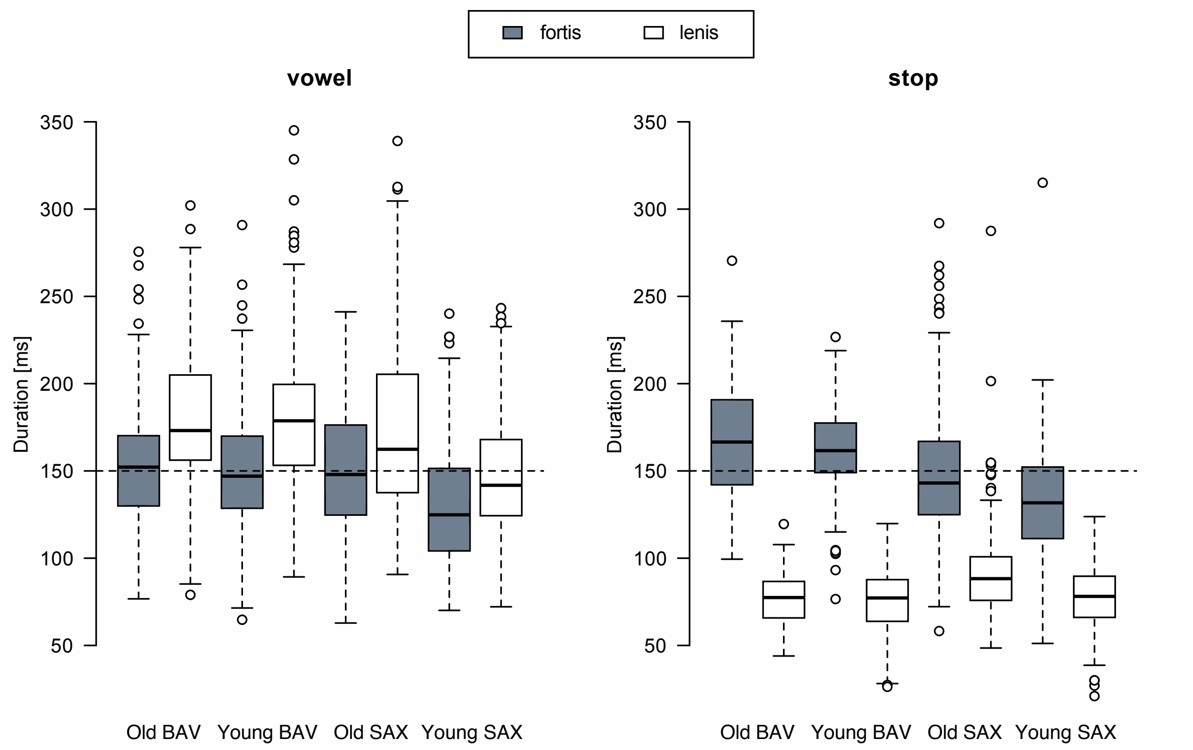


Figure 3.4: Absolute vowel (left) and stop (right) durations for fortis (grey) and lenis (white) stops across both minimal pairs separately for old and young, Bavarian (BAV) and Saxon (SAX) speakers. The dotted line at 150 ms signals visually the difference in vowel and stop duration in fortis and lenis sequences.

3.3.4 V:C duration ratio

Although there was a difference in the VC:Word duration ratio, we analysed the relative vowel duration as the main parameter that distinguishes lenis from fortis stops in Standard German. Ratios below 0.60 indicate the production of fortis stops and ratios at or above 0.60 signal lenis stops (cf. Table 3.3). A repeated measures MANOVA with the mean V:C duration ratio as the dependent variable revealed significant main effects for Voicing ($F[1,37]= 677.8$, $p < 0.001$) and POA ($F[1,37] = 76.5$, $p < 0.001$) as well as significant interaction effects for Voicing x Region ($F[1,37] = 25.1$, $p < 0.001$) and Voicing x POA ($F[1,37] = 8.3$, $p < 0.01$). The significant main effects for Voicing and POA shows that the V:C duration ratio differs with respect to the underlying voicing and the place of articulation of the stop; as is indicated by the mean ratios in Table 3.3 and by the data in Figures 3.5 and 3.6, fortis stops always had shorter V:C duration ratios than lenis stops and labial stops have longer closure durations than alveolar stops. That is, the voicing opposition was maintained or only incompletely neutralized by both dialect groups; this finding supports hypothesis H1.

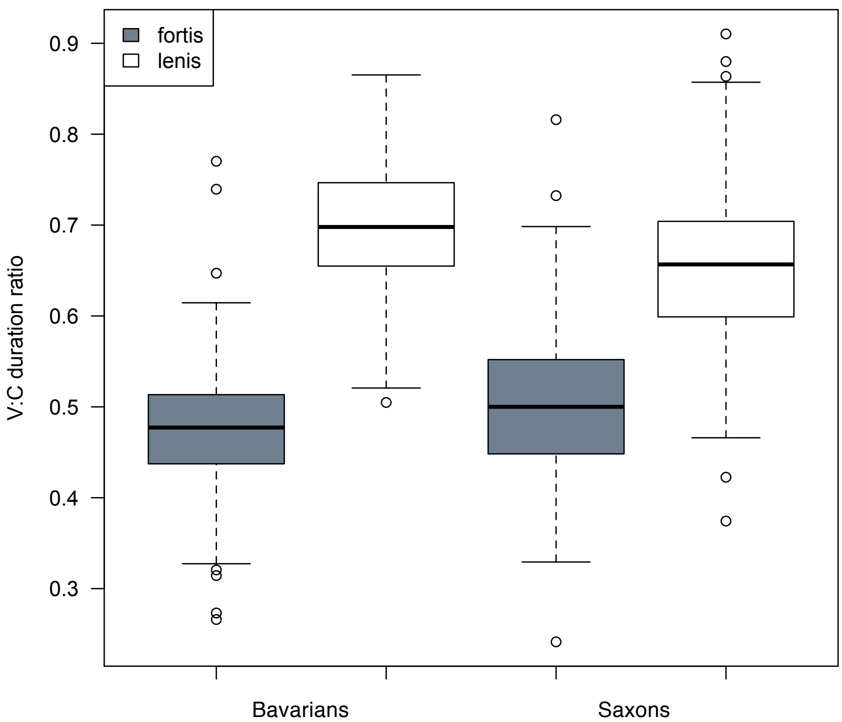


Figure 3.5: V:C duration ratios for fortis (grey) and lenis (white) stops across both minimal pairs and age groups separately for Bavarian and Saxon speakers.

However, as can be seen from Figure 3.5 Saxon speakers produced an ambiguous underlying lenis stop as the V:C duration ratio was on average below the ratio for Standard German lenis stops, i.e. there was a greater tendency for the fortition of lenis stops than for the lenition of fortis stops. The VC sequences differ in the intrinsic duration ratio depending on the stop's place of articulation. The significant interaction effect for Voicing x Region indicates that Saxon and Bavarian speakers differ with respect to the V:C duration ratio depending on the underlying voicing of the stop and the significant interaction between Voicing x POA shows that the V:C duration ratio is different for lenis and fortis stops depending on the place of articulation (cf. Figure 3.6). Post hoc tests confirmed that the V:C duration ratio was significantly different for lenis vs. fortis as well as labial vs. alveolar stops in all speaker groups. However, the two dialect groups differed only in the labial context in their V:C duration ratios but not in the alveolar context.

Table 3.3: Mean V:C duration ratios for words with lenis and fortis stops shown separately for young and old, Bavarian and Saxon speakers.

Word		**Bavarian**		**Saxon**	
		Young	**Old**	**Young**	**Old**
boten	mean	0.50	0.50	0.55	0.53
	sd	0.04	0.05	0.07	0.06
Oper	mean	0.45	0.45	0.45	0.48
	sd	0.07	0.06	0.09	0.07
Boden	mean	0.75	0.73	0.71	0.69
	sd	0.06	0.04	0.07	0.06
Ober	mean	0.66	0.67	0.60	0.62
	sd	0.05	0.05	0.06	0.07

To summarize, in both regional varieties and age groups the voicing contrast was maintained by means of the V:C duration ratio but also in terms of VC:Word duration ratio, as the overall duration for sequences containing fortis stops where significantly longer than sequences with lenis stops. Since there is a dialect dependent difference in the fine phonetic detail of lenis vs. fortis stops, one can ask whether there is a difference in the degree to which the contrast is maintained. The fact that there is a difference in the V:C duration ratio that depends on place of articulation and which can be attributed to the intrinsic duration differences as described in Kohler and Künzel (1978) does not contradict our measure of the neutralization degree – the mean V:C duration ratio difference between lenis and fortis stops – because lenis stops always have longer V:C duration ratios compared to fortis stops.

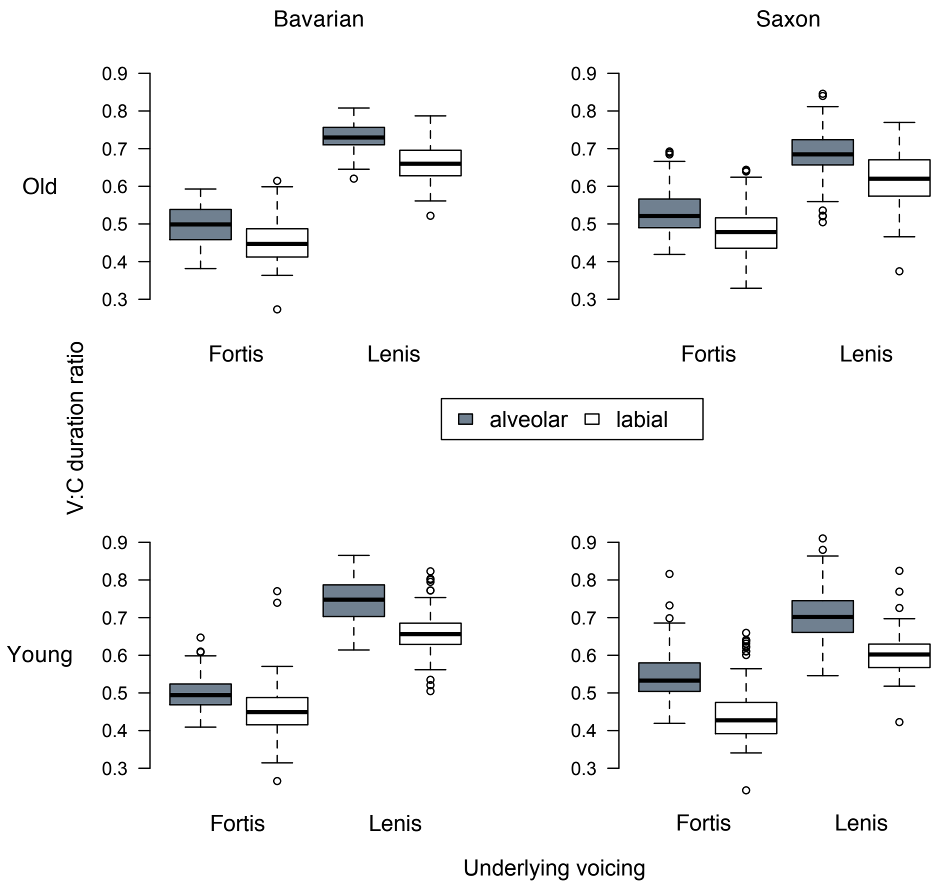


Figure 3.6: V:C duration ratios for fortis and lenis stops in minimal pairs *boten/Boden* (grey) and *Oper/Ober* (white) separately for old (above) and young (below), Bavarian (left) and Saxon (right) speakers.

3.3.5 *Schwa deletion*

The presence or absence of the schwa vowel may very likely influence the duration of the preceding stop and hence the duration of the word or the V:C duration ratio (cf. Davidson, 2006; Torreira & Ernestus, 2011). We therefore analyzed the test words ending in *–en* with respect to whether the schwa vowel in the second syllable was deleted or not (on the basis of the acoustic segmentation described in 1.6.2.3 and 3.2.2). In the latter case the postvocalic stop would be orally released, while in the former case the stop release would be nasal. The aim of this analysis was to eliminate any possible misinterpretations of duration differences between the minimal pairs and speaker groups that may be caused by schwa deletion.

Table 3.4: Number of tokens with and without schwa as well as the percentage of tokens with schwa deletion for each word with *en*-ending shown separately for young and old, Saxon and Bavarian speakers.

Word	Saxon						Bavarian					
	Young			Old			Young			Old		
	ə	Ø	%	ə	Ø	%	ə	Ø	%	ə	Ø	%
boten	65	15	19	115	12	9	37	83	69	79	1	1
Boden	54	26	33	99	27	21	43	76	64	77	2	3

Table 3.4 shows the distribution of schwa deletions in words with *en*-ending. Young Saxon speakers deleted schwa more often than old Saxon speakers (ca. 10%), but overall Saxon speakers realized the schwa vowel (ca. 70% to 90%). Furthermore, nasal releases occurred more frequently in the productions of Saxon speakers when the underlying stop was lenis (ca. 10%) On the other hand, there was a great difference between young and old Bavarian speakers regarding schwa deletion: while the latter produced all instances (with the exception of three cases) of *boten*- and *Boden*-Tokens with an oral release, the former produced the majority (64% and 69%) of these tokens with a nasal release.

Schwa deletion was highly speaker-dependent, i.e. with the exception of a few speakers in both dialect groups, speakers either deleted or realized schwa within all repetitions of a word. A table of the distribution of schwa realizations and deletions per speakers and word is given in Appendix B (Table B.1). Whether the schwa was deleted did not depend on the underlying voicing of the stop: in total 242 out of 811 analyzed tokens with *en*-ending were produced with a nasal release and this number was equally distributed among words with underlying lenis (131) and words with underlying fortis (111) stops.

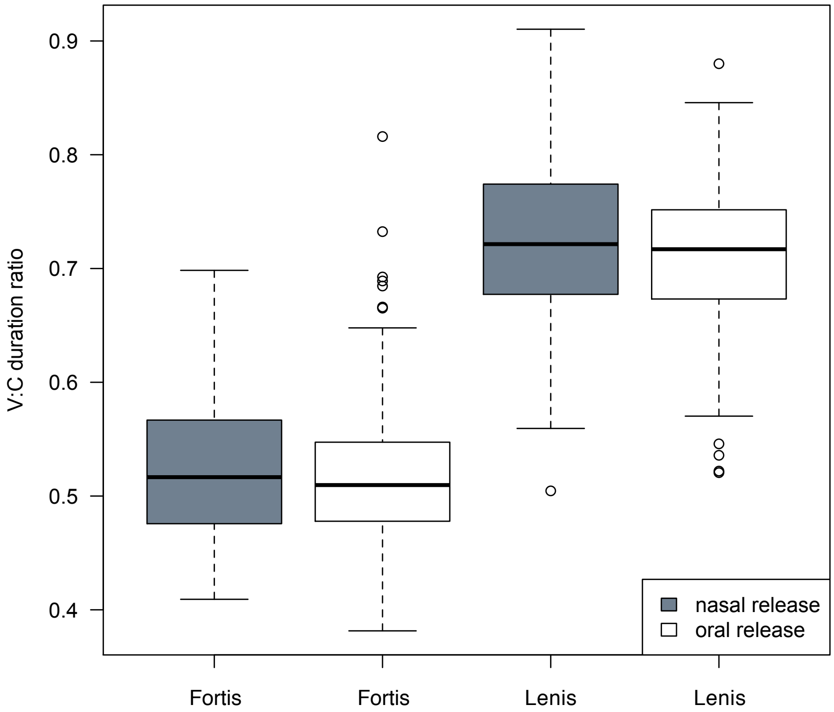


Figure 3.7: V:C duration ratios in words with schwa deletion (grey) and with schwa realization (white) across all four speaker groups.

As can be seen from Figure 3.7, the V:C duration ratio was slightly shorter in words with an oral release. However, two separate Wilcoxon rank sum tests with continuity correction (because the data was not normally distributed) revealed that the difference between tokens with schwa and tokens without schwa was not significant (lenis: $W = 16028$, $p = 0.092$; fortis: $W = 15226$, $p = 0.256$). That means that the V:C duration ratio was not affected by schwa deletion. Note, however, that schwa deletion was speaker-dependent and therefore we did not only compare nasally and orally released stops but also different speakers within each of the four speaker groups (i.e. this was no repeated measures analysis).

3.3.6 Neutralization degree

As can be seen in Figure 3.8, there is more voicing neutralization (though by far not complete) in the productions of labial vs. alveolar stops and in the production by Saxon as opposed to Bavarian speakers. A repeated measures MANOVA with the mean V:C duration ratio difference as the dependent variable revealed significant main effects for Region ($F[1,37] =$

25.1, $p < 0.001$) and POA ($F[1,37] = 8.3$, $p < 0.01$). The significant main effect for POA showed that the extent to which the contrast was neutralized also depends on the place of articulation. Except for young Saxon speakers, there was more neutralization when the stop's place of articulation was labial. This finding supports hypothesis H2. The difference in the neutralization degree was for both dialect and age groups the same as there were no significant interactions between Region x POA or Age x POA. This result contradicts hypothesis H3.

The significant effect for Region shows that Saxon and Bavarian speakers differ in their degree of neutralization; this result supports hypothesis H4. However, there is no difference in the neutralization degree with respect to Age, i.e. there is no evidence that young speakers neutralize the voicing contrast to a lesser extent than older speakers in neither of the two dialect groups. This finding does not support hypothesis H5.

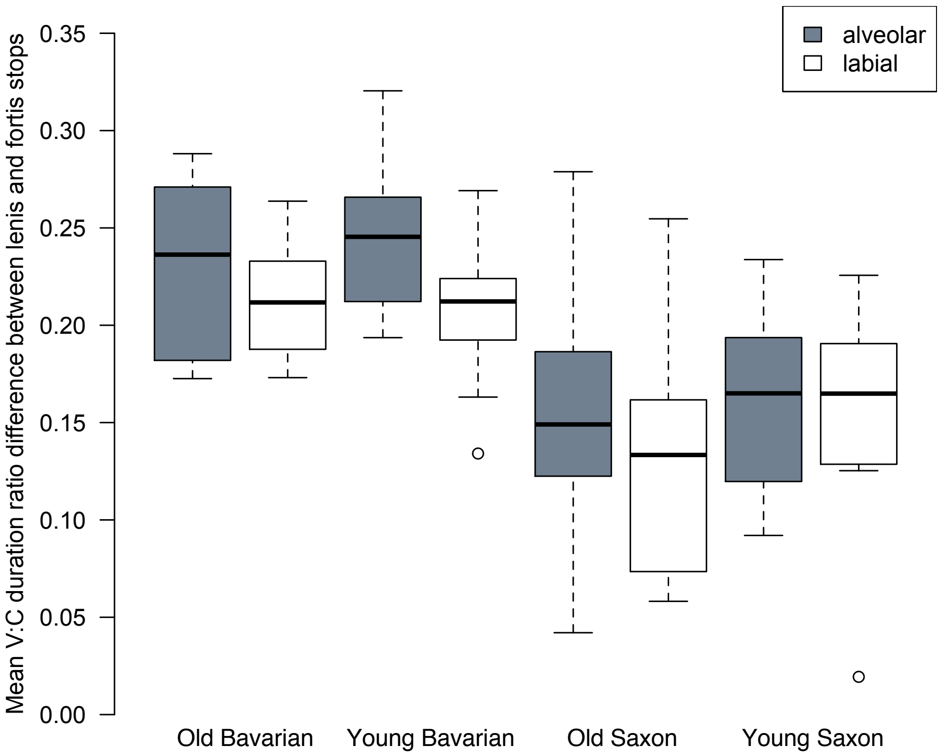


Figure 3.8: Mean V:C duration ratio differences between lenis and fortis stops for alveolar (grey) and labial (white) plosives separately for old and young, Bavarian and Saxon speakers.

3.4 General discussion

The aims of this study have been to determine whether the different statistical distribution for VC sequences in the lexicon affects the degree with which a contrast is maintained and whether Saxon speakers neutralize the voicing contrast in Standard German more than Bavarian speakers because this opposition does not exist in the phonological system of Upper Saxon. A further aim was to establish whether older speakers show more tendencies to preserve dialectal features in Standard German than young speakers of the same speech community.

Our first hypothesis H1 addressed the issue whether the voicing contrast in intervocalic position is maintained in the speech of Bavarian and Saxon speakers when they were asked to produce a variety of Standard German even though it is neutralized towards the lenis component in Upper Saxon as has been claimed in a number of dialect descriptions (Bergmann, 1990; Auer et al., 1993; Barden & Großkopf, 1998; Rues et al., 2007). Our results quite clearly showed that the voicing contrast is maintained by means of different word durations, VC:Word duration ratios and V:C duration ratios. Mean V:C duration ratios of 0.50 – 0.55 for alveolar and 0.45 – 0.48 for labial stops seem to be clear indicators for the production of fortis stops. On the other hand, V:C values of 0.69 – 0.75 for alveolar and 0.60 – 0.66 for labial stops were found in the production of lenis stops. Thus, the mean V:C duration ratios are similar to those given in Kohler (1979: 333), but they are overall smaller for Saxon and Bavarian speakers as compared to Standard speakers of a North German variety which probably comes about because Kohler (1977) and Kohler & Künzel (1978) measured the vowel to closure duration ratio (Kohler & Künzel, 1978: 121) and not the vowel to stop duration ratio. That is, for all age and dialect groups the relative vowel and stop durations were shorter and longer, respectively. Moreover, Saxon and Bavarian speakers did not only use different V:C duration ratios to signal distinct voicing categories but also varied the VC dyad. Words with fortis stops had significantly longer overall durations and significantly longer VC:Word duration ratios than words with lenis stops. This result is different from (1) the finding for Standard German speakers who showed no difference in the VC dyad duration for fortis vs. lenis stops (Kohler, 1977) and (2) findings for

English (where the duration cue is even more important than in German) which have shown that sequences with lenis stops are slightly longer than sequences with fortis stops, although on the whole the overall duration of vowel plus consonant is the same for both lenis and fortis sequences (Laeufer, 1992). Our results are also in line with a context-dependent variation in vowel length and intrinsic stop duration in the production of North German speakers who showed relatively longer vowel durations in vowels preceding alveolar stops compared to labial stops (Kohler, 1977; Kohler & Künzel 1978) and in the production of English stops (Stathopoulos & Weismer, 1983; Byrd, 1993). As was already discussed in 3.1, one reason for longer closure durations in labial than in alveolar stops could be that it requires more time to reach a high intraoral air pressure (which is necessary for the explosion) when the cavity behind the closure is larger (cf. Maddieson, 1997).

The finding that the contrast is preserved may not be very surprising since speakers produced a variety of Standard German and a substantial number of studies have shown incomplete neutralization of word-final voicing, a process which is probably much less marked than the dialectal feature of intervocalic, word-medial lenition. The finding that speakers (tend to) preserve or even emphasize a phonological contrast in an otherwise neutralizing position has often been claimed to be an artefact of laboratory speech (Fourakis & Iverson, 1984). That means, that speakers may have maintained the voicing contrast in this laboratory recording setting which they would then neutralize in other communicative situations. John (2004), for example, found neutralization of the voicing contrast in short stories read by Saxon speakers. However, the comparison of the data of two dialect groups, who are known to differ with respect to neutralization, allowed us to test whether the dialect groups showed different tendencies for voicing neutralization.

As far as the correlation of vowel length, place of articulation, and stop voicing in Standard German is concerned, H2 predicted more voicing neutralization in words with labial stops than with alveolar stops, accounting for the phonological frequency with which a stop following a tense vowel is lenis or fortis in the Standard German lexicon. More specifically, there is a bias for tense vowels to co-occur with lenis stops. Alveolar stops, on the other hand, can equally be lenis or fortis after lenis stops. Our results confirmed that the probabilistic co-occurrence of vowel tenseness and stop voicing in the German lexicon affects the amount with which a contrast is maintained even in German varieties where the primary dialect either does not have the contrast as in Upper Saxon, or shows no differences in the probabilistic co-occurrence of vowel plus stop sequences with respect to the stop's place of articulation as in Central Bavarian. The mean V:C duration ratio difference between lenis and fortis stops was greater for alveolar than labial stops in all speaker groups indicating that the contrast is diminished in labial as opposed to alveolar stops. Since VC sequences containing alveolar stops have a high number of lenis and fortis lexical competitors, it may be more important to make the underlying voicing clear in alveolar stops following tense vowels compared to labial stops where the probability of a lenis stop is highly predictable. Our results are compatible with findings showing that phonological frequency patterns in a language have an impact on fine phonetic detail in speech production and speech perception: speakers emphasize a contrast (e.g. Wright, 2003; Clopper & Pierrehumbert, 2008), that is otherwise at risk to be confused by the listener with a more frequent phonological competitor (cf. chapter II, Pitt & McQueen 1998; Hay et al., 2003).

However, our findings did not show longer V:C duration ratios for labial fortis stops, which would have signalled a more lenis-like production. That is, although the lenis/fortis contrast is diminished in pairs where one phonological category is more probable than another, the production of the less frequent category is not changed in the direction of the more frequent component, which would have suggested neutralization towards the more frequent component. One possible explanation for this result may be the connection between cavity size and intraoral air pressure (see above). Another explanation may be the frequency distribution of the words *Oper* and *Ober*. The latter word is dated[7] and thus less frequent than the former word.[8] Thus,

[7] Nowadays, the word *Kellner* is usally used to refer to a waiter.
[8] On the other hand, the morpheme *ober* is part of a number of other frequent German words such as e.g. *oberhalb* (‚above') and therefore it cannot be regarded as rare.

speakers may have indeed produced the more frequent pattern. A third explanation for this finding could be that the stops not only occur in intervocalic, but also in syllable-initial position. In chapter II, the possibility of re-syllabification was shown to enhance the discrimination performance of listeners. Speakers are used to producing labial fortis stops in syllable-initial position. Thus speakers may have interpreted the stop as syllable-initial and produced a fortis stop. On the other hand, for prevocalic or syllable initial labial stops we have an equal distribution of fortis and lenis stops; thus, if re-syllabification is the only driving force behind contrast maintenance, then we should not have found a difference in the neutralization degree between sequences with labial vs. alveolar stops (where the re-syllabification assumption also applies). As argued in chapter II, we presume that both the probabilistic distribution of the vowel tensity and stop voicing together with the potential for re-syllabification contributed to the extent with which a contrast is maintained or neutralized.

Before we turn to the discussion of age dependent differences concerning the place of articulation, let us consider the regional dependent differences. Hypothesis H4 predicted that there is a greater tendency for Saxon speakers to neutralize (though incompletely or only by means of a diminished difference with which the contrast is maintained) the voicing contrast than for Bavarian speakers, since the voicing contrast is lost in intervocalic position in Upper Saxon but not in Central Bavarian. Our results were consistent with this hypothesis. More specifically, the mean difference between the V:C duration ratio in sequences with lenis vs. fortis stops was smaller for Saxon speakers than Bavarian speakers. That means that Saxon and Bavarian speakers differ in fine phonetic detail when they produce a variety of Standard German and these differences are conditioned by the phonological systems of the dialects. This finding supports the idea that regional varieties of Standard German may only differ in fine phonetic detail and that those differences are gradient rather than categorical.

According to hypothesis H5, there should be more voicing neutralization in the production of old Saxon speakers than young Saxon speakers, because of dialect levelling towards a Standard German pronunciation in younger generations. Our results did not support this hypothesis. There were no significant age differences in either dialect group. There was nevertheless a tendency towards less neutralization in young vs. old Saxon speakers (Figure 3.8).

H3 predicted that the difference between the incomplete maintenance of the voicing contrast in labial vs. alveolar stops is greater for young than for old speakers, irrespective of the dialects (cf. Figures 3.1, 3.8). Our results did not confirm this, as there was no significant interaction of POA x Age. That is, the difference between the neutralization degrees of labial vs. alveolar stops was the same for all age groups. This result suggests that the bias within the vowel plus stop co-occurrence in the Standard German lexicon has the same effect on young vs. old speakers. There was no difference between each of the two age groups concerning the degree of dialect levelling (in terms of a diminished manifestation of a dialect feature), although the frequency bias does not exist in either of the two dialects. That means that the bias in the Standard German lexicon affects the production of all speaker groups. However, in the young Saxon group there is no or only a very small difference in the neutralization degree in the opposite direction, which indicates that (1) there is less neutralization for young vs. old Saxon speakers which would support hypothesis H5, but (2) the difference between the mean V:C difference for labial vs. alveolar stops is smaller for young vs. old Saxon speakers which contradicts hypothesis H3.

Descriptions of German dialects are very often based on fine auditory analyses by one investigator (and are therefore more prone to be subjective), which are then summarized in dialect features in which a dialect diverges from a reference language, e.g. Middle High German or Modern Standard German. That is, many dialect descriptions are carried out within a generative model. Within a generative framework features are either existent or not, phonological rules are either applied or not, and dialects and Standard German differ categorically in a feature. When we consider that all dialects and accents change gradually over time, which can be observed among others either in form of differences between different age groups (Harrington et al., 2008; Müller, 2010; Rowley, 1990b; Bergmann, 1990) or between dialect groups, than it is hard to model dialect levelling in a generative framework. The production results in this experiment demonstrate that differences – in this experiment fine differences in the V:C duration ratio – between dialects and age groups are gradient. One finding was that the voicing contrast is not neutralized in Saxon as the V:C duration ratio differed significantly for lenis and fortis stops. But it cannot be concluded that the contrast maintenance is the same for both Saxon and Bavarian speakers, as the differences in the neutralization degree between the two dialect groups show. In order to account for fine differences in dialectal features between age groups, it is important to include these gradient differences in models of present-day regional varieties of German and dialect levelling.

A probabilistic model such as the exemplar theory (Pierrehumbert, 2001, 2002, 2003) accommodates even small and gradient dialectal and age dependent differences. Although exemplar models were originally developed to model perception, they can also be used to model production data, given that every speaker is at the same time also a listener. Gradual (historical) sound changes such as lenition or contrast neutralization are easily predicted by the probability with which a token occurs in the lexicon (cf. Pierrehumbert, 2001; see also 1.4 above). It is thus important to account for frequency effects in a theory of sound change (or dialect levelling). In a usage-based model such as the exemplar theory, phonological categories are defined by the density distribution of accumulated exemplars of a category in an acoustic-perceptual space. New exemplars are collated with exemplars that are already memorized and then stored in the vicinity of the most similar exemplars. Thus, the distribution is constantly enlarged by means of new remembered tokens.

The findings for different degrees of contrast maintenance in alveolar vs. velar stops as well as Saxon vs. Bavarian speakers can both be explained by an exemplar model. With regard to the statistical distribution of vowel plus stop combinations, the density distribution of lexically frequent combinations should occupy much more perceptual-acoustic space than infrequent combinations. Thus, as far as the probabilistic co-occurrence of tense vowels plus labial stops vs. tense vowels plus alveolar stops is concerned, listeners should have an equally dense distribution of remembered lenis and fortis alveolar stops but much more stored exemplars of lenis vs. fortis labial stops given that combinations of tense vowels and fortis labials stops are rare in the Standard German lexicon. It follows from this distribution that speakers tend to neutralize the contrast to a greater extent in a labial than an alveolar context. Likewise, the density distribution of sequences with lenis stops should be enlarged across the perceptual-acoustic space for Saxon speakers commensurate with the neutralization of the voicing contrast in favour of lenis stops in broad Saxon. A consequence of this bias may be that the Saxon speaker/listener produces more lenis-like versions of tokens with underlying fortis stops. On the other hand, the density distribution should be equally large for lenis and fortis stops in Bavarian speakers because vowel duration is varied in order to maintain the voicing contrast in the broad Central Bavarian dialect (irrespective of the stop's place of articulation) and indeed Bavarian speakers show no such bias towards lenis stops, but maintain the contrast to a greater extent than Saxons. Thus, the phonological frequency of stop voicing in a speaker's native dialect seems to be stored in the mental lexicon of a speaker. However, our findings also suggested that the asymmetry of vowel tenseness and stop voicing in the Standard German lexicon is reflected in the production of Saxon and Bavarian speakers although there are no such asymmetries in broad Upper Saxon or Central Bavarian. This finding can be taken as evidence that phonological frequencies from several varieties contribute to the density distribution of the exemplar cloud in the acoustic-perceptual space. The predictions from exemplar theory are compatible with most of our findings except for the non-lenis-like production of labial stops (unless we regard word frequency as more influential than the frequency with which sound co-occur, see the discussion above).

Lindblom's (1990) Hypo- & Hyperarticulation (H&H) theory is another model that accounts for the different extent to which the contrast is maintained in alveolar vs. labial stops. According to the H&H theory, a speaker produces speech along a continuum from hypo- to hyperarticulation. Which style is chosen depends on the context that always affects the communicative situation. Speech production is hypoarticulated at points in the speech signal that are redundant for the listener as far as meaning is concerned. That is, in a noisy environment a speaker produces speech from the hyperarticulation end of the continuum while in a quiet environment s/he tends to a more hypoarticulated pronunciation. In order to communicate efficiently, it is important to prevent redundant effort. In this theory, the listener likewise accommodates the speaker with using all the meta-linguistic knowledge that is available. Thus, for example, words that are highly predictable, e.g. high frequent words with a low neighbourhood density may be hypoarticulated because they are easier to understand (Wright, 2003; Clopper & Pierrehumbert, 2008). For that reason, the voicing contrast in post-vocalic sequences containing an alveolar stop should be hyperarticulated because the underlying voicing is not predictable from the lexical probability. On the other hand, lenis labial stops are hypoarticulated because of the lexical bias for lenis labial stops following tense vowels. Thus our findings are also explained by Lindblom's H&H theory. Moreover, this theory accounts also for the non-lenis-like production of underlying /p/ in our data: on the one hand, fortis labial stops are likely to be hyperarticulated after tense vowels because this combination is rare, but at the same time the contrast between /p/ and /b/ is diminished because /b/ is hypoarticulated. That is, in order to reduce redundancy a contrast is only produced to such an extent as it is necessary for successful communication.

To conclude, we have presented evidence that speakers with different dialectal backgrounds including a dialect with no intervocalic voicing contrast are able to produce the voicing contrast when asked to produce a variety of Standard German. The extent to which a voicing contrast is maintained is not only dialect-dependent, but also conditioned by the probabilistic co-occurrence of vowel tensity and the stop's place of articulation. Once a combination is asymmetrically distributed in the lexicon, i.e. there is a lexical bias towards one voicing category, the extent to which the feature distinguished words in the lexicon is by definition reduced. Incomplete neutralization is thus not only a phenomenon that is restricted to contrasts in neutralizing positions but also in non-neutralizing positions, i.e. in non-neutralizing positions contrasts may be incompletely maintained depending on the statistical distribution in the lexicon and the dialectal background. Thus, phonological categories such as voiced or lenis and voiceless or fortis mark the endpoints of a phonetic continuum that covers both incomplete maintenance and incomplete neutralization of the voicing contrast.

Vowel length in regional varieties of German

Chapter IV

Abstract

The main purpose of this chapter was to investigate the extent to which the vowel length contrast is maintained in the speech of Saxon and Bavarian speakers when they produced speech of a Standard German variety. Another aim was to investigate whether there were age-dependent differences in the extent to which these variety-dependent features were manifested in their production of Standard German. To do this, we investigated the vowel length contrast in minimal pairs such as *bitten* vs. *bieten*. The prediction was that vowel quantity together with the type of post-vocalic stop voicing would be different in those two varieties. Commensurate with the way that stops are produced in their broad dialects, we expected Saxon speakers to maintain the length contrast but produce a more lenis-like obstruent: that is, we expected them to produce [bɪdn] and [biːdn], respectively. As far as the Bavarian speakers were concerned, there were two possible outcomes. The first was that they would completely neutralize the quantity contrast before the voiceless stop in *bitten* vs. *bieten*. The second was that they would maintain it indirectly by producing a lenis stop in the latter thereby causing a phonetic lengthening of the preceding vowel: that is under this second scenario the contrast would be [bɪtən] vs. [biːdən] in which the lengthened [i:] is a phonetic consequence of [d]. Our results support the second hypothesized outcome: Bavarian speakers maintained the vowel length contrast to a greater extent than did Saxon speakers and the mean difference between the vowel:stop ratio in words with underlying tense and lax vowels was significantly greater for old than young Bavarian speakers. Both Saxon speaker groups and young Bavarian speakers showed similar vowel:consonant duration ratio differences between tense and lax vowels. Our results suggest that there is a language change in progress especially in the Bavarian speaker group with young speakers having less pronounced dialect features.

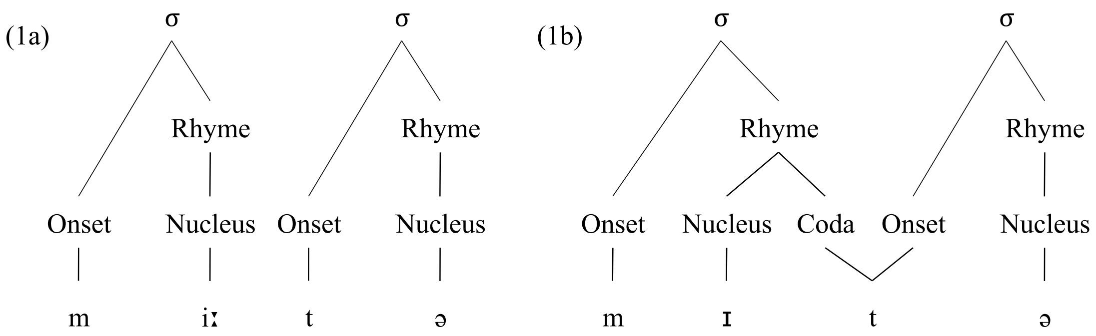
4.1 Introduction

In the previous chapter it was shown that Bavarian speakers indeed maintain the voicing contrast more than Saxon speakers. This finding was explained by the speakers' regional background of Central Bavarian. In this dialect, the Central Bavarian Lenition rule applies which states that vowels preceding lenis stops are long or tense while fortis stops are preceded by lax or short vowels. Such a correlation between vowel length and stop voicing is also observable in many vowel plus stop combinations in Standard German, but in Central Bavarian it has become phonologized. It thus influences the phonotactic probabilities of a language and has affected the structure of the Standard German lexicon to a great deal. The motivation for this chapter is thus to investigate whether the vowel length contrast is preserved by Bavarian speakers when they produce a variety of Standard German which at least before alveolar fortis stops has an equally distributed number of tense or long and lax or short vowels. But before we turn to the tense/lax distinction in German varieties in detail we will first review the literature on the tense/lax opposition in Standard German.

Standard German has a phonological vowel length contrast which is traditionally described under the term tense/lax distinction (Kohler, 1995). The terms *vowel length contrast* or *tense/lax opposition* are (similar to the fortis/lenis opposition in chapter III) much discussed concepts both in the phonological and phonetic literature especially since it remains an yet unresolved issue which feature(s) are primarily distinctive and thus most adequate to describe the contrast (for an overview see Mooshammer, 1998): vowel quality and/or vowel quantity. Both features are phonologically important to the contrast and are also phonetically established. We will briefly summarize (1) non-linear or auto-segmental phonological theories and the syllable cut theory and (2) present evidence from phonetic studies. For terminological simplification we use the tense/lax distinction when we refer to underlying, i.e. phonemic vowel length without addressing differences in vowel quality. The terms short and long are used for fine phonetic segment duration differences within one phonological category.

Vowel length is one of four contrastive parameters (besides tongue height, tongue position and lip rounding) in which vowels are distinguished in German. For almost each vowel category there is a vowel pair that is only distinguished by means of phonemic vowel length: there are eight tense vowels /iː yː eː øː ɛː aː oː uː/ and seven lax vowels /ɪ ʏ ɛ œ a ɔ ʊ/. Thus we have minimal pairs like *mieten* (/miːtən/, 'to rent') and *mitten* (/mɪtən/, 'in the middle of'), *Ofen* (/oːfən/, 'stove') and *offen* (/ɔfən/, 'open'), or *Staat* (/ʃtaːt/, 'state') and *Stadt* (/ʃtat/, 'town'). The front mid vowels behave slightly different since they are part of a three way distinction with /eː ɛː/ being tense and /ɛ/ being lax. From an articulatory-acoustic perspective, /eː/ is the most divergent sound in this triplet, but it is much more frequent in Standard German than open-mid /ɛː/ which is considered obsolete, its pronunciation only representing orthography mostly in subjunctives. Usually /eː ɛː/ are considered to merge into /eː/, resulting in a tense/lax pair /eː ɛ/.

The auto-segmental or non-linear phonological theory (for an overview see Goldsmith, 1990) operates (similar to the theory of generative phonology) with underlying forms that are transformed to various surface representations, which depend on different ordered rules that are applied during the transformation process. In non-linear theories, vowel quantity (as opposed to vowel quality) is often considered the primary feature that distinguishes tense and lax vowels (e.g. Ramers, 1988; Wiese, 1988, 1996; Hall, 1992; Yu, 1992; but see Giegerich, 1985 for a different account in which two primary features are considered – quality and tensity) in Standard German since it facilitates the syllable rules in Standard German and thus tense and lax vowels differ with respect to the syllable position in which they can occur. Tense vowels are freely distributed, i.e. they can occur in any syllable position, e.g. in syllables with neither onset nor coda as in *Ah* (/aː/, amazed expression); in open syllables as in *sah* (/zaː/, 'sb. saw'), in which the coda is empty, in no-onset syllables as in *Aas* (/aːs/, 'carrion'), or in syllables containing all three constituents as in *saß* (/zaːs/, 'sb sat'). On the other hand, lax vowels are limited to the nucleus of closed syllables, i.e. syllables with a coda, as for example in *mit* (/mɪt/, 'with'). Following the Maximum Onset Principle, an intervocalic consonant is aligned with the onset of the second syllable in disyllabic words. However, in trochaic words with a lax vowel as the first syllables' nucleus, the stop is – according to many theories (e.g. Wiese, 1988, 1996) – ambisyllabic[1], i.e. it is both the coda consonant of the first as well as the onset consonant of the second stop. Only in this way the syllable structure remains intact. Thus, minimal pairs such as *Miete* and *Mitte* also differ in syllable structure which is represented in (1a) and (1b), respectively:



In many of these non-linear phonological theories syllable structure is considered to be derived through a process called syllabification (e.g. Hall 1992, 2000), but not to be part of the underlying form. Inherent to all these models is that vowel length is phonemic which in turn determines the syllable structure.

One early model is the so-called *Silbenschnitt*-theory (henceforth syllable cut theory), which goes back to Sievers (1872), who distinguished so-called *Schallsilben* (as in /biːtən/) from so-called *Drucksilben* (as in /bɪtən/) based on this difference in syllable structure: in /biːtən/ the energy maximum within one syllable is reached before the intervocalic consonant, in /bɪtən/, however, the syllable's energy maximum is reached after the vowel at the stop. Trubetzkoy (1938, 1939/1958) differentiates between syllables with close and with loose contact (Fischer-Jørgensen & Jørgensen, 1969), the latter being abruptly and the former smoothly cut. The vowel pairs that have been differentiated by phonological length above can likewise be grouped according to the syllable cut: tense vowels are loosely and lax vowels are

[1] According to Giegerich (1985) and Yu (1992) ambisyllabicity depends on lexical stress. Hall (1992), on the other hand, considers ambisyllabicity to be a phonetic process.

abruptly cut. In this theory syllable cut is believed to be the primary feature from which the two secondary features vowel quantity and vowel quality are derived. Syllable cut, however, was initially an auditorily defined feature, which was difficult to measure. Vowel quality and quantity on the other hand were quantifiable by means of spectral differences and duration. For that reason the syllable cut theory was not considered for almost half a century. Only in the past two decades this phonological approach has been again in the focus of German researcher (Vennemann, 1991; Restle, 1998, 2002; Spiekermann, 2000, 2002; Auer, Gilles & Spiekermann 2002). Based on evidence from articulatory studies (see below), Restle (1998) developed a phonological theory of syllable oscillation in which syllables are not composed of phonemes but of opening and closing phases.

Spiekermann (2000, 2002) found differences in the course of the intensity over the accented vowel between trochaic words with loosely and abruptly cut syllables. Loosely cut syllables showed a (longer) high-energy plateau than abruptly cut syllables. He concludes that the vocalic intensity curve is an acoustic correlate of syllable cut. Spectral balance[2] was also found to correlate with syllable cut (Jessen, 2002). More phonetic evidence for this theory in form of an articulatory correlate was provided by Hoole, Mooshammer, and Tillmann (1994), Kroos, Hoole, Kühnert, and Tillmann (1997) as well as Hoole and Mooshammer (2002). They used electromagnetic midsagittal articulography to analyze (among others) tongue velocity in CVC syllables. CVC sequences were defined in terms of movement cycles which were subdivided into opening (CV), nucleus, and closing (VC) phases. The on- and offsets of the CV and VC gestures were defined by a 20% threshold of the overall peak velocity and the nucleus was always the interval between the CV offset and the VC onset. Lax and tense vowels differ phonetically on the articulatory level by means of a "tighter coupling between CV and VC phases for lax vowels" (Hoole & Mooshammer, 2002: 149) resulting in shorter nucleus durations in these vowels. In abruptly cut syllables the opening gesture is "truncated by the following closing gesture [and] vowel length and tenseness […] can be seen as an outcome of

[2] Spectral balance was initially offered as an alternative correlate, that captures better the linguistic use of amplitude (Sluijter, 1995; Sluijter & van Heuven, 1996; see Jessen, 2002: 154).

this truncation” (Restle & Mooshammer, 1999: 531). However, no articulatory evidence was found for a correlation between vowel length and the coupling of phases of the postvocalic consonants (Restle & Mooshammer, 1999). According to these dynamic approaches (as opposed to the static approach of non-linear theories), it is important to view vowel plus consonant sequences as an entity and tenseness has to be considered a feature of the syllable and not of the vowel alone (Mooshammer, 1998: 30). The syllable cut theory coincides in this respect with articulatory phonology (Browman & Goldstein, 1986, 1988, 1989, 1992), in which articulatory gestures are the units for phonological contrast. Duration is part of these gestures, though they are abstract units. Acoustically and perceptually, the vocalic tense/lax contrast is manifested in particular by different vowel durations and spectral patterns. Acoustic analyses have shown that lax vowels – at least in lexically stressed position (cf. Jessen, 1993) – have proportionally shorter vowel durations than tense vowels (e.g. Meyer, 1904; Zwirner, 1962; Ramers, 1988; Jessen 1993); this tense to lax vowel quotient, however, is different in the various studies (see Mooshammer, 1998), depending on speech style (read vs. spontaneous speech, sentences vs. isolated words), dialect (northern speakers showing longer quotients than southern speakers (Zwirner, 1962), and age with younger speakers showing smaller quotients than older speakers, i.e. the tense/lax contrast was preserved to a lesser extent by young versus old speakers (Bethge, 1963). Furthermore, lax vowels are more centralized in the vowel space than their tense counterparts, i.e. the formant frequencies are more extreme in the latter case (cf. Figure 4.1; Meyer, 1911, 1913a, b; Jørgensen, 1969; Ramers, 1988; Jessen, Marasek, Schneider, & Claßen, 1995). Perception experiments confirmed that both vowel quantity and quality were also the dominant cues in the perception of the tense/lax contrast (Bennet, 1968; Heike, 1970; Ramers, 1988; Strange & Bohn, 1998). Other potential acoustic cues such as for example intrinsic fundamental frequency (f0)[3] were not found to be robust parameters.

[3] Since it is well-documented for a number of languages that intrinsic fundamental frequency (f0) increases with vowel height, i.e. more closed vowels have a higher f0 than more open vowels (Peterson & Barney, 1952; Steele, 1986; Whalen & Levitt, 1995; Whalen, Gick, Kumada & Honda, 1999) one would also expect that intrinsic f0 is lower for lax than for tense vowels. However, intrinsic f0 is an acoustic parameter that is difficult to interpret when it comes to the tense/lax distinction in Standard German. As pointed out by Fischer-Jørgensen (1990) the intrinsic f0 is surprisingly high and even, though only

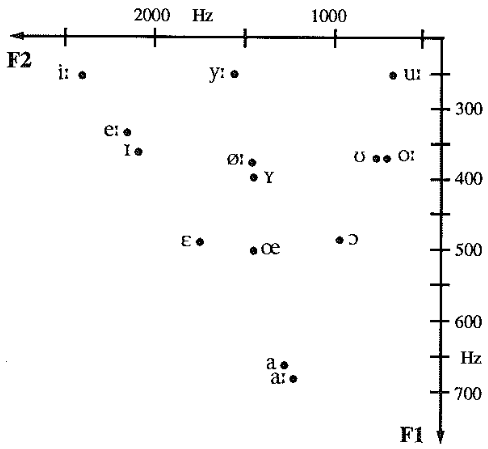


Figure 4.1: The German monophthongs (except for /ɛː/) in a F1 x F2 formant plane (adapted from Pompino-Marschall, 1995).

German dialects differ from Standard German with respect to the realization of tense and lax vowels and some varieties even in the contrast maintenance and phonological status of the tense/lax distinction (e.g. Central Bavarian, Wiesinger, 1990). Spiekermann (2000) has investigated whether German dialects (spoken in the 1950s and 1960s) can be classified by means of the phonetic syllable cut correlate intensity. His results suggest that Northern and Central German but not Southern dialects use syllable cut in terms of intensity. Southern varieties use to a greater extent durational parameters to signal a difference in phonemic vowel length.

According to the literature Upper Saxon does not differ from Standard German with respect to vowel quantity (as opposed to vowel quality which is overall more centralized (Auer et al., 1993; Hirschfeld, 1999; Iivonen, 1996), i.e. both tense and lax vowels are more centralized). Voiceless stops tend to be neutralized towards the voiced category (cf. results in chapter III; Bergmann, 1990; Auer et al., 1993), but irrespective of the stops' underlying voicing, the phonological vowel length contrast is maintained. Spiekermann (2000) found that in Upper Saxon (as spoken in Wurzen) both syllable cut (in terms of the intensity contour) as

slightly, higher for lax as opposed to tense vowels. Hoole & Mooshammer (2002) corroborated this finding for speakers from other parts of Germany.

well as vowel quality are used to differentiate tense and lax vowels. Thus, Upper Saxon may be classified as a variety with a syllable cut contrast.

Central Bavarian, on the other hand, is distinct from Standard German as there is a correlation between vowel quantity and the underlying voicing of the subsequent obstruent: voiced obstruents are preceded by tense or long vowels while voiceless consonants are always preceded by lax or short vowels. Thus, there are two possible realizations of Standard German /beːtən/: /eː/ is either realized as a short /e/ or /ɛ/ in combination with a fortis stop as in /beten/ or it is produced with a long vowel but then it is followed by a lenis stop as in /beːden/ (cf. König & Renn, 2005: 67). According to this so-called Central Bavarian lenition rule[4] (henceforth CBL), the phonological status either of vowel quantity or consonant voicing is at issue in Central Bavarian: A consonantal approach, which considers the stop voicing contrast to be phonological and vowel quantity allophonic is favoured by many dialectologists (e.g. Kufner, 1964; Hinderling, 1980; Scheutz, 1983; Rowley, 1990b; Wiesinger, 1990), i.e. according to this account the pre-consonantal vowel quantity depends on the voicing status of the stop. The variation regarding the vowel quality (/betən/ vs. /bɛtən/) but not quantity may be taken as an argument for this consonantal approach. Bannert (1976), on the other side, rejects the phonemic fortis/lenis approach and proposes complementary length within VC sequences as the contrastive feature. Duration measurements and perception tests, in which the vowel and the obstruent were progressively shortened and lengthened, respectively, showed that the proportional shift in vowel and stop duration is sufficient to perceive either a $V_{tense}C_{lenis}$ or $V_{lax}C_{fortis}$ sequence. That is, in this suprasegmental account a prosodic quantity feature causes either the vowel or the stop to be long resulting in the shortening of the adjacent sound. Voicing is cued by vowel duration relative to word duration and this results in the reduction of either one of the two contiguous segments. Spiekermann's (2000) analysis revealed that in the Central Bavarian dialect tense and lax vowels were mainly differentiated in terms of the following stop but neither by means of vowel quality nor quantity, i.e. syllable cut features. This means that it

[4] The CBL rule applies also word-finally given that the preceding vowel is a tense vowel (e.g. Standard German /diːp/, 'thief', Central Bavarian /diɐb/, Wiesinger, 1990).

is not the vowel but the following stop that is differentiated and that is irrespective of the underlying voicing (since loosely and abruptly cut syllables contained both voiced and voiceless segments). Thus, Bavarian may not be classified as a syllable cut variety. However, consonant lengthening leads to proportional shorter vowel duration, which facilitates the tense/lax contrast (or the syllable cut contrast for that matter). This finding seems to support Bannert's (1976) suprasegmental account of vowel length and obstruent voicing in Bavarian.[5] Uncontroversial in all accounts, however, is that vowel length and obstruent voicing always correlate in Bavarian.

A similar correlation between vowel length and stop voicing holds true in Standard German, but with one exception: The phonological vowel length contrast in Standard German is mainly preserved in vowels preceding alveolar fortis stops. Stops at the other places of articulation favour either lax or tense vowels depending on their underlying voicing. This frequency distribution is shown in Table 4.1 (cf. also 2.1 and 3.1.3).

Table 4.1: Frequency distribution for vowel plus stop sequences.

Vowel	Stop					
	Fortis			Lenis		
	labial	**alveolar**	**velar**	**labial**	**alveolar**	**velar**
Lax	frequent	frequent	frequent	infrequent	infrequent	infrequent
Tense	infrequent	frequent	infrequent	frequent	frequent	frequent

To summarize, there is a phonemic vowel length contrast in Standard German and in Upper Saxon, but the postvocalic stop is neutralized towards lenis in the latter and thus the vowel to stop duration ratio is longer before underlying fortis stops than in Standard German. In Central Bavarian, by contrast, vowel length is complementarily distributed since vowel length varies depending on the following stop voicing. Thus, there are two possible Bavarian realizations of the Standard German vowel length contrast before fortis stops: (1) the vowel

[5] Bannert (1976), on the one hand, rejects syllable cut as a distinctive feature arguing that syllable cut is the result of the temporal structure of the vowel plus consonant sequence and Spiekermann (2000), on the other hand, does not discuss his findings with respect to Bannert's suprasegmental approach.

length contrast is either neutralized towards short or (2) maintained in terms of a greater vowel to stop duration ratio difference which either signals a lenis or a fortis stop.

We are foremost interested in the durational parameter vowel quantity (both in terms of the vowel to word as well as the vowel to consonant ratio) and its contribution to the tense/lax contrast in regional varieties of Standard German. Duration has been shown to be a very important and seemingly robust cue to the phonological tense/lax contrast. At the same time, however, the intrinsic phonetic vowel duration is affected by the phonetic environment such as for example the manner of articulation (obstruent vs. sonorant) or the underlying voicing of the following consonant or polysyllabic shortening. Vowels are relatively shorter in polysyllabic vs. monosyllabic words (Lindblom, 1968 for Swedish, Nooteboom, 1972 for Dutch; Lehiste, 1972 and Port, 1981 for English) and before fortis as opposed to lenis stops (Chen, 1970; Kohler, 1977; Ladefoged & Maddieson, 1996). While research on the fortis/lenis distinction has extensively focussed on the effect of duration of the preceding vowel (e.g. Chen, 1970; Mitleb, 1984; Kluender, Diehl, & Wright, 1988; Davis & Van Summers, 1989; Fowler, 1992), there are considerably fewer studies that have investigated the effect of voicing cues on the tense/lax distinction (e.g. Parker & Walsh, 1981; Braunschweiler, 1997) even though a vowel plus stop sequence is phonetically co-dependent with respect to the phonological interpretation of both segments. Braunschweiler (1997), for example, found “evidence for an anticipatory vowel-lengthening effect before voiced stops” (p. 353) and that closure duration varies depending only on the underlying voicing of the stop but not on the phonemic vowel length of the preceding vowel.

The aim of this chapter is to investigate whether Saxon and Bavarian speakers differ with respect to the degree with which the Standard German tense/lax opposition is maintained before underlying fortis stops when they produce a variety of Standard German because of the divergent manifestation of vowel length in the primary dialects. A second aim is to test whether there are age dependent differences in the production of the tense/lax contrast based on the assumption that dialectal features are found to a lesser extent in the productions of young than old speakers of both varieties. Furthermore, Bethge (1963) has provided evidence for a smaller tense:lax quotient in young vs. old speakers (regardless of the underlying voicing of the following stop). Given that speakers in the corpus were asked to produce a variety of Standard German and that according to the CBL two outcomes are possible we formulate the first hypothesis as follows:

(H1) Bavarian speakers maintain the tense/lax distinction before obstruents to a greater extent than Saxon speakers.

In chapter II it was shown that Saxon speakers neutralize the voicing contrast after tense vowels to a greater extent than Bavarian speakers. Accordingly, we want to test whether there is more lenition of fortis stops in Saxon vs. Bavarian speakers.

(H2) Saxon and Bavarian speakers differ markedly in the realisation of the post-vocalic stop after lax vowels: Bavarian speakers realise the alveolar stop as fortis and Saxon speakers produce a more lenis-like stop.

(H3) The difference between the underlying fortis stops following lax and tense vowels is much greater for Bavarian than for Saxon speakers: Bavarians show a vowel to stop duration ratio that indicates a /t, d/ difference while Saxon speakers show a vowel to stop duration ratio that indicates a more lenis-like stop.

(H4) The Bavarian-Saxon difference in H1 – H3 is more pronounced for old than for young speakers (on the assumption that there is a greater tendency that young speakers of both varieties tend towards a more standard pronunciation).

In order to test whether the hypothesized greater contrast maintenance between tense and lax vowels preceding obstruents in Bavarian is a consequence of a Bavarian co-dependency of vowel length and stop voicing, we formulate the following hypothesis:

(H5) Before sonorant consonants, Bavarian speakers maintain the tense/lax contrast to a lesser extent than Saxon speakers because vowels are always long before voiced segments.

4.2 Method

4.2.1 Speech materials and speakers

From the corpus described in chapter I all five minimal pairs were chosen that differ in Standard German with respect to the phonological vowel quantity of the lexically accented syllables' nucleus: *beten* (/beːtən/, 'to pray'), *betten* (/bɛtən/, 'to bed'), *bieten* (/biːtən/, 'to offer), *bitten* (/bɪtən/, 'to request), *Hüte* (/hyːtə/, 'hats'), *Hütte* (/hʏtə/, 'hut'), *Höhle* (/hølə/, 'cave'), *Hölle* (/hœlə/, 'hell'), *Sohne* (/soːnə/, 'son', 3 sg.), and *Sonne* (/sɔnə/, 'sun'). Both tense and lax vowels occur frequently before alveolar fortis stops. The test words were recorded together with 36 other words (cf. 1.6.2.1 and Table A.1 in Appendix A), which served as distractors (although some of them were analyzed in other studies).

Table 4.2: Number of tokens analyzed in this experiment for young, old, Saxon, and Bavarian speakers shown separately for each vowel.

Vowel	Stop	Saxon		Bavarian		∑
		Young	Old	Young	Old	
Tense						
/eː/	/t/	80	130	108	80	**398**
/iː/	/t/	80	131	110	80	**401**
/yː/	/t/	74	124	104	81	**383**
/oː/	/n/	78	126	107	80	**391**
/øː/	/l/	80	140	102	80	**402**
Lax						
/ɛ/	/t/	73	127	106	80	**386**
/ɪ/	/t/	79	130	109	79	**397**
/ʏ/	/t/	82	139	103	78	**402**
/ɔ/	/n/	79	126	110	80	**395**
/œ/	/l/	79	122	109	80	**390**
∑						**3945**

The speakers were the same as in the corpus described in chapter I (cf. 1.6.2.2) and the experiment described in chapter III (cf. 3.2.1) and the same speaker groups (Saxon vs. Bavarian and Young vs. Old) were tested. In the present experiment, however, the complete data sets of two speakers were excluded because for at least one of the ten test words, none of the ten repetitions were read correctly. Thus, the Saxon speaker group comprised 13 old speakers (five female and eight male, between the age of 50 and 80) and eight young speakers (five female and three male, aged between 18 and 40) and in the Bavarian speaker group, there were eleven young speakers (8 female and 3 male, 10 speakers in their twenties and one speaker at the age of 43) and eight old speakers (five female and three male, between 50 and 80 years of age). In cases where it was obvious that the subject misread the presented test words, the tokens were attributed to the opposing tenseness category. This method was chosen when the speaker was either from the Saxon group (because Saxon – according to the literature – distinguishes lax and tense vowel irrespective of the following stop's voicing) or when the dialectal background was Bavarian (where neutralization is possible), when all other repetitions of one token where unambiguously pronounced as tense or lax. 55 tokens of the remaining 40 speakers had to be excluded, because the presented tokens were misread or uttered too late in the course of the recording. The distribution of the 3945 tokens that were included in the analyses presented below is given in Table 4.2.

4.2.2 Segmentation and data analysis

The data processing was the same as in the experiment described in chapter III (cf. 2.2.2 and 1.6.2.3). The data was first automatically segmented with MAUS (Schiel, 2004) and in a second step segment boundaries were checked manually and corrected when necessary in Praat (Boersma & Weenink, 2008). The on- and offset of the vowel were always set to the beginning and end of a clearly visible second formant, the stop's offset was placed at the beginning of the voicing of the following vowel, i.e. burst, aspiration phase and VOT were included in the stop duration. Additionally, we calculated the first four formants in Emu (Harrington, 2010) with the following parameters: LPC order of 10, a pre-emphasis of 0.95, and a 30 ms Blackman window with a frame shift of 5 ms. The formant data was checked manually and corrected if necessary both by inspection of the spectrograms and identification of outliers from duration plots of the first, second and third formant over the course of the entire vowel.

Duration and formant data were extracted in Emu/R. The normalization method for duration was the same as in chapter III, i.e. the vowel duration of the accented syllable was – depending on the analysis – normalized for (1) word duration (henceforth V:W duration ratio) and (2) for VC duration (henceforth V:C duration ratio). Additionally, the intervocalic stop duration was also normalized for word duration (henceforth C:W duration). Formant data was measured at the temporal midpoint of the vowel and converted to Bark using the formula in Traunmüller (1990). The statistical analyses depended on the different data subsets that were compiled to test the research questions formulated in the introduction. All statistical analyses will be described in detail in the corresponding result sections below.

4.3 Results

4.3.1 The tense/lax opposition before obstruents

In the first analysis, we were interested in whether Bavarian speakers maintain the tense/lax distinction before obstruents to a greater extent than Saxon speakers. To test this research question, we analyzed the V:W duration ratio in the following minimal pairs /biːtən, bɪtən/, /beːtən, bɛtən/, and /hyːtə, hʏtə/.

A mixed model with V:W duration ratio as the dependent variable, Length (two levels: lax vs. tense), Region (two levels: Bavarian vs. Saxon), and Age (two levels: old vs. young) as fixed factors and Subject and Vowel as random factors revealed significant main effects for Length ($F[1,60] = 6747.8$, $p < 0.001$) and Region ($F[1,60] = 6.0$, $p < 0.05$) as well as significant interaction effects for Length x Region ($F[1,60] = 49.8$, $p < 0.001$) and Length x Age ($F[1,60] = 50.7$, $p < 0.001$). All effects due to age will be described and discussed below in section 3.5. The significant main effects for Length and Region indicate commensurate with Figure 4.2

(left) that the V:W duration ratio was in both dialect groups shorter for lax versus tense vowels and that the two dialect groups differed with respect to the V:W duration ratio. However, the regional difference in that parameter depends on the underlying length of the vowel as the significant interaction effect for Length x Region shows. As can be seen in the left panel of Figure 4.2 the V:W duration ratio was smaller for lax vowels produced by Bavarian vs. Saxon speakers, but it was about the same for tense vowels. This finding already suggests that the contrast between tense and lax vowels is maintained to a lesser extent by Saxons as opposed to Bavarians.

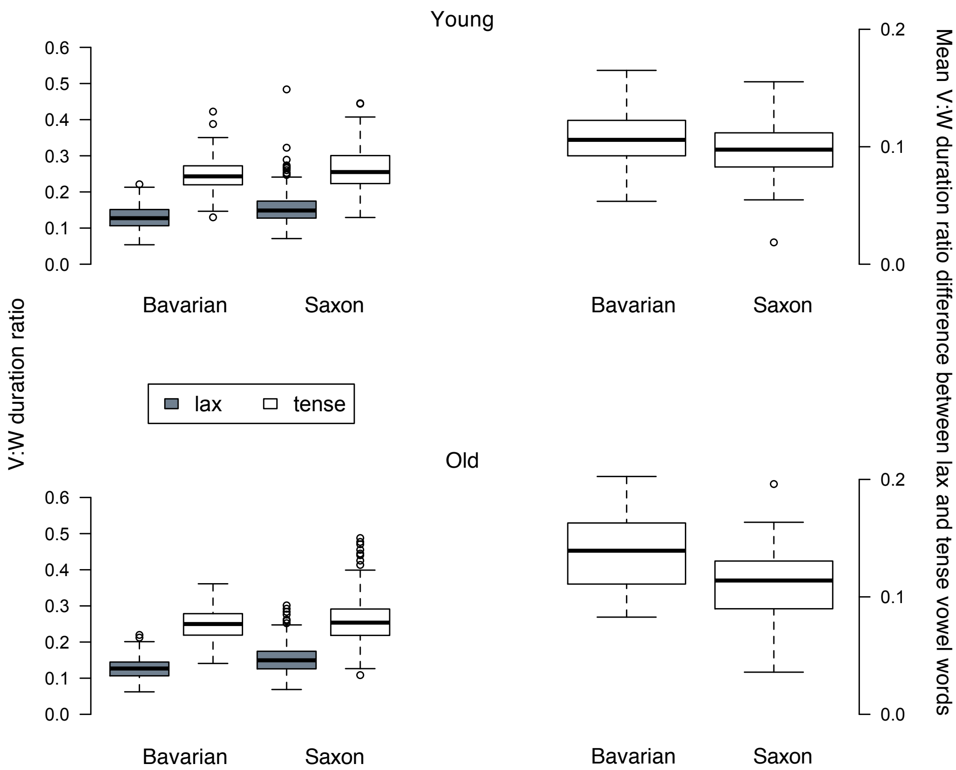


Figure 4.2: V:W duration ratios for lax (grey) and tense (white) vowels before obstruents (left) and mean V:W duration differences between tense and lax vowels before obstruents (right) across all minimal pairs separately for young (above) and old (below), Bavarian and Saxon speakers.

To test whether the differentiation of tense and lax vowels by means of vowel duration was different for Saxon vs. Bavarian speakers, we calculated the mean difference between the V:W duration ratio in all tense tokens and the V:W duration ratio in all lax tokens for each minimal pair and speaker resulting in 120 values (three minimal pairs x 40 speakers). The three values per speaker were subsequently averaged. An ANOVA with the mean V:W duration ratio difference as the dependent variable and Length, Region and Age as independent variables showed significant main effects for Region ($F[1,36] = 6.6$, $p < 0.05$) and Age ($F[1,36] = 6.5$, $p < 0.05$). The significant main effect for Region confirms that the tense/lax distinction before obstruents is maintained to a greater extent in the tokens produced by Bavarian than by Saxon speakers (cf. Figure 4.2, right panel).

A third analysis of the spectral differences between tense and lax vowels aimed to verify whether the tense/lax opposition was maintained by differences in vowel quality in these two dialects and if so, whether the two dialect groups differed in that parameter. A mixed model with F1 (Bark) at the temporal midpoint of the vowel as the dependent variable and Length (two levels: lax vs. tense), Region (two levels: Bavarian vs. Saxon), and Age (two levels: young vs. old) as fixed factors and Subject and Vowel as random factors revealed a significant main effect for Length ($F[1,60] = 5633.0$, $p < 0.001$) and a significant interaction between Length x Age ($F[1,60] = 60.7$, $p < 0.001$). A second mixed model with the same fixed and random factors, but with F2 (Bark) at the vowel's temporal midpoint as the dependent variable showed again a significant main effect for Length ($F[1,60] = 2248.4$, $p < 0.001$) and a significant interaction between Length x Age ($F[1,60] = 11.0$, $p < 0.01$). The significant effects for Length in both models confirm that in both age and dialect group tense and lax vowels were also differentiated by means of vowel quality, and in this case in both acoustic fronting and height. As can be seen from Figures 4.3, there is a great deal of overlap[6] between lax and tense vowels (but also between close and close-mid vowels), but at the same time, two categories emerge for the two members of each pair of vowels that differ in phonological vowel quantity: front, close, lax vowels have a more central, i.e. open and back vowel quality as opposed to their tense counterparts. The non-significant effect for Region indicates that there are no dialectal

[6] Which may also come about because of inter-speaker variation.

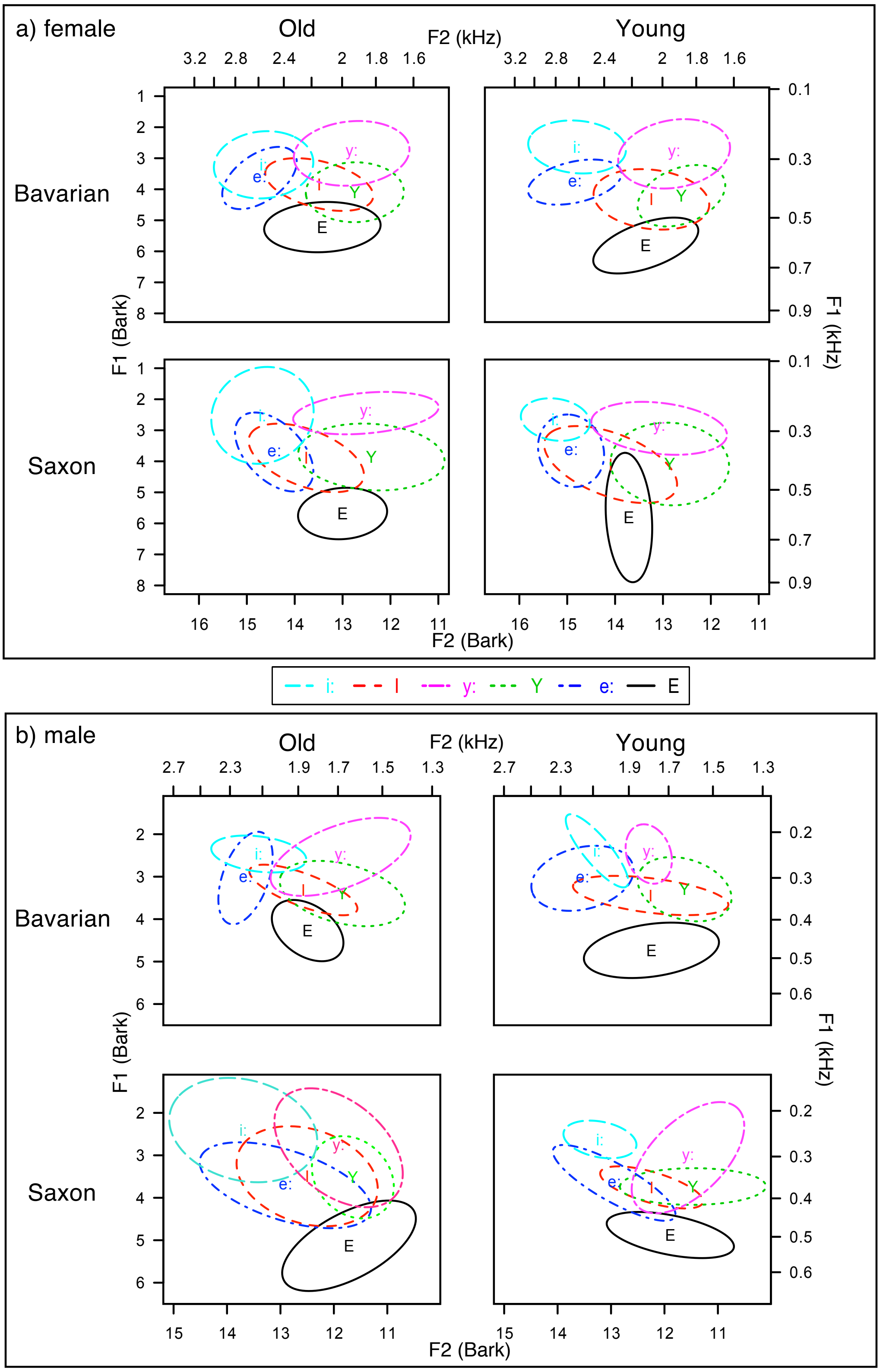


Figure 4.3: 95% confidence ellipses for /iː ɪ yː ʏ eː ɛ/ in a Bark-scaled formant plane shown separately for female (a) and male (b) age and dialect groups. The formant data was extracted at the vowels' temporal midpoints.

differences with respect to vowel quality and the tense/lax distinction.[7] To summarize, both dialect groups maintained the tense/lax opposition by use of both vowel length and quality, but they differed only with respect to the (proportional) duration of tense vs. lax vowels. Therefore we need not consider vowel quality any further in this investigation.

4.3.2. /t/-lenition after lax vowels

According to the Central German Lenition rule, voiceless consonants are voiced in Upper Saxon and we have provided evidence in the experiment discussed in chapter III that Saxon speakers tend to produce underlying voiceless stops after tense vowels with a greater V:C duration ratio (i.e. more lenis-like) than Bavarian speakers. The aim of the second analysis was to investigate whether Saxon and Bavarian speakers also differed in the realisation of the post-vocalic stop after lax vowels, i.e. whether fortis stops are lenited in all contexts. To test this, a subset consisting of the three test words *bitten*, *betten*, and *Hütte* was analyzed with regard to the V:C duration ratio, which is one of the most important acoustic cues to signal stop voicing.

Table 4.3: Mean V:C duration ratios for words with lax vowels shown separately for young and old, Bavarian and Saxon speakers.

Word		Saxon		Bavarian	
		Young	Old	Young	Old
betten	mean	0.38	0.35	0.30	0.27
	sd	0.09	0.06	0.05	0.03
bitten	mean	0.32	0.28	0.24	0.20
	sd	0.09	0.06	0.05	0.03
Hütte	mean	0.31	0.30	0.25	0.21
	sd	0.08	0.07	0.04	0.03

In chapter III a ratio below 0.60 was found to be a clear indicator for the production of fortis stops after tense vowels. This finding was in line with the results presented in Kohler

[7] The non-significant effect for Region may be somewhat surprising since one would expect differences in vowel quality between the two varieties (cf. 1.4). One of the reasons for this effect may be the fact that vowels were analyzed that do not differ to a great extent in the two varieties. Another might be that regional characteristics are reduced in the laboratory speech used for this analysis.

(1979), Kohler and Künzel (1978) as well as Braunschweiler (1997). Kohler and Künzel (1978) and Braunschweiler (1997) found lax vowels to be approximately half as long as tense vowels. Furthermore, the mean V:C duration ratios in words containing short /a/ were shown to be 0.39 when the stop was fortis and 0.56 when the stop was lenis (cf. Table 1 in Braunschweiler, 1997: 363; Vowel / (Vowel + Closure + Release)). In Table 4.3, the mean V:C duration ratios for the test words of the present analysis are shown: all ratios were below 0.39, i.e. all speaker groups produced the stops as fortis stops. The mean V:C duration ratios differed slightly with respect to the underlying vowel quality: ratios were higher for open-mid /ɛ/ vowels than for close /ɪ ʏ/ vowels (cf. Table 4.3). But more importantly, there was a regional difference between the V:C duration ratios of the two dialect groups. Saxon speakers always showed longer mean V:C duration ratio values than Bavarian speakers. That is, for Saxons as opposed to Bavarians, the vowel and stop durations were longer and shorter, respectively. This indicates that the underlying fortis stops were realized more lenis-like when Saxons produced them.

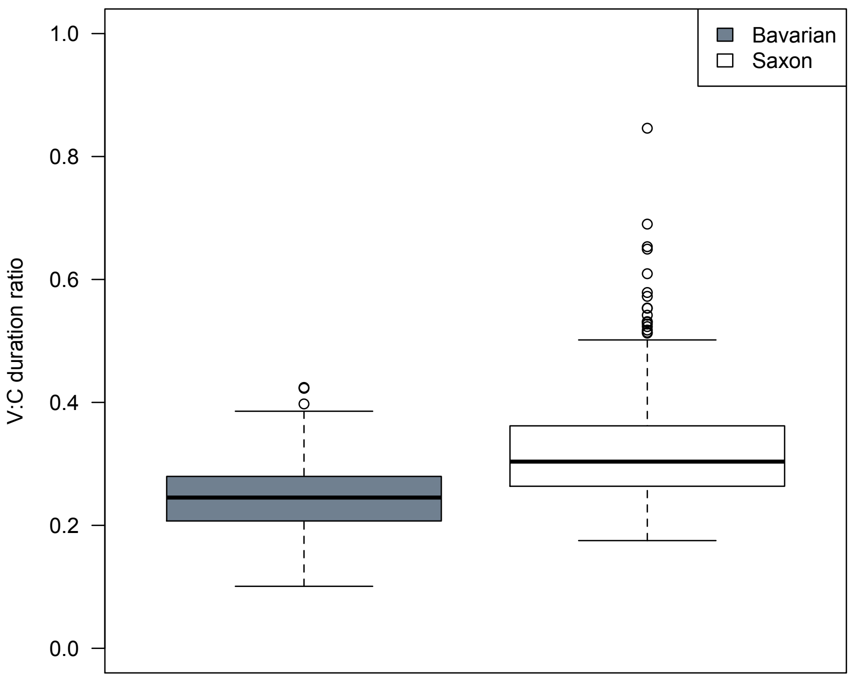


Figure 4.4: V:C durations ratios for lax vowels before obstruents across all minimal pairs and age groups shown separately for Bavarian (grey) and Saxon (white) speakers.

A mixed model with V:C duration ratio as the dependent variable and Region (two levels: Bavarian vs. Saxon) and Age (two levels: young vs. old) showed a significant main effect for Region ($F[1,60] = 26.6$, $p < 0.001$) and no other significant main or interaction effects. The significant main effect for Region corroborates the finding that the V:C duration ratio was significantly longer for Saxon vs. Bavarian speakers. The results show, therefore, that the alveolar fortis stops in the test words were produced by both dialect groups as fortis stops whereas there was a greater tendency for Saxon as opposed to Bavarian speakers to lenite them, i.e. the fortis stops produced by Saxon speakers were more lenis-like than the stops produced by Bavarian speakers (cf. Figure 4.4).

4.3.3 Schwa deletion

Before we turn to the next research question it is important to consider schwa deletion in our data because the realization or deletion of the schwa vowel may affect the duration of the preceding stop and hence the V:C duration ratio (cf. 3.3.5). We therefore analyzed all test words with *en*-ending with respect to whether the postvocalic stop was orally or nasally released. The purpose of this analysis was to eliminate any possible misinterpretations of duration differences between the speaker groups that are caused by schwa deletion. Since the analysis is not part of the investigation of the above formulated research question, we will present age-dependent differences with respect to schwa deletion in this section and not in 4.3.5. Table 4.4 presents the distribution of schwa deletions in words with *en*-ending.

Overall, young speakers in both dialect groups deleted the schwa vowel more often than old speakers, i.e. a nasal release was found more often in the young group whereas an oral release was found more often in the old group. However, the difference between young and old speakers was much greater in the Bavarian than in the Saxon group. While most tokens were produced with an oral release in both Saxon groups, and even more so in the old Bavarian group, the majority of tokens produced by young Bavarians contained a nasal release. In the Bavarian group, schwa deletion was highly speaker dependent: speakers either deleted or realized the schwa across all ten repetitions of each word, that is the inter-speaker variation was high and the intra-speaker variation was low. The distribution among the Saxon speakers revealed a slightly higher intra-speaker variation, i.e. within speakers both variants were found. A table of the distribution of schwa deletions per speaker and word is given in Appendix B (Table B.1). Apart from these speaker-dependent differences, there were no word-dependent differences: speakers realized or deleted the schwa vowel irrespective of whether the nucleus of the first syllable was an underlying tense or lax vowel or whether it was a close or a close-mid vowel. Out of 1582 words with *en*-ending, 459 tokens are produced without a schwa vowel and this number was almost equally distributed among lax vowel words (216) and tense vowel words (243).

Table 4.4: Number of tokens with and without schwa as well as the percentage of tokens with schwa deletion for each word with *en*-ending shown separately for young and old, Saxon and Bavarian speakers.

Word	Saxon						Bavarian					
	Young			Old			Young			Old		
	ə	Ø	%	ə	Ø	%	ə	Ø	%	ə	Ø	%
beten	54	26	33	110	20	15	31	77	71	80	0	0
betten	52	21	29	111	16	13	29	77	73	79	1	1
bieten	53	27	34	112	19	15	37	73	66	79	1	1
bitten	61	18	23	118	12	9	38	71	65	79	0	0

Schwa deletion caused a longer V:C duration ratio as can be seen from Figure 4.5, that is the stop is shorter in words without a schwa than in words with a schwa. Because the data was not normally distributed, we ran two separate Wilcoxon rank sum tests with continuity correction on lax and tense vowel words. Both tests revealed that the V:C duration ratio was significantly longer in words with a nasal release than in words in which the schwa vowel was realized (lax: $W = 49241$, $p < 0.001$; tense: $W = 51565$, $p < 0.001$).The difference between nasally vs. orally released stops was smallest for the young Bavarian and greatest for the young Saxon group as can be seen in Figure 4.6. A mixed model with only the young speakers' V:C duration ratio as the dependent variable was calculated. Region (two levels: Bavarian vs. Saxon), Length (two levels: tense vs. lax), and Release (two levels: nasal vs. oral) were entered as fixed factors and Speaker and Vowel quality were entered as random factors. The significant main effects for Region ($F[1,60] = 7.1$, $p < 0.01$), and Length ($F[1,60] = 4107.1$, $p < 0.001$) as well as the significant interaction for Region x Length ($F[1,60] = 8.2$, $p < 0.01$) were commensurate with our findings described in the previous section 4.3.2. More important in this analysis were the

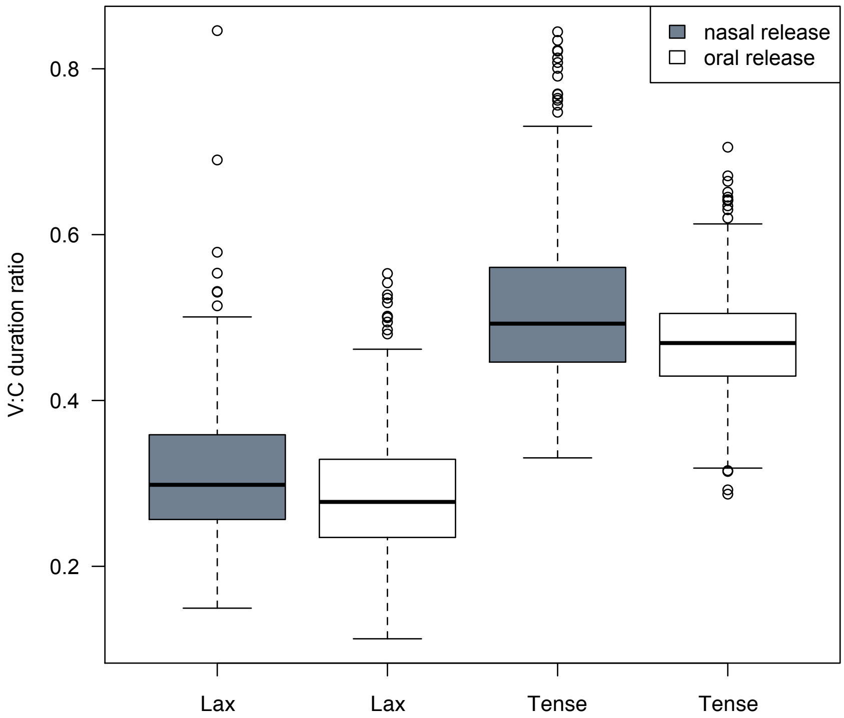


Figure 4.5: V:C duration ratios in *en*-words with schwa deletion (grey) and with schwa realization (white) across all four speaker groups.

effects for Release: there was a significant main effect for Release ($F[1,60] = 14.9$, $p < 0.001$), but no significant interactions for Region x Release nor Length x Release. That means that the significant shortening of the stop due to schwa deletion was neither affected by the vowel's underlying Length nor the speakers' regional background. However, the interaction between Region and Release showed a significant trend ($F[1,60] = 3.7$, $p = 0.06$), i.e. commensurate with Figure 4.6 the V:C duration ratio difference between nasally and orally released stops was greater in young Saxons vs. young Bavarians. Recall that the presence or absence of a Schwa was speaker dependent, i.e. when we compare nasally vs. orally released stops within one speaker group we also compare for the most part different speakers within that group.

Since neither schwa deletion nor the difference in V:C duration ratio depended on the preceding vowel's underlying Length, the calculation of the V:C duration ratio difference is an adequate normalizing method for testing the next research question which addresses the implication of the tense/lax opposition for a possible fortis/lenis contrast. The V:C duration ratio remains a valid parameter in both words with and without schwa deletion, because the stop shortening caused by schwa deletion appears in both underlying tense and lax vowels to the same extent. Thus, interpretations that possible significant effects of the independent variables Length, Region and Age are caused by schwa deletion can be excluded at this point.

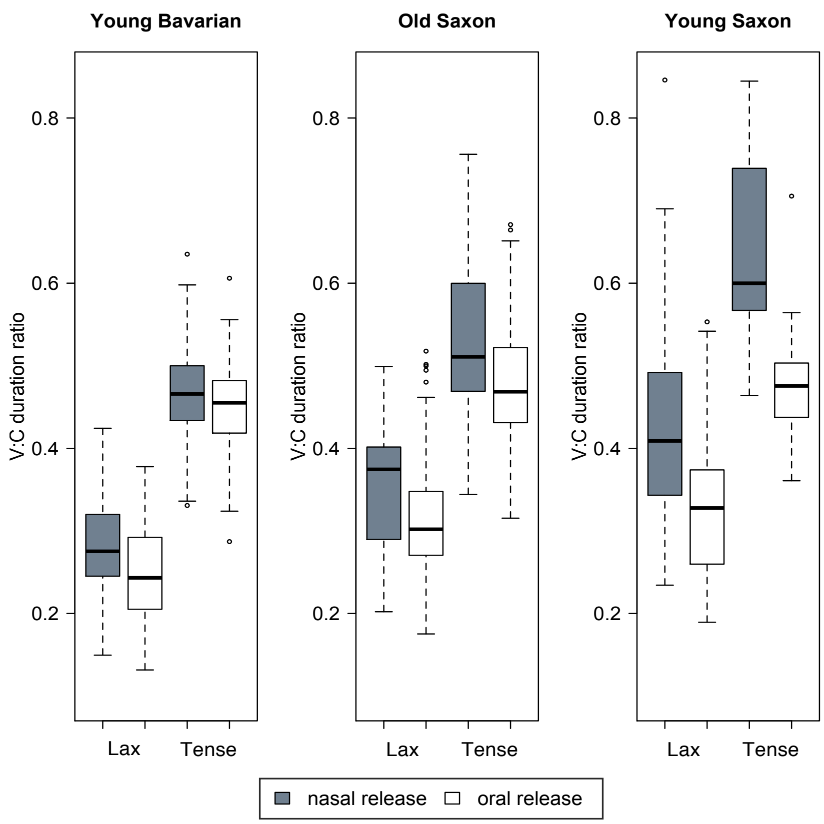


Figure 4.6: V:C duration ratios in *en*-words with schwa deletion (grey) and with schwa realization (white) separately for young Bavarian (left), old Saxon (middle) and young Saxon (right) speakers. The old Bavarian group was excluded from this Figure since in this speaker group only two out of 317 words were produced without a schwa vowel.

4.3.4 The tense/lax opposition before obstruents by means of the V:C duration ratio

The results so far have confirmed that the tense/lax opposition was maintained to a greater extent by Bavarian vs. Saxon speakers and that Saxon speakers lenited fortis stops following lax vowels to a greater extent than Bavarian speakers. The next question to be asked then is whether the difference between underlying fortis stops following lax and tense vowels is greater for Bavarians than for Saxon speakers, on the assumption that the V:C duration ratio is effected by a fortis/lenis contrast for Bavarians but a neutralization tendency towards lenis stops for Saxons. Before we turn to this analysis, we present the results for /t/-lenition after tense vowels which are partly a replication of the findings described in chapter III and which are necessary for the further analysis of the contrast maintenance between tense and lax vowels. The V:C duration ratio in the three test words *bieten*, *beten*, and *Hüte* were analyzed.

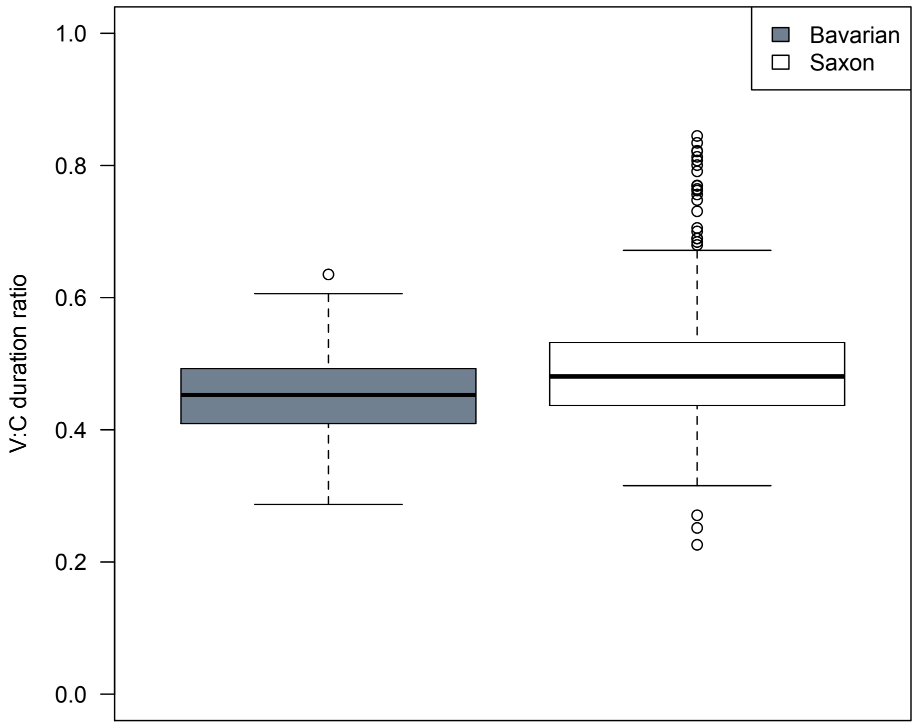


Figure 4.7: V:C duration ratios for tense vowels before obstruents across all minimal pairs and age groups shown separately for Bavarian (grey) and Saxon (white) speakers.

Table 4.5: Mean V:C duration ratios for words with tense vowels shown separately for young and old, Bavarian and Saxon speakers.

Word		Saxon		Bavarian	
		Young	Old	Young	Old
beten	mean	0.54	0.51	0.48	0.48
	sd	0.10	0.07	0.05	0.04
bieten	mean	0.51	0.46	0.44	0.45
	sd	0.11	0.07	0.06	0.07
Hüte	mean	0.47	0.47	0.43	0.42
	sd	0.07	0.08	0.06	0.06

Table 4.5 presents the mean V:C duration ratios for these tests. All ratios were again distinctly below 0.60, i.e. all speaker groups produced the stops as fortis stops. A mixed model with the V:C duration ratio as the dependent variable and Region (two levels: Bavarian and Saxon) and Age (two levels: young vs. old) as fixed factors and Subject and Vowel as the dependent variable revealed a significant main effect for Region ($F[1,60] = 6.1$, $p < 0.05$) and no other significant main nor interaction effects. That is, Saxons and Bavarians differed significantly in the realization of fortis stops after tense vowels commensurate with Figure 4.7. Again, Saxon speakers tended to produce fortis stops with greater lenition than Bavarian speakers, but after tense vowels the effect for Region was smaller compared to lax vowels (cf. 4.3.2). This implies either that Saxon speakers lenite fortis stops to a greater extent after lax than tense vowels or that Bavarian speakers tend to lenite fortis stops only in the context of tense vowels.

To determine the degree to which the contrast was maintained by the V:C duration ratio, we calculated the mean V:C duration ratio difference by subtracting the mean V:C duration ratio of all ten repetitions of the lax vowel words from the mean V:C duration ratio of all ten repetitions of the tense vowel words per speaker and per minimal pair resulting in 120 values (3 minimal pairs x 40 speakers, cf. 3.1). We then averaged over the three values per speaker. Note, however, that the V:C duration ratio contains both information on vowel length and the fortis/lenis contrast. The results are presented in Figure 4.8.

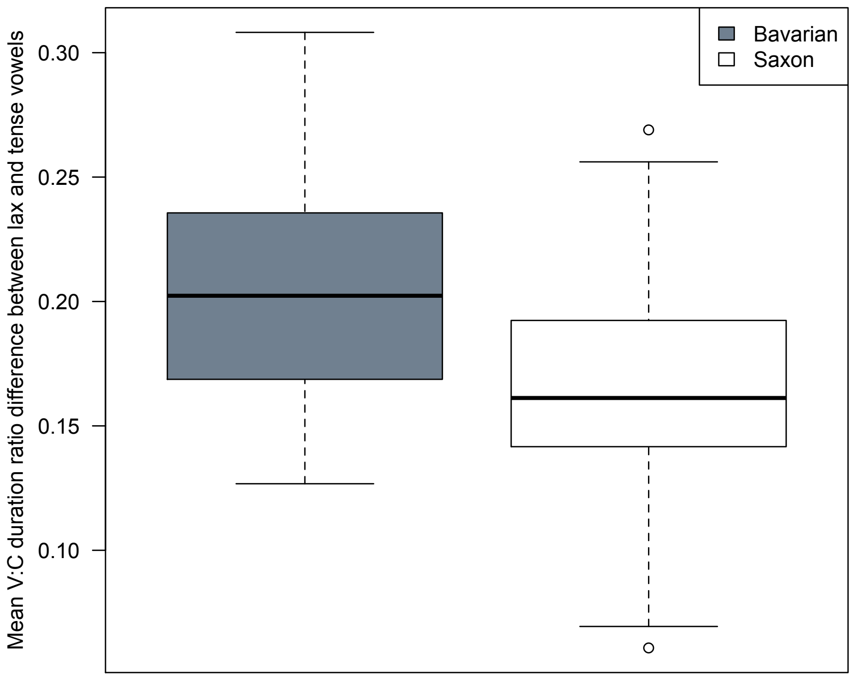


Figure 4.8: Mean V:C duration ratio differences between tense and lax vowels before obstruents across all three minimal pairs and age groups shown separately for Bavarian (grey) and Saxon (white) speakers.

An ANOVA[8] with V:C duration ratio difference as the dependent variable and the independent variables Region and Age revealed a significant main effect for Region ($F[1,36] = 14.8$, $p < 0.001$) and a significant interaction between Region x Age ($F[1,36] = 4.9$, $p < 0.05$), but no significant main effect for Age. A post-hoc test (Tukey HSD) showed that old Bavarian speakers differed significantly from all other speaker groups but neither young Bavarian and young or old Saxon speakers nor young and old Saxon speakers (cf. section 4.3.5, Figure 4.9).

[8] An ANOVA was chosen because the number of analyzed tokens was too small for a mixed model analysis and although we analyzed different minimal pairs, the only within subject factor Vowel (i.e. closed unrounded, closed rounded, and close-mid unrounded) in this analysis has to be considered a random and not a fixed factor, since there are no predictions made about a difference in V:C duration ratio that depends on vowel height or lip rounding. In addition, a post-hoc test with Bonferroni correction (as is used with repeated measures ANOVAs) is not appropriate since it is very conservative and the high number of comparisons (which is multiplied with the p-value of each factor level) in our model leads to non-significant results even between pairs that are obviously different.

4.3.5 Age-dependent differences

The fourth hypothesis was that the predicted Bavarian-Saxon differences in the first three analyses (H1 – H3) are more pronounced for old than for young (on the assumption that there is a greater tendency that young speakers of both varieties tend towards a standard pronunciation). In this section we will describe in detail the results for the fixed factor Age in each of the above analyses following the same order.

4.3.5.1 The tense/lax opposition before obstruents

In the ANOVA with the mean V:W duration ratio difference between lax and tense vowels before obstruents as the dependent variable there was a significant main effect for Age and in the mixed model with the V:W duration ratio as the dependent variable there was a significant interaction effect for Length x Age. These results support the hypothesis that the degree with which the tense/lax contrast is maintained by means of vowel duration depends on Age. As can be seen from the left column in Figure 4.2 old Bavarian speakers maintain the vowel length contrast by means of the V:W duration ratio to a greater extent than young Bavarians and young and old Saxon speakers.

There were no main effects for Age in the two mixed models with either F1 (Bark) or F2 (Bark), which implies that young and old members of the same dialect group do not differ overall in their spectral patterns. However, the significant interaction effects for Length x Age in both mixed models indicate that the age groups differ in vowel quality, i.e. they show different centralization tendencies with respect to the underlying tenseness of the vowel. There may have been slightly more overlap between tense and lax vowels in the F1/F2 vowel space for old than for young speakers in both dialect groups (cf. Figure 4.3). This could mean that young speakers differentiate the two vowel categories to a greater extent by means of spectral patterns than old speakers. The aim of the spectral analysis was to test for contrast maintenance with respect to vowel quality. It is important to note that all speaker groups distinguish the tense/lax opposition by means of both acoustic correlates. The extent to which the two age groups differ in their spectral overlap between lax and tense vowels, however, is not relevant for the present analysis. Thus, we will not further discuss this finding.

4.3.5.2 /t/-lenition after lax and tense vowels

There were no significant main or interaction effects for Age in the two mixed models with V:C duration ratio in words with underlying lax vowels or in words with underlying tense vowels as the dependent variables. That means that young and old speakers of both dialect groups lenite (or not) fortis stops to the same extent.

4.3.5.3 The tense/lax opposition before obstruents by means of the V:C duration ratio

In the analysis of the mean V:C duration ratio difference between tense and lax vowels, there was again no significant main effect for Age. However, there was a significant interaction effect for Region x Age ($F[1,36] = 4.9$, $p < 0.05$) and a post-hoc test showed – as can be seen in Figure 4.9 – that young and old Bavarian speakers differ significantly in their mean V:C duration ratio differences between tense and lax vowels.

A repeated measures MANOVA with the mean V:C duration ratio difference as the dependent variable and Region (between subject factor with two levels: Bavarian vs. Saxon), Age (between-subject factor with two levels: young vs. old) and Vowel (within-subject factor with three levels: iː/ɪ vs. eː/ɛ vs. yː/ʏ) was calculated to test whether this pattern/interaction was the same for all three minimal pairs and whether the frequently occurring schwa deletion especially in the productions by young Bavarian speakers potentially causes this difference between young and old Bavarians. The repeated measures MANOVA revealed significant main effects for Region ($F[1,36] = 14.8$, $p < 0.001$) and Vowel ($F[2,35] = 15.7$, $p < 0.001$) as well as a significant tendency for Age ($F[1,36] = 3.1$, $p = 0.088$) and a significant interaction effect for Age x Region ($F[1,36] = 4.9$, $p < 0.05$).[9] As can be seen from Figure 4.10 and as is indicated by

[9] Thus, the results between the ANOVA in 4.3.4 and the repeated measures MANOVA are overall the same.

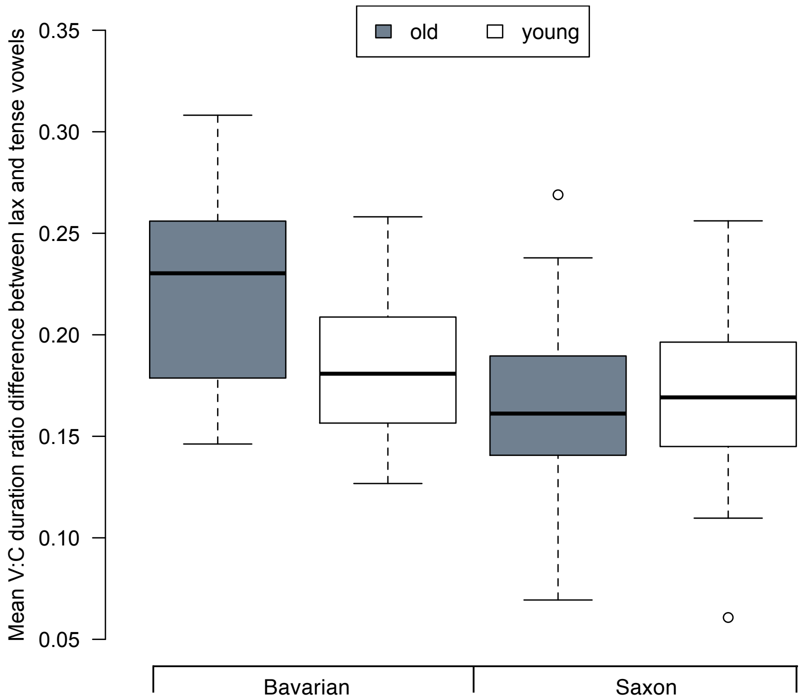


Figure 4.9: Mean V:C duration ratio differences between tense and lax vowels before obstruents across all three minimal pairs shown separately for old (grey) and young (white), Bavarian and Saxon speakers.

the significant main effect for Region, the difference between the V:C duration ratio for lax and tense vowels was greater for Bavarian than for Saxon speakers. This finding again supports hypotheses H2 and H3. The mean V:C duration ratio difference indicates more lenition in Saxon vs. Bavarian. Moreover, the significant tendency for Age partly supports hypothesis H4 that the Bavarian–Saxon differences are more pronounced for old than for young, especially in the minimal pairs *bieten/bitten* and *Hüte/Hütte* (cf. Figure 4.10). However, the significant interaction for Region x Age suggests that the difference between age groups was not the same for Saxons and Bavarians. As can be seen in Figures 4.9 and 4.10 and as the results of the earlier post-hoc Tukey test showed, only the Bavarian age groups (but not old and young Saxons) differed significantly: the tense/lax contrast before underlying fortis stops was diminished for young vs. old Bavarians. One possible interpretation of this finding could be that the large Bavarian age difference comes about because young speakers tend to delete the schwa

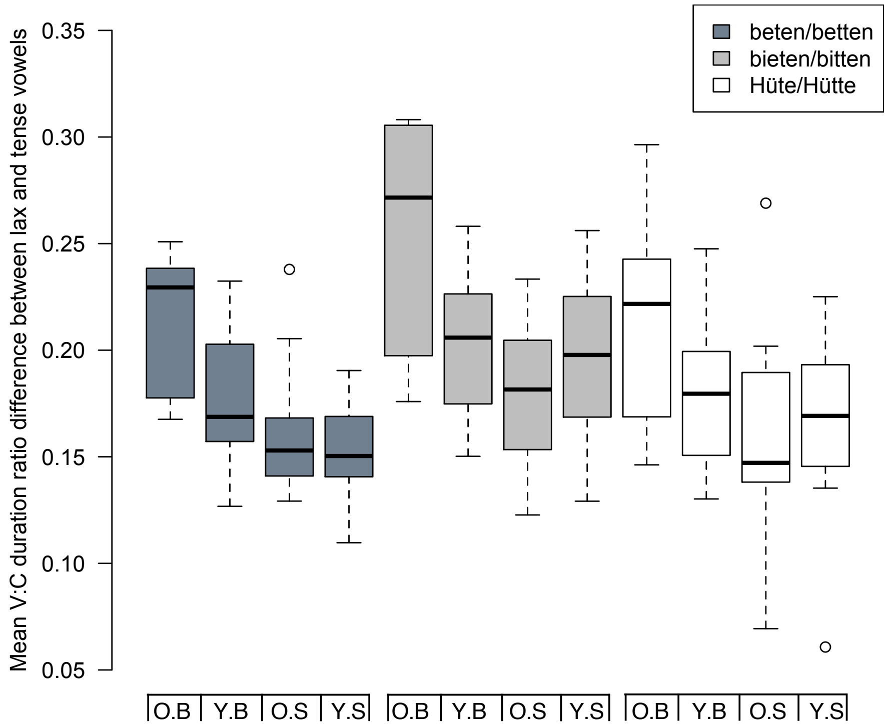


Figure 4.10: Mean V:C duration ratio differences between tense and lax vowels before obstruents across all minimal pairs shown separately for old (O) and young (Y), Bavarian (B) and Saxon (S) speakers.

vowel much more frequently than old speakers. And indeed there was a significant main effect for Vowel ($F[2,35] = 15.7$, $p < 0.001$), which implies that the mean V:C duration ratio difference between tense and lax vowels was different for the pairs /i: ɪ/, /y: ʏ/, and /e: ɛ/. But, there was neither a significant interaction between Age x Vowel nor Region x Vowel. The effect for Vowel possibly comes about because of intrinsic vowel duration differences especially in closed unrounded vowels. Schwa deletion, however, which occurred only in words with an *en*-ending, cannot be the reason, since the pattern of contrast maintenance between the four speaker groups, was about the same for *Hüte*/*Hütte*, in which the schwa vowel is always realized, and *beten*/*betten*, in which schwa was deleted by some speakers. On the contrary, the mean V:C duration ratio difference between *bieten*/*bitten* was greater than between *beten*/*betten,* even though the two minimal pairs did not differ in the extent with which schwa was deleted (cf. Table 4.4). The non-significant interaction effects for Vowel x Age and Vowel x Region confirm – commensurate with Figure 4.10 – that the vowel dependent divergence in the V:C duration ratio difference was about the same for both age and dialect groups.

4.3.6 The tense/lax distinction before sonorants

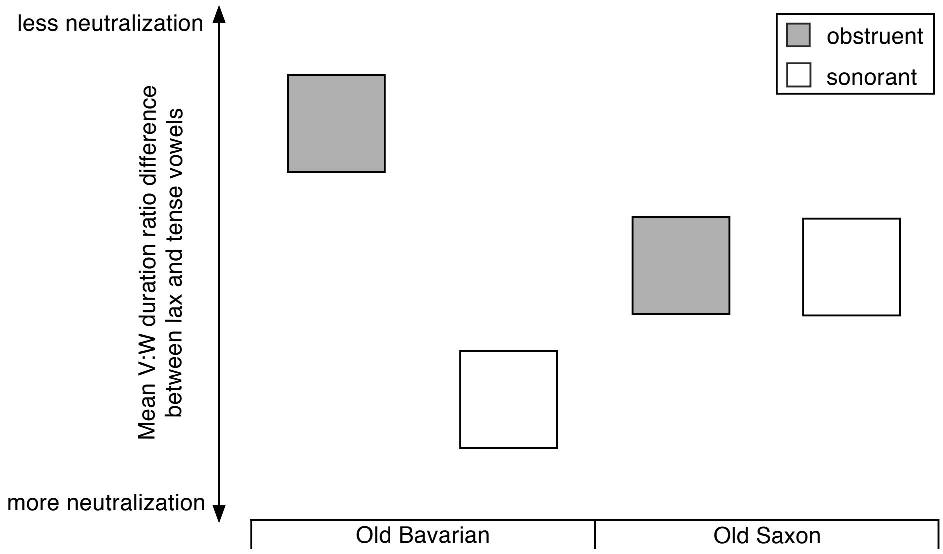


Figure 4.11: Hypothetical mean V:W duration ratio differences between lax and tense vowels preceding obstruent (grey) and sonorant (white) consonants separately for old Bavarian and old Saxon speakers.

If the greater contrast maintenance between tense and lax vowels was a consequence of a Bavarian co-dependency of vowel length and stop voicing then one possible prediction is that Bavarians and Saxons maintain the tense/lax distinction to the same extent when a sonorant consonant follows based on the assumption that only long vowels can precede voiced segments. The purpose of this section is to address this issue. Since the results of the preceding section showed that only old Bavarian speakers differed in their degree of contrast maintenance from the other three speaker groups we first compared only the old speakers' V:W duration ratios in test words with a sonorant consonant, i.e. *Höhle*/*Hölle* and *Sohne*/*Sonne*. The prediction was that Bavarian speakers preserve the vowel length contrast before sonorants to a lesser extent than Saxon speakers based on the assumption that (1) lax vowels are lengthened when they occur before voiced segments and (2) there is no possibility of contrast maintenance in terms of an additional voicing contrast in the postvocalic consonant (cf. Figure 4.11).

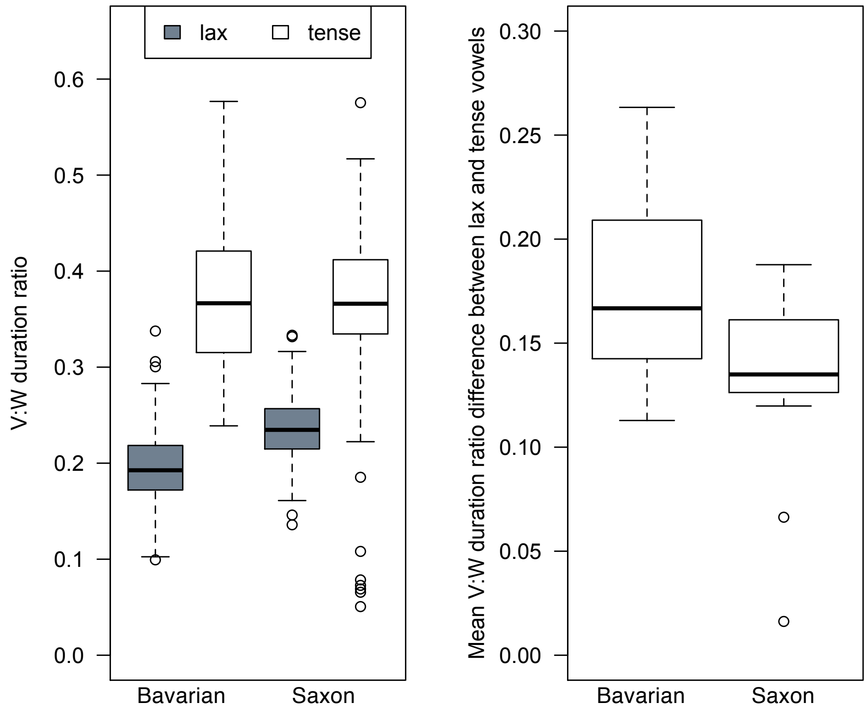


Figure 4.12: V:W duration ratios for lax (grey) and tense (white) vowel words containing sonorants (left) and mean V:W duration ratio differences between tense and lax vowels preceding sonorants (right) separately for old Bavarian and old Saxon speakers.

A mixed model with the old speakers' V:W duration ratio in words only with sonorant consonants as the dependent variable, and Length (two levels: tense vs. lax), and Region (two levels: Bavarian vs. Saxon) as fixed factors and Subject and Vowel as random factors revealed significant main effects for Length ($F[1,60] = 2212.6$, $p < 0.001$) and Region ($F[1,60] = 9.3$, $p < 0.01$) as well as a significant interaction effect for Length x Region ($F[1,60] = 218.1$, $p < 0.001$). The significant main effects for Length and Region indicate that the tense/lax contrast before sonorants was preserved by both speaker groups and that old Saxons and old Bavarians differed with respect to the proportional vowel duration. The significant interaction effect for Length x Region implies that the tense/lax contrast was maintained to a different extent by old Bavarian vs. old Saxon speakers. As can be seen from Figure 4.12 Saxon and Bavarian speakers differed in the proportional vowel duration in words containing lax but not tense vowels.

In order to test the extent to which the contrast is preserved by old Bavarians and old Saxons we again calculated the difference between the mean V:W duration ratio of all ten repetitions of tense vowels and the mean V:W duration ratio of all ten repetitions of lax vowels per minimal pair and speaker resulting in 80 values (2 minimal pairs x 40 speakers). The two values per speaker were again averaged. An ANOVA with the mean V:W duration ratio difference as the dependent variable and Region (two levels: Bavarian vs. Saxon) as the independent variable showed a significant main effect for Region ($F[1,19] = 4.4$, $p < 0.05$). This means that the contrast preservation was not the same for Bavarian and Saxon speakers. As can be seen from Figure 4.12, the tense/lax opposition was maintained to a greater extent by old Bavarian than old Saxon speakers and this is because lax (but not tense) vowels were shorter in the productions of Bavarian vs. Saxon speakers. This finding does not support the above-formulated prediction.

The next question to be asked then was whether the consonant duration was the same for Bavarian and Saxon speakers or whether the dialect groups differed also with respect to sonorant duration. A mixed model with the old speakers' C:W duration ratio in words only with sonorant consonants as the dependent variable, and Length (two levels: tense vs. lax) and Region (two levels: Bavarian vs. Saxon) as fixed factors as well as Subject and Vowel as random factors showed significant main effects for Length ($F[1,60] = 931.7$, $p < 0.001$) and Region ($F[1,60] = 8.9$, $p < 0.01$) as well as a significant interaction effect for Length x Region ($F[1,60] = 213.2$, $p < 0.001$). The significant main effects for Length and Region indicate that the proportional sonorant duration was – commensurate with Figure 4.13 – longer after lax as opposed to tense vowels and that it was dialect-dependent. The significant interaction effect for Length x Region implies that – again commensurate with Figure 4.13 – the lengthening effect of sonorants following lax vowels was much greater for old Bavarian vs. old Saxon speakers.

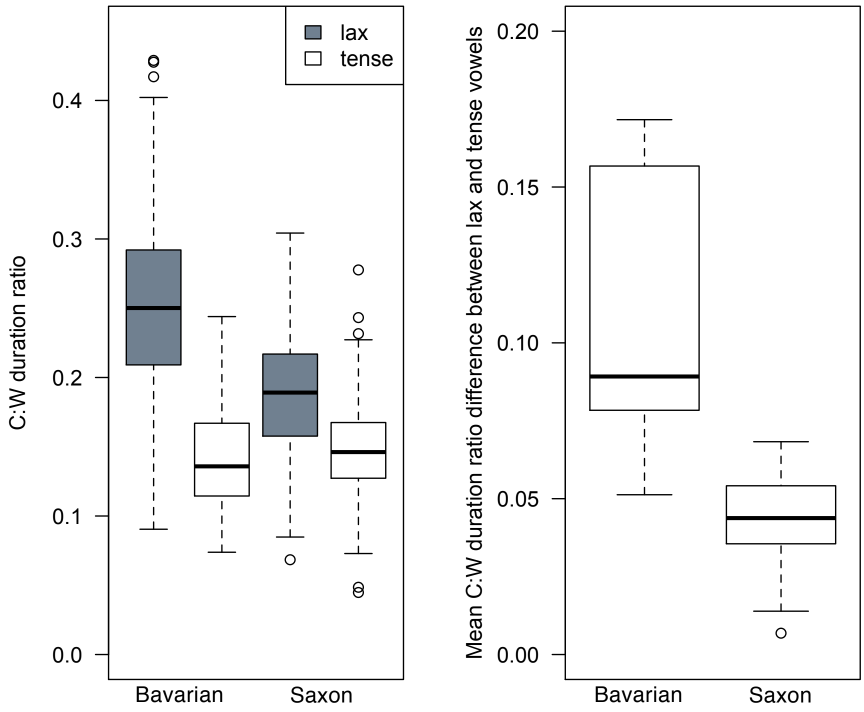


Figure 4.13: C:W duration ratios for lax (grey) and tense (white) vowel words containing sonorants (left) and mean C:W duration ratio differences between tense and lax vowel words containing sonorants (right) separately for old Bavarian and old Saxon speakers.

To determine the extent to which old Bavarians and old Saxons differed with respect to the sonorant duration after lax vs. tense vowels we calculated the mean C:W duration difference by subtracting the mean C:W duration ratio of all ten repetitions of tense vowel words from the mean C:W duration ratio of all ten repetitions of lax vowel words per minimal pair and speaker resulting in 80 values (2 minimal pairs x 40 speakers). The two values per speaker were again averaged. A one-way analysis of means (because the variances of the two levels Bavarian vs. Saxon of the independent variable Region were significantly different, $F(7,12) = 5.8$, $p < 0.01$) confirmed – commensurate with Figure 4.13 – that the mean C:W duration ratio difference between tense and lax vowel words was significantly greater ($F(1,9) = 15.5$, $p < 0.01$) for old Bavarian than for old Saxon speakers.

Given that young Bavarian speakers' V:C duration ratio did not differ significantly from that of young (and old) Saxon speakers in the analysis described in section 4.3.5, the two young age groups should also behave similarly with respect to the V:W and C:W duration ratios in words containing sonorants.

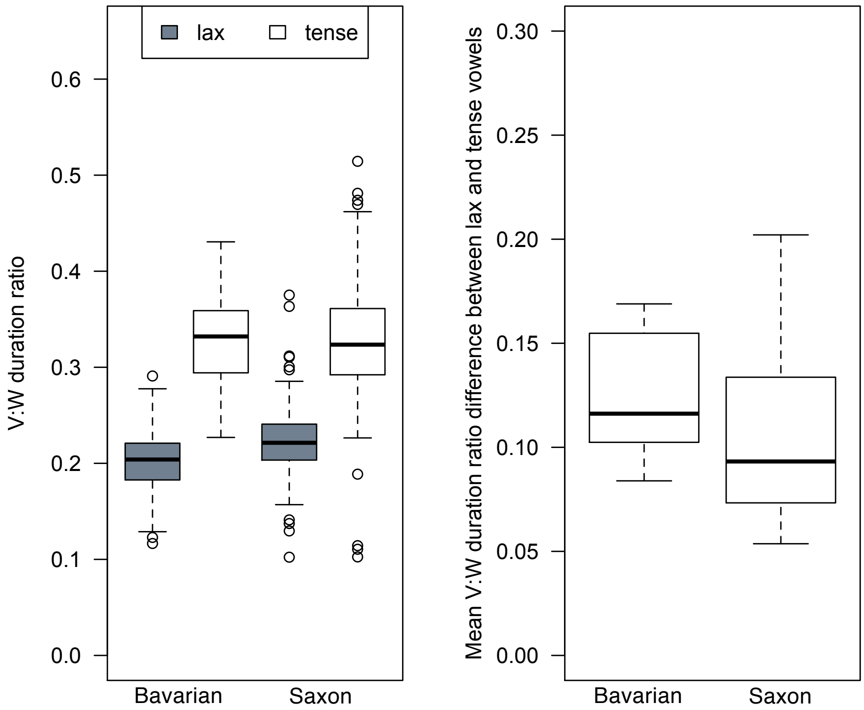


Figure 4.14: V:W duration ratios for lax (grey) and tense (white) vowel words containing sonorants (left) and mean V:W duration ratio differences between tense and lax vowels preceding sonorants (right) separately for young Bavarian and young Saxon speakers.

A mixed model with the young speakers' V:W duration ratio in words only with sonorant consonants as the dependent variable, and Length (two levels: tense vs. lax), and Region (two levels: Bavarian vs. Saxon) as fixed factors and Subject and Vowel as random factors showed a significant main effect for Length ($F[1,60] = 1473.5$, $p < 0.001$) and a significant interaction effect for Length x Region ($F[1,60] = 16.8$, $p < 0.001$). The significant main effect for Length confirmed that young Saxons and young Bavarians maintained the tense/lax contrast before sonorants. The non-significant effect for Region implies that the proportional vowel duration was about the same for both dialect groups, although the significant interaction effect for Length x Region indicates that young Bavarian as opposed to young Saxon speakers maintained the tense/lax contrast to a different extent as can be seen from Figure 4.14. An ANOVA with the mean V:W duration ratio difference (which was derived following the methodology described above for the old age groups) as the dependent variable and Region (two levels: Bavarian vs. Saxon) as the independent variable revealed no significant main effect for Region ($F(1,17) = 0.9$, $p = 0.3$). This result suggests that the contrast maintenance is about the same for young Bavarians and young Saxons.

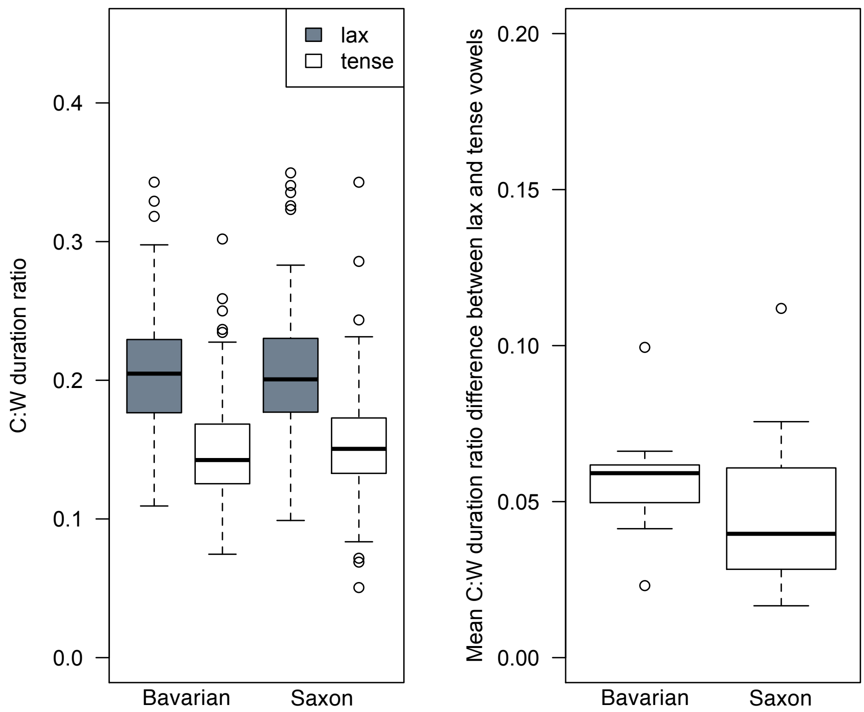


Figure 4.15: C:W duration ratios for lax (grey) and tense (white) vowel words containing sonorants (left) and mean C:W duration ratio differences between tense and lax vowel words containing sonorants (right) separately for young Bavarian and young Saxon speakers.

A mixed model with the young speakers' C:W duration ratio in words only with sonorant consonants as the dependent variable, and Length (two levels: tense vs. lax), and Region (two levels: Bavarian vs. Saxon) as fixed factors and Subject and Vowel as random factors revealed a significant main effect for Length ($F[1,60] = 521.8$, $p < 0.001$). All other effects were not significant. The significant main effect for Length suggests that the proportional sonorant duration is longer after lax as opposed to tense vowels (cf. Figure 4.15). However, the non-significant effects for Region and Length x Region imply that the lengthening effect of sonorants following lax vowels was about the same for young Bavarian and young Saxon speakers. An ANOVA with the mean C:W duration ratio difference (which was derived following the methodology described above for the old age groups) as the dependent variable and Region (two levels: Bavarian vs. Saxon) as the independent variable revealed again no significant main effect for Region ($F(1,17) = 0.6$, $p = 0.4$). This result confirms that young Bavarians and young Saxons preserved the tense/lax contrast in vowels preceding sonorants to the same extent.

To summarize, only young Bavarians and young Saxons maintained the tense/lax contrast before sonorants to the same extent; old Bavarians, on the other hand, showed a greater difference in the proportional duration of lax and tense vowels than old Saxons. However, this difference in vowel duration was accompanied by significantly longer sonorant durations for old Bavarian than for old Saxon speakers and there were again no differences with respect to sonorant durations between the two young age groups. These findings suggest that the vowel length contrast was indeed maintained by means of an additional consonant length contrast, but only in the productions of old Bavarians: short vowels were followed by long consonants and long vowels preceded short consonants.

4.3.7 Postvocalic lengthening of obstruents

In 4.3.1 it was shown that young and old, Bavarian and Saxon speakers distinguish tense and lax vowels in terms of vowel duration and the age effect reported in 4.3.1 and discussed in 4.3.5.1 indicated – commensurate with Figure 4.2 – that the preservation of the vowel length opposition in terms of the mean V:W duration ratio was greater for old Bavarian speakers than for all other speaker groups. Furthermore, the analysis in 4.3.4 revealed that old Bavarian speakers maintained the vowel length contrast by means of the V:C duration ratio to a greater extent. In the last analysis we found postvocalic sonorant lengthening when the preceding vowel was lax in the production of all speaker groups but the effect was more marked in old Bavarian speakers as opposed to all other speaker groups. Thus, in our last analysis we return to the lax/tense contrast before obstruents and examine the role of stop lengthening. That is we tested whether the contrast maintenance in terms of V:C duration ratio came about because of a greater variation in vowel or consonant length (or both). The same three minimal pairs that were analysed in 4.3.1 were included in this analysis.

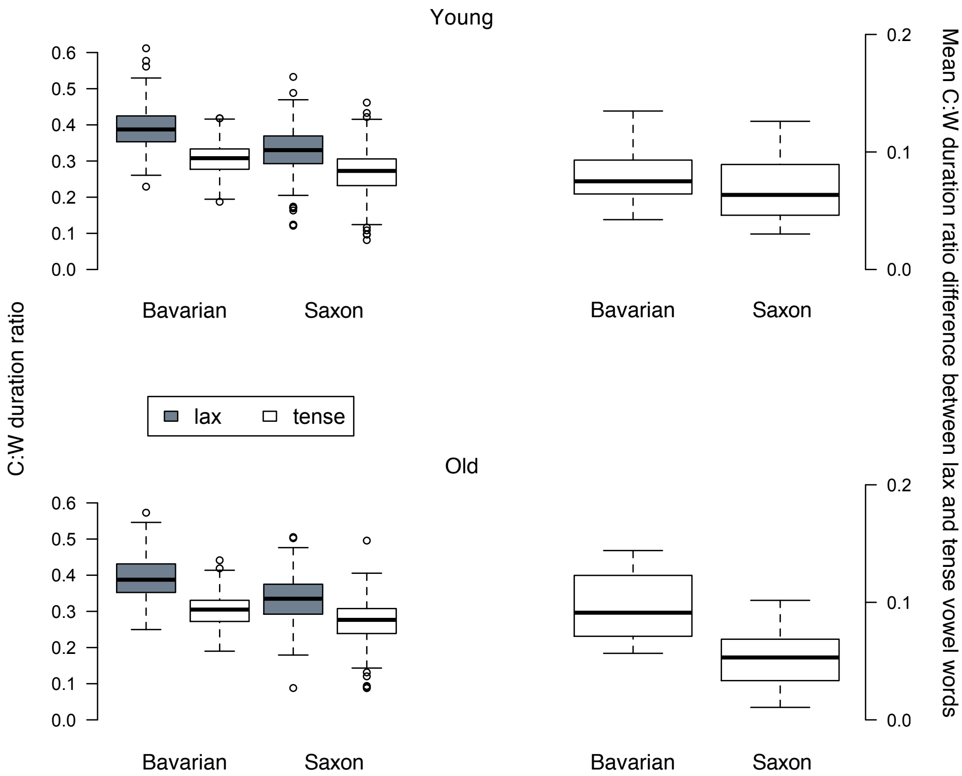


Figure 4.16: C:W duration ratios (left) for lax (grey) and tense (white) vowel words containing obstruents and mean C:W duration differences between tense and lax vowel words containing obstruents (right) separately for young (above) and old (below), Bavarian and Saxon speakers.

A mixed model with the C:W duration ratio as the dependent variable, and Length (two levels: tense vs. lax), Region (two levels: Bavarian vs. Saxon) and Age (two levels: young vs. old) as fixed factors as well as Subject and Vowel as fixed factors showed significant main effects for Length ($F[1,60] = 2111.1$, $p < 0.001$) and Region ($F[1,60] = 20.3$, $p < 0.001$) as well as significant interaction effects of Length x Age ($F[1,60] = 5.4$, $p < 0.05$) and Length x Region ($F[1,60] = 71.5$, $p < 0.001$). As can be seen in the left column of Figure 4.16 and as the significant effect for Length indicates, underlying fortis consonants have proportional longer durations after lax vowels than after tense vowels. This finding suggests that the vowel's underlying tensity exerts a lengthening effect on the postvocalic stop. This effect, however, is different for Saxons and Bavarians as the significant effects for Region and Length x Region show and to some extent for old and young speakers as the interaction effect for Length x Age demonstrates.

An ANOVA with the mean C:W duration ratio difference (which was derived by the same method as described above in 4.3.1) as the dependent variable and Region (two levels: Saxon vs. Bavarian) and Age (two levels: young vs. old) as the independent variables revealed a significant main effect for Region ($F[1,36] = 22.7$, $p < 0.001$) and a significant interaction for Region x Age ($F[1,36] = 8.7$, $p < 0.01$). This significant effect for Region indicates – c commensurate with the right column of Figure 4.16 – that the stop duration difference between tense and lax vowel words was greater for Bavarian than Saxon speakers, but this dialect-dependent divergence was much more marked for old than for young speakers as the interaction effect for Region x Age shows. That is, the tense/lax contrast in terms of postvocalic stop lengthening is maintained to the greatest extent by old Bavarian speakers and to the least extent by old Saxon speakers. By contrast, the ANOVA with the mean V:W duration ratio difference as the dependent variable revealed no significant interaction effect for Region x Age (see 4.3.1 above).

4.4 General discussion

The results of the experiments in chapter III showed that Bavarian speakers preserve the voicing contrast more than Saxon speakers. One possible interpretation of this finding was that the acoustic cue V:C duration ratio is used to a greater extent in Central Bavarian than in Upper Saxon to distinguish voiced from voiceless stops by means of an additional lengthening and shortening of the preceding vowel, respectively. In fact, many researchers consider vowel length to be complementarily distributed (Rowley, 1990b; Wiesinger, 1990): when a stop is voiceless then the preceding vowel is short, but only long vowels precede voiced stops. The motivation for the present analysis then was to determine the extent to which Bavarian and Saxon speakers maintain the vowel length contrast in Standard German when they produce a regional variety of Standard German. A second aim was to establish whether old speakers differ to a greater extent on that parameter based on the assumption that older speakers show more tendencies to preserve dialectal features in Standard German than young speakers of the same speech community.

There are two possible realizations of the vowel length contrast before a fortis obstruent according to the CBL: Bavarian speakers either neutralize the vowel length contrast towards lax vowels and produce a fortis stop or they preserve the contrast but additionally lenite the stop when it occurs after a tense vowel. Given that speakers produced a regional variety of Standard German, we hypothesized that Bavarian speakers maintain the tense/lax contrast before obstruents to a greater extent than Saxon speakers. Our results support this hypothesis. The tense/lax contrast in /beːtən/ and /bɛtən/, /biːten/ and /bɪten/ as well as /hyːte/ and /hʏte/ was maintained by means of different V:W duration ratios as well as formant values in both dialects. The dialect and age groups, however, differed only with respect to the duration parameter and not in quality.

As far as the tendency for /t/-lenition in the production of Saxon speakers is concerned, hypothesis H2 predicted that Saxon speakers tend to lenite the underlying fortis stop after lax vowels to a greater extent than Bavarian speakers. Our results support hypothesis H2. Saxon speakers showed proportionally longer vowel and proportionally shorter stop durations than Bavarian speakers. This greater V:C duration ratio indicates more /t/-lenition after lax vowels in Saxon vs. Bavarian. In an additional analysis, we analyzed the degree of /t/-lenition after tense vowels. In this context too, Saxon speakers lenited the fortis stop to a greater extent than Bavarian speakers (cf. also the results in 3.3.4). The difference between Saxon and Bavarian speakers, however, is much smaller than in the context of lax vowels. There are two possible interpretations to these divergent, context-dependent results: (1) Saxon speakers differ in their degree of /t/-lenition after tense vs. lax vowels or (2) Bavarian speakers maintain the vowel length contrast partly by using V:C duration ratios that indicate either lenis or fortis stops. We will come back to these two interpretations below.

With respect to the tense/lax contrast before fortis stops, hypothesis H3 predicted that the difference between the underlying fortis stops following lax and tense vowels is much greater for Bavarian than for Saxon speakers (again on the assumption that Bavarian speakers maintain the tense/lax contrast when they speak a variety of Standard German). That is, we predicted that Bavarian speakers use a greater range of the V:C duration ratio that is also indicative of a fortis/lenis contrast. Saxon speakers, on the other hand, were predicted to show V:C duration ratio characteristics of a more lenis-like stop. Our results partly support hypothesis H3. Recall that Kohler (1977, 1979) and Braunschweiler (1997) have found V:C duration ratios above 0.7 and below 0.6 to be clear indicators for lenis and fortis stops after tense vowels, respectively. On the other hand, lax vowels preceding fortis and lenis stops were shown to have mean V:C duration ratios of 0.39 and 0.56, respectively (Braunschweiler, 1997). Thus both dialect groups produced fortis stops in terms of the V:C duration ratio, but the proportional vowel duration difference between tense and lax vowels preceding underlying fortis stops was greater for Bavarians than Saxons.

As far as age-dependent differences are concerned, Hypothesis H4 predicted that the Bavarian-Saxon differences in hypothesis H1 – H3 are more pronounced for old than for young (on the assumption that there is a greater tendency that young speakers of both varieties tend towards a standard pronunciation). Our results support the hypothesis partly. Young and old speakers differ with respect to the tense/lax contrast on both acoustic parameters vowel quantity and vowel quality, but they did not differ with respect to /t/-lenition. The durational differences between age groups depended also on region. Young Bavarian speakers maintained the contrast to a lesser extent than old Bavarian speakers. The mean V:C duration ratio difference between tense and lax vowels before fortis stops was about the same for young Bavarians as well as young and old Saxons. These results support the interpretation that the maintenance of the vowel-length contrast was context-dependent for Bavarian speakers. Since there were no age dependent differences in /t/-lenition (neither after tense nor lax vowels) between young and old Bavarians we interpret our findings as follows: the vowel length contrast in terms of the mean V:C duration ratio difference between tense and lax vowels is preserved to a greater extent by old Bavarian speakers than by all other speaker groups. Their V:C duration ratios indicated clear fortis stops in lax vowel words, but implied a more lenis-like production in tense vowel words. The degree of /t/-lenition, however, also depended on the underlying tensity of the preceding vowel. This finding will be discussed below.

The last hypothesis H5 predicted that before sonorant consonants, Bavarian speakers maintain the tense/lax contrast to a lesser extent than Saxons, because (1) according to the literature vowels are lengthened before voiced segments and (2) the greater contrast maintenance between tense and lax vowels preceding stops found especially in the production of old Bavarians was hypothesized to be a consequence of a Bavarian correlation of vowel length and stop voicing. Our findings did not support this prediction with respect to the proportional vowel duration. Old Bavarians preserved the vowel length contrast also before sonorants to a greater extent than old Saxons. However, an additional analysis revealed that the proportional sonorant duration was significantly longer after lax vowels and, more importantly, that the difference in sonorant duration after lax vs. tense vowels was significantly greater for old Bavarian vs. young Bavarian speakers. This finding supports our hypothesis H3 that the greater contrast preservation between tense and lax vowels is a consequence of a Bavarian co-dependency of vowel length and stop voicing.

Irrespective of the dialectal background, sonorant and obstruent consonants were lengthened after lax vowels; that is, there was a progressive consonant-lengthening effect in our data. This result is in line with Braunschweiler's (1997) finding that the duration (here closure duration) of underlying fortis stops were longer after lax vowels than after tense vowels; that is, closure duration varied as a function of the preceding underlying vowel tensity. The extent to which consonants were lengthened in our data was dialect and age dependent. The significant difference between consonant lengthening by old Bavarian speakers and all other speaker groups may be interpreted as that consonant lengthening is phonetic in Saxon and young Bavarian speakers, but phonological in old Bavarian speakers.

We cannot and we will not resolve the controversial issue of the phonological status of vowel length in Central Bavarian (and neither was this controversy part of our hypotheses), simply because our laboratory data of regional varieties of Standard German does not allow for this. However, our findings for old Bavarians provide some evidence in favour of Bannert's (1976) prosodic account of the correlation between vowel length and stop voicing. According to Bannert (1976), the length of vowels and following consonants are complementarily distributed. His account contradicts the consonantal approach (Kufner, 1964; Hinderling, 1980; Rowley, 1990b; Wiesinger, 1990) which assumes a phonemic fortis/lenis contrast and allophonic vowel length, the latter being complementary distributed. In Standard German, where the vowel length contrast is preserved irrespective of the underlying voicing (or length for that matter) of the following stop, the vowel duration is decisive. All speaker groups in our data did preserve the vowel length contrast by means of vowel duration. Additionally, old Bavarian speakers lengthened the following consonant. However, they did so in order to maintain the vowel length contrast and not a voicing contrast. This conclusion can be drawn from two points. For one thing, the Standard German words only differed in vowel length and not in underlying voicing. And for another thing, old Bavarians showed considerable consonant lengthening in words containing sonorants in intervocalic position. Note that lateral and nasal consonants occur only in prevocalic position in the broad Central Bavarian dialect (cf. 1.4.2); in postvocalic position, however, nasals are deleted or the preceding vowel becomes nasalized (e.g. Standard German /ʃtaɪn/, 'stone', Central Bavarian /ʃtoã/) and laterals are vocalized (e.g. Standard German /hals/, 'neck', Central Bavarian /hɔɪz/). In order to produce the Standard German vowel plus sonorant sequence and to maintain the phonological vowel length contrast, old Bavarian speakers apparently use the same strategy as for the voicing distinction in terms of a post-vocalic consonant lengthening, although there is no voicing contrast in sonorants. Thus, our data shows that vowel and consonant length correlate to a greater extent in the production of old Bavarians than in all other speaker groups and this correlation does not necessarily depend on obstruent voicing as our findings for the vowel length contrast before sonorants have shown. Therefore, our findings support Bannert's account.

Previous research (e.g. Spiekermann, 2000) showed that, in Central Bavarian, loosely and abruptly cut syllables differ primarily in the post-vocalic consonant's duration but not with respect to vowel quality and quantity. Spiekermanns' analysis, however, did not factor in the manner of articulation and phonological voicing of the following consonant.[10] Our results support his finding of the post-vocalic lengthening effect that the underlying vowel length exerts on voiced vs. voiceless obstruents as well as in voiced sonorants. However, our results did not support his finding of only marginal vowel duration (and vowel quality) differences between underlying tense vs. lax vowels. In the present study, old Bavarian speakers maintain the vowel length opposition by means of both vowel duration and stop duration (as well as vowel quality). One reason for this divergence could be that Spiekermann (2000) analyzed spontaneous dialectal speech (interviews from the *Lautbibliothek der deutschen Mundarten*) while we included only laboratory standard speech in our analysis (although as spoken by dialect speakers). On the other hand, the divergence could also be due to a changing dialect feature. Most of the old speakers in our analysis belong to a younger generation compared to most of the speakers of the interviews in the *Lautbibliothek der deutschen Mundarten* (who were adults at the time of the recordings during the 1950s). Assuming that there is an increase in standard pronunciation (i.e. a decrease in dialect features) in younger generations (e.g. Trudgill, 1986; Wagener, 2002; Lameli, 2004) we would then predict that the next generation, i.e. the young speakers in our analysis, would differ even more with regard to both the extent and also the acoustic cues with which the vowel length contrast is preserved. Indeed, there were no differences between the two young speaker groups in the degree of contrast maintenance of

[10] Although he found that the intensity curve, i.e. his acoustic correlate of syllable cut, differed depending on the underlying voicing of a stop (Spiekermann, 2000: 55)

Standard German phonemic vowel length neither in vowel nor sonorant duration. Young as opposed to old Bavarian speakers, maintain the vowel length contrast by means of vowel duration only. This finding suggests, that the dialect feature of phonemic correlation of vowel length and stop voicing is lost in younger Bavarian generations.

Already Bethge (1963: 207) has claimed that dialectologists should use quantitative methods to further explore impressionistic descriptions of dialect levelling that young speakers avoid extra long vowels that are characteristic for German dialects and to test whether there are regional differences. He measured the durations of tense vowels and compared them to the durations of lax vowels and found that young speakers had smaller tense/lax quotients than old speakers. That is, Bethge (1963) found a diminished vowel length contrast in the productions of young as opposed to old speakers. However, he averaged across speakers that came from all parts of Germany. Our data confirm the diminished contrast maintenance in young vs. old speakers, but only for speakers of a dialect that has a phonemic correlation of vowel length and voicing.

The finding of an age-dependent difference in contrast maintenance within the Bavarian speaker group is the most important result of this analysis. Internal or external (or both) factors may have been the reason for the decrease in contrast preservation in young speakers. Recall from the introductory chapter I that lenition of intervocalic fortis obstruents is among the frequently attested diachronic sound changes reported in the world's languages (e.g. Kohler, 1984; Crowley, 1998). Moreover, neutralization of the intervocalic voicing contrast is a feature that characterizes a number of German varieties, including Upper Saxon. Internal factors such as synchronous consonant gradation in spontaneous speech are often considered to trigger diachronic lenition of fortis stops (Kohler, 1984; Bauer, 2008). Intervocalic lenition in younger speakers, however, would lead to a less standard-like pronunciation. We also considered external factors as potential triggers of sound change. It was argued that external factors such as the prestige of a dialect might lead to a less regional, more standard-like pronunciation. Although Bavarian is not a low-prestige dialect like Saxon, we assumed that young speakers of both dialects would show some dialect levelling towards the standard language (cf. Lameli, 2004; Wagener, 2002).

Let us first consider the results for Saxon speakers. Our results indicated that Saxon speakers realized Standard German fortis stops as fortis although they showed a greater tendency for /t/-lenition than Bavarian speakers. This finding is in line with the results presented in chapter III showing that Saxons preserved the intervocalic voicing contrast but to a lesser extent than Bavarians. However, we found no age-dependent differences for /t/-lenition in the productions of Saxon speakers (neither in this chapter nor in chapter III). That is, both age groups showed the same degree of contrast preservation as well as the same tendency for stop lenition and thus the same degree of dialect levelling.

Now let us shift the attention to the Bavarian results. Recall that, in Central Bavarian, the vowel length contrast may be neutralized as in the realization [bɛtn] for Standard German /beːtən/ (König & Renn, 2005: 67; Wiesinger, 1990: 449), but not necessarily. Alternatively, the vowel length contrast can be preserved by an additional lengthening of the consonant as in [beː(d)n] (König & Renn, 2005: 67). However, in the latter case, the long vowel is – according to the consonantal account of Central Bavarian lenition (e.g. Kufner, 1964; Hinerdeling, 1980) – a phonetic consequence of the lenis stop. Just like intervocalic lenition, vowel length neutralization is an attested sound change (Bauer & Warren, 2004), in particular in domain-final position (see Kubozono, 2002; Myers & Hansen, 2005), caused by internal factors. But neutralization of the vowel length contrast would again result in a more dialectal pronunciation. A more standard-like pronunciation, however, implies that the vowel length contrast is maintained without producing an additional fortis/lenis contrast. In our data, Bavarian speakers maintained the vowel length contrast, but old speakers preserved it to a greater extent (by means of both the V:W and the V:C duration ratio). On the one hand, the diminished vowel length contrast may have been caused by internal factors. On the other hand, external factors may have prevented speakers from vowel length neutralization. In any case, the finding suggests that there is a sound change in progress that leads to a more standard-like pronunciation, but we cannot state whether external or internal factors are the trigger for this change in progress.[11]

The result for age-group dependent differences is in line with previous findings showing less dialect features in younger generations than older generations (Lameli, 2004; Wagener, 2002). To our knowledge, this is the first acoustic study that presents results for a dialect change in terms of a gradual decrease in contrast preservation. We have discussed this change in progress with respect to internal and external factors that may have caused such a change. However, it remains an open question which of the factors triggered this change. Further experiments, such as perception experiments, are necessary to test the (phonetic) origin of this change (cf. 1.5). Models such as Lindblom's (1990) H&H theory account for such a gradual language change. The H&H theory predicts that a speaker deliberatively chooses from a continuum between hypoarticulated and hyperarticulated speech depending on the listener's needs. Although there is a correlation between vowel length and stop voicing in Standard German – as the lexical frequency distribution of vowel tensity and stop voicing in Standard German (cf. chapters II and III) as well as the complimentary lengthening effect (that underlying voiced/voiceless stops and tense/lax vowels exert on each other) show – this correlation is by no means phonemic in Standard German. The vowel length opposition is perceptible by means of vowel duration only (Weiss, 1976), i.e. without an additional difference in consonant duration. Schaeffler (2005) in his investigation of vowel quantity in Swedish dialects has argued that "the additional change in consonant duration observed with complementary quantity could aid quantity perception but it is not as important as a 'ratio account' of complementary quantity would suggest[;] consequently, there are no obvious or perceptual reasons for complimentary quantity" (p. 127). Thus, following Lindblom (1990), speakers may loose a contrastive feature, such as consonant length, when it does not serve any function such as quantity distinction.

[11] A different interpretation of the findings – which is independent from the factors and which has been suggested by Phil Hoole – is that there is no decrease in contrast preservation, but instead a change from a one-contrast system in which complementary length is the decisive feature (long vowel + short stop or short vowel + long stop) to a two-contrast system in which vowel length and voicing (or length for that matter) can be freely combined.

To conclude, speakers from different age and dialect groups maintain the Standard German vowel length opposition, however, they differ in their extent of contrast preservation. Once again, it was shown that speakers do not differ in terms of categorical contrast neutralization or preservation. Instead their results show gradual differences that depend on regional background and age group. These gradual differences between age-groups provide insight into a potential sound change in progress. Models of speech communication that incorporate both speakers and listeners needs in communicational situations explain these gradient differences better than traditional phonological accounts that treat neutralization or complementary distribution of vowel and consonant length as structural changes.

Summary and Conclusion

Chapter V

Standard German differentiates between voiced and voiceless obstruents as well as long and short vowels. The latter phonemic opposition is commonly referred to as the tense/lax contrast and the former as the fortis/lenis contrast, particularly since stops differ to a greater extent in articulatory force and timing than in vocal-fold vibration (Kohler, 1984; Barbour & Stevenson, 1990: 93 – 94).[1] Thus, Standard German distinguishes minimal pairs like /baːdən/ and /baːtən/ or /biːtən/ and /bɪtən/. The duration ratio between the vowel and the following stop is one of the most important cues to the fortis/lenis contrast (e.g. Kohler, 1977) and the vowel duration is (among others) a crucial parameter in the tense/lax opposition (e.g. Jessen, 1993). However, the voicing opposition is neutralized towards fortis in domain-final position in Standard German (and most varieties of German) and towards lenis in intervocalic position in some varieties of German, e.g. Upper Saxon (Bergmann, 1990). In Central Bavarian the voicing contrast is preserved in terms of a complementary distribution of vowel length: long vowels precede lenis stops and short vowels precede fortis stops (Wiesinger, 1990). Within phonological theory, neutralizations of phonemic contrasts were long considered to cause complete changes at the surface structure. Yet phonetic studies have presented evidence that domain-final underlying voiced obstruents were only gradually devoiced in the production of Standard German and thus the phonemic voicing contrast only incompletely neutralized (e.g. Port & O'Dell, 1985; Port & Crawford, 1989; Charles-Luce, 1985, 1993).

The aim of this thesis was to (1) investigate the extent to which these phonological contrasts are incompletely neutralized or maintained in perception and production and (2) to examine the effect that (extra-)linguistic factors potentially exert on the degree of contrast

[1] In this thesis, we used the voice/voiceless distinction for domain-final stops (commensurate with the literature on incomplete neutralization) and the fortis/lenis distinction for intervocalic stops (commensurate with the literature on German dialectology).

maintenance. Therefore we extended our research to (other) contrasts in other positions that have been treated as complete neutralizations of phonemic oppositions in the literature. We also included conditions such as phonological frequency, regional background and age group in our analysis. The latter factor was included as a predictor of contrast neutralization because we expected phonological processes that cause only gradual differences at the surface structure to be a potential source for language change.

In chapter II the focus was on the (in)complete perception of the domain-final voicing contrast in Standard German. Chapter III addressed the question whether and to what extent old and young, Saxon and Bavarian speakers neutralized the voicing contrast in intervocalic position. The vowel length contrast was investigated in chapter IV, taking into account the interplay of vowel and stop length. In this chapter we will briefly summarize the experiments described in chapters II – IV and discuss each of the factors we assumed to influence the degree of contrast maintenance.

5.1 Incomplete neutralization and maintenance of phonemic contrasts

Incomplete neutralization of phonological contrasts was the overall issue of all three studies described in this thesis. In chapter II we addressed the question of incomplete neutralization of the domain-final voicing contrast in perception. In two two-alternative forced choice tasks, 19 listeners were presented with resynthesized stimuli and asked to identify the words that differed only in the underlying voicing of the stop in syllable-final position. Listeners labelled more stimuli as voiced than it was expected in syllable-final position. This result may have come about because the stimuli were derived from underlying voiced stimuli. But more importantly, the proportion of voiceless responses as a function of decreasing V:C duration ratio increased slightly, but gradually from more voiced responses to more voiceless responses. There was no indication for the perception of two distinct voicing categories as there were no abrupt changes in the psychometric curves. Thus, the results in chapter II supported our hypothesis that listeners only incompletely neutralize the domain-final voicing contrast in perception. This finding is in line with results from other perception studies showing that listeners make use of fine phonetic detail in sound or word recognition (Hawkins, 2003; Hawkins & Nguyen, 2003, 2004).

Chapters III and IV dealt with phonemic contrasts that are maintained in Standard German but are neutralized in broad German dialects. The focus of chapter III was on the intervocalic voicing contrast which is neutralized towards lenis in Upper Saxon (and other varieties of German), but preserved in Central Bavarian. From the speech corpus that contained isolated minimal pair words read by 42 young and old, Saxon and Bavarian speakers, the minimal pairs /boːdən – boːtən/ and /oːbɐ - oːpɐ/ were chosen for analysis. All speaker groups maintained the contrast in terms of the V:C duration ratio (and other duration parameters) but they differed in the extent of contrast maintenance depending on (extra-)linguistic factors (see 5.2. – 5.4 below). A similar result was obtained in the analysis of the Standard German vowel length contrast described in chapter IV. All speaker groups preserved the tense/lax contrast in minimal pairs such as /biːtən – bɪtən/ or /høːlə - hœlə/ by means of vowel duration (and other acoustic parameters) but again the degree to which the phonemic opposition is maintained was different for the age and dialect groups (see 5.3 and 5.4 below)

In summary, the acoustic parameter V:C duration ratio was incompletely neutralized in the perception of the domain-final voicing contrasts. The intervocalic voicing as well as the vowel length contrast were maintained in production by means of various duration parameters. Nevertheless, the neutralization degree in terms of the mean duration ratio difference varied depending on (extra-) linguistic factors and thus, we may refer to this process as incomplete maintenance of phonemic contrasts. We will recap the findings for each of the (extra-)linguistic factors in the following paragraphs.

5.2 Linguistic factors

The degree of contrast maintenance was shown to depend on linguistic factors such as phonetic environment and sentence position (Charles-Luce, 1985) as well as on semantic information (Charles-Luce, 1993). We extended our research on incomplete neutralization to the investigation of the linguistic factors phonological frequency and the potential for re-syllabification. These factors were included as predictors for incomplete neutralization in the perception of the domain-final voicing contrast that was investigated by means of two perception experiments reported in chapter II. We used resynthesized trochaic words as stimuli. The accented vowel differed in underlying tensity and the stop in syllable final position was either velar or alveolar. The onset consonant(s) of the second syllable was either /l/ or /ʃt/. For each of the six words we derived seven stimuli which differed with respect to the V:C duration ratio. 19 Northern Standard German listeners participated in the two experiments.

In the first perception experiment reported in chapter II, we tested whether the phonological frequency with which vowel tensity and stop voicing co-occur in the lexicon of Standard German affect the identification performance of listeners. Only words in which the second syllable's onset consonant was /l/ were tested in this experiment. Lax vowels are almost always followed by voiceless stops and tense vowels are frequently followed by voiced stops. Only in sequences of a tense vowel plus alveolar stop the stop voicing is equally distributed.

The hypothesis was that listeners would show a bias towards one voicing category in their responses to sequences that frequently co-occur with one but not the other voicing category. Listeners indeed labelled more stimuli from the V:C duration ratio continua as voiced when the vowel was tense than lax. Thus, our results supported this hypothesis. Furthermore, we predicted that the underlying voicing was better identified in sequences in which phonological voicing is equally distributed in the lexicon. The results revealed a complex relationship between acoustic cues and phonological frequency distributions and both contributed to the voicing perception. On the one hand, the perceptibility of the voicing contrast was greater in tense vowel plus alveolar stop sequences than in lax vowel plus alveolar stop sequences, probably because of the lexical bias towards voiceless stops in the latter but not in the former sequence. On the other hand, listeners perceived the voicing contrast to a greater extent in tense plus velar stop sequences than in tense plus alveolar sequences, although the voicing contrast is equally distributed in the latter but not in the former.

One explanation for this contradictory latter result could be the potential for resyllabifying the final velar stop with the /l/-onset of the second syllable. This was examined in the second perception experiment reported in chapter II, in which only words with a tense vowel in the first syllable were tested. The potential for re-syllabification indeed improved (in terms of a steeper psychometirc curve) listeners' identification of the underlying voicing in velar stops as opposed to other non-resyllabifiable stops. However, the results also revealed an increase in the perceptibility of the voicing contrast when the stop was followed by a sonorant, irrespective of the potential for resyllabification. Thus, we found some evidence for the idea that the voicing contrast is more perceptible in resyllabifyable than in non-resyllabifiable sequences, which is in line with the *licensing-by-prosody* approach (Itô, 1986, 1989; Goldsmith, 1990; Rubach, 1990). Nevertheless, our results also support the *licensing-by-cue* account (Steriade, 1997, 1999, 2000) which predicts that the acoustic properties of the neutralizing environment are crucial for an increase in contrast maintenance. The finding that the sonorant environment in non-resyllabifiable sequences (e.g. /dl/) facilitates the perceptibility of the voicing contrast is not compatible with the *licensing-by-prosody* account. Our findings were best explained by a usage-based model of speech perception such as the exemplar theory (Pierrehumbert, 2002, 2003) in which linguistic factors such as phonological frequency are incorporated and which also accounts for the potential of resyllabification. We concluded that the perception of the underlying voicing contrast in final stops depends on a complex interaction between acoustic information and phonological knowledge based on linguistic experience and the lexicon. After all, the voicing perceptibility was enhanced in sequences in which a combination of factors favouring lenis perception contributed to the contrast maintenance.

Phonological frequency was also included as a predictor of incomplete maintenance of the intervocalic voicing contrast in the production of two regional varieties of German, which was investigated in chapter III. The prediction was that speakers of both dialects tend to neutralize labial stops to a greater extent than alveolar stops because labial stops are almost always lenis when following tense vowels, but no such prediction can be made for alveolar stops. In an acoustic analysis, we measured various duration parameters in the minimal pairs *Boden*-*boten* and *Ober*-*Oper* and found that the maintenance of the voicing contrast in terms of the V:C duration ratio was indeed diminished for labial stops as opposed to alveolar stops. However, the V:C duration ratio of labial stops was more fortis-like, although we had predicted a more lenis-like realization commensurate with the lexical bias towards lenis stops after tense vowels. Our results are partly explained by an exemplar model (Pierrehumbert, 2002, 2003), which would predict a diminished contrast for less frequent combinations of vowel tensity and stop voicing. The more fortis-like realization of labial as opposed to alveolar stops in the production of all speaker groups, however, was not explained by exemplar theory, because the speaker/listener should have only very few stored exemplars of tense plus labial fortis stop sequences (since this combination is rare). This outcome, however, was compatible with Lindblom's (1990) H&H theory, which predicts hyperarticulation in rare combinations in order to enhance the perceptibility for the listener. The finding for longer labial stop than alveolar stop durations is in line with Kohler's (1977) results for German as well as Stathopoulos and Weismer's (1983) and Byrd's (1993) outcomes for English showing intrinsic stop duration differences depending on the place of articulation. One reason for this difference may lie in the production of plosives: the cavity behind the closure is larger for labial than for alveolar stops and consequently, it takes longer during a labial than an alveolar closure phase to reach a high intraoral air pressure which leads to an audible release (cf. Maddieson, 1997).

5.3 Regional variation

Chapters III and IV addressed the question whether speakers of regional varieties neutralized or maintained contrasts that are phonemic in Standard German (but may not be in the broad dialects) and whether they differed in the degree of contrast maintenance. Speakers from two dialect backgrounds, Upper Saxon and Central Bavarian, were asked to read Standard German words – thus the speakers produced a regional variety of Standard German.

In chapter III we investigated the realization of the intervocalic voicing contrast, which is neutralized in the broad dialect Upper Saxon and maintained in Bavarian (by means of variation in vowel length). We predicted that Saxon and Bavarian speakers maintain the contrast when they produce a regional variety of Standard German, but that Saxon speakers should show a greater tendency for neutralization towards the lenis category commensurate with the intervocalic stop lenition in the broad dialect of Upper Saxon. Our acoustic analysis of two minimal pairs differing only in the underlying voicing of the intervocalic stop (e.g. /boːtən/ and /boːdən/) revealed that the voicing contrast was maintained by various duration parameters, including the V:C duration ratio. Nevertheless, Saxon speakers maintained the voicing contrast to a lesser extent.

The interplay of vowel and stop length and their contribution to the vowel length distinction was the objective of the acoustic analysis in chapter IV. In Central Bavarian, vowel length and stop voicing have complementary distribution (long vowel plus lenis stop versus short vowel plus fortis stop) and many researchers consider vowel length to be allophonic and stop voicing phonemic (Wiesinger, 1990; Rowley, 1990b). We measured vowel and stop duration in words that differed in underlying tensity of the accented vowel. The stop was either a fortis stop or a sonorant (e.g. /biːtən/ vs. /bɪtən/ or høːlə/ vs. /hœlə/). We predicted that Saxon speakers would differ in vowel duration and show a greater tendency for stop lenition. Two outcomes were possible with respect to Bavarian vowel plus stop realizations: they could either neutralize the vowel length contrast in order to produce a fortis stop or they could enhance the tense/lax contrast by producing a more lenis-like stop after underlying tense vowels. Furthermore, we predicted that Bavarian speakers lengthen Standard German short vowels when they precede voiced segments such as sonorants. The results showed that Bavarian speakers maintained the vowel length contrast to a greater extent than Saxon speakers in obstruent and sonorant contexts. The greater contrast preservation was partly achieved by an additional lengthening of the consonant, irrespective of the manner of articulation. This result was interpreted as further evidence for Bannert's (1976) account which proposes complementary quantity of either the vowel or the consonant. That is, the vowel is not only lengthened or shortened in order to maintain the underlying voicing distinction, but the consonant is likewise lengthened in order to enhance a phonemic length contrast in the vowels. Therefore, vowel length is not neutralized in our data and cannot be regarded allophonic. The lengthening effect in consonants depending on the underlying tensity of the preceding vowel was found for all speaker groups, but it was greater for the Bavarian group.

Speakers from different dialect areas do not differ as much in their regional realizations of Standard German contrasts as descriptions of German dialects might suggest. Present day regional varieties of German show phonetic and phonological characteristics of both the standard language and the broad dialect. Dialect speakers produce contrasts which may be neutralized in their broad dialects when asked to speak a regional variety of Standard German, however, they may do it to a lesser extent. This may be explained by the observation that speakers/listeners are exposed to various standard and regional varieties and that this diverse linguistic experience contributes to the variability in speakers' realizations (Pierrehumbert, 2002, 2003). The results demonstrate that phonological accounts that consider neutralization to be categorical cannot model regional variation adequately. A model of present day regional variation in German should incorporate gradual differences between regional realizations, and not only count the number of realized dialect features: this is because present day regional varieties are very likely to exhibit many Standard German characteristics, but they certainly differ from each other and the standard language in fine phonetic detail.

5.4 Dialect levelling and sound change

Chapters III and IV also investigated the idea that decreasing or increasing contrast maintenance provide insight into potential sound changes in progress. Neutralization of the intervocalic voicing contrast in the form of lenition is a frequently reported diachronic sound change that may emerge from a synchronous process of decreasing articulatory strength in consonants in spontaneous speech (cf. Kohler, 1984). According to such internal factors, the predicted direction for the intervocalic voicing contrast in sound change would be complete neutralization. External factors, however, such as the prestige of a dialect may lead to sound changes that take place in the opposite direction (Torgensen & Kerswill, 2004). Our prediction was that speakers from neutralizing varieties produce Standard German contrasts when they speak a regional variety of Standard German. For example, Saxon is a low-prestige dialect and Saxon speakers may avoid dialect features that they are aware of, such as the lenition of fortis stops. Therefore, this thesis contained two apparent time analyses of the voicing and vowel length neutralization. Saxon and Bavarian speakers were assigned to one of two age groups: speakers above the age of 50 belonged to the Old group and speakers younger than 50 were in the Young group. Age group was included in the statistical analyses as a predictor of incomplete contrast maintenance.

In the analysis of the intervocalic voicing opposition reported in chapter III, we expected old Saxon speakers to neutralize the voicing contrast to a greater extent than young Saxons based on the assumption that young speakers tend to a more standard pronunciation than old speakers. However, the results revealed no significant differences between the age groups. That is, old and young, Bavarian and Saxon speakers maintained the contrast in terms of the V:C duration ratio, and both young and old Saxon speakers showed a greater tendency for lenition (this trend was statistically confirmed in the analysis described in chapter IV). Moreover, both age groups (irrespective of the regional background) tended to maintain the voicing contrast to a lesser extent in minimal pairs containing a labial stop than in pairs with an alveolar stop. Consequently, the extent to which the speakers are influenced by both Standard German (including the frequency distribution in the German lexicon) as well as their native dialects is the same for young and old speakers. Thus, we conclude that Saxon speakers maintain the contrast, which supports the idea of dialect levelling, but the degree of dialect levelling with respect to the voicing opposition is the same for both age groups.

On the other hand, the results reported in chapter IV indeed showed diminished contrast maintenance for young Bavarian speakers. We predicted that Bavarian speakers preserve the vowel length contrast in minimal pairs like /biːtən - bɪtən/ to a greater extent than Saxon speakers because in Central Bavarian (proportional) consonant duration is varied with vowel duration in order to maintain the vowel length (or voicing for that matter) contrast. However, our results corroborated this hypothesis only for old Bavarians. Young Bavarians preserved the vowel length contrast by means of the V:C duration ratio to the same extent as young and old Saxon speakers. This finding suggests that the Central Bavarian dialect feature of complementary distributed vowel length depending on the underlying voicing of the postvocalic stop is still maintained in the production of old Bavarian speakers but lost in the speech of young Bavarian speakers.

To conclude, speakers from neutralizing dialect areas are able to produce Standard German phonemic contrasts, but they show neutralization tendencies that originate from their dialectal backgrounds. Young speakers do not tend *per se* to a more standard-like pronunciation, i.e. the external factor age did not influence the degree of stop lenition in the production of young Saxon than old Saxon speakers. Then again, internal as well as external factors may have contributed to the diminished maintenance of the vowel length contrast in the productions of young as opposed to old Bavarian speakers. That is, internal factors may have been responsible for the greater tendency to neutralize the vowel length contrast. At the same time, however, the diminished contrast maintenance between long and short vowels leads to a more standard-like pronunciation. Thus, we cannot tell apart whether internal or external factors (or both) may have been the reason for decreasing contrast preservation. The vowel length contrast is diminished as vowel length is not complementary distributed with respect to the following stop, but it is maintained without a greater consonant lengthening.

5.5 Conclusion, implications and future directions

Within phonological and dialectological theories, phonemic contrasts have long been considered to be either completely maintained or neutralized. The findings presented in this thesis show further evidence that phonological oppositions are neither completely maintained nor neutralized: instead, speakers choose from a continuum which may be marked by endpoints that are equivalent to phonological categories, but they rarely produce variants identical to these endpoints. The extent to which contrasts are neutralized or maintained depends on linguistic factors such as phonological frequency and resyllabification as well as on the extra-linguistic factors age and regional background. Thus, this thesis adds to previous work on incomplete neutralization (e.g. Port & O'Dell, 1985; Port & Crawford, 1989; Charles-Luce, 1985, 1983; Piroth & Jancker, 2004) and extends it to contrast maintenance in other phonetic positions by young and old speakers of regional varieties of German. Our results demonstrate that it is important to incorporate the gradation of contrast maintenance in a model of present day regional varieties of German, given that many regional varieties diverge from the standard language in terms of contrast maintenance or neutralization. Oppositions may appear to be completely preserved but they are influenced by a number of conditioning factors and thus may be maintained either to a greater or to a lesser degree.

The extent to which a contrast is preserved also seems to depend on language experience. In intervocalic position the contrast is apparently more easily maintained, and this may be because speakers from a neutralizing variety are exposed to non-neutralizing varieties. That means, they gain experience with an unfamiliar contrast and, following exemplar theory (Pierrehumbert, 2002, 2003) store exemplars in the mental lexicon that lead to a modified distribution of tokens and, ultimately, to the production of the contrast. Pilot perception experiments based on the findings in this thesis have shown that Saxon listeners also perceive the intervocalic stop contrasts in resynthesized stimuli in which the V:C duration ratio cue is present. But for contrasts where such experience is missing, as for example the voicing contrast in final position, the cue (even if present in the acoustic signal as in our perception experiments in chapter II) cannot be applied as successfully as in intervocalic position. Therefore, the results for the final voicing contrast may have revealed "only" incomplete neutralization in production (Port & O'Dell, 1985; Port & Crawford, 1989; Charles-Luce, 1985, 1993) and perception (chapter II of this thesis) and not incomplete maintenance as in the analyses of the intervocalic voicing or the vowel length contrast. This finding is supported by speech perception studies showing that exposure to and experience with other languages help to perceive contrasts (e.g. Flege & Wang, 1989; Broersma, 2005).

Moreover, the investigation of age-dependent differences in contrast maintenance has proven to be a promising starting point for research on German sound changes in progress. More sociolinguistic and sociophonetic studies should include apparent time analyses using instrumental techniques to quantify the degree of dialect levelling in German and to uncover sound changes in progress. Further insight into the magnitude of the sound change in progress found in the analyses presented in this thesis may come from additional investigations focussing on other acoustic parameters. In this thesis, findings for differences and similarities between age and dialect groups are restricted to duration parameters (except for the subsidiary spectral analysis of vowel quality in 4.3.1) and so are the implications drawn from these results. The reason for this was the relatively high number of independent variables and hypotheses tested. The investigation concerning the extent to which old and young, Saxon and Bavarian speakers of Standard German differ, remains an open research question, since it may well be that the speaker groups diverge on other acoustic cues such as for example duration of aspiration. It could be that distinct acoustic cues are used to a different extent by speakers from two generations to distinguish between phonemes (as the analysis in chapter IV has shown) without actually really participating in a sound change; it might also be the case that the modified use of cues could mark the beginning of a sound change.

The acoustic analyses of this thesis were restricted to two age groups and two regional varieties. One of the dialect groups always served as a reference group: e.g. in the analysis reported in chapter III, the Bavarian speakers served as reference since the voicing contrast is maintained in Central Bavarian. In chapter IV, on the other hand, the Saxon speaker group acted as a reference group because the vowel length contrast is maintained in Upper Saxon. However, it is important to record Standard German and other dialect speakers with the same materials used in this corpus in order to further evaluate the amount of dialect levelling. This would be relevant to understanding the degree to which young and old Standard German speakers maintain the voicing and the vowel length contrast in terms of proportional vowel duration. We did compare our results with Kohler's (1977) findings for Northern Standard German. Yet, it is crucial to verify that there are no age dependent-differences between Standard German speakers either. Acoustic analyses should be carried out with a different division of age groups into old, middle-aged and young speakers, in order to determine potential intermediate stages of a sound change in progress.

Furthermore, the analyses focussed either on the voicing contrast or on the vowel length opposition in order to disentangle the way these parameters are used for contrast maintenance. As a result, there may have been a more flexible usage of the duration parameters because only one phonemic opposition was contrasted at the same time. Thus the question remains of the degree to which the proportional vowel duration cue is used in the realization of minimal triplets such as *Hagen* (/haːgən/, 'a name'), *Haken* (/haːkən/, 'hook') and *hacken* (/hakən/, 'to chop') which include both the vowel length and the voicing contrast. Perception experiments are necessary for further investigation of the interaction between vowel and consonant length and its implication for a potential sound change in progress. One important issue, for example, is whether young as opposed to old Bavarian listeners show a categorical change in their responses to a vowel length continuum from *Haken* to *hacken* because young but not old speakers maintain the vowel length contrast in terms of vowel duration only. The prediction for old Bavarian speakers would be a more gradual change in the response curve, as the vowel duration is lengthened but not the stop duration.

The studies described in this thesis were a first attempt to study dialect levelling using instrumental techniques and thus to investigate a possible sound change in progress. Much work remains to be done in this field. Further experiments are necessary to investigate the phonetic origin of this sound change (cf. Ohala, 1993, 2003) and to test whether internal or external factors (or a combination of both) cause the change (see Labov, 1994, 2001; Lindblom et al., 1995; Torgensen & Kerswill, 2004). The results have shown a complex relationship between linguistic and extralinguistic factors that were all shown to contribute to the degree of contrast maintenance. Usage-based models of speech communication such as exemplar theory (Pierrehumbert, 2001, 2002, 2003) or Lindblom's (1990) H&H model were found to explain such a complex interaction better than traditional accounts that treat neutralization as a structural change. A model of present day regional variation in German should include these gradual differences in the preservation of phonological contrast since they resemble the speaker/listener's knowledge of various phonological systems in which contrasts may be neutralized or maintained to a different extent.

Zusammenfassung

Einleitung

Die Produktion und Perzeption phonologischer Kontraste ist ein wichtiger Bestandteil sprachlicher Kommunikation; Kontraste helfen Wörter und deren semantische Bedeutungen zu unterscheiden und zu dekodieren. Viele Sprachen (Deutsch, Englisch, etc.) haben z.B. einen phonologischen Stimmhaftigkeitskontrast bei Obstruenten, d.h. die Stimmhaftigkeit kann zu Bedeutungsunterscheidung zwischen zwei Wörtern führen (z.B. *baden*, /baːdən/ vs. *baten*, /baːtən/). Die Opposition zwischen stimmhaften und stimmlosen Plosiven in postvokalischer Position kann durch diverse akustische Parameter aufrecht erhalten werden, z.B. dem Dauerverhältnis zwischen Vokal und Konsonant oder der *voice onset time*, d.h. dem Einsetzen der Stimmhaftigkeit relativ zur Verschlusslösung. In bestimmten phonetischen Kontexten oder auch sprachlichen Varietäten können phonologische Kontraste jedoch neutralisiert werden.

Neutralisierung ist ein phonologischer Prozess, bei dem der Unterschied zwischen zwei verschiedenen zugrundeliegenden Repräsentationen an der Oberflächenstruktur neutralisiert wird. Ein Kontrast, der in mehreren Sprachen (z.B. Bulgarisch, Deutsch, Holländisch, Katalanisch, Polnisch, Russisch, Tschechisch, Türkisch) neutralisiert wird, ist die Stimmhaftigkeitsopposition. Obstruenten werden z.B. in wortmedialer, intervokalischer oder in silbeninitialer Position hinsichtlich der Stimmhaftigkeit unterschieden, aber in silbenfinaler Position wird dieser Stimmhaftigkeitskontrast häufig zugunsten des stimmlosen Obstruenten neutralisiert. Die Neutralisierung der Stimmhaftigkeit in silbenfinaler Position wird auch als *Auslautverhärtung* bezeichnet. Ein bekanntes Beispiel für Auslautverhärtung im Deutschen ist das Wortpaar *Rad* vs. *Rat*, das mit Bezug zur Oberflächenstruktur homophon ist, trotz unterschiedlicher zugrundeliegender Formen: Während die Realisierung des zugrundeliegenden stimmhaften Plosivs, je nachdem, ob er in finaler oder medialer Position steht, variiert, wird der zugrundeliegende Plosiv grundsätzlich stimmlos produziert.

In einer Reihe mittel- und oberdeutscher Dialekte wird auch der intervokalische Stimmhaftigkeitskontrast neutralisiert, allerdings zugunsten der stimmhaften Variante, weswegen dieser Prozess auch Lenisierung genannt wird. Wörter wie *baden* und *baten* werden in den Dialekten, die zum Gebiet der sogenannten *Binnendeutschen Konsonantenschwächung* gehören, jeweils als [baːdən] realisiert. Die stimmhaft/stimmlos-Opposition wird häufig auch als Lenis/Fortis-Kontrast bezeichnet, da sich Plosive vielmehr in artikulatorischer Stärke und temporaler Struktur als im Ausmaß der Phonation, d.h. der eigentlichen Stimmlippenschwingung unterscheiden (Kohler, 1984). Insbesondere in der Literatur zu deutschen Dialekten wird die Lenis/Fortis-Terminologie verwendet (Barbour & Stevenson, 1990).

In vielen phonologischen Theorien, wie z.B. der Generativen Phonologie (Chomsky & Halle, 1968), wurde lange Zeit angenommen, dass phonologische Prozesse eine kategoriale Änderung der Oberflächenform hervorrufen. Auslautverhärtung im Deutschen wurde wahrscheinlich ebenso lange als das prototypische Beispiel für den phonologischen Prozess der Neutralisierung betrachtet. Viele Dialektuntersuchungen beschreiben die intervokalische Lenisierung ebenfalls als eine kategoriale Änderung der Stimmhaftigkeit. Experimentalphonetische Untersuchungen haben allerdings gezeigt, dass zugrundeliegende stimmhafte und stimmlose Plosive in silbenfinaler Position Unterschiede in relevanten akustischen Parametern, so z.B. der präkonsonantischen Vokaldauer, aufweisen. Zwar wurden zugrundeliegende stimmhafte Plosive nicht als stimmhafte Plosive realisiert, aber sie hatten signifikant längere Vokaldauern als zugrundeliegende stimmlose Plosive. Das bedeutet, dass sich stimmhafte und stimmlose Plosive in finaler Position in den selben Merkmalen und auch in ähnlichen (wenn auch abgeschwächten) Ausprägungen unterschieden wie in medialer Position. Daraufhin wurde geschlussfolgert, dass der domänen-finale Stimmhaftigkeitskontrast nicht vollständig neutralisiert, sondern zum Teil beibehalten wird. Das Phänomen der unvollständigen

Neutralisierung ist dabei nicht unbedingt sprachspezifisch. Unvollständige Neutralisierung des finalen Stimmhaftigkeitskontrast wurde für Sprachen wie Deutsch (Port & O'Dell, 1985; Port & Crawford, 1989; Charles-Luce, 1985, 1993), Katalanisch (Dinnsen & Charles-Luce, 1984), Polnisch (Slowiaczek & Dinnsen, 1985) und zum Teil auch für Holländisch (Warner et al., 2004, 2006; Ernestus & Baayen, 2006) gefunden. In Spontansprache konnte unvollständige Neutralisierung allerdings nicht nachgewiesen werden, weswegen immer wieder kritisiert wurde, dass es sich hierbei um ein Artefakt von Lesesprache, die in einem Labor aufgenommen wurde, handeln würde (Fourakis & Iverson, 1984; Jassem & Richter, 1989). Den Kritikern zufolge tritt unvollständige Neutralisierung nur in einem Kontext auf, in dem Sprecher zu Hyperartikulation neigen, insbesondere, wenn der Kontrast in der Orthographie hervorgehoben wird.

Diese Dissertation untersucht, ob der finale zugrundeliegende Stimmhaftigkeitskontrast auch in der Perzeption nur unvollständig neutralisiert wird und ob der intervokalische Stimmhaftigkeitskontrast möglicherweise auch in Varietäten des Deutschen wie dem Sächsischen nur partiell neutralisiert wird. Des Weiteren wird geprüft, inwiefern das teilweise Aufrechterhalten von Kontrasten abhängig ist von linguistischen Faktoren wie z.B. phonologischer Häufigkeit oder extra-linguistischen Faktoren wie Alter und regionalem Hintergrund eines Sprechers. Der Einfluss extra-linguistischer Faktoren wurde getestet, um zu prüfen, ob der Grad der Kontrastaufrechterhaltung in verschiedenen regionalen Varietäten auch mögliche Lautwandeltendenzen aufdecken kann.

Die Neutralisierung des intervokalischen Stimmhaftigkeitskontrastes zugunsten von Lenis ist ein häufig vorkommender diachroner Lautwandel. Kohler (1984) geht davon aus, dass Fortis-Obstruenten zunächst in Spontansprache lenisiert werden und dass aus diesem synchronen Prozess ein diachroner Lautwandel entstehen kann. Solche internen Faktoren können also dazu führen, dass jüngere Sprecher den phonologischen Kontrast immer stärker neutralisieren. In der dialektologischen Literatur wird die Lenisierung von Fortis-Plosiven in Dialekten aus dem Gebiet der *Binnendeutschen Konsonantenschwächung* als vollständige

Lenisierung beschrieben, d.h. man nimmt für diese Dialekte bereits eine vollständige Neutralisierung des Stimmhaftigkeitskontrastes an (Bergmann, 1990 für Sächsisch; Wiesinger, 1990 für Bairisch). Im Falle der intervokalischen Lenisierung wäre aber auch eine gegenläufige Entwicklung vorstellbar, bei der jüngere Sprecher den Kontrast stärker aufrechterhalten als ältere Sprecher. Externe Faktoren, wie z.B. das Prestige eines Dialektes, können zu Dialektabbau und auch zu Lautwandel führen (Torgensen & Kerswill, 2004). Solchen externen Faktoren zufolge streben jüngere Sprecher eine weniger dialektale, dafür standardnähere Aussprache an, die das Aufrechterhalten von standarddeutschen phonologischen Kontrasten zur Folge hätte. Es gibt so gut wie keine experimentalphonetischen Untersuchungen zu aktuellen Lautwandeltendenzen im Deutschen. Allerdings wurde in einigen Untersuchungen gezeigt, dass jüngere Generationen weniger Dialektmerkmale aufweisen im Vergleich zu älteren Generationen (Wagener, 2002; Lameli, 2004). Ein Ziel dieser Dissertation ist daher zu prüfen, ob charakteristische Dialektmerkmale in der Aussprache jüngerer Sprecher zweier regionaler Varietäten des Deutschen weniger stark ausgeprägt sind. Diese Forschungsfrage wurde mithilfe einer *apparent time* Analyse untersucht, bei der zeitgleich Sprachdaten von Sprechern aus zwei Altersgruppen erhoben und verglichen werden.

Unvollständige Neutralisierung des finalen Stimmhaftigkeitskontrastes in der Perzeption

Mehrere empirische Studien, v.a. aus den 1980er Jahren, haben gezeigt, dass der Stimmhaftigkeitskontrast in silbenfinaler Position, der verschiedenen phonologischen Modellen zufolge vollständig neutralisiert wird, teilweise beibehalten wird (Port & O'Dell, 1985; Charle-Luce, 1985, 1993). Im zweiten Kapitel wurden zwei Perzeptionsexperimente vorgestellt, in denen untersucht wurde, ob die Neutralisierung des finalen Stimmhaftigkeitskontrastes auch in der Perzeption unvollständig ist. Das vorrangige Ziel war zu testen, ob akustische Reize, die im Signal vorhanden sind, auch wenn der Plosiv in einer neutralisierenden Umgebung steht, wahrgenommen werden. Ein weiteres Ziel war zu prüfen, inwiefern linguistische Faktoren wie phonologische Häufigkeit oder eine mögliche Resilbifizierung des finalen Plosivs mit dem

Onset-Konsonanten der zweiten Silbe den Grad der Wahrnehmung des Stimmhaftigkeitskontrastes beeinflussen.

Es wurden insgesamt sechs Kontinua erstellt, die jeweils ein trochäisches Wortpaar umfassten, das sich nur in der zugrundeliegenden Stimmhaftigkeit des finalen Plosivs der ersten Silbe unterschied. Der Nukleus war entweder ein /iː/ oder ein /ɪ/, der finale Konsonant der ersten Silbe war entweder ein alveolarer oder ein velarer Plosiv und die Onset-Konsonanten der zweiten Silbe waren entweder /l/ oder /ʃt/. Für jedes Wortpaar wurden, ausgehend vom Wort mit dem zugrundeliegenden stimmhaften Plosiv, sechs Stimuli erstellt, indem die Vokaldauer sukzessive gekürzt und die Plosivdauer gelängt wurde (d.h. die Reimdauer blieb gleich). Die Stimuli aller sechs Kontinua (d.h. beider Experimente) wurden 19 norddeutschen Standardsprechern in einer Sitzung auditiv präsentiert. Die Aufgabe der Versuchspersonen war es zu beurteilen, ob sie entweder das Wort mit dem zugrundeliegenden stimmhaften oder stimmlosen Plosiv wahrgenommen haben (*two-alternative forced choice task*); hierfür wurden ihnen die Stimuli zusätzlich orthographisch präsentiert.

Im ersten Perzeptionsexperiment wurde mithilfe der Wörter, deren zweite Silbe mit /l/ begann, der Effekt der phonologischen Häufigkeit getestet. Ungespannte Vokale kommen im Deutschen überwiegend vor Fortis-Plosiven vor. Kombinationen aus ungespannten Vokalen und Lenis-Plosiven entstammen fast alle dem niederdeutschen Sprachraum (z.B. *Ebbe*) und sind somit eigentlich kein Teil des deutschen Lexikons (Féry, 2003: 150). Auf der anderen Seite folgen auf gespannte Vokale fast ausnahmslos stimmhafte Plosive, wenn der Artikulationsort labial oder velar ist. Im Falle von alveolaren Plosiven nach gespannten Vokalen ist die zugrundeliegende Stimmhaftigkeit ungefähr gleich verteilt. Die Hypothese war, dass Versuchspersonen in den Sequenzen, in denen es einen lexikalischen *Bias* zu einer der beiden Stimmhaftigkeitskategorien gibt, die wahrscheinlichere Stimmhaftigkeitskategorie wahrnehmen. Die stimmlosen Konsonanten in finaler Position wurden häufiger als erwartet als stimmhaft wahrgenommen. Ein möglicher Grund hierfür ist, dass die Stimuli ausgehend von zugrundeliegenden stimmhaften Konsonanten erstellt wurden. Dennoch nahm der Anteil der 'stimmlos'-Urteile mit zunehmend kürzeren Vokalen bzw. längeren Plosiven kontinuierlich zu, d.h. der finale Stimmhaftigkeitskontrast wurde in der Perzeption nur partiell neutralisiert, und Hörer perzipierten den Plosiv häufiger als stimmlos, wenn dem Plosiv ein ungespannter Vokal vorausging (im Gegensatz zu einem gespannten Vokal), d.h. die Ergebnisse bestätigten diese Hypothese.

Des Weiteren stellten wir die Hypothese auf, dass Hörer den Kontrast besser wahrnehmen (d.h. kategorialer) in Sequenzen, in denen die Stimmhaftigkeit aufgrund der Verteilung im Lexikon nicht vorhersagbar ist, da dann womöglich das akustische Signal ausschlaggebend für die Kategorisierung ist. Diese Hypothese wurde nur zum Teil bestätigt. Auf der einen Seite wurde – wie von der Hypothese vorhergesagt – der zugrundeliegende Stimmhaftigkeitskontrast in Sequenzen mit einem gespannten Vokal und einem alveolaren Plosiv besser wahrgenommen als in Sequenzen mit ungespanntem Vokal und alveolarem Plosiv. Auf der anderen Seite wurde nach gespannten Vokalen der Stimmhaftigkeitskontrast besser in velaren als in alveolaren Plosiven wahrgenommen. Dieses Ergebnis spricht gegen unsere Hypothese. Eine Erklärung für das letztgenannte Ergebnis ist die Möglichkeit der Resilbifizierung des finalen /k/ mit dem Onset-Konsonanten /l/ der zweiten Silbe. Diese Hypothese wurde im zweiten Perzeptionsexperiment geprüft.

Die Kontinua unterschieden sich im zweiten Experiment nicht mehr in der Gespanntheit des ersten Vokals (es handelte sich immer um einen gespannten Vokal), sondern im Onset der zweiten Silbe. Auch im zweiten Experiment wurde der Kontrast perzeptiv nur unvollständig neutralisiert (d.h. Zunahme der 'stimmlos'-Urteile von Stimulus 1-7). Des Weiteren nahmen Hörer den zugrundeliegenden Stimmhaftigkeitskontrast in Sequenzen mit velarem Plosiv gefolgt von einem Lateral besser wahr als in allen anderen Kontinua, in denen der letzte Konsonant nicht resilbifizierbar war. Dieses Ergebnis bestätigt unsere Hypothese. Darüber hinaus stimmen die Ergebnisse mit dem sogenannten *licening-by-prosody* Ansatz überein (Itô, 1986, 1989; Goldsmith, 1990, Rubach, 1990), bei dem angenommen wird, dass die Position innerhalb des Wortes für die Neutralisierung des Stimmhaftigkeitskontrastes ausschlaggebend ist. Die Wahrnehmung des Stimmhaftigkeitskontrastes wurde allerdings auch erhöht, wenn der Onset-Konsonant ein Sonorant war, unabhängig davon, ob er resilbifiziert werden konnte oder nicht (z.B. /dl - tl/). Dieses Ergebnis spricht eher für den *licensing-by-cue* Ansatz (Steriade, 1997, 1999, 2000). Bei diesem Ansatz wird davon ausgegangen, dass die akustischen Eigenschaften der neutralisierenden Umgebung ausschlaggebend für eine verbesserte Wahrnehmung des Stimmhaftigkeitskontrastes sind. Dem *licensing-by-prosody* Ansatz zufolge spielen die akustischen Eigenschaften der Umgebung keine Rolle für eine mögliche Stimmhaftigkeitswahrnehmung.

Es wurde gezeigt, dass es graduelle Unterschiede in der Wahrnehmung des finalen Stimmhaftigkeitskontrastes gab. Das bedeutet, dass man den Prozess der Neutralisierung nicht als kategoriale Strukturänderung generalisieren darf. Die Ergebnisse konnten am besten mithilfe der Exemplartheorie (Pierrehumbert, 2001, 2002, 2003) modelliert werden, da diese auch linguistischen Faktoren wie phonologischer Häufigkeit und Resilbifizierungsmöglichkeiten Rechnung trägt. Die Ergebnisse dieser Untersuchung demonstrieren, dass die Perzeption des zugrundeliegenden Stimmhaftigkeitskontrastes in silbenfinalen Konsonanten auf einer komplexen Interaktion zwischen akustischer Information und phonologischem Wissen (das wiederum auf linguistischer Erfahrung beruht) basiert. In Vokal-Konsonant-Folgen, in denen mehrere Faktoren die Stimmhaftigkeitswahrnehmung begünstigten, war die Perzeption des zugrundeliegenden Stimmhaftigkeitskontrastes am größten.

Der intervokalische Stimmhaftigkeitskontrast in regionalen Varietäten des Deutschen

Im Fokus des dritten Kapitels stand die Realisierung des intervokalischen Stimmhaftigkeitskontrastes nach gespannten Vokalen in zwei regionalen Varietäten des Deutschen. Im Standarddeutschen werden in intervokalischer Position stimmhafte und stimmlose Laute unterschieden. Dieser Kontrast wird im Sächsischen neutralisiert (Bergmann, 1990; Auer et al., 1993, 1997), im Bairischen jedoch beibehalten, da für das Bairische komplementäre Länge von Vokal und Konsonant angenommen werden (Bannert, 1976;

Wiesinger, 1990). Ein zweites Ziel war, zu prüfen, ob der Grad der Neutralisierung vom Alter abhängt. Ausgehend von der Annahme, dass jüngere Sprecher zu einer weniger dialektalen Aussprache neigen (Lameli, 2004), insbesondere wenn es sich um einen stigmatisierten Dialekt wie dem Sächsischen handelt, stellten wir die Hypothese auf, dass ältere sächsische Sprecher eher zu einer Neutralisierung des intervokalischen Stimmhaftigkeitskontrastes neigen als jüngere sächsische Sprecher. Eine weitere Annahme war, dass der Stimmhaftigkeitskontrast eher in labialen Plosiven neutralisiert wird als in alveolaren Plosiven, da nach gespannten Vokalen sehr viel häufiger zugrundeliegende stimmhafte labiale Plosive vorkommen als stimmlose labiale Plosive. Für den alveolaren Artikulationsort gibt es nach gespannten Vokalen keine solche *Bias*; d.h. stimmhafte und stimmlose Plosive sind nach gespannten Vokalen gleichermaßen frequent.

Für die akustischen Analysen, die in Kapitel III und IV beschrieben werden, wurde ein Corpus erstellt, das aus 46 Wörtern bestand, darunter viele Minimalpaare. Jedes Wort wurde den Versuchspersonen zehnmal in randomisierter Reihenfolge präsentiert. 20 Sprecher aus Bayern und 22 Sprecher aus Sachsen lasen das Sprachmaterial. Die Versuchspersonen wurden einer von 2 Altersgruppen zugeteilt: Sprecher, die älter als 50 Jahre alt waren, wurden der älteren Gruppe zugeordnet, Sprecher unter 50 Jahren der jüngeren. Die Aufteilung der Altersgruppen war wie folgt: 13 ältere und 9 jüngere sächsische Sprecher sowie 8 ältere und 12 jüngere bayrische Sprecher. Die Laute der Testwörter wurden in Praat (Boersma & Weenink, 2008) segmentiert und die ersten drei Formanten innerhalb des akzentuierten Vokals, wenn notwendig, in EMU (Harrington, 2010) korrigiert. Die statistische Analyse der Segmentdauern und Formanten wurde anschließend in EMU/R durchgeführt. Vokal- und Konsonantendauern wurden zwischen den Segmentgrenzen extrahiert und entweder auf die Wort- oder Reim-Dauer (Vokal + Konsonant) normalisiert.

Bei dem in Kapitel III analysierten Sprachmaterial handelte es sich um die Minimalpaare *Boden-boten* und *Ober-Oper.* Es wurden Vokal- und Konsonantendauern gemessen. Der Vokal wurde anschließend auf die Vokal-Konsonant-Dauer normalisiert (V:K).

Kohler (1977, 1979) sowie Kohler und Künzel (1978) hatten gezeigt, dass im Standarddeutschen der Stimmhaftigkeitskontrast u.a. im Dauerverhältnis von Vokal zu Reim kodiert ist. V:K-Dauerverhältnisse über 0,7 zeigten einen deutlichen Lenis-Konsonanten und Dauerverhältnisse unter 0,6 waren ein eindeutiges Maß für Fortis-Konsonanten. Die Ergebnisse zeigten für alle Sprechergruppen V:K-Werte über oder um 0,7 für /d/ bzw. unter 0,6 für /t/ (vgl. Tabelle 3.3), d.h. die Versuchspersonen haben den Fortis/Lenis-Kontrast produziert. Die Konsonantendauern waren länger bei labialen Plosiven und die V:K-Werte somit niedriger. Ein möglicher Grund für die artikulationsortabhängigen Verschlussdauern, die auch von Kohler (1977) und Kohler und Künzel (1978) gefunden wurden, liegt in dem Verhältnis von Luftdruck und der Größe des Mundraumes hinter dem Verschluss. Bei einem labialen Plosiv ist der Mundraum hinter dem Verschluss größer als bei einem alveolaren Plosiv. Daher kann es länger dauern bis ein supraglottaler Überdruck entsteht, der zur Verschlusslösung des Plosivs führt (siehe Maddieson, 1997). Der Stimmhaftigkeits-Kontrast wurde stärker neutralisiert, wenn der Artikulationsort labial anstatt alveolar war. Dieses Ergebnis spricht für unsere Hypothese. Des Weiteren zeigten sächsische Sprecher eine größere Tendenz zur Neutralisierung, d.h. die Differenz zwischen dem V:K-Verhältnis für zugrundeliegende Fortis- und Lenis-Plosive war kleiner als bei bayrischen Sprechern. Die V:K-Verhältnisse waren innerhalb der Dialekte gleich, d.h. es gab keine signifikanten Altersunterschiede (vgl. Abbildung 3.8). Die Ergebnisse deuten darauf hin, dass auch Sprecher aus einem neutralisierenden Dialekt einen phonologischen Kontrast im Standarddeutschen aufrechterhalten, jedoch zu einem geringeren Grad. Die nicht signifikanten Altersgruppenunterschiede wurden wie folgt interpretiert: Sowohl jüngere als auch ältere sächsische Sprecher tendieren gleichermaßen zum Dialektausgleich.

Der Vokallängenkontrast in regionalen Varietäten des Deutschen

In Kapitel III wurde davon ausgegangen, dass bayrische Sprecher den intervokalischen Stimmhaftigkeitskontrast aufrechterhalten, da im Bairischen besondere Quantitätsverhältnisse herrschen: Vokale vor Lenis-Plosiven sind grundsätzlich lang und vor Fortis-Plosiven kommen nur Kurzvokale vor. Das Hauptziel des Kapitels IV war es zu untersuchen, inwieweit der Vokallängenkontrast von sächsischen und bayrischen Sprechern beibehalten wird, wenn sie Standarddeutsch produzieren. Ein weiteres Ziel war wiederum zu prüfen, ob es altersabhängige Unterschiede im Grad der dialektalen Ausprägung gibt. Aus dem oben beschriebenen Corpus wählten wir Minimalpaare aus, die sich nur in der zugrundeliegenden phonologischen Vokallänge unterschieden. Der postvokalische Konsonant war entweder ein alveolarer Fortis-Plosiv (z.B. /biːtən/ vs. /bɪtən/) oder ein Sonorant (z. B. /høːlə/ vs. /hœlə/). Vor alveolaren Fortis-Plosiven kommen sowohl gespannte als auch ungespannte Vokale gleichermaßen häufig vor, so dass der linguistische Faktor der phonologischen Häufigkeit ausgeklammert war.

Die Hypothese war, dass sich die beiden Dialektgruppen in der Realisierung des Vokallängenkontrastes unterscheiden. Mit Bezug auf die sächsische Sprechergruppe nahmen wir an, dass der Vokallängenkontrast realisiert wird, aber der Fortis-Plosiv stärker lenisiert wird (vgl. Kapitel III). Hinsichtlich der bayrischen Sprechergruppe gab es zwei mögliche Realisierungsmuster: Entweder wird der Vokallängenkontrast zugunsten eines ungespannten Vokals neutralisiert und der zugrundeliegende Fortis-Plosiv wird als solcher realisiert. Die andere Möglichkeit ist, dass der Vokallängenkontrast indirekt durch die Produktion eines zusätzlichen Stimmhaftigkeitskontrastes aufrechterhalten wird; d.h. bei der zweiten Möglichkeit würde ein Kontrast zwischen [bɪtən] und [biːdən] realisiert, wobei das gelängte [iː] ein phonetisches Resultat von [d] wäre. Des Weiteren nahmen wir an, dass bayrische Sprecher den ungespannten Vokal vor sonoranten Konsonanten längen und somit den Vokallängenkontrast neutralisieren.

Die Sprecher sowie ihre Zuordnung zu vier Sprechergruppen waren die gleichen wie in Kapitel III. Die Datenanalyse entsprach ebenfalls der in Kapitel III.

Sowohl bayrische als auch sächsische Sprecher kontrastierten gespannte und ungespannte Vokale sowohl durch verschiedene Vokalquantitäten als auch –qualitäten. Ungespannte Vokale waren kürzer als gespannte Vokale und wiesen eine im Vergleich zu gespannten Vokalen zentralisiertere Vokalqualität auf. Die Unterschiede zwischen den Dialekt-und Altersgruppen wurden nur für den Parameter Vokaldauer gefunden, nicht aber für die Qualität. Sächsische (wie auch bayrische) Sprecher produzierten einen Fortis-Plosiv hinsichtlich des Parameters V:K-Dauerverhältnis, tendierten aber – wie auch die Analyse in Kapitel III ergeben hatte – stärker zur Lenisierung als die bayrischen Sprecher. Der Unterschied im Lenisierungsgrad (V:K-Dauerverhältnis) zwischen bayrischen und sächsischen Sprechern war nach ungespannten Vokalen größer als nach gespannten Vokalen. Dieses Ergebnis konnte folgendermaßen interpretiert werden: (1) Sächsische Sprecher lenisieren /t/ nach ungespannten Vokalen anders als nach gespannten oder (2) bayrische Sprecher erhalten den Vokallängenkontrast mittels V:K-Dauerverhältnissen aufrecht, die entweder eher in Richtung eines Lenis- oder eher in Richtung eines Fortis-Plosivs weisen.

Um den Grad der Kontrastaufrechterhaltung mittels V:K-Dauerverhältnissen zu quantifizieren, haben wir die durchschnittliche V:K-Differenz zwischen gespannten und ungespannten Vokalen berechnet. Die Ergebnisse zeigten, dass ältere bayrische Sprecher den Kontrast in größerem Maße produzieren als die drei anderen Sprechergruppen, und zwar sowohl in Wörtern mit einem postvokalischen Obstruenten als auch in Wörtern mit einem postvokalischen Sonoranten. Es gab keinen Unterschied im Kontrastierungsgrad zwischen jüngeren Sachsen und jüngeren Bayern und älteren sächsischen Sprechern (vgl. Abbildung 4.9). Alle Sprechergruppen wiesen längere Konsonantendauern nach ungespannten Vokalen auf; dieser Effekt war aber wiederum am größten für ältere bayrische Sprecher.

Dieses Ergebnis wurde als Beweis für Bannerts (1976) Ansatz der komplementären Länge im Bairischen interpretiert. Selbst in Vokal+Sonorant-Sequenzen längen ältere bayrische Sprecher den postvokalischen Konsonanten stärker als die anderen Sprechergruppen. Wir interpretieren dieses Ergebnis als eine phonetische postvokalische Längung, um den phonemischen Vokalkontrast aufrechtzuerhalten, und nicht als einen phonemischen Stimmhaftigkeitskontrast, vor dem die Vokallänge komplementär distribuiert ist. Das Dialektmerkmal der komplementären Länge im Bairischen scheint aber verloren zu gehen, da die jüngeren bayrischen Sprecher den Vokallängenkontrast nicht stärker aufrechterhalten und sich hinsichtlich des V:K-Verhältnisses nicht von den sächsischen Sprechern unterscheiden.

Schlussfolgerung

Innerhalb phonologischer und dialektologischer Theorien wurden phonemische Kontraste lange Zeit als entweder vollständig aufrechterhalten oder vollständig neutralisiert angesehen. Die Ergebnisse der in dieser Dissertation vorgestellten Untersuchungen haben gezeigt, dass phonologische Oppositionen weder vollständig neutralisiert noch aufrechterhalten werden. Sprecher wählen stattdessen Varianten von einem Kontinuum, das durch Endpunkte markiert wird, die den phonologischen Kategorien entsprechen. Sie produzieren aber wahrscheinlich nur selten Varianten, die mit den Endpunkten übereinstimmen. Der Grad, in dem Kontraste neutralisiert oder aufrechterhalten werden, hängt von linguistischen Faktoren (z.B. phonologische Häufigkeit und/oder Resilbifizierungs-Möglichkeiten) und extra-linguistischen Faktoren (z.B. Alter und/oder regionaler Hintergrund) ab. Die Untersuchungen dieser Dissertation tragen zu früheren Arbeiten zur unvollständigen Neutralisierung bei und wurden durch die Frage nach der Aufrechterhaltung phonologischer Kontraste in anderen Positionen durch jüngere und ältere Sprecher regionaler Varietäten des Deutschen erweitert. Die Ergebnisse zeigen, wie wichtig es ist, graduelle Unterschiede in der Kontrastaufrechterhaltung in ein Modell moderner Regionalsprachen des Deutschen aufzunehmen.

Das Ausmaß, mit dem Kontraste aufrechterhalten werden können, scheint u.a. von linguistischer Erfahrung abzuhängen. In intervokalischer Position wird der Stimmhaftigkeits-Kontrast eher aufrechterhalten als in domänen-finaler Position. Dies kann darauf zurückzuführen sein, dass Sprecher einer neutralisierenden Varietät mit Sprechern nicht-neutralisierender Varietäten in Kontakt kommen und dadurch diese Kontraste (kennen)lernen. In domänen-finaler Position wird der Kontrast in (fast) keiner regionalen Varietät aufrechterhalten, so dass Sprecher auch keine linguistische Erfahrung mit dem Stimmhaftigkeits-Kontrast in domänenfinaler Position sammeln können.

Die Ergebnisse lassen sich besser durch die Exemplar-Theorie (Pierrehumbert, 2002, 2003) oder Lindblom's (1990) H&H Modell erklären als durch traditionelle phonologische Theorien modellieren, da letztere von einer kategorialen Änderung ausgehen. Sowohl die Exemplartheorie als auch das H&H-Modell tragen feinen phonetischen und somit auch graduellen Unterschieden Rechnung, die durch linguistische Erfahrung und/oder der kommunikativen Situation begründet werden können. Im vierten Kapitel wurde ein möglicher Sprachwandel gefunden. Nun muss geklärt werden, wodurch der Lautwandel zustande kommt (interne vs. externe Faktoren) und wie er sich in der Perzeption dieser Kontraste bemerkbar macht.

Appendix A: Speech materials

Table A.1: List of all test and filler words recorded for the corpus.

Contrast	Sound	No.	Word	Pronunciation	Translation
Tense vowel contrasts	iː	1.	bieten	/biːtən/	to offer
	eː	2.	beten	/beːtən/	to pray
	ɛː	3.	bäten	/bɛːtən/	if they requested
	aː	4.	baten	/baːtən/	requested
		5.	Staat	/ʃtaːt/	state
	oː	6.	Gote	/goːtə/	Goth
		7.	Sohne	/zoːnə/	son
	uː	8.	gute	/guːtə/	good
		9.	schmust	/ʃmuːst/	sb. cuddles
	yː	10.	Güte	/gyːtə/	goodness
		11.	Hüte	/hyːtə/	hats
	øː	12.	böten	/bøːtən/	if they offered
		13.	Goethe	/gøːtə/	a name
		14.	Höhle	/høːlə/	cave
Lax vowel contrasts	ɪ	15.	bitten	/bɪːtən/	to request
	ɛ	16.	Betten	/biːtən/	beds
		17.	hätte	/hɛtə/	sb. had (subjunctive)
		18.	helle	/hɛlə/	bright
	a	19.	hatte	/hatə/	sb. had
		20.	kannte	/kantə/	sb. knew
		21.	Stadt	/ʃtat/	town
	ɔ	22.	konnte	/kɔntə/	could
		23.	Sonne	/sɔnə/	sun
		24.	musst	/mʊst/	must
	ʊ	25.	Kunde	/kʊndə/	customer
	ʏ	26.	Hütte	/hʏtə/	hut
		27.	Hülle	/hʏlə/	cover
		28.	Bütten	/bʏːtən/	laid paper
	œ	29.	Hölle	/hœlə/	hell
		30.	könnte	/kœntə/	could
Initial voicing contrasts	p	31.	Pass	/pas/	passport
	b	32.	Bass	/bas/	opera
	t	33.	Tasse	/tasə/	cup
	d	34.	das	/das/	the
	k	35.	Kasse	/kasə/	cash till
	g	36.	Gasse	/gasə/	alley
Medial voicing contrasts	p	37.	Oper	/oːpɐ/	bass
	b	38.	Ober	/oːbɐ/	waiter
	t	39.	Boten	/boːtən/	messengers
	d	40.	Boden	/boːdən/	floor
Other filler words		41.	pflücken	/pflʏkən/	to pick
		42.	Regen	/reːgən/	rain
		43.	Flieger	/fliːgɐ/	aviator
		44.	Meer	/meːɐ/	sea
		45.	spräche	/ʃprɛːçə/	would speak
		46.	Milch	/mɪlç/	milk

Appendix B: Schwa deletion

Table B.1: Number of oral vs. nasal releases in all test words with *en*-ending analyzed in chapter III.

Subject	Region	Age	Gender	*boten*		*Boden*	
				oral	nasal	oral	nasal
AFO	Bavarian	old	female	10	0	10	0
ASI	Bavarian	young	female	0	10	0	9
CHE	Bavarian	young	male	0	10	10	0
CLU	Bavarian	young	female	10	0	9	1
CMA	Bavarian	young	female	0	10	0	10
DME	Bavarian	young	male	10	0	10	0
ENE	Bavarian	old	female	10	0	10	0
FPO	Bavarian	young	female	1	9	3	7
GBL	Bavarian	old	male	10	0	10	0
GFO	Bavarian	old	female	10	0	8	2
HWI	Bavarian	young	male	10	0	10	0
KDE	Bavarian	young	female	1	9	0	10
LFO	Bavarian	young	female	0	10	0	10
LPR	Bavarian	young	female	4	6	1	9
MBE	Bavarian	young	male	0	10	0	10
RLE	Bavarian	old	female	10	0	10	0
RNE	Bavarian	old	male	10	0	10	0
SFO	Bavarian	old	male	9	1	9	0
TBL	Bavarian	old	female	10	0	10	0
VSE	Bavarian	young	female	1	9	0	10
AMI	Saxon	young	female	5	5	0	10
BBO	Saxon	old	female	2	5	0	7
BFU	Saxon	old	female	10	0	10	0
CKL	Saxon	young	male	10	0	10	0
CON	Saxon	old	female	10	0	9	0
COT	Saxon	young	female	10	0	10	0
DKL	Saxon	old	female	10	0	10	0
DWE	Saxon	young	male	10	0	10	0
DWU	Saxon	old	male	10	0	10	0
FWE	Saxon	young	female	0	10	2	8
GKL	Saxon	old	male	10	0	10	0
HRU	Saxon	old	male	7	3	0	10
IGE	Saxon	young	female	10	0	10	0
JGE	Saxon	young	male	2	8	0	9
KLI	Saxon	old	male	10	0	10	0
PBL	Saxon	old	male	10	0	10	0
RHE	Saxon	old	female	10	0	10	0
SLI	Saxon	young	female	10	0	10	0
URE	Saxon	young	female	10	0	2	8
WDI	Saxon	old	male	10	0	10	0
WRU	Saxon	old	male	6	4	0	10
WWE	Saxon	old	male	10	0	10	0

Table B.2: Number of oral vs. nasal releases in all test words with *en*-ending analyzed in chapter IV.

Subject	Region	Age	Gender	***beten***		***betten***		***bieten***		***bitten***	
				oral	**nasal**	**oral**	**nasal**	**oral**	**nasal**	**oral**	**nasal**
AFO	Bavarian	old	female	10	0	10	0	10	0	10	0
ASI	Bavarian	young	female	0	9	0	9	0	10	0	10
CHE	Bavarian	young	male	10	0	8	2	10	0	10	0
CLU	Bavarian	young	female	9	0	7	1	10	0	10	0
CMA	Bavarian	young	female	0	10	0	10	1	9	1	9
DME	Bavarian	young	male	10	0	10	0	10	0	9	0
ENE	Bavarian	old	female	10	0	10	0	10	0	10	0
FPO	Bavarian	young	female	1	9	3	7	3	7	3	7
GBL	Bavarian	old	male	10	0	10	0	10	0	10	0
GFO	Bavarian	old	female	10	0	10	0	10	0	9	0
HWI	Bavarian	young	male	10	0	10	0	10	0	10	0
KDE	Bavarian	young	female	0	10	0	9	2	8	0	10
LFO	Bavarian	young	female	0	10	0	10	0	10	0	10
LPR	Bavarian	young	female	0	10	1	9	0	10	4	6
MBE	Bavarian	young	male	0	10	0	10	1	9	0	10
RLE	Bavarian	old	female	10	0	10	0	10	0	10	0
RNE	Bavarian	old	male	10	0	10	0	9	1	10	0
SFO	Bavarian	old	male	10	0	9	1	10	0	10	0
TBL	Bavarian	old	female	10	0	10	0	10	0	10	0
VSE	Bavarian	young	female	1	9	0	10	0	10	1	9
AMI	Saxon	young	female	4	6	0	7	3	7	2	8
BBO	Saxon	old	female	4	6	5	2	5	6	5	5
BFU	Saxon	old	female	10	0	10	0	10	0	10	0
CKL	Saxon	young	male	9	0	10	0	10	0	9	0
CON	Saxon	old	female	10	0	10	0	10	0	10	0
COT	Saxon	young	female	11	0	9	0	10	0	10	0
DKL	Saxon	old	female	10	0	10	0	10	0	10	0
DWE	Saxon	young	male	10	0	10	0	10	0	10	0
DWU	Saxon	old	male	10	0	10	0	10	0	10	0
FWE	Saxon	young	female	0	11	5	5	0	10	7	3
GKL	Saxon	old	male	10	0	10	0	10	0	9	1
HRU	Saxon	old	male	5	5	5	5	5	5	8	2
IGE	Saxon	young	female	7	0	11	0	9	0	10	0
JGE	Saxon	young	male	1	8	1	6	0	10	5	5
KLI	Saxon	old	male	10	0	10	0	10	0	10	0
PBL	Saxon	old	male	10	0	10	0	10	0	10	0
RHE	Saxon	old	female	10	0	10	0	10	0	10	0
SLI	Saxon	young	female	10	0	10	0	10	0	10	0
URE	Saxon	young	female	9	1	7	3	10	0	8	2
WDI	Saxon	old	male	10	0	10	0	10	0	10	0
WRU	Saxon	old	male	1	9	1	9	2	8	6	4
WWE	Saxon	old	male	10	0	10	0	10	0	10	0

Appendix C: Publications

The author of this thesis has been among the authors of the following works:

Harrington, J., Hoole, P., Kleber, F., & Reubold, U. (2011). The physiological, acoustic, and perceptual basis of high back vowel fronting: evidence from German tense and lax vowels. *Journal of Phonetics*, 39, 121-131.

Harrington, J., Kleber, F., & Reubold, U. (2007). U-fronting in RP: a link between sound change and diminished perceptual compensation for coarticulation? *Proc. of the 16th ICPhS*, *Saarbrücken,* 1473-1476.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: an acoustic and perceptual study. *Journal of the Acoustical Society of America*, 123, 2825-2835.

Harrington, J., Kleber, F., & Reubold, U. (2011). The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English. *Journal of the International Phonetic Association*, 41, 137-156.

Kleber, F. (2006). Form and function of falling pitch contours in English. In *Proc. 3rd Conference on Speech Prosody*, 61-64.

Kleber, F., Harrington, J., & Reubold, U. (in press). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech.* (doi:10.1177/0023830911422194)

Kleber, F., John, T., & Harrington, J. (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 38, 185-196.

Kleber, F., & Niebuhr, O. (2010). Semantic-context effects on lexical stress and syllable prominence. In *Proc. 5th Conference on Speech Prosody, Chicago*.

Kleber, F., Rathcke, T. (2008). More on the "segmental anchoring" of prenuclear rises: Evidence from East Middle German. In *Proc. 4th Conference on Speech Prosody, Campinas*.

Kleber, F., Reubold, U., & Harrington, J. (2010). /u/-fronting in RP and the implications of perceptual integration of lip gestures for sound change processes. *Laboratory Phonology*, 12 (Abstract only).

Mády, K., & Kleber, F. (2010). Variation of pitch accent patterns in Hungarian. In *Proc. 5th Conference on Speech Prosody, Chicago*.

Reichel, U.D., Kleber, F., & Winkelmann, R. (2009). Modelling similarity perception of intonation. In *Proc. 10th Interspeech*, *Bristol,* 1711-1714.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: a longitudinal analysis in adult speakers. *Speech Communication*, 52, 638-651.

Bibliography

Auer, P. (1988). MHG î and û in the city dialect of Constance. In: P. Auer and A. di Luzio (eds.), *Variation and convergence*. Berlin: de Gruyter, 44 – 75.

Auer, P., Barden, B., & Großkopf, B. (1993). Dialektwandel und sprachliche Anpassung bei "Übersiedlern" und "Übersiedlerinnen" aus Sachsen. *Deutsche Sprache*, 21, 80 – 87.

Auer, P., Barden, B., & Großkopf, B. (1997). Long-term linguistic accommodation and its sociolinguistic interpretation: evidence from the inner migration after the Wende. In: K. Mattheier (ed.), *Dialect Migration in a Changing Europe*. Frankfurt: Peter Lang, 79 – 98.

Auer, P., Gilles, P., & Spiekermann, H. (eds.) (2002). *Silbenschnitt und Tonakzente.* Tübingen: Niemeyer. (Linguistische Arbeiten ; 463)

Auer, P., Hinskens, F., & Kerswill, P. (2005). *Dialect Change: Convergence and Divergence in European Languages.* Cambridge: Cambridge University Press.

Baayen, H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bannert, R. (1976). *Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage*. PhD dissertation, Lund University. CWK Gleerup: W. Fink (Traveaux de l'institut de linguistique de Lund ; 10)

Barbour, S. & Stevenson, P. (1990). *Variation in German: A critical approach to German sociolinguistics.* Cambridge: Cambridge University Press.

Barden, B. & Großkopf, B. (1998). *Sprachliche Akkomodation und soziale Integration: Sächsische Übersiedler und Übersiedlerinnen im rhein-/moselfränkischen und alemannischen Sprachraum.* Tübingen: Niemeyer. (Phonai ; 43)

Bauer, L. (2008). Lenition revisited. *Journal of Linguistics*, 44, 605 – 624.

Bauer, L. & Warren, P. (2004). New Zealand English: phonology. In: In B. Kortmann, E. Schneider, K. Burridge, R. Mesthrie and C. Upton (eds.), *A Handbook of Varieties of English.* Berlin/New York: de Gruyter, 580 – 602.

Baugh, A. & Cable, T. (2002). *A History of the English Language*. 5th revised edition. London: Routledge.

Baumann, M. (1995). *The production of syllables in connected speech*. Unpublished PhD dissertation, University of Nijmegen.

Beckman, J. (1997). Positional faithfulness, positional neutralisation, and Shona vowel harmony. *Phonology*, 14, 1 – 46.

Bennet, D. (1968). Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech*, 11, 65 – 81.

Bergmann, G. (1990). Upper Saxon. In: C. Russ (ed.), *The dialects of Modern German.* London: Routledge, 290 – 312.

Bergmann, G. & Becker, H. (1969). *Sächsische Mundartenkunde. Entstehung, Geschichte und Lautstand der Mundarten des obersächsischen Gebiets.* Revised and edited by G. Bergmann. Halle.

Besch, W. (1983). Entstehung und Ausprägung der binnensprachlichen Diglossie im Deutschen. In: W. Besch, U. Knoop, W. Putschke, and H. Wiegand (eds.), *Dialektologie: ein Handbuch zur deutschen und allgemeinen Dialektforschung, vol. 2.* Berlin/New York: de Gruyter, 1399 – 1411. (Handbücher zur Sprach- und Kommunikationswissenschaft ; 1)

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In: J. Goodman and H. Nusbaum (eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge, MA: MIT, 167 – 224.

Bethge, W. (1963) Beziehungen der Generation zur Quantität in den deutschen Mundarten. *Phonetica*, 9, 200 – 208.

Blevins, J. & Grawunder, S. (2009). *KL > TL sound change in Germanic and elsewhere: Descriptions, explanations, and implications. *Linguistic Typology*, 13, 267 – 303.

Boersma, P. & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.27) [Computer program]. Retrieved June 28, 2008, from http://www.praat.org/

Booij, G. (1981). *Generatieve fonologie van het Nederlands*. Utrecht: Het Spectrum.

Booij, G. & Rubach, J. (1987). Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry*, 18, 1 – 44.

Braun, A. (1988). *Zum Merkmal "Fortis/Lenis": Phonologische Betrachtungen und instrumentalphonetische Untersuchungen an einem mittelhessischen Dialekt.* Wiesbaden: Steiner. (Zeitschrift für Dialektologie und Linguistik ; Beiheft 55)

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

Brockhaus, W. (1995). *Final Obstruent Devoicing in the Phonology of German*. Tübingen: Niemeyer.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117, 3890 – 3901.

Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219 – 252.

Browman, C. & Goldstein, L. (1988). Some notes on syllable structure. *Phonetica*, 45, 140 – 155.

Browman, C. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 219 – 252.

Browman, C. & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: J. Kingston and M.Beckman (eds.), *Papers in Laboratory Phonology 1: Beyond the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, 341 – 376.

Browman, C. & Goldstein, L. (1991). Gestural structures: distinctiveness, phonological processes, and historical change. In: I. Mattingly and M. Studdert-Kennedy (eds.), *Modularity and the Motor Theory of Speech Perception.* Hillsdale: Erlbaum, 313 - 338

Browman, C. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49, 155 – 180.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, J. (2004). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261 – 290.

Byrd, D. (1993). 54,000 stops. *UCLA Working Papers in Phonetics*, 83, 97 – 115.

Charles-Luce, J. (1985). Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics*, 13, 309 – 324.

Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, 50, 28 – 43.

Charles-Luce, J. & Dinnsen, D. (1987). A reanalysis of Catalan devoicing. *Journal of Phonetics*, 15, 187 – 190.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129 – 159.

Cho, T. (1998). *The specification of intergestural timing and overlap: EMA and EPG studies.* Unpublished MA thesis, University of California, Los Angeles.

Cho, T. & McQueen, J. (2005). Prosodic influence on consonant production in Dutch: effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33, 121 – 157.

Chomsky, N. & Halle, M. (1968). The sound pattern of English. New York: Harper & Row.

Clark, J. & Yallop, C. (1990). *An Introduction to Phonetics and Phonology*. Oxford, UK/Cambridge, USA: Basil Blackwell.

Clopper, C. & Pierrehumbert, J. (2008). Effects of semantic predictability and regional dialect on vowel spece reduction. *Journal of the Acoustical Society of America*, 124, 1682 – 1688.

Clopper, C. & Pisoni, D. (2006). Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change*, 18, 193 – 221.

Coleman, J. (2003). Commentary: probability, detail and experience. In: J. Local, R. Ogden, and R. Temple (eds.), *Papers in Laboratory Phonology 6*. Cambridge: Cambridge University Press, 88 – 100.

Cox, F. (1999). Vowel change in Australian English. *Phonetica*, 56, 1 – 27.

Croft, W. (2000). *Explaining Language Change. An Evolutionary Approach*. London: Longman.

Crowley, T. (1998). *An Introduction to Historical Linguistics.* 3rd edition. Melbourne: Oxford University Press.

Cutler, A. (2002). Phonological processing: comments on Pierrehumbert, Moates et al., Kubozono, Peperkamp & Dupoux, and Bradlow. In: C. Gussenhoven and N. Warner (eds.). *Laboratory Phonology 7*. Cambridge: Cambridge University Press, 275 – 296.

Davidson, L. (2006). Schwa elision in fast speech: segmental deletion or gestural overlap? *Phonetica*, 63, 79 – 112.

Davis, M., Marslen-Wilson, W., & Gaskell, G. (2002). Leading up the lexical garden-path: segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 218 – 244.

Davis, S. & Van Summers, W. (1989). Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics*, 17, 390 – 353.

Delvaux, V. & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, 64, 145 – 173.

Dingeldein, H. (1997). Sprachvarietäten in "Mitteldeutschland": Gebrauch und Räumlichkeit. In: G. Stickel (ed.), *Varietäten des Deutschen.* Berlin: de Gruyter, 109 – 141.

Dinnsen, D. & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49 – 60.

Docherty, G. & Foulkes P. (1999). Instrumental phonetics and phonological variation: case studies from Newcastle upon Tyne and Derby. In: P. Foulkes and G. Docherty (eds.), *Urban Voices: Accent Studies in the British Isles*. London: Arnold, 47 – 71.

Docherty, G. & Foulkes, P. (2000). Another chapter in the story of labiodental variants in British English. *Journal of Sociolinguistics*, 4, 30 – 59.

Draxler, C. & Jänsch, K. (2004). SpeechRecorder - A universal platform independent multi-channel audio recording software. *Proc. of the IV. International Conference on Language Resources and Evaluation, Lisbon*, 559 – 562.

Ernestus, M. & Baayen, H. (2006). The functionality of incomplete neutralization in Dutch: the case of past-tense formation. In: L. Goldstein, D. Whalen, and C. Best (eds.), *Laboratory Phonology 8*. Berlin: de Gruyter, 27 – 49.

Evans, B. & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America*, 121, 3814 – 3826.

Farrar, K. & Jones, M. (2002). Introduction. In: K. Farrar and M. Jones (eds.), *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*. New York: de Gruyter, 1 – 16.

Féry, C. (2003). Final devoicing and the stratification of the lexicon in German. In: J. van de Weijer, V. van Heuven, and H. van der Hulst (eds.). *The Phonological Spectrum: Segmental Structure*. Amsterdam: Benjamins, 145 – 170.

Fischer-Jørgensen, E. (1990). Intrinsic F0 in tense and lax vowels with special reference to German. *Phonetica*, 47, 99 – 140.

Fischer-Jørgensen, E. & Jørgensen, H. (1969). Close and loose contact ("Anschluß") with spezial reference to North German. *Annual Report of the Institut of Phonetics of the University of Copenhagen (ARIPUC)*, 4, 43 – 80.

Flege, J. & Wang, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/ – /d/ contrast. *Journal of Phonetics*, 17, 299 – 315.

Fougeron, C. & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101, 3728 – 3740.

Fourakis, M. & Iverson, G. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, 4l, 140 – 149.

Fowler, C. (1992). Vowel duration and closure duration in voiced and unvoiced stops: there are no contrast effects here. *Journal of Phonetics*, 20, 143 – 165.

Fridland, V. (1998). *The Southern vowel shift: linguistic and social factors.* PhD dissertation, Michigan State University.

Fridland, V. (2000). The Southern shift in Memphis, Tennessee. *Language Variation and Change*, 11, 267 – 285.

Frisch, S., Large, N., & Pisoni, D. (2000). Perception of wordlikeness: effects of segment probability and length on processing of nonword sound patterns. *Journal of Memory and Language*, 42, 481 – 496.

Gafos, A. (2006). Dynamics in grammar: Comment on Ladd and Ernestus & Baayen. In: L. Goldstein, D. H. Whalen, and C. T. Best (eds.), *Laboratory phonology 8*. Berlin: de Gruyter, 51 – 79.

Geiger, S. & Salmons, J. (2006). Reconstructing variation at shallow time depths: The historical phonetics of 19th-century German dialects in the U.S. In: T. Cravens (ed.), *Variation and Reconstruction.* Amsterdam: Benjamins, 37 – 58.

Giegerich, H. (1985). *Metrical Phonology and Phonological Structure: German and English.* Cambdrige: Cambridge University Press. (Cambridge Studies in Linguistics ; 43)

Giegerich, H. (1992). Onset maximisation in German: the case against resyllabification rules. In: P. Eisenberg, K. H. Ramers and H. Vater (eds.), *Silbenphonologie des Deutschen*. Tübingen: Narr, 134 – 171. (Studien zur deutschen Grammatik ; 42)

Giles, H. (1973). Accent mobility: a model and some data. *Anthropological Linguistics*, 15, 87 – 105.

Giles, H. & Powesland, P. (1997). Accommodation theory. In: N. Coupland and A. Jaworski (eds.), *Sociolinguistics: A reader*. Basingstoke: Macmillan, 232 – 239. (Reprinted from H. Giles & Powesland, P. (1975). *Speech Style and Social Evaluation*. London: Academic Press, 154 – l70.)

Gilles, P. (2005). *Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin/New York: de Gruyter.

Goldinger, S. & Van Summers, W. (1989). Lexical neighborhoods in speech production: a first report. *Research on Speech Perception Progress Report. Speech Research Laboratory, Indiana University*, 15, 331 – 342.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology.* Oxford: Blackwell.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., & Trudgill, P. (2004). *New Zealand English: its origins and evolution*. Cambridge: Cambridge University Press.

Gussenhoven, C. & Jacobs, H. (1998). *Understanding phonology.* London [et al.]: Arnold.

Greenberg, J., Ferguson, C., & Moravscik, E. (eds.). (1978). *Universals of human language.* Stanford: Stanford University Press.

Hall, T. (1992). *Syllable Structure and Syllable-related Processes in German*. Tübingen: Niemeyer. (Linguistische Arbeiten ; 276)

Hall, T. (2000). *Phonologie: Eine Einführung*. Berlin/New Yourk: de Gruyter.

Harrington, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In: J. Cole and J. Hualde (eds.), *Laboratory Phonology 9*. Berlin: de Gruyter, 125 – 143.

Harrington, J. (2010). *The Phonetic Analysis of Speech Corpora*. Oxford: Wiley-Blackwell.

Harrington, J., Hoole, P., Kleber, F., & Reubold, U. (2011). The physiological, acoustic, and perceptual basis of high back vowel fronting: Evidence from German tense and lax vowels. *Journal of Phonetics,* 39, 121-131.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: an acoustic and perceptual study. *Journal of the Acoustical Society of America*, 123, 2825 – 2835.

Harrington, J., Palethorpe, S., & Watson, C. (2000a). Does the Queen speak the Queen's English? *Nature*, 408, 927 – 928.

Harrington, J., Palethorpe, S., & Watson, C. (2000b). Monophthongal vowel changes in Received Pronunciation: An Acoustic Analysis of the Queen's Christmas Broadcasts. *Journal of the International Phonetic Association*, 30, 63 – 78.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31, 373 – 405.

Hawkins, S. & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, 35, 183 – 199.

Hawkins, S. & Nguyen, N. (2003). Effects on word recognition of syllable-onset cues to syllable-coda voicing. In: J. Local, R. Ogden, and R. Temple (eds.), *Papers in Laboratory Phonology 6.* Cambridge: Cambridge University Press, 38 – 57.

Hawkins, S. & Nguyen, N. (2004). Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics*, 32, 199 – 231.

Hay, J., Pierrehumbert, J., & Beckman, M. (2003). Speech perception, well-formedness and the statistics of the lexicon. In: J. Local, R. Ogden, and R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology 6*. Cambridge: Cambridge University Press, 58 – 74.

Heike, G. (1970). Lautdauer als Merkmal der wahrgenommenen Quantität, Qualität und Betonung im Deutschen. *Proc. of the VIth International Congress of Phonetic Sciences, Prague*, 1967. The Hague/Paris: Mouton, 725 – 729.

Hinderling, R. (1980). Lenis und Fortis im Bairischen. Versuch einer morphophonemischen Interpretation. *Zeitschrift für Dialektalogie und Linguistik*, 47, 25 – 51.

Hirschfeld, U. (1999). Phonetische Merkmale des Sächsischen und das Fach Deutsch als Fremdsprache. In: B. Skibitzki and B. Wotjak (eds.), *Linguistik und Deutsch als Fremdsprache. Festschrift für Gerhard Helbig zum 70. Geburtstag.* Tübingen: Niemeyer, 109 – 120.

Hoole, P. & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In: P. Auer, P. Gilles, and H. Spiekermann (eds.), *Silbenschnitt und Tonakzente.* Tübingen: Niemeyer.

Hoole, P., Mooshammer, C., & Tillmann, H.G. (1994). Kinematic analysis of vowel production in German. *Proc. ICSLP 94, Yokohama*, 53 – 56.

Hulst, H. van der. (1985). Ambisyllabicity in Dutch. In: H. Bennis and F. Beukema (eds.), *Linguistics in the Netherlands.* Dordrecht: Foris, 57 – 67.

Hyman, L. (1976). Phonologization. In: A. Juilland, A. Devine, and L. Stephens (eds.), *Linguistic studies offered to Joseph Greenberg, vol. II*. Saratoga: Anma Libri, 407 - 418

Iivonen, A. (1996). Unterschiede des Vokalismus und der Intonation in der ostmitteldeutschen und wienerdeutschen Standardaussprache. In: E.-M. Krech and E. Stock (eds.), *Beiträge zur deutschen Standardaussprache.* Hanau/Halle: Dausien, 225 – 240.

Itô, J. (1986). Syllable theory in prosodic phonology. PhD dissertation, University of Massachusetts at Amherst.

Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory*, 7, 217 – 259.

Iverson, G. & Salmons, J. (2006). On the Typology of Final Laryngeal Neutralization: Evolutionary Phonology and Laryngeal Realism. *Theoretical Linguistics*, 32, 205 – 216.

Iverson, G. & Salmons, J. (2007). Domains and directionality in the evolution of German final fortition. *Phonology*, 24, 121 – 145.

Janson, T. (1983). Sound change in perception and production. *Language*, 59, 18 – 34.

Janson, T. (1986). Sound change in perception: an experiment. In: J. Ohala and J. Jaeger (eds.), *Experimental phonology.* Orlando: Academic Press, 253 – 260.

Janson, T. & Schulmann, R. (1983). Non-distinctive features and their use. *Journal of Linguistics*, 19, 321 – 336.

Jassem, W. & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, 17, 317 – 325.

Jessen, M. (1993). Stress conditions on vowel quality and quantity in German. *Working Papers of the Cornell Phonetics Laboratory*, 8, 1 – 27.

Jessen, M. (1998). *Phonetics and Phonology of Tense and Lax Obstruents in German.* Philadelphia: Benjamins.

Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In: T. Hall (ed.), *Distinctive Feature Theory*. Berlin: de Gruyter.

Jessen, M. (2002). Spectral balance in German and its relevance for syllable cut theory. In: P. Auer, P. Gilles, and H. Spiekermann (eds.), *Silbenschnitt und Tonakzente.* Tübingen: Niemeyer.

Jessen, M. & Ringen, C. (2002). Laryngal features in German. *Phonology*, 19, 189 – 218.

Jessen, M., Marasek, K., Schneider, K., & Claßen, K. (1995). Acoustic correlates of word stress and tense/lax opposition in the vowel system of German. *Proc. of the XIIIth International Congress of Phonetic Sciences*, *Stockholm*, 428 – 431.

John, T. (2004). *Eine akustische Analyse der Lenis/Fortis-Opposition in Varietäten des Sächsischen.* Unpublished MA thesis, University of Kiel.

Jongman, A., Sereno, J. A., Raaijmakers, M., & Lahiri, A. (1992). The phonological representation of [voice] in speech perception, *Language and Speech*, 35, 137 – 152.

Jørgensen, H. (1969). Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantfrequenzen. *Phonetica*, 19, 217 – 245.

Jun, S.-A. (1993). *The Phonetics and Phonology of Korean Prosody*. Unpublished PhD dissertation, Ohio State University.

Keating, P. (1979). *A Phonetic Study of a Voicing Contrast in Polish*. Unpublished PhD dissertation, Brown University.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In: J. Local, R. Ogden, and R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology 6*. Cambridge: Cambridge University Press, 143 – 161.

Kerswill, P. (1986). *A sociolinguistic study of rural immigrants in Bergen, Norway*. PhD dissertation, University of Cambridge.

Kerswill, P. (2003). Dialect levelling and geographical diffusion in British English. In: D. Britain and J. Cheshire (eds.), *Social dialectology. In honour of Peter Trudgill*. Amsterdam: Benjamins, 223 – 243.

Kiparsky, P. (1982). Lexical phonology and morphology. In: I.-S. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hanshin, 3 – 91.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonological Yearbook*, 2, 85 – 138.

Kleber, F., Harrington, J., & Reubold, U. (in press). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*. (doi:10.1177/0023830911422194)

Kleber, F., John, T., & Harrington, J. (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 38, 185 – 196.

Kloeke, W. van Lessen. (1982). *Deutsche Phonologie und Morphologie: Merkmale & Markiertheit.* Tübingen: Niemeyer.

Kluender, K., Diehl, R., & Wright, B. (1988). Vowel-length differences before voiced and voiceless consonants. *Journal of Phonetics*, 16, 153 – 169.

Kohler, K. (1977). The production of plosives. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel*, 8, 30 – 110.

Kohler, K. (1979). Dimensions in the perception of fortis and lenis plosives. *Phonetica*, 36, 332 – 343.

Kohler, K. (1982). F0 in the production of lenis and fortis plosives. *Phonetica*, 39, 199 – 218.

Kohler, K. (1984). Explanation in phonology: the feature fortis/lenis. *Phonetica*, 41, 150 – 174.

Kohler, K. & Künzel, H. (1978). The temporal organization of closing-opening movements for sequences of vowels and plosives in German. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel*, 10, 117 – 165.

Kohler, K. (1995) *Einführung in die Phonetik des Deutschen.* 2nd edition. Berlin: Schmidt.

König, W. & Renn, M. (2005). *Kleiner Bayerischer Sprachatlas.* München: DTV.

Kopkalli, H. (1993). *A Phonetic and Phonological Analysis of Final Devoicing in Turkish*. Unpublished PhD dissertation, University of Michigan.

Kroos, C., Hoole, P., Kühnert, B., & Tillmann, H.-G. (1997). Phonetic evidence for the phonological status of the tense-lax distinction in German*. Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation, München*, 35, 17 – 25.

Kubozono, H. (2002). Temporal neutralization in Japanese. In: C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7*. Cambridge: Cambridge University Press, 171 – 202.

Kufner, H. (1964). *München*. Göttingen: Vandenhoeck & Ruprecht. (Lautbibliothek der deutschen Mundarten ; 35)

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening effects on the duration and voicing assimilation of German fricatives. *Journal of Phonetics*, 35, 301 – 320.

Labov, W. (1991). The three dialects of English. In: P. Eckert (ed.), *Quantitative Analyses of Sound Change*. New York: Academic Press, 1 – 44

Labov, W. (1994) *Principles of Linguistic Change. Vol. 1: Internal Factors*. Oxford: Basil Blackwell.

Labov, W. (2001). *Principles of Linguistic Change. Vol. II: Social Factors*. Oxford: Blackwell.

Labov, W. & Ash, S. (1997). Understanding Birmingham. In: C. Bernstein, T. Nunnally, and R. Sabino (eds.), *Language Variety in the South Revisited*. Tuscaloosa: University of Alabama, 508 – 573.

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: Phonology and Phonetics*. Berlin: de Gruyter.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A Quantitative Study of Sound Change in Progress.* Philadelphia: U. S. Regional Survey.

Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Laeufer, C. (1992). Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics*, 20, 411 – 440.

Lameli, A. (2004). Hierarchies of dialectal features in a diachronic view – implicational scaling of real time data. In: B. Gunnarsson [et al.] *Language Variation in Europe. Papers from the 2nd Int. Conf. Language Variation in Europe*, *Uppsala*, 253 – 266.

Langstrof, C. (2006). Acoustic evidence for a push-chain shift in the Intermediate Period of New Zealand English. *Language Variation and Change*, 18, 141 – 164.

Lass, R. (1997). *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51, 2018 – 2024.

Lindblom, B. (1968). Temporal organisation of syllable production. *Quarterly Progress and Status Report*, *Speech Transmission Laboratory, Stockholm*, 1 – 5.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In: W. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modelling.* Dordrecht: Kluwer, 403 – 39. (NATO ASI Series D: Behavioural and Social Sciences ; 55)

Lindblom, B., Guion, S., Hura, S., Moon, S., & Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7, 5 – 37.

Lombardi, L. (1999). Positional faithfulness and voicing assimilation in optimality theory. *Natural Language and Linguistics Theory*, 17, 267 – 302.

Luce, P. & Pisoni, D. (1998). Recognizing spoken words: the neighbourhood activation model. *Ear & Hearing*, 19, 1 – 36.

Maclagan, M. & Hay, J. (2007). Getting *fed* up with our *feet*: Contrast maintenance and the New Zealand English 'short' front vowel shift. *Language Variation and Change*, 19, 1 – 25.

Maclagan, M., Watson. C., Harlow, R., King, J., & Keegan, P. (2009). /u/-fronting and /t/ aspiration in Māori. *Language Variation and Change*, 21, 175 – 192.

Maddieson, I. (1997). Phonetic universals. In: W. Hardcastle and J. Laver (eds.), *The Handbook of Phonetic Sciences*. Oxford: Blackwell, 619 – 639.

Mangold, M. (ed.) (1990). *Duden: Aussprachewörterbuch der deutschen Standardaussprache*. 3th edition. Mannheim: Dudenverlag.

Mann, V. & Repp, B. (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics*, 28, 213 – 228.

Mansell, P. (1979). The articulation of German plosives. *Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation der Universität München*, 11, 1 – 207.

Manuel, S. (1995). Speakers nasalize /o/ after /n/, but still hear /o/. *Journal of Phonetics*, 23, 453 – 476.

Martinet, A. (1955). *Économie des changements phonétiques*. Bern: A. Francke.

McMahon, A. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.

Meyer, E. (1904). Zur Vokaldauer im Deutschen. *Nordiska Studier tillegnade A. Noreen*. Uppsala: K. W. Appelbergs Boktryckeri, 347 – 356.

Meyer, E. (1911). *Untersuchungen über Lautbildung*. Marburg: Elwertsche Verlagsbuchhandlung.

Meyer, E. (1913a). Das Problem der Vokalspannung. *Die neueren Sprachen*, 21, 65 – 86.

Meyer, E. (1913b). Das Problem der Vokalspannung. *Die neueren Sprachen*, 21, 145 – 171.

Mitleb, F. (1984). Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics*, 12, 23 – 27.

Mohanan, K. (1986). *Lexical Phonology*. Dordrecht: Foris.

Mooshammer, C. (1998). Experimentalphonetische Untersuchungen zur artikulatorischen Modellierung der Gespanntheitsopposition im Deutschen. *Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation, München*, 36, 3 – 191.

Müller, V. (2010). *Neutralisierung des Stimmhaftigkeitskontrasts im ostfränkischen Dialekt.* Unpublished MA thesis, University of Munich.

Munson, B. (2007). Lexical access, lexical representation, and vowel articulation. In: J. Cole and J. Hualde (eds.), *Laboratory Phonology 9.* New York: de Gruyter, 201 – 228.

Myers, S. & Hansen, B. (2005). The origin of vowel length neutralisation in vocoid sequences: Evidence from Finnish speakers. *Phonology*, 22, 317 – 344.

Nooteboom, S. (1972). *Production and Perception of Vowel Duration: A Study of Durational Properties in Dutch*. PhD dissertation, University of Utrecht.

O'Dell, M. & Port, R. (1983). Discrimination of word-final voicing in German. *Journal of the Acoustical Society of America,* 73(S1), S31(A).

Ohala, J. (1974). Experimental historical phonology. In: J. Anderson and C. Jones (eds.), *Historical linguistics II. Theory and description in phonology. Proc. of the 1st Int. Conf. on Historical Linguistics*, *Edinburgh*, *1973*. Amsterdam: North Holland, 353 – 389.

Ohala, J. (1981). The listener as a source of sound change. In: C. Masek, R. Hendrick, and M. Miller (eds.), *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Linguistic Society, 178 – 203.

Ohala, J. (1983). The direction of sound change. In: A. Cohen and M.P.R. van den Broecke (eds.), *Abstracts of the 10th Int. Congress of Phonetic Sciences, Utrecht.* Dordrecht: Foris, 253 – 258.

Ohala, J. (1989). Sound change is drawn from a pool of synchronic variation. L. Breivik and E. Jahr (eds.), *Language Change: Contributions to the Study of its Causes.* Berlin: de Gruyter, 173 – 198.

Ohala, J. (1990). The phonetics and phonology of aspects of assimilation. In: J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech.* Cambridge: Cambridge University Press, 258 – 275.

Ohala, J. (1993). The phonetics of sound change. In: C. Jones (ed.), *Historical Linguistics: Problems and Perspectives.* London: Longman, 237 – 278.

Ohala, J. (2003). Phonetics and historical phonology. In: B. Joseph and R. Janda (eds.), *Handbook of Historical Linguistics*. Malden/Oxford: Blackwell, 669 – 686.

Parker, F. & Walsh, T. (1981). Voicing cues as a function of the tense/lax distinction in vowels. *Journal of Phonetics*, 9, 353 – 358

Peterson, G. & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175 – 184.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In: J. Bybee and P. Hopper (eds.), *Frequency and the Emergence of Lexical Structure*. Amsterdam: Benjamins, 137 – 157.

Pierrehumbert, J. (2002). Word-specific phonetics. In: C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7*. Berlin: de Gruyter, 101 – 139.

Pierrehumbert, J. (2003). Probabilistic phonology: discrimination and robustness. In: R. Bod, J. Hay, and S. Jannedy (eds.), *Probabilistic Linguistics*. Cambridge: MIT Press, 177 – 228.

Piroth, H.G. & Janker, P. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32, 81 – 109.

Pitt, M. & McQueen, J. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, 39, 347 – 370.

Pompino-Marschall, B. (1995). *Einführung in die Phonetik.* Berlin/NewYork: de Gruyter.

Port, R. (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America*, 69, 262 – 274.

Port, R. & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257 – 282.

Port, R. & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455 – 471.

Port, R., Mitleb, F., & O'Dell, M. (1981). Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America*, 70, S13, F10.

Ramers, K. (1988). *Vokalquantität und –qualität im Deutschen.* Tübingen: Niemeyer. (Linguistische Arbeiten ; 213)

Rankin, R. (2003). The comparative method. In: B. Joseph and R. Janda (eds.), *Handbook of Historical Linguistics*. Malden/Oxford: Blackwell, 183 – 212.

Restle, D. (1998). *Silbenschnitt – Quantität – Kopplung. Zur Geschichte, typologischen Einordnung und Repräsentation der Anschlußprosodie unter dem Blickwinkel einer Oszillationssilbentheorie.* PhD dissertation, University of Munich.

Restle, D. (2002). Normierung der Silbenqualität. Ein typologischer Beitrag zur Charakteristik des Silbenschnitts in und außerhalb der Germania. In: P. Auer, P. Gilles, and H. Spiekermann (eds.), *Silbenschnitt und Tonakzente.* Tübingen: Niemeyer, 35 – 66. (Linguistische Arbeiten ; 463)

Restle, D. & Mooshammer, C. (1999). The influence of the tense-lax contrast in vowels on the production of post-vocalic consonants in Standard German. *Proc. of the 14th International Congress of Phonetic Sciences, San Francisco*, 531 – 534.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: a longitudinal analysis in adult speakers. *Speech Communication*, 52, 638 – 651.

Rowley, A. (1990a). East Franconian. In: C. Russ (ed.), *The dialects of Modern German*. London: Routledge, pp. 394 – 416.

Rowley, A. (1990b). North Bavarian. In: C. Russ (ed.), *The dialects of Modern German*. London: Routledge, pp. 417 – 437.

Rubach, J. (1990). Final devoicing and cyclic syllabification in German. *Linguistic Inquiry*, 21, 79 – 94.

Rues, B., Redecker, B., Koch, E., Wallraff, U., & Simpson, A. (2007). *Phonetische Transkription des Deutschen*. Tübingen: Narr.

Schaeffler, F. (2005). Phonological quantity in Swedish dialects: typological aspects, phonetic variation and diachronic change. *Phonum*, 10, 1 – 142.

Scheutz, H. (1983). Quantität und Lenis/Fortis im Mittelbairischen. In: P. Wiesinger (ed.), *Beiträge zur bairischen und ostfränkischen Dialektologie. Ergebnisse der Zweiten Bairisch-Österreichischen Dialektologentagung*, *Wien.* Göppingen: Kümmerle Verlag, 13 – 33.

Schiel, F. (2004). MAUS Goes Iterative. *Proc. of the IV. International Conference on Language Resources and Evaluation, Lisbon*, 1015 – 1018.

Schlobinski, P. (1987). *Stadtsprache Berlin. Eine soziolinguistische Untersuchung.* Berlin/New York: de Gruyter.

Schmidt, W. (2000). *Geschichte der deutschen Sprache.* 8th edition. Stuttgart: S. Hirzel.

Sievers, E. (1872). *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen.* Reprint 1980. Hildesheim/New York: Olms.

Slowiaczek, L. & Dinnsen, D. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13, 325 – 341.

Slowiaczek, L. & Szymanska, H. (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics*, 17, 205 – 212.

Sluijter, A. (1995). *Phonetic Correlates of Stress and Accent*. The Hague: Holland Academic Graphics.

Sluijter, A. & van Heuven, V. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100, 2471 – 2485.

Spencer, A. (1996). *Phonology: Theory and Description.* Malden, USA / Oxford, UK: Blackwell.

Spiekermann, H. (2000). *Silbenschnitt in deutschen Dialekten*. Tübingen: Niemeyer. (Linguistische Arbeiten ; 425)

Spiekermann, H. (2002). Ein akustisches Korrelat des Silbenschnitts: Formen des Intensitätsverlaufs in Silbenschnitt- und Tonakzentsprachen. In: P. Auer, P. Gilles, and H. Spiekermann (eds.), *Silbenschnitt und Tonakzente.* Tübingen: Niemeyer. (Linguistische Arbeiten ; 463)

Stathopoulos, E. & Weismer, G. (1983). Closure duration of stop consonants. *Journal of Phonetics*, 11, 395 – 400.

Steele, S. (1986). Interaction of vowel F0 and prosody. *Phonetica*, 43, 92 – 105.

Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. UCLA manuscript copy at: http://www.humnet.ucla.edu/humnet/linguistics/people/steriade/steriade.htm

Steriade, D. (1999). Alternatives to the syllabic interpretation of consonantal phonotactics. In: O. Fujimura, B. Joseph, and B. Palek (eds.) *Proc. of the 1998 Linguistics and Phonetics Conference*. Prague: The Karolinum Press, 205 – 242.

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. In: M. Broe and J. Pierrehumbert (eds.), *Papers in Laboratory Phonology 5.* Cambridge: Cambridge University Press, 313 – 334.

Strange, W. & Bohn, O. (1998). Dynamic specification of coarticulated German vowels: perceptual and acoustical studies. *Journal of the Acoustical Society of America*, 104, 488 – 504.

Torgersen, E. & Kerswill, P. (2004). Internal and external motivation in phonetic change: dialect levelling outcomes for an English vowel shift. *Journal of Sociolinguistics*, 8, 24 – 53.

Torreira, F. & Ernestus, M. (2011). Vowel elision in casual French: the case of vowel /e/ in the word c'était. *Journal of Phonetics*, 39, 50 – 58.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88, 97 – 100.

Trubetzkoy, N. (1938). Die phonologischen Grundzüge der sogenanten "Quantität" in den verschiedenen Sprachen. In: *Scritti in onore di Alfredo Trombetti.* Milano: Ulrico Hoepli ditore, 155 – 174.

Trubetzkoy, N. (1939/1958): *Grundzüge der Phonologie*. 2nd edition. Göttingen: Vandenhoeck & Ruprecht.

Trudgill, P. (1986). *Dialects in Contact.* Oxford: Basil Blackwell.

Veith, W. (2005). *Soziolinguistik: Ein Arbeitsbuch*. 2nd edition. Tübingen: Narr.

Vennemann, T. (1972). On the theory of syllabic phonology. *Linguistische Berichte*, 18, 1 – 18.

Vennemann, T. (1982). Zur Silbenstruktur der deutschen Standardsprache. In: T. Vennemann (ed.), *Silben, Segmente, Akzente.* Tübingen: Max Niemeyer, 261 – 305. (Linguistische Arbeiten ; 126)

Vennemann, T. (1991). Syllable structure and syllable cut prosodies in Modern Standard German. In: P. M. Bertinetto, M. Kenstowicz, and M. Loporcaro (eds.), *Certamen Phonologicum II: Papers from the 1990 Cortona Phonology Meeting*, *Torino*, 211 – 243.

Vitevitch, M., Luce, P., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactic and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, 40, 47 – 62.

Vitevitch, M. & Luce, P. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374 – 408.

Wagener, P. (2002). German dialects in real time change. *Journal of Germanic Linguistics,* 14, 271 – 285.

Warner, N., Good, E., Jongman, A., & Sereno, J. (2006). Orthographic versus morphological incomplete neutralization effects. *Journal of Phonetics*, 34, 285 – 293.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251 - 276.

Watson, C., Maclagan, M., & Harrington, J. (2000). Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change*, 12, 51 - 68.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics.* Austin: University of Texas Press, 95 – 188.

Weiss, R. (1976). *The perception of vowel length and quality in German: an experimental-phonetic investigation.* Hamburg: Buske. (Hamburger phonetische Beiträge ; 20)

Wells, J. (1982). *Accents of English, vol. 1: An introduction.* Cambridge: Cambridge University Press.

Werker, J. (1995). Age-related changes in cross-language speech perception: standing at the crossroads. In: W. Strange (ed.), *Speech Perception and Linguistic Experienc*e. Baltimore: York Press, 155 – 169.

Werker, J. & Tess, R. C. (1998). Influences on infant speech processing: toward a new synthesis. *Annual Review of Psychology*, 50, 509 – 535.

Whalen, D. & Levitt, A. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, 23, 349 – 366.

Whalen, D., Gick, B., Kumada, M., & Honda, K. (1999). Cricothyroid activity in high and low vowels: Exploring the automaticity of intrinsic F0. *Journal of Phonetics*, 27, 125 – 142.

Wiese, R. (1988). *Silbische und lexikalische Phonologie: Studien zum Chinesischen und Deutschen.* Tübingen: Niemeyer. (Linguistische Arbeiten ; 211).

Wiese, R. (1996). *The Phonology of German*. Oxford: Oxford University Press.

Wiesinger, P. (1990). The Central and Southern Bavarian Dialects in Bavaria and Austria. In: C. Russ (ed.), *The dialects of Modern German.* London: Routledge, 438 – 519.

Wright, R. (2003). Lexical competition and reduction in speech. In J. Local, R. Ogden, and R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology 6*. Cambridge: Cambridge University Press, 75 – 87.

Wurzel, W. (1970). Studien zur deutschen Lautstruktur. *Studia grammatica*, 8. Berlin: Akademie Verlag.

Yu, S.-T. (1992). *Unterspezifikation in der Phonologie des Deutschen*. Tübingen: Niemeyer. (Linguistische Arbeiten ; 274)

Zwirner, E. (1962). Phonometrischer Beitrag zur Geographie der prosodischen Eigenschaften. *Proc. of the IVth International Congress of Phonetic Sciences, Helsinki*, 1961. The Hague: Mouton, 500 – 518.