The Ongoing (Un)merging of Stops in the Swabian Dialect of Isny im Allgäu

Robert Felty
University of Michigan
Qualifying Research Paper

TABLE OF CONTENTS

1. INTRODUCTION

Most standard handbooks place the Swabian dialect of German in the area affected by the *binnenhochdeutsche Konsonantenschwächung* ‘internal High German consonant lenition’, resulting in a neutralization of /b d/ and /p t/. I see this as a byproduct of the High German consonant shift, in which fortis consonants /p t k/ became affricates /pf ts kx/ or fricatives /f s x/ (but the shift from /k/ to /x/ does not occur in initial position, and likewise /g/ and /k/ are not neutralized). According to this lenition, one would expect Standard German (SG) [pʰakən] “pack” and [bakən] ”bake” to both be realized as [pakən] in Swabian. As evidence of this neutralization the *schwäbisches Wörterbuch* (SW) ‘Swabian dictionary’ (Fischer 1904-1936) lists these letters together (but not <g> and <k>). In spite of the combined listing, the phonetic transcriptions in the SW show that while the great majority of words spelled with <b p> are lenis (unaspirated), several words are listed as fortis (aspirated). Most of the words listed with a fortis pronunciation are loan words, borrowed either from other languages, dialects, or from SG. For the purposes of this study, the relevant loan words are those borrowed after the completion of the High German consonant shift, i.e. after around 900 a.d. The wordlist in Appendix A includes pronunciations from the SW and dates of first attestation according to Kluge’s (2002) etymological dictionary. The fortis stops in Germanic words have almost all affricated or spirantized during the High German consonant shift (from the 7th to the 9th century). Words with fortis stops entering the language after this period have not shifted. These loan words will either maintain the fortis stops as exceptions to the dialect, or the stops will become deaspirated, and merge with the lenis stops. It appears that newer loan words tend to remain fortis, whereas older words have undergone lenition, but there does not seem to be a precise cut-off point in time. A pilot study of one speaker from Isny im Allgäu confirmed these hypotheses.

In order to gain a better understanding of the distribution of fortis and lenis stops in Swabian, I carried out an acoustic study of the speech of 18 speakers from the town of Isny im Allgäu. Though Swabian is widely documented, few studies of Swabian have included an acoustic analysis. As I will show, an acoustic analysis provides a more detailed and accurate account of the phenomenon in question. The acoustic parameter of voice-onset time (VOT) was used to quantify the fortis/lenis (voiced/voiceless) distinction. Though VOT is not the only acoustic parameter of the fortis/lenis distinction, it has been used in numerous studies (e.g. Lisker & Abramson 1964, Cho & Ladefoged 1999), and has been shown to be more perceptually salient than other parameters such as f_0 perturbation or F1 transition (Lisker 1975, Benki 2001). In the present study, each speaker was recorded reading a wordlist, and the closure duration and VOT of each token was measured. The main results show that the words used can be grouped into three categories: (1) low mean VOT with little variation, (2) high mean VOT with large variation, and (3) low mean VOT with large variation. It is this third group which is most interesting. The combination of a low mean VOT with a large variation suggests a great amount of inter- and intra-speaker variation. There are also words that are fairly consistently pronounced with a high VOT, yet occasionally are pronounced with a low VOT. Such examples are further evidence for the variation in this dialect. These findings shed light on what seems to be a current change in progress in the Swabian dialect, that the contrast between fortis and lenis stops, which was lexically specific before, is spreading throughout the entire lexicon. In my analysis I focus on this variation, particularly on the words exhibiting a low mean VOT with a large standard deviation.

In section §2 I discuss the historical background for this phenomenon, including a review of literature treating the High German consonant shift; I analyze the lenition process as a byproduct of the High German consonant shift. In section §3 I discuss various aspects of the production and perception of fortis and lenis consonants. In section §4 I outline the methods used for collecting and measuring the data. In section §5 I present the results of the study and discuss their relevance. Section §6 includes a brief recapitulation of the paper, and comments on further study.

2. Historical Background

This section includes a brief description of the language situation in Isny im Allgäu and the characteristic linguistic features of Swabian, as well as the historical development of the High German consonant shift and the subsequent lenition.

2.1. Isny im Allgäu

This study investigates the dialect of Swabian as spoken in Isny im Allgäu. Isny is a small city in the southeastern corner of the state of Baden-Württemberg, with approximately 14,000 inhabitants. It was founded in the 11th century, and was at one time part of the kingdom of Swabia. In 1806 the border between Bavaria and Baden-Württemberg was firmly established, which has led to a sharp distinction between the Swabian dialect of Isny and the Bavarian dialects spoken just a few kilometers away. Approximately 50 kilometers from Switzerland, the dialect of Isny purportedly also has some influence from Swiss dialects, e.g. the retention of older monophthongs in words such as [hus] for SG *Haus* [haʊs]; I found little evidence of this however. Since the collapse of the linen industry in the Allgäu region in the 16th century, dairy farming has dominated the region. Since the 1950s Isny im Allgäu has also been a tourist destination, officially classified as a ‘clean air spa resort’. These various factors distinguish the dialect of Isny im Allgäu from that of a rural dialect; however, most speakers seem to agree that towns as close as 10-20 kilometers away have noticeably different dialects. Given this situation, special care was taken to find speakers who grew up in the close vicinity of Isny.

2.2. A Brief linguistic description of Swabian

Most traditional dialect grammars of German trace the dialect's development from Middle High German (MHG), the common ancestor to most southern German dialects. While such comparisons can be useful, Berroth (2001: 96) makes the perspicuous observation that current dialect speakers do not have access to an internal MHG grammar, while most of them do have access to an internal Standard German (SG) grammar; that is, most current dialect speakers in southern Germany are to a varying extent bidialectal, and are cognizant of many of the differences between their dialect and SG. Berroth shows that influence from SG can have an effect on the dialect, particularly in the case of over-generalizing distinctions between the dialect and SG. Berroth gives an example from the Swabian dialect of Ruppertshofen in which younger speakers are starting to merge the diphthongs [uɪ] and [eɪ], which are also merged in SG (2001: 102-103, 110).[1]

SG	MHG	Ruppertshofen	English gloss
[ɔɪ] *Kr<u>eu</u>zworträtsel*	[iʊ]	[uɪ, eɪ] [gʀ<u>eɪ</u>tsvɔɐtʀɛːsl]	crossword puzzle
[ɔɪ] *L<u>eu</u>te*	[y]	[eɪ] [l<u>eɪ</u>t]	people
[ɔɪ] *H<u>eu</u>schrecke*	[øʊ]	[aɪ] [h<u>aɪ</u>]	grasshopper

These three MHG categories have all been merged in SG, but have been kept distinct in the dialect of Ruppertshofen. However, the first category in Ruppertshofen seems to be merging

[1] For the sake of clarity, I have converted all transcriptions into their closest IPA equivalents. Throughout the paper square brackets [ ] indicate phonetic transcriptions, slashes / / indicate phonemic descriptions (including hypothesized pronunciations for which no phonetic transcriptions are available, e.g. OHG), angled brackets < > indicate spellings, and backslashes \ \ indicate pronunciations listed in the SW.

with the second; some older speakers still use [uɪ], but younger speakers all use [eɪ]. This cannot be explained as a direct influence from SG, since SG has [ɔɪ]. It can be explained by way of analogy with SG though. Let us hypothetically consider a child learning both the Ruppertshofen dialect and SG simultaneously. The child may very well notice that SG [ɔɪ] often corresponds to Ruppertshofen [eɪ], and fail to notice that it sometimes corresponds to [uɪ]. Berroth does not discuss why the older three-way distinction is collapsing into a two-way distinction, and not collapsing into one phoneme altogether as is the case for SG. It could be that this is due to lexical frequency, though I know of no measures of lexical frequency for MHG or this particular dialect. Some of the data from Isny also shows similar patterns.

In order to better understand the fortis/lenis distinction in question, a comparison of common features of Swabian with SG should prove to be fruitful. By informally investigating to what degree speakers exhibit other characteristics of Swabian, we can determine whether the speakers are shifting towards SG in all respects, or only with regards to several features. Some characteristic differences in which Swabian differs from SG are:

(1) [ʃ] in syllable coda clusters as well onset clusters – cf. SG *Post* [pʰost] 'mail' vs. Swabian [poʃt]

(2) Loss of final /n/ – cf. SG *backen* [bakən] 'to bake' vs. Swabian [bakə]

(3) Reduction or loss of past participle ge- prefix – cf. SG *gebacken* [gəbakən] 'baked' vs. Swabian [bakə], SG *gespielt* [gəʃpilt] 'played' vs. Swabian [kʃpilt]

(4) Unrounding of front rounded vowels – cf. SG *schön* [ʃøn] 'beautiful' vs. Swabian [ʃe]

(5) Retention of MHG diphthongs *ie*, *uo* – cf. SG *lieben* [libən] ‘to love’ vs. Swabian [liəbə];

SG *Gruß* [gʀus] ‘greeting’ vs. Swabian [gʀuəs]

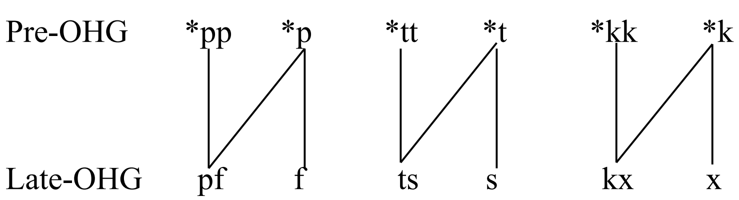
(6) Lenition (de-aspiration) of initial /p t/ - cf. SG *Tag* ‘day’ [tʰak] vs. Swabian [tak]

I will refer back to some of these features in the results and discussion sections.

2.3. The High German Consonant Shift

The lenition evident in many High German dialects can be seen as a byproduct of the High German consonant shift. The High German consonant shift was first described by Jacob Grimm in 1822. Grimm found that Germanic /pp p/, /tt t/, /kk k/ shifted to /pf/, /ts/, /kx k/ in initial position and in medial and final position to /pf f/, /ts s[2]/, /x ç/ in OHG. This shift, often referred to as the fortis shift,[3] can be seen in Figure 1. An additional shift, known as the lenis shift, is also reported, in which /b v/ ; /d/ ; /g ɣ/ shift to /p b/ ; /t/ ; /k g/. It is the failure of the lenis shift which results in lenition. This will be discussed separately in §2.4.

Figure 1. The High German Consonant Shift (the fortis shift)



Various theories of the cause of the High German consonant shift exist, but only the relevant facts will be handled here. In Pre-OHG there were two distinctions between obstruents– geminate/simplex, and fortis/lenis (Naiditsch 1997:256). In the High German consonant shift geminates /pp tt kk/ became affricates, and fortis /p t k/ became affricates or fricatives; if the

[2] This sound is actually not [s], but some sound with unknown phonetic quality, which is transcribed as <ʒ>. In Old and Middle High German manuscripts <s> represents Gmc. **s*, whereas <ʒ> represents the shifted sound from Gmc. **t*. Since these sounds eventually merge in modern German, I have used /s/ for simplicity throughout the paper

[3] Traditional grammars (e.g. Braune 1987, Sonderegger 1987) use the terms *Tenues* and *Medien*. I follow Davis et al. (1999) in translating this as ‘fortis’ and ‘lenis’

fortes /p t k/ became affricates or fricatives, then it would seem that the distinctions /b/:/p/, /d/:/t/, /g/:/k/ were lost. This lack of distinction results in the consonant mergers we find in many southern German dialects.

The picture is a bit more complicated, though. It is also important to note the spread of the shift. Evidence from older texts and from modern dialects shows that the shift probably started somewhere in Switzerland and gradually moved northwards. The spread of the shift depends not only upon geography, however; it is also dependent upon word position and (most relevant here) place of articulation (POA). When examining evidence from SG and from modern dialects, we find different distributions of the shifted voiceless fortis stops. Generally we find that coronals are most likely to shift, followed by labials, and velars are least likely to shift. In fact, velars in initial position have shifted only in Swiss dialects, e.g. SG *Kind* 'kid' [kɪnt] vs. Swiss German [kxɪnt]. This distribution is illustrated in Table 1. If we see lenition as a byproduct of the High German consonant shift, then it follows that in instances where /k/ has failed to shift, the distinction between /g/ and /k/ would be maintained, which is exactly what is found.

Table 1. Spread of the High German Consonant Shift (adapted from Sonderegger 1987: 159). The dialects are arranged more or less in a north-south continuum.

	Coronal					**Labial**						**Velar**			
Pre-OHG	t-	-tt-	C+t	-t-	-t	p-	-pp-	mp	lp	rp	-p(-)	k-	-kk-	C+k	-k(-)
Old Saxon	t	tt	t	t	t	p	pp	mp	lp	rp	p	k	kk	k	k
Middle Franconian	ts	ts	ts	s	t/s	p	pp	mp	lp	rp	f(f)	k	kk	k	x
Rhine-Franconian	ts	ts	ts	s	s	p	pp	mp	lp/lpf	rp/rpf	f(f)	k	kk	k	x
SRhine-Franconian	ts	ts	ts	s	s	p	pf	mpf	lpf	rpf	f(f)	k	kk	k	x
East Franconian	ts	ts	ts	s	s	pf	pf	mpf	lpf	rpf	f(f)	k	kk	k	x
Bavarian	ts	ts	ts	s	s	pf	pf	mf	lf	rf	f(f)	kχ	kχ	kχ	x
Alemannic	ts	ts	ts	s	s	pf/f	pf/ff	mf	lf	rf	f(f)	x	kχ	x	x
Langobardian	ts	ts	ts	s(s)	s	p	p(p)	mpf	lpf	rpf	p/f(f)	k	kk	k/kχ	x

It is generally assumed that Germanic*p, *t, *k were voiceless aspirated fortes, and that *b, *d, *g were voiced unaspirated lenes. The road to the state of the shifted OHG consonants had several stages along the way. The first stage, according to Fourquet (1954: 12), was a change from the distinctive feature [voice] in Proto-Indo-European to [aspiration] in Germanic. These aspirated stops then eventually became affricates or fricatives by several other changes. Braune (1963: 84 [1886]) first introduced the idea that aspirated stops could be regarded as a sequence of two phonemes, i.e. /p/ and /h/, in order to account for the High German Consonant Shift.

Davis et al. (1999: 181-182) incorporate several different prior analyses into their account of the shift. They begin with Vennemann's (1988) Syllable Weight Law. According to Vennemann, there is a general tendency for stressed syllables to be longer than unstressed syllables; thus in words such as pre-OHG [ó.p^{h}an], the short vowel is stressed, which is in contrast to this tendency. In order to fit the pattern of syllable weight, the syllable must be lengthened, either by lengthening the vowel, or by creating a closed syllable.[4] The High German Consonant Shift adheres to the Syllable Weight Law by creating closed syllables. Therefore, following Braune 1987, Davis et al. analyze the aspirated [p^{h}] as two segments [ph], resulting in [óp.han]. The [h] is then strengthened and assimilated to [f], and finally the [p] is weakened to [f] to give OHG [óffan]. The shift spreads to initial and final positions later. They also provide evidence from the North-Rhenish dialect in Wermelskirchen, where the change has taken place only intervocalically, thus resulting in words such as [pɛfər] 'pepper' (SG [pfɛfər]). The shift does not even take place in verb paradigms, e.g. [esən, ɔ:t, jɛsən] 'eat, ate, eaten' (SG [ɛsən, a:s, gəgɛsən]).

[4] Vowel lengthening occurs in some other Germanic languages in this environment.

2.4. LENITION

Several other changes in the consonant system of High German dialects took place in the eighth to tenth centuries. The most notable of these is the so-called lenis shift. This involves the shifting of /b d g/ to /p t k/ after the fortis shift had taken place. As Kranzmayer notes (1956:76), the lenis shift occurred to varying degrees according to geography and POA, much like the fortis shift. He dates the alveolar shift to 750, covering the entire High German dialect area, while the labial shift can be first seen in 770, but only in Bavarian and Alemannic; in Alemannic, the shift appears to be reversed in the ninth century. The case of Alemannic undergoing and then reversing the shift from /b/ to /p/ within the span of one hundred years seems quite implausible. A much more likely story is that <p> spellings during this time are due to scribal peculiarities, or perhaps the texts found in the Alemannic area at this time were written by Bavarian monks. Even in the Bavarian dialects, <p> spellings are no longer found after 1050.

It is also possible that the distinction between /b d g/ and /p t k/ was simply allophonic at this point. Notker's *Anlautgesetz* 'onset law' points strongly toward this conclusion. Notker Labeo was a monk at the monastery at St. Gallen in the10th century. In his writings there is a systematic alternation between <b d g> and <p t k> depending on the environment in which they occur; <b d g> are found following vowels and sonorants, while <p t k> are found after obstruents and pauses. Richard Page (1999) notes the controversy over whether Notker's <b d g> are voiced or not, but declares this to be irrelevant to understanding the alternation, and concentrates only on the distinction between fortis and lenis. Page notes that this alternation takes place between consonants stemming from the same Germanic root, thus representing a change in progress – /p~b/ < **b*, /t~d/ < **þ*, /k~g/ < **g*.[5] Page remarks that the lenes are thus the

[5] Note that /t/ < **d* is always fortis.

unspecified consonants. This makes perfect sense, for if the old fortes have become affricates or fricatives, then the fortis/lenis distinction has been lost. Stephen Clausing (1980) discusses Notker's Law in more depth. One important conclusion he draws is that Notker's Law is found throughout the High German area, not just in Notker's writing. The second conclusion he draws is that Notker's Law is a manifestation of the incomplete lenis shift. The complementary distribution of Notker's spellings implies that the difference is representing allophonic variation of a single phoneme. Thus Notker could vary freely between the fortis and lenis representations without signaling a change in meaning.[6] Page argues against this analysis on the grounds that /t/ < **d* does not show this alternation. I do not see this as a problem however. It is certainly possible that the /t/ < **d* and the <t> spellings from **þ* had slightly different pronunciations, perhaps alveolar vs. dental. We find a similar situation in the separation between <s> from Gmc. **s* and <ʒ> Gmc. **t*. These graphs are kept separate throughout the Old and Middle High German periods, but eventually merge in modern German.

From these facts we can conclude that the lenis shift never took place in dialects such as Swabian.[7] Thus the process known as *die binnenhochdeutsche Konsonantenschwächung* 'the internal High German consonant lenition' is actually not a process at all, but rather a failure to undergo the lenis shift. What more appropriately can be termed lenition is the adaptation of loan words coming in after the completion of the High German consonant shift. This study quantitatively investigates the degree to which these loan words have been adapted.

---

6 Goblirsch (1997: 139) sees Notker's writings as one reason to claim that there is no phonological opposition between fortis and lenis stops in Upper German.

7 Another possibility would be to see this time of flux in orthography as a reanalysis of the fortis/lenis distinction, that is, a shift from a voiced/voiceless distinction to an unaspirated/aspirated distinction. However, most analyses of the High German consonant shift assume that this took place before the shift.

3. Phonetic Background

In this section I discuss some of the various theories for describing the fortis/lenis distinction, and theories about language change which are relevant to the variation I have found in Isny.

3.1. The Fortis/Lenis Distinction

Most traditional grammars of German use the terms fortis and lenis to distinguish between the obstruents /b d g z v/ and /p t k s f/. These terms were first introduced by the German dialectologist J. Winteler in 1876. In his investigation of the German dialect in the Swiss canton of Glarus, he found that the two opposing stop series differed neither in voicing nor aspiration, and therefore he proposed to use the terms fortis and lenis (22). He described fortis consonants as exhibiting more “expiratory and articulatory energy” and a longer duration than lenis consonants (25). Since then these terms have often been used more generally, also incorporating sounds which are distinguishable by voicing or aspiration. These terms are at the same time useful and problematic; they are useful when one does not wish to specify which articulatory or acoustic attributes are at play, but they are also problematic for the same reason. Though we know that phonological features can often subsume a variety of different articulatory and acoustic variables, many linguists have nevertheless attempted to find one variable which best defines a particular category. Many different features have been proposed as the true distinction between fortis and lenis consonants, which I will briefly review here.[8]

Voice: Voicing is produced by the periodic vibration of the vocal folds. Consonants produced with voicing during oral closure are considered [+voice]. While voicing during closure does provide a very accurate measure for the distinction between fricatives, it is much less reliable for stops, especially in most of the Germanic languages, in which initial stops are

[8] Here I discuss relevant literature for both English and German. In regards to the fortis/lenis distinction English and German are very similar

frequently not voiced during closure. According to Kohler (1977: 87), the German stops /b, d, g/ are only fully voiced intervocalically. (Docherty (1992) reports this for English as well).

Aspiration: Traditionally, aspiration has been categorized in many ways, from “a puff of air” upon release of closure to a following voiceless glottal approximant /h/. Iverson & Salmons (1995) base their definition on glottal spread, i.e. the width of the glottal opening. Kohler (1977: 160) describes the difference between /p t k/ and /b d g/ by means of Voice Onset Time (VOT), which is roughly the duration of aspiration. For fortes it is greater than for lenes in standard German.

Intra-oral pressure and muscular tension: Chomsky & Halle (1968: 324-325) also describe the feature tense/lax (essentially synonymous with fortis/lenis) as one involving heightened subglottal pressure and increased muscular tension. Tenseness arises from a build-up of pressure in the oral cavity, which, if it exceeds sub-glottal pressure, will not allow air to flow through the glottis, and thus inhibits voicing. Whereas in voiceless stops the vocal tract is rigid (tense), in voiced stops the vocal tract is relaxed, which creates a difference between intra-oral and subglottal pressure, thus allowing voice (but not necessitating it). A major difficulty in using this definition is that these parameters are very difficult (if not impossible) to measure accurately. Moreover, simply using muscular tension as a parameter is problematic for physiological reasons. While it may be true that the walls of the vocal tract are relaxed during voiced stops, there is also an active displacement of the forward wall of the pharynx as a result of contraction of the lower fibers of the genioglossus muscle (Stevens 1999: 469). If muscular tension and/or articulatory effort are to be used as distinctive features, they must be specific and relative features, i.e., they must be of the type “the production of A and B are alike in all aspects, except that the production of A requires more tension of muscle X”. Such cases are probably rare or nonexistent, for the contraction of one muscle usually coincides with the relaxation of another.

Duration: Alexander (1983) and Naiditsch (1997) both argue that duration is the key distinction between Germanic stops. Geminates are essentially very long (or doubly long) consonants. Since in most German dialects geminates no longer exist, longer consonants are often termed fortis and shorter ones lenis. Whereas Alexander and Naiditsch both view the fortis/lenis distinction from a historical and phonological perspective, Jessen (1997) and Willi (1996) provide phonetic data in support of duration as the primary distinction between fortis and lenis stops. Both Jessen and Willi include several durational measurements in their analyses, including both closure duration and aspiration duration.[9] Jessen concludes that aspiration duration is the basic correlate of the fortis/lenis distinction for SG (1997: 217). Willi's study of the Swiss German dialect in Zurich finds the exact opposite result. While total consonantal duration is also the distinguishing factor in this dialect, the difference is mostly in closure duration. Willi finds the closure duration of intervocalic fortis stops to be over twice as long as lenis stops, with the aspiration duration approximately equal (1996: 77-84). One should note that Willi only studies intervocalic stops. According to Keller, the fortis/lenis opposition in Zurich is "somewhat tenuous and rare and probably comparatively recent" (1961: 46). For some reason, many Swiss dialectologists have been reluctant to call this distinction one of simplex/geminate, but the facts certainly lead more towards such an analysis.

First formant transition: Stevens & Klatt (1974) suggest that the first formant (F1) transition is the distinguishing feature between homorganic stops. The F1 transition is defined as

[9] It is important to note the difference between VOT and aspiration duration. Jessen defines aspiration duration as "... the temporal interval between the beginning of an abrupt increase in energy indicating stop release and the onset of F2 of the following vowel" (1997: 59). This definition includes any period of breathy voicing, whereas VOT does not. Nevertheless, VOT and aspiration duration are fairly comparable measures. The reader should keep in mind that VOT measures might be slightly lower than aspiration duration.

the time required for F1 to reach its steady-state frequency in the vowel. Unaspirated stops exhibit a shorter F1 transition than aspirated stops. Stevens & Klatt find that listeners can perceive a distinction based solely on this measure. They favor using F1 transition over VOT, because they find that F1 transition varies little according to POA, in contrast to VOT. Lisker (1975) argues against this theory for several reasons. One reason is that F1 transition is very difficult to measure in natural speech. More importantly, he presents data which show that listeners can distinguish between aspirated and unaspirated (or voiced and voiceless) stops without any F1 transition as a cue.

In summary, there seem to be two different schools of thought on what should be considered the distinctive feature for consonant strength in German. Alexander (1983), Naiditsch (1997), Jessen (1997), Willi (1996), and Goblirsch (1994a, 1994b, 1997) argue that duration is the distinctive feature, while Braun (1988), Braune (1886), Davis et al(1999), Iverson & Salmons (1995) argue for aspiration. The cause of the High German consonant shift can be explained more elegantly if we take aspiration to be the relevant feature (see Iverson/Salmons 1995, Davis/Iverson/Salmons 1999). Furthermore, we will see that the data from this study strongly support aspiration (as measured by VOT) as the main feature distinguishing fortis and lenis plosives.

3.2. Perception, Language Change and Dialect Mixing

For many years phoneticians were most concerned with describing the sounds of languages as precisely as possible. At some point, however, they began to discover that speakers of these languages did not make such fine distinctions, but instead recognized phonemes. Experimental work in speech perception in the last fifty years has emphasized this fact even more. Numerous studies have shown that listeners perceive distinctions categorically (e.g.

Liberman et al. 1957, Lisker & Abramson 1970). Given a stimulus set varying in one parameter incrementally, the listeners find some boundary to divide the stimulus set. Many perceptual experiments have been carried with VOT as the parameter. As an example, a stimulus set consisting of the syllable /aBa/ may be constructed with VOTs ranging from 0-50 ms at 5 ms steps. Such experiments have shown that (for English speakers) listeners are relatively insensitive to differences in VOT between 5 and 15 milliseconds (ms), and above 40 ms, but are very sensitive to differences in VOT between 15 and 25 ms. Thus we can say that the category boundary is at approximately 20 ms.

When doing phonetic analysis, one must consider both acoustic/articulatory and perceptual factors, even when one is only doing acoustic analysis, as in this study. This is important because we want to separate acoustic differences due to physiological traits from differences due to phonology. For example, many traditional grammars (including much of the current work in German dialectology) make distinctions between fortis, semi-fortis, and lenis consonants (e.g. Berroth 2001). Berroth notes that intervocalic stops /p t k/ in the dialect of Ruppertshofen are realized as semi-fortis. Berroth does not do an acoustic analysis, but it is very likely that she is referring to the seemingly universal fact that plosives exhibit shorter VOTs in intervocalic position than in word-initial position (see Jessen 1997). While it is crucial to know such facts, it is very unlikely that the listener distinguishes among three categories.

In the last thirty years, linguists have also begun to consider perception as a contributing factor to language change. Work in this vein was pioneered by Ohala (1974), and is currently being carried out by numerous others (Wright 2001, Flemming forthcoming). Much of this work focuses on the conflict between two opposing forces of language: using the least articulatory effort vs. maximizing perceptual contrasts (see especially Lindblom 1990). As mentioned earlier in the discussion of the High German consonant shift, these two factors can be seen at work. After the fortis plosives had shifted to affricates or fricatives, the contrast between fortis and lenis stops was lost. As foreign words with aspirated plosives were then borrowed into German, these words were de-aspirated, in line with the principal of least effort. As we will later see however, it does appear that certain words have escaped this lenition process. Not by chance, these words happen to contrast with another word solely by aspiration, e.g. *backen* 'to bake' vs. *packen* 'to pack, leave'.

Several scholars have used anti-homophony constraints to explain contrast maintenance (Steriade 2000, Crosswhite 1999). In these studies, data is presented in which various sound changes fail to take place only where the application of the sound change would result in homophony. For example, Crosswhite shows that in the Trigrad dialect of Bulgarian, vowel reduction does not apply in cases where it would create homophony within a paradigm. In unstressed syllables, the plural neuter noun morpheme -o is reduced[10] to [a] in most cases, but not when this would produce homophony. Compare the following words:

singular	plural	gloss
varˈzala	varzaˈla	mooring point
kaˈpita	kapiˈta	hoof
ˈpetalo	ˈpetala	horseshoe
ˈblago	ˈblaga	good, blessing

In the first two words the stress shift prevents homophony, so the neuter singular ending –o is reduced; in the last two words there is no stress shift, so –o is not reduced, in order to avoid homophony.

[10] Reduced is probably not the correct term here, since reduced vowels are usually [ə], however I use Crosswhite's terminology here.

In these studies the anti-homophony constraints apply only to contrasts within paradigms. To the best of my knowledge no studies have shown such resistance to homophony for individual lexical items. This should not come as a surprise, for many languages contain homophonous lexical items, and homophony seems to pose few if any problems for speakers and listeners. Such lexical contrasts may play a role, however, when a language is in contact with another. Speakers often have a heightened awareness of differences between languages or dialects, and this heightened awareness could result in an avoidance of homophony when borrowing words.

In this case, Swabian is in a high contact situation with Standard German. While Swabian is spoken widely in Isny im Allgäu, SG is used in schools and in most media. Edward Flemming (p.c. 2004) has noted that in research on the /a/ - /ɔ/ merger currently taking place in much of America, he has found that many speakers are not aware of the merger until a minimal pair is pointed out to them. It is quite possible that aspirated stops are being introduced into the dialect in Isny first only in words that are contrastive, and then are slowly spreading to other words. The best example of this is *packen* [pʰakə] vs. *backen* [pakə]. According to Kluge (2002), *packen* is first attested in German in the 16th century. One of the other words used in the study, *Post* [poʃt] is also listed as being attested first in the 16th century, however it does exhibit lenition; thus it is not the case that *packen* is simply a recent loanword, and therefore did not undergo lenition. It is also interesting to note that in the neighboring dialects of Bavarian and East Franconian, *packen* and *backen* are also distinct, however not by aspiration of the initial stop, but rather by spirantization of the medial stop – *packen* [pɔkn̩] vs. *backen* [pɔχn̩].

In addition to minimally contrastive words, another set of words is also typically aspirated in Isny. This set of words involves recent loan words. It appears that the increasing influence of Standard German (and perhaps English as well, evidenced by words such as *Toast* 'toasted bread') is introducing a general contrast between aspirated and unaspirated consonants which was only marginal before. It is of course difficult to prove dialect mixing and leveling, but many scholars have discussed how it might be explained. One of the main factors which leads dialect mixing and leveling is an increase in mobility. Chambers (2002) explains at length how an increase in mobility leads to increased contact with speakers of other dialects, and therefore also dialect leveling. This is certainly the case in Germany. Until 1871, Germany did not exist as a nation; it consisted of several hundred small kingdoms and principates loosely tied to together by a common culture (including a common language to a certain extent). Only after 1871 did Standard German begin to emerge as a widely used norm. At the same time Germany transformed itself from a mostly agricultural society to one of the leading industrial nations in Europe. The increase of mobility can be seen by the shift from a rural to an urban society. In 1895, approximately 69% of the German population lived in rural settings (Wells 1987: 346); in 1995 approximately 35% of the population lived in cities of 100,000 or more people (Barbour & Stevenson 1998: 108). After World War II, mobility increased even more as a large influx of Germans living in Poland and Czechoslovakia were forced back to Germany; 9 million came into West Germany and 4.5 million into East Germany. Isny has also seen a dramatic increase in tourism since the 1950s, when it was declared a spa resort, resulting in yet more contact with speakers of different dialects. These factors combine to provide a convincing picture for the increasing influence of SG on Swabian.

4. Methods

4.1. Speakers

Since the sole speaker in my pilot study was a 45 year old woman, I decided that the participants in the study should be in a comparable age range. Using the concept of life-stages

(Milroy & Gordon 2003: 38-39), I chose to seek participants in the working life-stage, i.e. post-secondary education and pre-retirement. An equal number of male and female participants were also sought.

Using these criteria, I solicited participants via the friend-of-a-friend method. This method worked to a certain extent, yet did not yield as many participants as I had hoped. I therefore also posted flyers in public spaces asking for participants, and asked people I met in the town, e.g. shopkeepers, librarians. These methods also resulted in finding several participants, and I was able to further use the friend-of-a-friend technique. Men were much less willing to participate than women in general, which ultimately resulted in fewer men taking part. Using these methods and criteria, ten women and eight men took part in the study, ranging in age from 25-66. All speakers grew up within ten miles of Isny, in the state of Baden-Württemberg. Care was taken not to choose speakers from close-lying Bavaria.

4.2. MATERIALS

The same wordlist materials used in the pilot study were used in this study. The wordlist consisted of 28 target words and 14 filler words, each repeated five times, embedded in the carrier phrase *Sagen Sie* ____ *lauter* 'say ____ louder'. Filler words were chosen that did not begin with stops. This was done in order to mask the aims of the experiment. If the participants knew I was interested in the neutralization of alveolar and bilabial stops, they might artificially accentuate a distinction or a neutralization. The target words contained an equal number of bilabial and alveolar stops, with fewer velar stops, since the velar stops are undisputedly not merged in Swabian. Words with differing vowels were also chosen. One word with an intervocalic stop was also chosen for each condition. (See appendix for full materials list.)

A semi-random list with 210 items was created, with the following criteria used as a guide: (1) the same word was not repeated twice in a row, (2) filler words were evenly distributed throughout the list, (3) the five repetitions of each token were evenly distributed throughout the list, and (4) minimally contrastive pairs did not occur directly adjacent to one another. The same list was used with all the participants. The wordlist was typed on a computer and printed with a large, easy-to-read font. One difficulty with using a wordlist in this case is that Swabian is usually not written, and has no standard orthography. In the pilot study I originally attempted to use a quasi-phonetic orthography, but this was confusing for the participant; therefore standard German orthography was used for the wordlist, and the participants were instructed to "say each word on the following list in Swabian, as you usually speak Swabian". The danger with using standard German orthography is that the neutralized consonants in question are differentiated in standard German orthography. It seems that most participants were not very affected by the orthography, e.g. speakers regularly did not pronounce final <n> in words such as *lieben* 'to love' and *schwätzen* 'to talk'. (See feature (2) in §2.2).

4.3. PROCEDURES

In order to elicit the most natural speech possible, recordings were made in a comfortable environment of the choice of the participant. In most cases this was the kitchen or living room of the participant's home; however several participants preferred to do the recording at their workplace. This was the case for three of the recordings, which took place in a watch shop, the town library, and a sport hall. All of the recordings were made using a Sony portable DAT TCD-D8 sampled at 44.1 KHz with a small Sony ECM 717 condenser microphone. The microphone was equipped with a small clip, and was clipped to the leather carrying case of the DAT recorder and placed on a level surface (usually a table) such that both the interviewer and the speaker could be heard well. These environments all provided a fairly quiet environment for recording, with a signal-to-noise ratio of 35– 50 dB, as measured between peak vowel intensities versus intensity during non-speaking portions of the recording. Some of the recordings had a slight echo, either due to microphone placement (on a table) or because of hard surfaces in the surrounding environment, e.g. tiled floors in kitchens, and hardwood floors in living rooms. Additional sources of noises such as babies talking and clocks ringing also impaired the recordings somewhat. In cases where a disturbance was noticed during the recording, I asked the speaker to repeat the word at the end of the recording. In instances where a disturbance was noticed during analysis, the token was discarded if a reliable measurement could not be made. In a few instances the speaker skipped a word, or chose a different Swabian word synonymous to the SG word. These tokens were also discarded. Approximately 1.2% of the tokens were discarded due to noise, and approximately 0.3% were discarded due to word choice. In some instances a speaker stuttered or repeated a word. In this case the last repetition was used for measurement, since prior repetitions were often incomplete.

4.4. MEASUREMENTS

All of the recordings were imported into an iBook using the SoundStudio 2.0 program and then were analyzed using the acoustic phonetics program Praat 4.1. For each token, the VOT and the closure duration was measured, though the closure duration is only reliable for the intervocalic tokens. (The word-initial tokens were preceded by a vowel, but some speakers paused slightly before some or all of the target words).

4.4.1. MEASURING VOT

VOT can be defined articulatorily as the time between the release of oral closure and the onset of glottal vibration. This can be measured directly using a variety of techniques such as electropalatography and electroglottography (Fourcin & Abberton 1971). Unfortunately these methods are often not practical for doing field research. VOT can be measured indirectly from recorded speech however. Various criteria have been used for distinguishing the oral release and onset of voicing in waveforms and spectrograms. Francis et al. (2003) review several criteria for determining the onset of voicing, including f_0, F1, F2, F3 and oscillation in the waveform. By comparing these measures with direct measurements of glottal vibration made using electroglottography, they conclude that measuring voice onset using the waveform is the most accurate method. This does not entirely solve the problem however. One can still use different criteria to measure the waveform. Francis et al. define the onset of voicing as "... the time of the zero crossing preceding the upward-going portion of the first cycle of oscillation visible in the acoustic waveform" (2003: 1027). This differs slightly from Cho & Ladefoged's definition: "VOT was measured as the interval between cursors placed at the onset of release (the final release, if there was more than one) and the onset of the first complete vibration of the vocal folds as indicated on the waveform" (1999: 215). It is unclear what is meant by "complete vibration"; however one can infer that this is meant to exclude any partial periods, which are often found. While Cho & Ladefoged are correct to consider the final release as the one which matches the articulatory definition of VOT best, it is often very difficult to distinguish additional releases from aspiration noise. Thus using the final release could introduce additional uncertainty into the measurements. For these reasons I have chosen to measure VOT as the period between the first burst or appearance of aspiration and the first downward pointing peak of the first oscillation in the waveform with a period at least 70% of the period during the steady-state portion of the vowel.

Two sample waveforms and spectrograms are provided here, with the VOT marked. In Figure 2, one can see a clear burst at t=0.0149 seconds; there is another peak in the period of aspiration shortly thereafter, but as mentioned above, it is difficult to determine whether this is another burst, or simply aspiration noise. A small downward peak is visible at t=0.0251 seconds, which begins the first period of oscillatory motion in the waveform. This also corresponds to the onset of F1, as can be seen by the dark spot around 700 Hz in the spectrogram. The first period of oscillation is somewhat shorter than subsequent periods, however it meets the 70% criterion. Figure 3 is similar, except that in this case there is an upward pointing peak in the waveform just before the measurement at time t=0.0229. This could be construed as the first period of oscillation, however it does not meet the 70% criterion.

Figure 2. *Tag* #4 as produced by speaker #5.

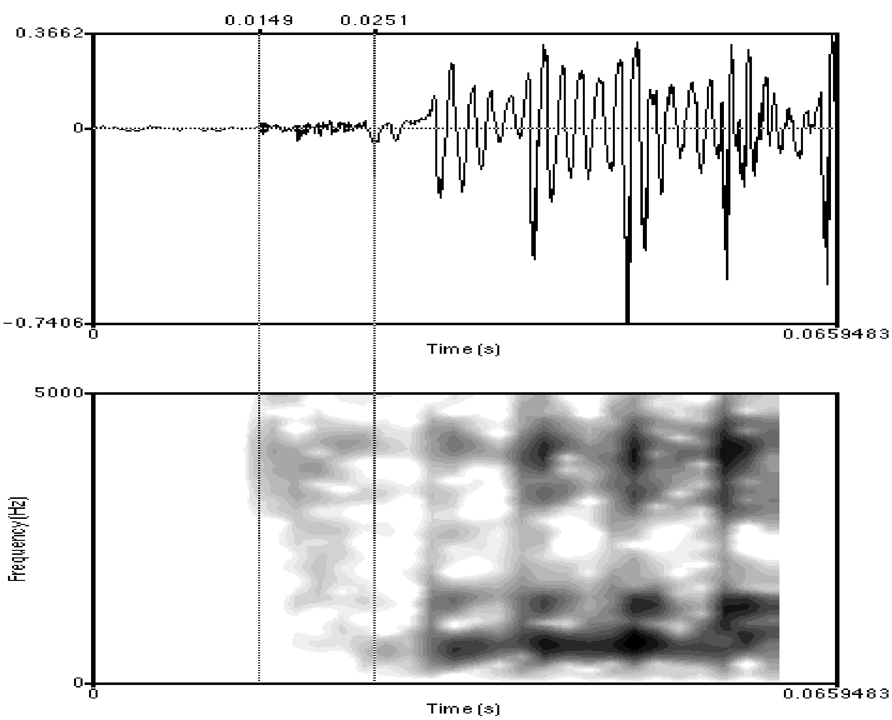
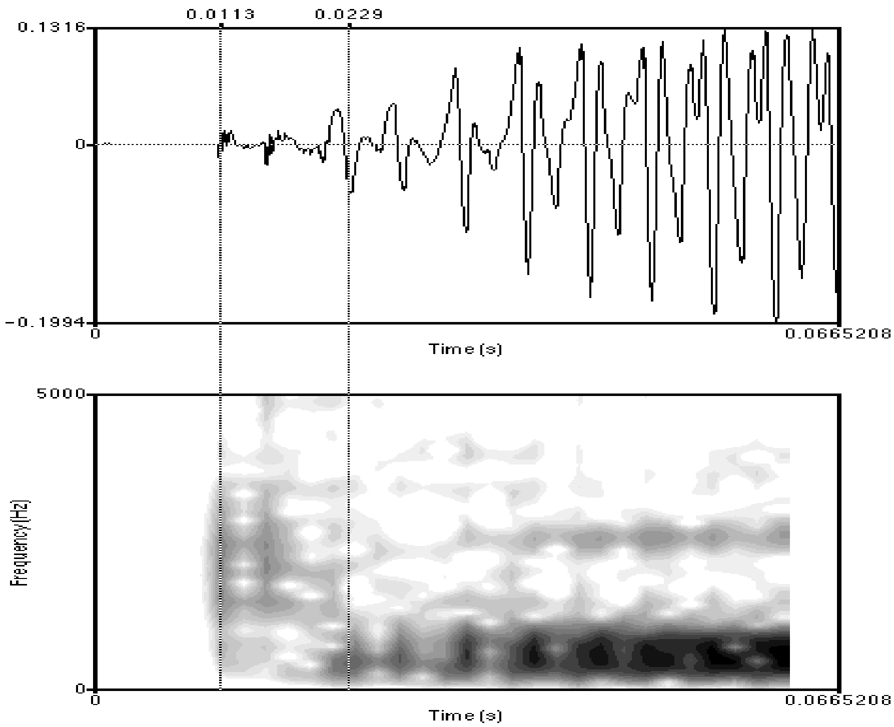


Figure 3 *Boot* #4 as produced by speaker #6



A small number of tokens (0.6%) were found to be pre-voiced. As noted before, some of the recordings had a slight echo, which sometimes resembled voicing during closure. In such cases spectra were also sampled from the preceding vowel and from the closure period. Tokens with an echo exhibit spectra during the closure period that closely resemble the spectrum of the preceding vowel. This can be seen in Figure 4. All three spectra exhibit strong peaks around 300 Hz and 2100 Hz, indicating F1 and F2 respectively. This is in strong contrast to the example shown in Figure 5. In this example the spectrum of the vowel shows strong peaks corresponding to the formant frequencies, whereas the spectrum sampled during closure shows a strong peak corresponding to f_0 but no peaks indicating formant frequencies.

Figure 4. *Bahre* #1 as spoken by speaker #14. The left side displays the waveform and spectrogram corresponding to the preceding vowel [i] and the beginning of the word [pa]. The right side displays three spectra, the first sampled during [i] and the remaining ones sampled during the consonant closure.

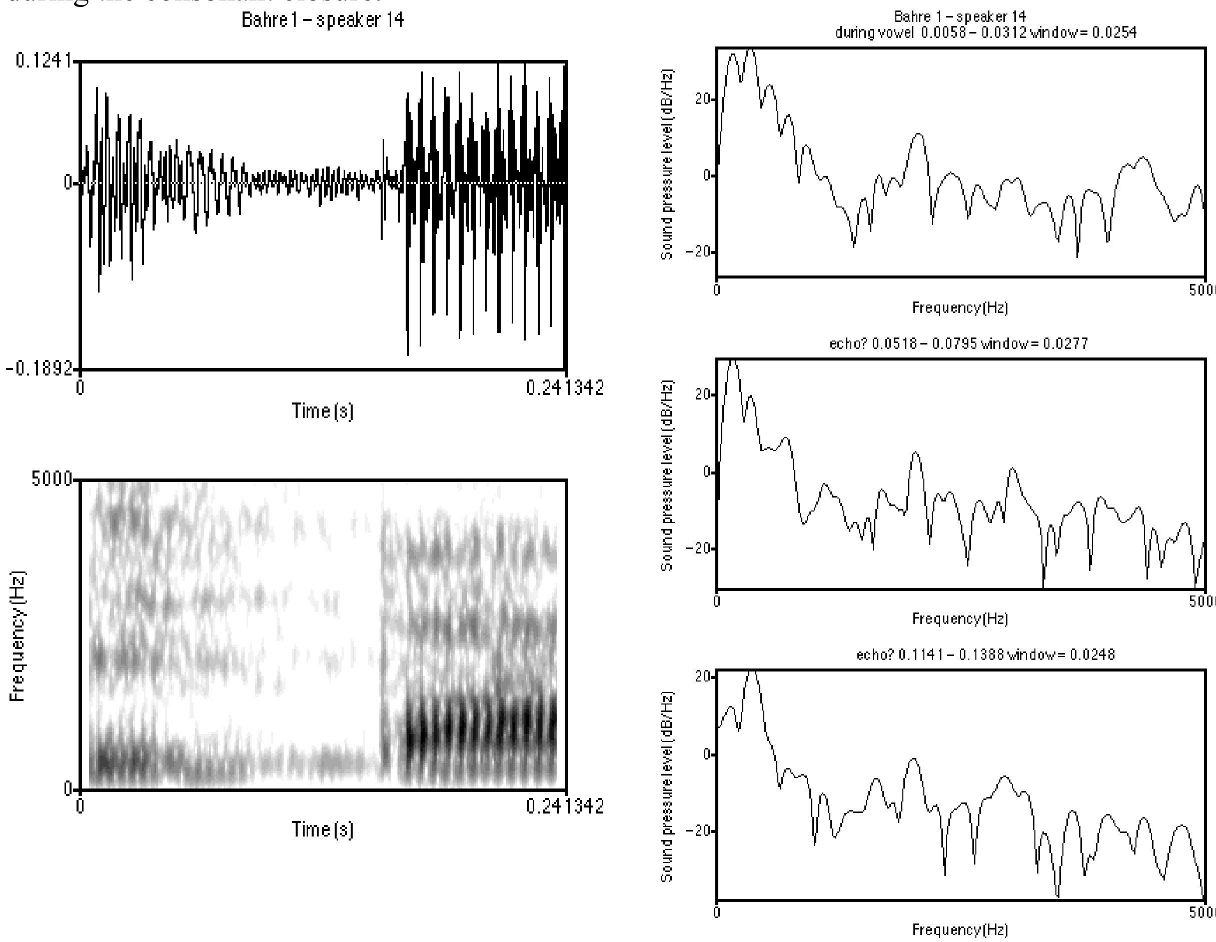
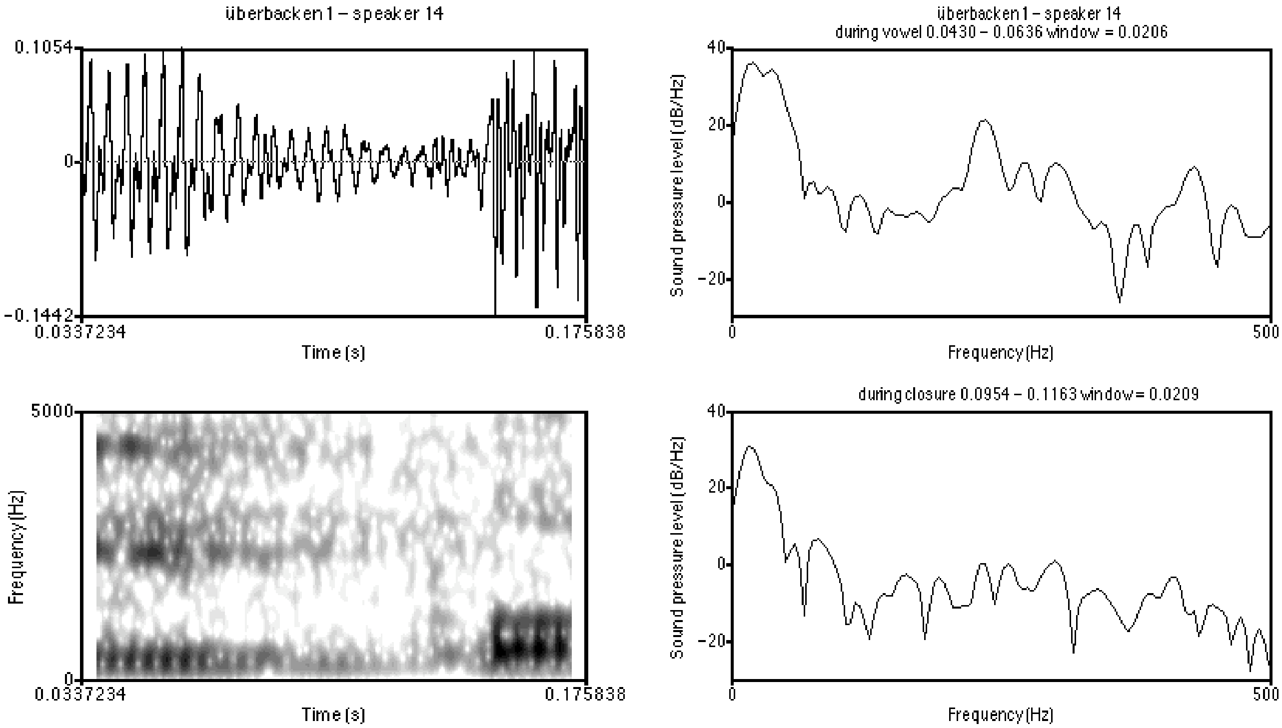


Figure 5. *überbacken* #1 as spoken by speaker #14. The left side displays the waveform and spectrogram corresponding to the initial vowel [y] and the second syllable [pɐ]. The right side displays two spectra, the first sampled during [y] and the second sampled during the consonant closure.



4.4.2. Speaker Exclusion

Speaker 18 differed from all other speakers substantially. His speech exhibited none of the features listed in §2.2, and was therefore left out of the analysis.

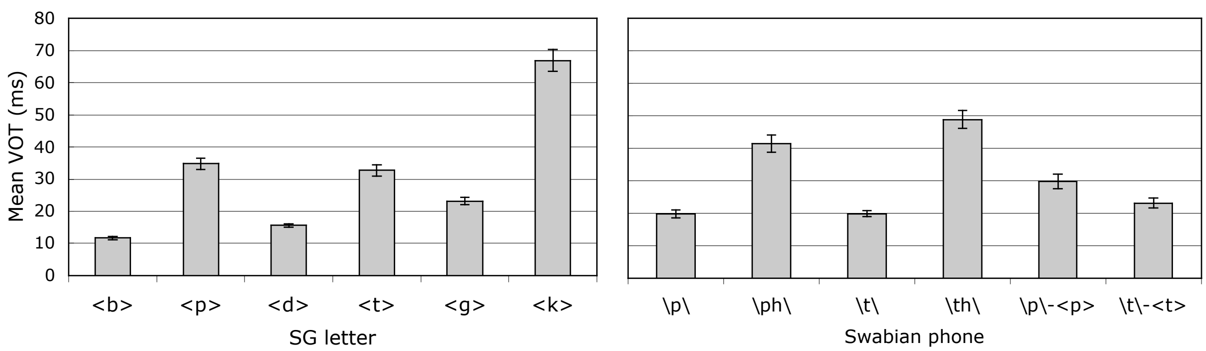
4.4.3. Token Exclusion

As mentioned in §4.3, some of the tokens were discarded due to noise or word choice. Given the extremely low proportion of pre-voiced stops, these were also discarded from analysis. A total of 2.1% of the data was excluded, leaving 2328 tokens for analysis.

5. Results and Discussion

The general results of the VOT analysis are displayed in Figure 6. Several trends can be found. As in many other studies of VOT (e.g. Lisker & Abramson 1964, Cho & Ladefoged 1999), bilabials tend to exhibit the lowest VOTs, while velars exhibit the highest. For this reason they will be handled separately throughout the majority of the analysis. The left plot in Figure 6 also shows a difference in means between words spelled <b d g> and <p t k>, though the difference for velars is much greater.

Figure 6. The left plot displays mean VOT according to SG spelling; the right plot displays mean VOT according to the pronunciation listed in the SW. In the Swabian phone plot, the \p\ column represents words spelled with both <b> and <p>, whereas the \p\-<p>column represents only words spelled with a <p>, but listed as having a lenis pronunciation (the same relationship applies to the \t\ and \t\-<t> columns). These results are pooled over all speakers. Error bars represent 95% confidence intervals.[11]

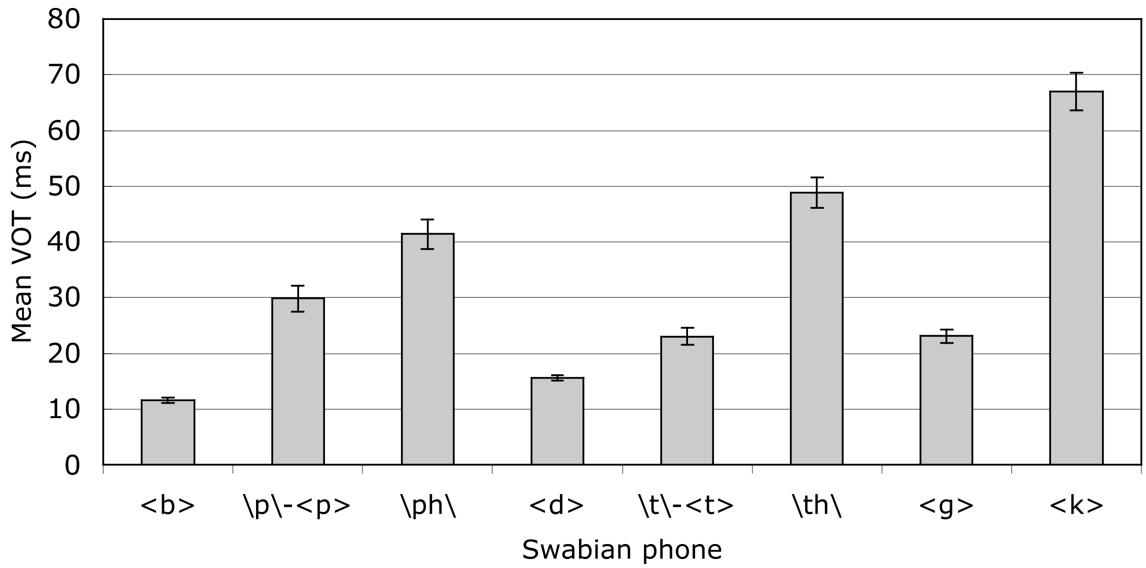


The right plot displays a slightly more detailed picture. Instead of grouping the sample by spelling, the first four columns are grouped according to the pronunciation listed in the *SW*. One can see that the difference between \p\ and \ph\ in the right plot is greater than the difference between <b> and<p> in the left plot (even more so for the difference between \t\ and \th\). One also notes that the values for \p t\ are slightly higher than those for <b d>. This reflects the fact that \p t\ include not only words spelled with <b d>, which are all expected to have low VOTs, but also the words spelled with <p t> reported as lenis according to the SW. The last two

[11] Confidence intervals were computed as 1.96 times standard error.

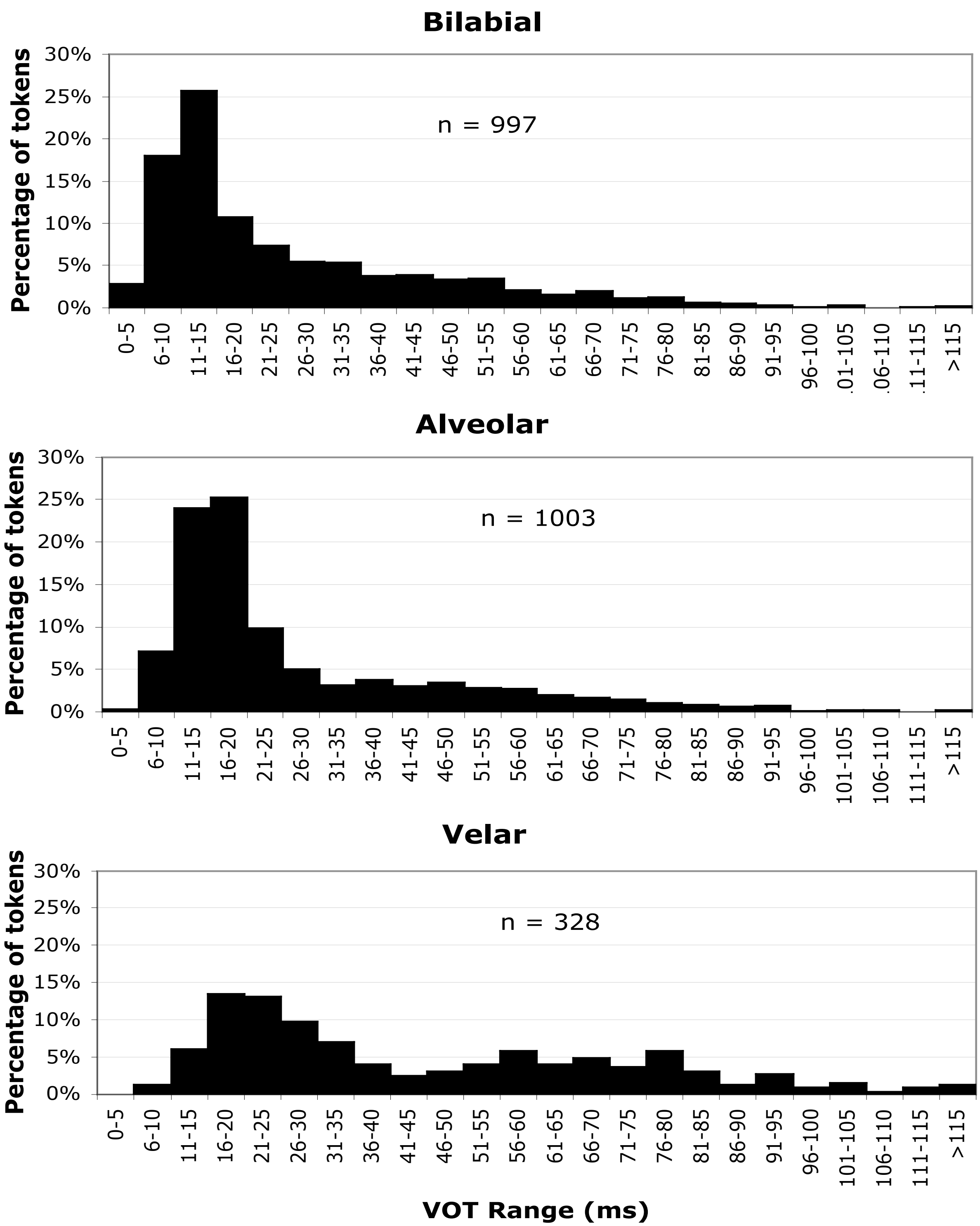
columns of the right plot include only these words (*Paare* 'couple, pair', *Pech* 'bad luck', *Pelz* 'fur, pelt', *Post* 'mail, letter', *Tag* 'day', *Tanne* 'fir tree, forest' *Teller* 'plate', *Tod* 'death', *vertauschen* 'exchange, swap'). By looking at both plots, one can see a three-way relationship for the bilabial and alveolar stops. The stops spelled with <b d> have the lowest mean VOTs; stops spelled with <p t> but listed as lenis have slightly higher VOTs, and stops listed as fortis have the highest VOTs. This pattern is reproduced in Figure 7 for easier comparison.

Figure 7. Mean VOT results from Figure 6, displaying the three way relationship of bilabials and alveolars. The velars are also included for comparison.



One disadvantage of relying on means to analyze data is that they do not tell us whether the distributions are modal or not. For example, the \p\ - <p> column could be comprised of a bi-modal distribution of tokens with high and low VOTs, not a modal distribution centered around thirty. Looking at the data in histograms will shed some light on this topic.

Figure 8. VOT histograms for each POA, pooled over all speakers.



As can be seen from the histograms, the shape of the distribution for bilabials and alveolars is similar, but differs from the shape of the distribution for the velars. The velar distribution displays a large peak in the 16-25 ms range, and a smaller peak in the 60-70 ms range. The bilabial and alveolar distributions only display one strong peak, with a long trailing tail. This is evidence for the partial merger of the bilabial and alveolar stops. Viewing the histograms give us a better idea of the distribution of the data, but statistical tests will provide more information yet. To test whether the three-way effect shown in Figure 7 is real, two-sample paired t-tests were carried out. A disadvantage with using a paired t-test is that the test assumes that the standard deviation of the two populations is the same, which in this case is highly unlikely. As will be shown later, the standard deviation of VOT for words spelled with <p t k> was much larger than that of words spelled with <b d g>. For this reason, a two-sample Kolmogorov-Smirnov test (which does not make the assumption that the standard deviations are the same) was also carried out. This statistic returns the maximum difference of the two distributions as calculated by

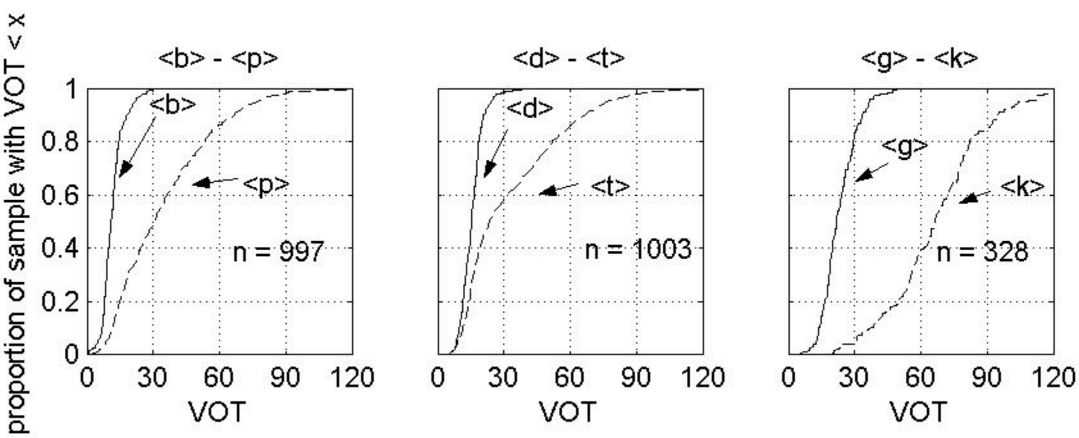
$$\max(|F1(x) - F2(x)|)$$

This is the difference of two proportions, and will always be between 0 and 1. A score of 0 means that the two distributions are exactly the same; a score of 1 means the two distributions share none of the same values. The associated p-value gives the probability that the result is due to chance. The results of the t-tests and the K-S tests are displayed in Table 2, with the respective cumulative distribution function plots in Figures 9 and 10.

Table 2. Results of two-sample paired t-tests and Kolmogorov-Smirnov tests, pooled over all speakers.

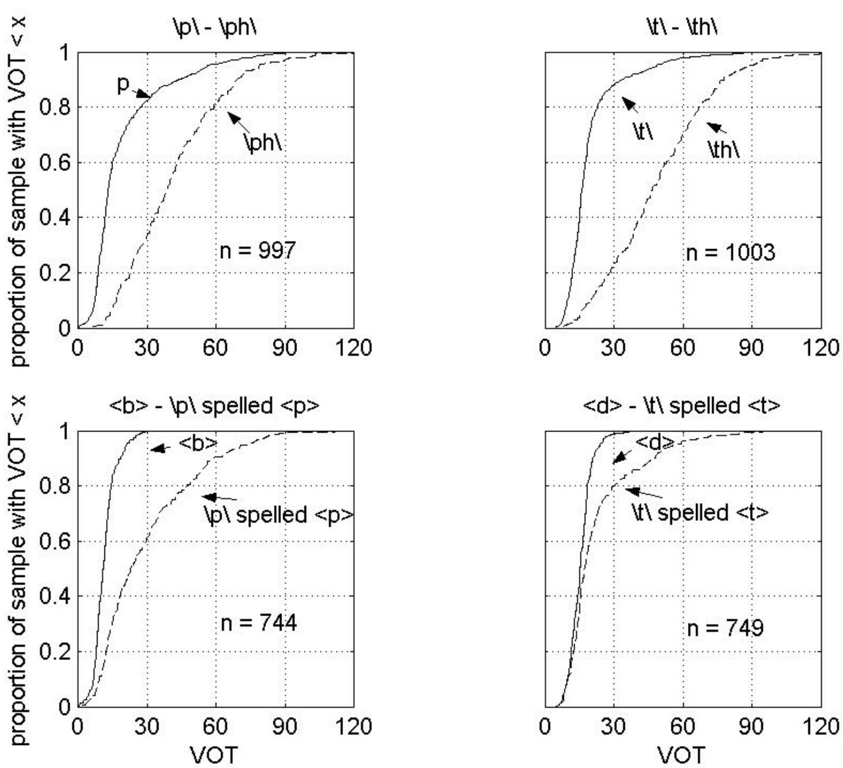
	t-test results					Kolmogorov-Smirnov results	
				difference conf. int.			
pairing	df	t	p	lower	upper	p	K-S-statistic
<b> - <p>	995	21.134	0.000	21.037	25.344	0.000	0.636
<d> - <t>	1001	13.646	0.000	14.696	19.633	0.000	0.459
<g> - <k>	326	23.816	0.000	40.165	47.398	0.000	0.866
\p\ - \p^{h}\	995	16.299	0.000	19.030	24.239	0.000	0.561
\t\ - \t^{h}\	1001	25.083	0.000	26.775	31.319	0.000	0.681
<b> - \p\ spelled <p>	742	16.924	0.000	16.078	20.298	0.000	0.533
<d> - \t\ spelled <t>	747	8.105	0.000	5.657	9.272	0.000	0.261

Figure 9. Cumulative distribution function plots for each POA, grouped by SG spelling. The y-axis values represent the proportion of the distribution less than the respective x-axis value. For example, in the left plot 100% of the <b> distribution has a VOT less than 50, but only approximately 76% of the <p> distribution has a VOT less than 50. The solid lines represent <b d g>, the dashed lines <p t k>



From Figure 9 it is apparent that the distribution of VOT for the velars is quite different than that of the bilabials and alveolars. The <p> curve is very close to the <b> curve until approximately 10 ms; the <t> curve is extremely close to the <d> curve until approximately 15ms. In contrast, the <k> curve is never close to the <g> curve. This is also reflected in the much higher K-S statistic (0.866 for the velars vs. 0.636 for the bilabials and 0.459 for the alveolars). Thus, though the distributions for the bilabials and alveolars significantly different, the differences are much less than the difference for the velars. The complicated distributions of the alveolars and bilabials is illustrated even more in Figure 10, where the distributions are grouped according to pronunciation listed in the SW. The top two plots show that the difference between the fortis and lenis consonants is much greater than what was seen for the difference between the spelling <b p> and <d t> in Figure 9. The bottom two plots show that the distributions between words spelled with <b d> and those spelled <p t>, but listed with lenis pronunciations, is not very great, especially for the alveolars. This is also reflected in the low K-S statistics (0.533 for bilabials, 0.261 for alveolars). This shows that it is not the case that the two distributions have equal variances, but slightly different means. Rather the \p\ spelled <p> and \t\ spelled <t> distributions are almost identical to the <b d> distributions, except that they also include some tokens with higher VOTs, which raises the overall mean. These tokens with high VOTs are the result of the variation in question.

Figure 10. Cumulative distribution function plots for bilabials and alveolars, grouped by pronunciation listed in the SW. Solid lines represent the left variable; dashed lines the right.



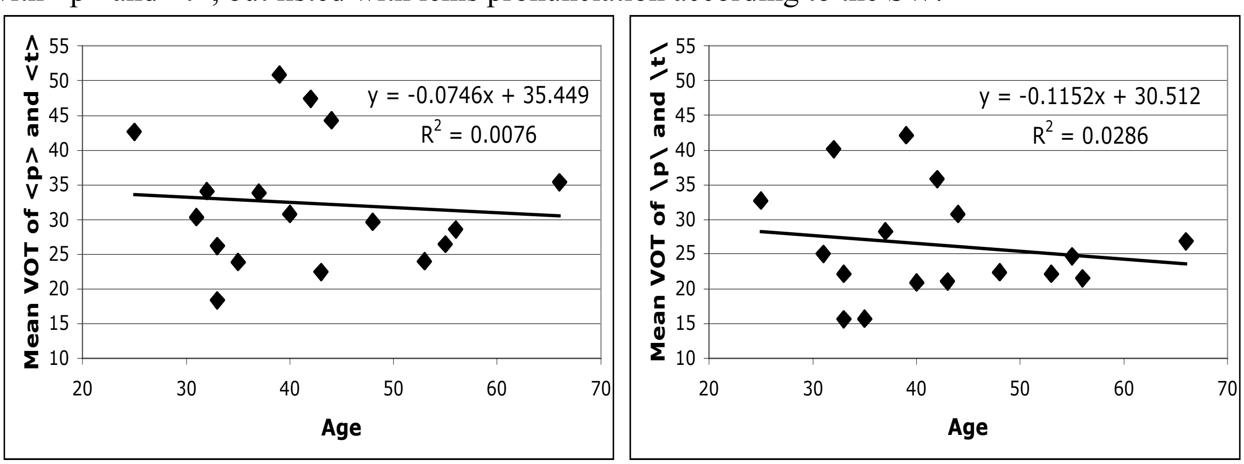
5.1. SOURCES OF VARIATION

In order to discover the origin of this variation, it will be useful to investigate any possible correlations between VOT and social as well as linguistic and experimental factors. The social factors used will be age and sex[12], the linguistic factors mean VOT and talker rate, and the experimental factors list order and orthography.

5.1.1. SOCIAL FACTORS OF VARIATION

Though I sought a homogenous age group for the study, the participants age covered a fairly wide range (25-66, mean = 41.3), it is possible that there could be an age effect within this group. If aspirated stops are becoming more common in Swabian, one would expect that older speakers would have lower mean VOTs than younger speakers. In order to test this hypothesis, mean VOTs of these words was plotted as a function of age. As one can see in Figure 11, there is a very slight downward sloping trend, however this accounts for at most 3% of the variation found.

Figure 11. Mean VOT as a function of age. The left plot displays the mean VOT all of the words spelled with <p> and <t>. The right plot displays the mean VOT of the words spelled with <p> and <t>, but listed with lenis pronunciation according to the SW.



[12] I use the term sex here and not gender, as recent research on language and gender uses a much more complex definition of gender. I do not intend to make any claims about gender in this paper.

There is also the possibility of a correlation between sex and VOT. Such a correlation could actually result either from physiology or from social factors. If the mean VOTs of <b d g> differ between males and females, this is most likely the result of physiological differences. One would predict that VOT for females might be slightly shorter than for males, due to the shorter vocal tract. If there is a sociolinguistic difference in VOT between males and females, one would expect this to be found in a difference between VOT values for words spelled with <p t>, or perhaps in the narrower subset, words spelled with <p t>, but pronounced as \p t\. In order to test whether the mean VOTs of males and females differed, two-sample t-tests were carried out for each group. The results of these are displayed in Table 3. None of the means were significantly different except for <b d g>. As predicted, females displayed slightly shorter VOTs than males. The difference for <p t> is approaching significance, however the confidence interval overlaps with 0; also, given the high number of tokens, it is improbable that more data would lead to a significant difference.

Table 3. Mean VOTs according for males and females, with two sample t-test results.

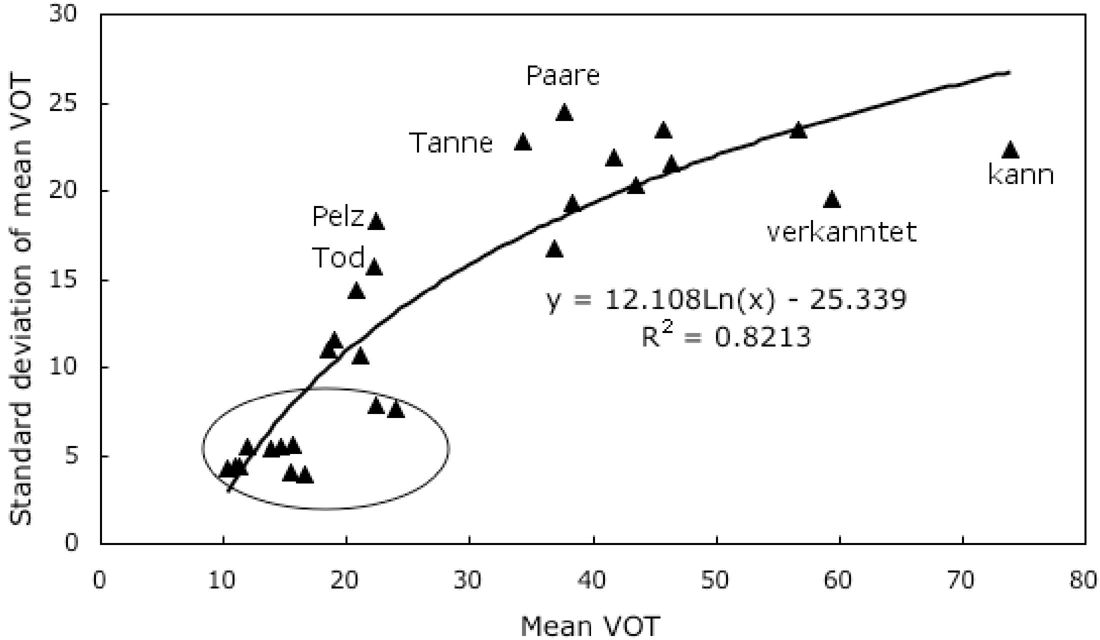
	Mean VOT		Two sample t-test results				
						confidence interval	
	female	male	df	tstat	p	lower	upper
total	28.989	28.693	2326	0.311	0.756	-1.569	2.161
<b d g>	13.946	16.957	901	-6.574	0.000	-3.909	-2.112
<p t>	34.629	32.437	1259	1.725	0.085	-0.301	4.686
\p\ <p> \t\ <t>	25.846	26.351	752	-0.363	0.717	-3.233	2.224

5.1.2. Linguistic Factors of Variation

Part of the great variation found in the VOTs for the words spelled with <p t k> must come as a result of physiology and/or broader categories, and not as a result of sociolinguistic variation. Statistically speaking, one would expect the standard deviation for a distribution with a high mean to be greater than one with a low mean, simply because the raw numbers are larger.

From a linguistic point of view, aspirated stops usually have a bigger category than unaspirated stops, i.e. unaspirated stops generally display VOTs in the range of 0-20 ms, whereas aspirated stops generally exhibit a range from 30-100 ms. In order for aspirated stops to be perceived as distinct from unaspirated stops, there must be some lower limit of VOT for aspirated stops, but there need be no upper limit. Unaspirated stops are the opposite, in that there is an upper limit, but theoretically no lower limit. However, as Kingston & Diehl (1994: 428) note, there is an articulatory driven lower limit of VOT for initial unaspirated stops in languages which contrast between unaspirated and aspirated stops – the transglottal pressure required to initiate voicing is greater than that required to maintain voicing; in other words, it requires additional effort to produce voicing during closure in initial stops. Thus for such languages there is a lower limit of VOT for unaspirated stops of 0 ms. In Figure 12, the standard deviation of the mean VOT for each word is displayed as a function of mean VOT. Though there is indeed a strong increasing trend, one should also note that the two words with the highest VOTs, *verkanntet* 'jammed' and *kann* 'can.3rd.sg' have standard deviations substantially below the model. Thus words such as *Paare* 'pair, couple', *Pelz* 'fur, pelt', *Tanne* 'fir tree, forest', and *Tod* 'death' exhibit large standard deviations, even when this is taken into account.

Figure 12. Standard deviation of VOT plotted as a function of mean VOT for each word used. Words spelled with <b d g> are circled. A logarithmic best-fit curve is also displayed. Words deviating substantially from this curve are labeled.



The other linguistic factor which could be a source of variation is talker rate. Since VOT is a measure of duration, one would expect it to also be somewhat correlated to other measures of duration. Allen et al. (2003) found that talker rate accounted for 82% of the variation in VOT in their study of eight speakers reading a list of 17 words (30 repetitions of each). Though my experiment was not designed to account for talker rate, an informal measure of talker rate was achieved by measuring parts of the carrier phrase *Sagen Sie* ____ *lauter* 'say ___ louder'. Five random tokens of the syllables [sa] and [lau] were measured for each speaker, and compared to mean VOT of words spelled with <b d g> for each speaker. Words spelled with <b d g> were chosen because any variation found in these words could not be a result of sociolinguistic variation, but would most likely be due to differences in talker rate. There is a modest correlation for [sa], but it is not significant at the .05 level. The correlation for [lau] is very small and not significant. There is a fairly strong correlation between [sa] and [lau], indicating that these are adequate measures for talker rate. Therefore we can conclude that talker rate has only at best a small effect on VOT.

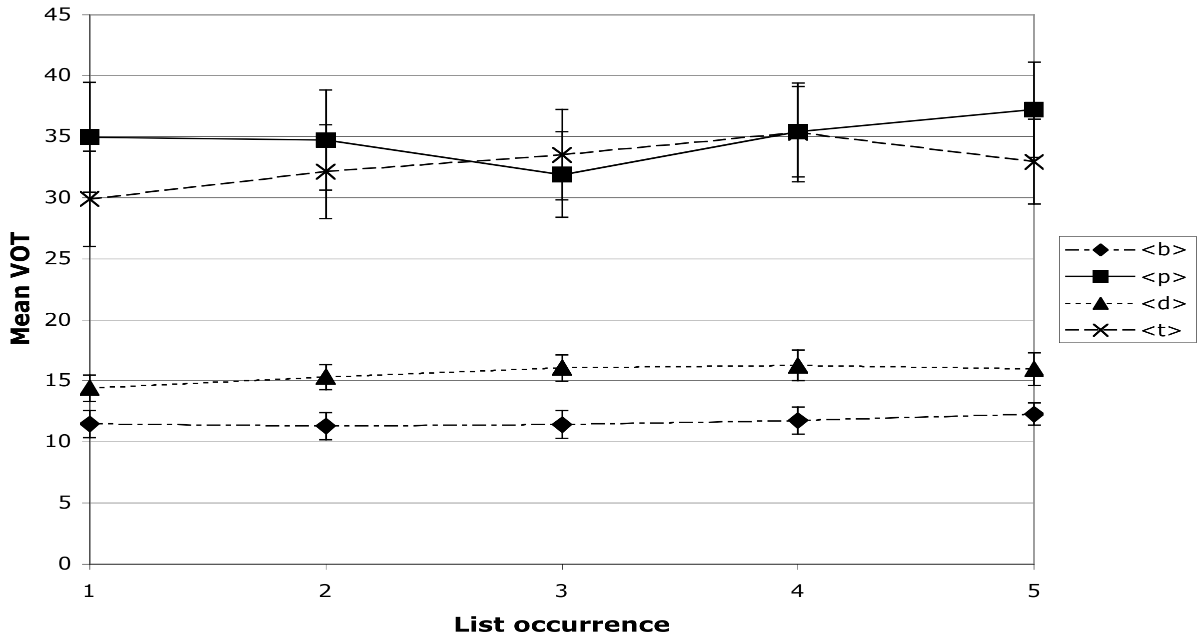
Table 4. Mean VOT as a function of talker rate. The correlation statistic r and the significance level p are shown.

pairing	r	p
[sa] – VOT <b d g>	0.3552	0.162
[lau] – VOT <b d g>	0.0345	0.896
[sa] – [lau]	0.495	0.044

5.1.3. Experimental Factors of Variation

Two possible experimental influences on VOT were recognized during recording and analysis. As I listened to the word list recordings, it seemed that speakers often spoke in a style somewhat closer to SG near the beginning of the list, and later began to use more Swabian features. That is, the speakers were being influenced by the SG orthography, or perhaps simply by the wordlist task in general, which usually elicits a more formal style than unscripted speech. In order to test this hypothesis, I re-grouped the data according to list occurrence. Thus the first occurrence of *backen* ‘to bake’ was grouped with the first occurrence of *Teller* ‘plate’ and so forth. Since the wordlist was semi-random, this is not an exact measure of time elapsed from the beginning of the experiment. In fact, the first occurrence of *backen* comes much later than the first occurrence of *Teller*. Nevertheless, this measure provides a consistent method across all speakers for any possible list effect. If my original hypothesis is correct, we would expect that words spelled with <p> and <t> would exhibit higher VOTs earlier in the list than later. As Figure 13 shows however, little if any effect of list occurrence was found.

Figure 13. Mean VOT as a function of list occurrence. Error bars represent 95% confidence intervals.



The other possible experimental factor which may have influenced the results is orthography in general. Though no list order effect was found, one could note the irregularity of the <p t> lines in Figure 13, suggesting that orthography could be playing a role. Though I have no method to test this more directly, I can remark that speakers generally had little troubling speaking Swabian words which differed in pronunciation from the SG spelling. These include several of the differences listed in §2.2, e.g. loss of final /n/. Many of the words on the listed ended in <n> and these were consistently produced without final /n/.

In order to more rigorously examine these effects, a general linear model was applied with fixed factors fortis/lenis and POA, random factor of speaker, and list occurrence as a covariate. Only main effects were used in the model, since no possible interpretation of interaction effects was possible at this stage. These results are shown in Table 5. The factors of age, sex, and talker rate were not included in this model because they are all consumed by the subject factor. Age and sex were already shown to be insignificant in §5.1.1. To investigate the effects of talker rate a separate model was carried out with fortis/lenis and POA as fixed factors and talker rate (measured as the mean duration of [sa]; see § 5.1.2) as a covariate. This model found a very slight positive trend ($B = 0.00656$, $p = 0.000$, $eta^2 = 0.015$) The B-value is an estimate of the change in the dependent variable that can be attributed to a change of one unit in the independent variable. That is for a 1 ms increase in talker rate, the model predicts an increase of 0.00656 ms in VOT.

Table 5. Results of general linear model.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	Hypothesis	499405.461	1	499405.461	404.135	0.000	0.944
	Error	29411.907	23.801	1235.738[a]			
Speaker	Hypothesis	96501.105	16	6031.319	22.531	0.000	0.135
	Error	617548.195	2307	267.685[b]			
Fortis	Hypothesis	380723.363	1	380723.363	1422.284	0.000	0.381
	Error	617548.195	2307	267.685[b]			
POA	Hypothesis	38692.696	2	19346.348	72.273	0.000	0.059
	Error	617548.195	2307	267.685[b]			
List	Hypothesis	984.63	1	984.63	3.678	0.055	0.002
	Error	617548.195	2307	267.685[b]			

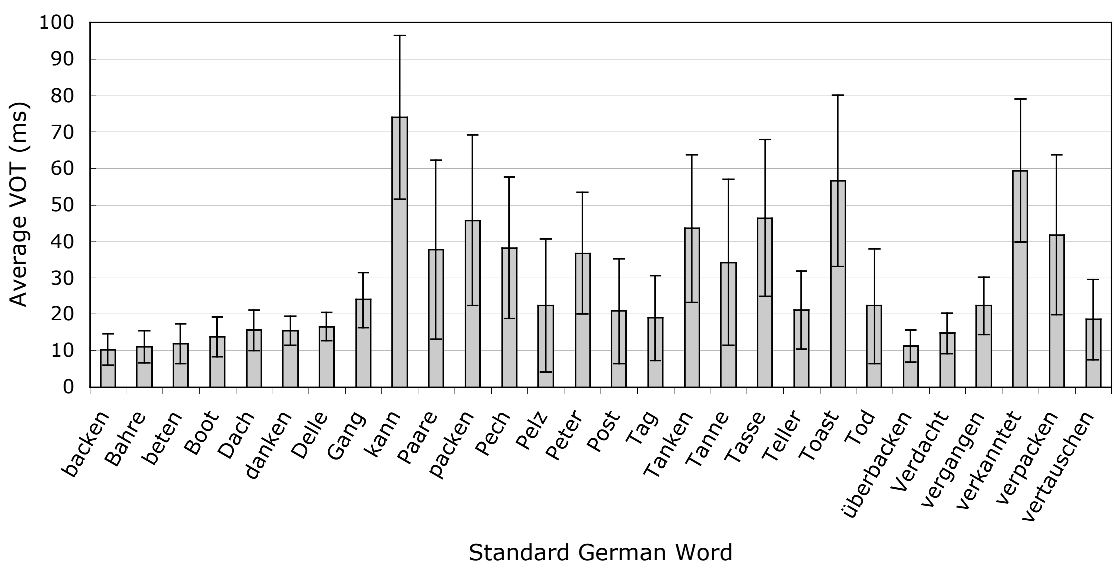
a. .168 MS(Subject) + .832 MS(Error)
b. MS(Error)

Since age, sex, and list occurrence were found to be insignificant, and talker rate accounted for at most 2% of the variation, the rest of the variation must be attributed to differences in individual speakers and lexical items. Since lexical item is not orthogonal to fortis/lenis and POA, this factor could not be included in the overall model. Instead, the data was grouped by orthography. Since the variation of interest here is contained in the words spelled with <p t>, only these words were used. A separate model was carried out for <p> and <t>, with lexical item and speaker as random factors. The results for <p> showed that the lexical item accounted for 21.7% of the variation, and speaker accounted for 30.5% of the variation. The results for <t> showed that lexical item accounted for 43.8% of the variation, and speaker accounted for 22.4% of the variation, all significant at $p < 0.001$.

5.2. LEXICALLY SPECIFIC AND FREE VARIATION

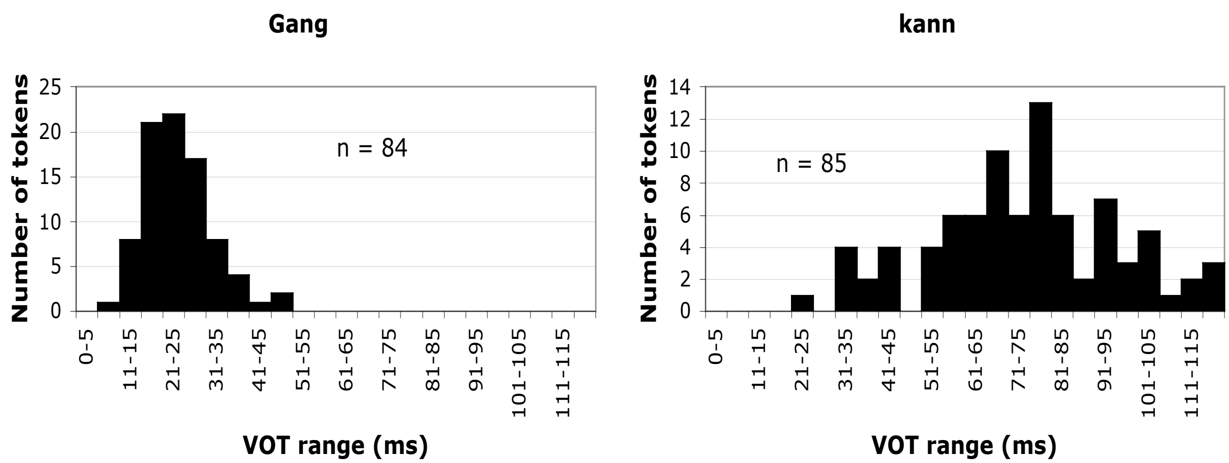
Since less than 2% of the variation in VOT was accounted for by sex, age, talker rate, and list occurrence, and a great deal was explained by lexical item and individual speaker differences, a more thorough investigation of lexically specific and free variation is called for. Much of this remaining variation can be explained by viewing individual words. The mean VOT and standard deviation for each word is displayed in Figure 14. One can distinguish between three different types of words in Figure 14: words with (1) low mean VOT with little variation, e.g. *backen* 'to bake' and *danken* 'to thank', (2) high mean VOT with large variation, e.g. *Toast* 'toasted bread' and *packen* 'to pack, leave', and (3) low mean VOT with large variation, e.g. *Pelz* 'fur, pelt' and *Tanne* 'fir tree, forest'. This third group includes the same words that deviated substantially from the model in Figure 12.

Figure 14. Mean VOT for each word used, pooled over all speakers. Error bars represent one standard deviation.



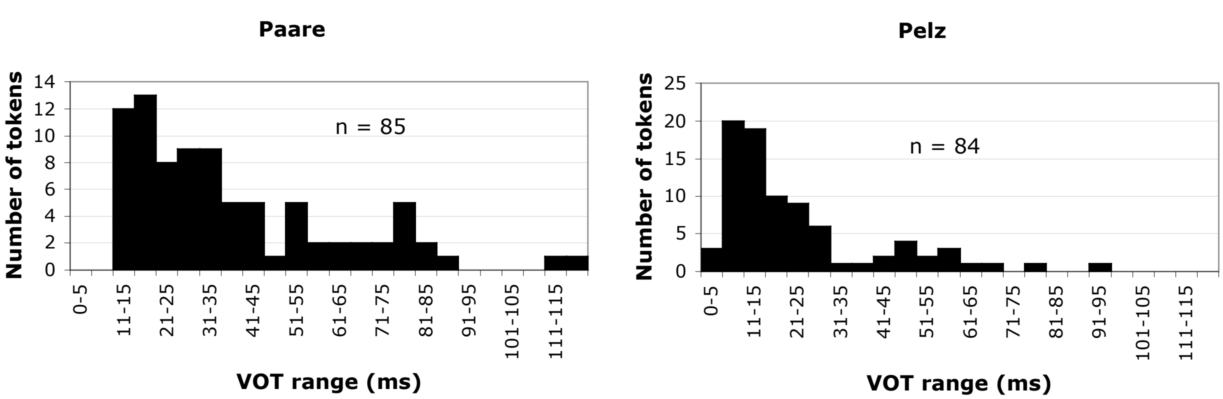
As noted earlier however, the mean and standard deviation do not give us any information about modality. Viewing histograms of the individual words will shed some light on this. Figure 15 displays histograms for *Gang* ‘walkway’ and *kann* ‘can 3rd.sg’, which both appear fairly modal, as one would expect.

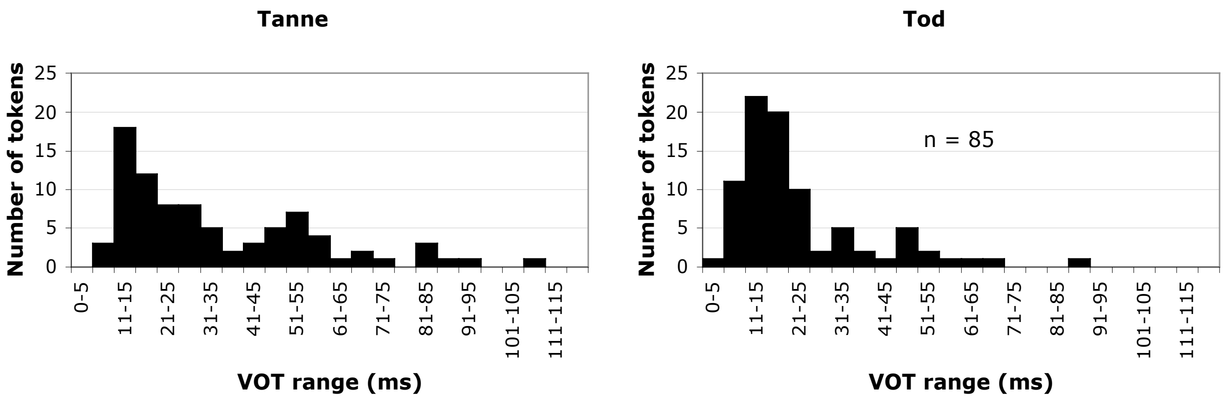
Figure 15. VOT histograms for *Gang* ‘walkway’, and *kann* ‘can 3rd. sg.’, pooled over all speakers.



A much different picture is found in the histograms displayed in Figure 16. Though the total number of tokens is somewhat too low to show clear bi-modality, the histograms certainly point in that direction, especially those for *Paare* and *Tanne*.

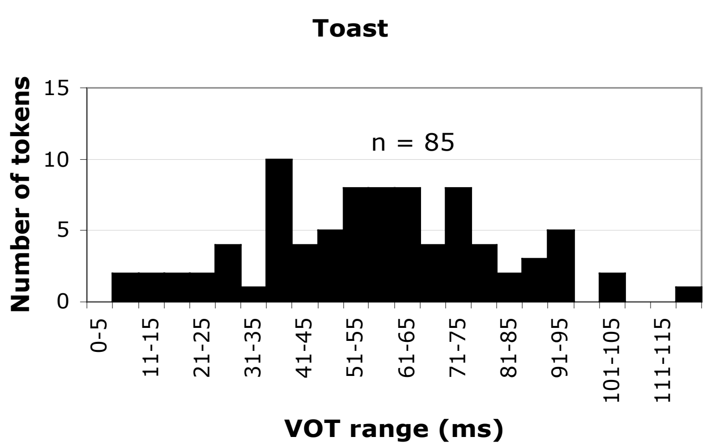
Figure 16. VOT histograms for *Paare* ‘pair, couple’, *Pelz* ‘fur, pelt’, *Tanne* ‘fir tree, forest’, *Tod* ‘death’, pooled over all speakers





It is still nevertheless possible that the tokens with high VOTs are merely spelling pronunciations. That is, these words are normally pronounced as unaspirated, but the experimental conditions resulted in the speakers occasionally pronouncing them as aspirated. As mentioned in §5.1.3, orthography did not seem to pose a problem for many other features. Additional evidence for the lack of orthographic effect comes from the distribution for *Toast* 'toasted bread', displayed in Figure 17. As can be seen, 85% of the tokens had VOTs greater than 35 ms, well within the aspirated range. However, several tokens were pronounced with very low VOTs. This cannot be a result of spelling pronunciation.

Figure 17. VOT histogram for *Toast* 'toasted bread', pooled over all speakers.



While the histograms show that much of the variation in VOT is due to individual words, they do not tell us whether the variation is inter- or intra-speaker. By looking at the raw VOT values for each speaker, we can tease apart this variation. From Table 6, one can both inter- and intra-speaker variation. Inter-speaker variation can be found in *Tod* - while most speakers exhibit fairly low VOTs, speakers 4 and 12 exhibit consistently high VOTs (shaded in light grey). Intra-speaker variation is also evident in *Tod* however. Note the varied values for speakers 9 and 15 (shaded in dark grey). Similar findings can be seen in the other words in Table 6.

Table 6. Raw VOT values (ms) for each speaker for *Paare, Pelz, Tanne, Tod,* and *Toast.*

Word	token	speaker																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Paare																		
	1	76	21	13	11	15	52	34	16	33	77	21	15	57	30	43	21	11
	2	18	38	29	19	15	81	37	74	18	58	17	19	32	30	42	21	25
	3	122	22	11	23	51	63	37	28	11	71	19	28	31	36	43	15	25
	4	112	28	17	28	19	70	39	52	31	77	12	25	17	64	32	16	42
	5	88	78	14	31	14	55	32	80	49	82	19	31	13	54	67	23	44
Pelz																		
	1	0	11	14	23	11	32	13	10	8	4	16	20	27	8	21	57	10
	2	59	22	13	30	14	47	9	10	11	65	9	12	8	9	13	77	15
	3	8	17	22	41	23	11	9	19	25	10	19	15	9	14	27	27	17
	4	9	21	12	27	47	52	10	20	8	59	23	17	5	8	67	18	26
	5	53	35	12	25	19	42	47	10		9	8	90	8	10	12	50	9
Tanne																		
	1	60	19	17	33	30	13	12	12	24	69	14	38	10	29	15	81	15
	2	55	31	28	15	45	57	47	19	29	59	12	17	9	29	19	84	63
	3	108	23	15	35	23	25	52	13	14	81	14	20	12	35	14	42	33
	4	86	55	15	27	25	37	15	20	28	69	20	49	29	51	8	52	71
	5	95	52	17	48	19	14	44	45	13	54	11	19	22	57	14	49	21
Tod																		
	1	11	32	15	48	17	25	39	12	11	11	5	66	8	10	17	42	15
	2	15	19	17	45	24	23	16	11	17	12	12	26	10	10	14	23	13
	3	20	49	30	34	19	20	6	20	52	11	12	35	10	9	14	18	12
	4	12	21	17	59	22	19	9	11	15	18	11	24	12	9	86	14	38
	5	16	63	20	47	17	20	29	15	13	52	7	49	12	6	15	24	34
Toast																		
	1	102	80	61	42	26	46	38	52	56	68	46	43	27	21	70	94	40
	2	103	93	65	48	47	76	36	9	41	95	40	57	38	73	16	92	82
	3	58	37	74	57	17	58	54	47	38	76	12	40	89	53	62	73	86
	4	76	24	73	68	26	60	63	73	60	67	15	54	74	51	9	122	88
	5	91	65	67	58	30	44	51	59	62	61	51	38	83	38	72	53	33

Since much of the variation seems to be lexically specific, an analysis of lexical frequency might explain some of this. An increasing body of research has shown that lexical frequency can have a large effect of sound change, most notably by Bybee (1985, 2000 etc.) In her 2000 paper, she notes the complex relationship between lexical frequency and sound change. Very frequent words and phrases are often subject to reduction, e.g. *going to* becoming *gonna*, but are "more resistant to change on the basis of other patterns" (Bybee 2000: 251-252) One such example is the verb *to be*, which tends to be highly irregular in a number of languages. In this study, one would expect that high frequency words would be more susceptive to lenition than low frequency words. However, the change in progress here does not seem to be lenition, but rather a reversal of lenition, as a result of influence from SG. The reversal of lenition would be a change based on other patterns, that is, on SG, and would therefore also predict that high frequency words would be more resistant to this change. In order to test this hypothesis, the lexical frequencies of the words showing this variation in VOT (those spelled <p t>) were compared..

A measure of lexical frequency was obtained from Ruoff (1981). Though this database is not as widely known as other databases such as CELEX, it has the advantage of being sampled from spoken corpora of freely spoken conversations from the Alemannic dialect area, including the dialects of Badish, Swabian, and several Swiss dialects. Thus this sample should be a good estimate of lexical frequency for Isny. The sample for this corpus was 500,000 words. In order to make it more comparable with other commonly used measures of lexical frequency (e.g. Kucera & Francis 1967), the frequencies were doubled. Since many studies have found linguistic effects to be correlated with the natural log frequency, this was then computed. The log frequency of the 15 words spelled with <p t> was compared with the mean VOT for these words. The word frequencies are listed in Appendix A. A slight negative correlation ($r = -0.241$, $p = 0.388$) was found, but was not significant. Though this result was not significant, the general trend is consistent with the hypothesis. That is, one would expect that high frequency words and/or older loan words would have lower VOTs, if the fortis plosives are becoming (or have become) deaspirated. Conversely, low frequency words and/or more recent loan words would have higher VOTs, as they have had less time to become deaspirated. In order to better understand the effects of lexical frequency, a study designed for this purpose should be carried out in the future.

5.3 Comparison to other studies

While identifying the sources of variation in VOT in Isny is the main goal of this study, it is also important to compare these results with other studies. Here I discuss my results for VOT and for closure duration as compared to other similar studies.

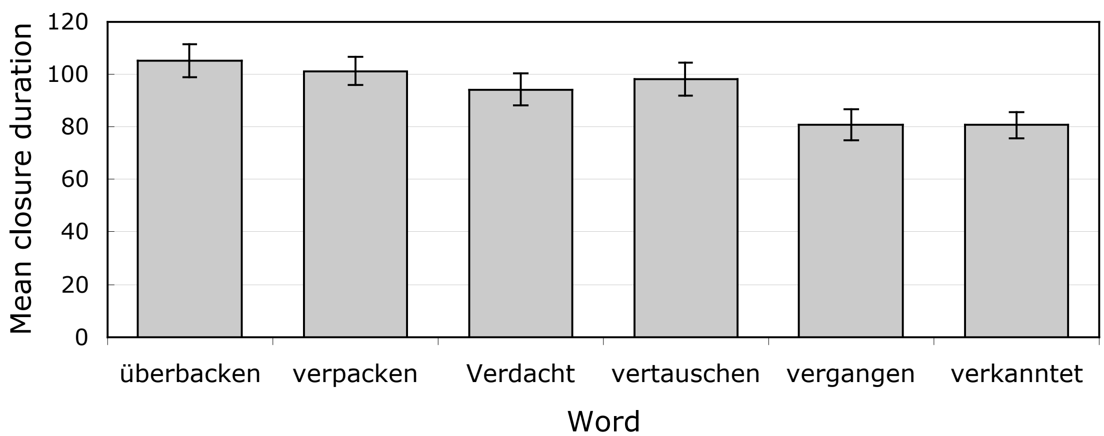
5.3.1. VOT compared to other studies

The total mean VOT of all fortis stops was 58 ms, and the total for lenis stops was 24 ms. Grouped according to orthography, the means are 13 ms for <b d g> and 35 ms for <p t k>. The grouping according to fortis/lenis is fairly comparable to Jessen's results of 60 ms and 22 ms in the environment V́_V, but the fortis average is somewhat lower compared to his results for the ##_V environment of 74 ms for fortis and 21 ms for lenis consonants (1997:253). These differences could be due to several factors. The tokens in this study were uttered in a carrier phrase, not in isolation as in Jessen's study, and also Jessen measured aspiration duration, not VOT, which he claims is consistently one to two voicing periods longer than VOT (1997: 41). In general it is evident that the VOT of words used in this study as grouped according to the SW is comparable to SG, and that SG orthography is not a good predictor of this contrast.

5.3.2. CLOSURE DURATION AS A MEASURE OF FORTIS/LENIS

Both Jessen (1997) and Willi (1996) conclude that duration is the key distinguishing factor between fortis and lenis consonants. Jessen finds that for SG aspiration duration is the basic correlate, with voicing duration and closure duration as secondary correlates (1997: 112). Closure duration was only found to be significant in the V́_V environment, with a mean for fortis stops of 84 ms, and a mean for lenis stops of 58 ms. Willi's study of the Zurich dialect shows that closure duration for fortis consonants is two to three times as long as lenis consonants - 82.375 ms for /d̥/ vs. 198.750 ms for /t/ (1996: 77). While I did not investigate stops in this environment, it is still worthwhile to ask whether closure duration is a correlate of the fortis/lenis distinction in pre-tonic stops. For this purpose I investigated only word-medial pre-tonic stops, since the word-initial tokens sometimes had slight pauses before them. As can be seen in Figure 18, there is no significant difference in closure duration in these tokens.

Figure 18. Mean closure duration for words with intervocalic stops, pooled over all speakers. Error bars represent 95% confidence intervals.



6. General Discussion

This study had several goals: (1) to investigate the historical development of lenition in Swabian, (2) to study the variation in the fortis/lenis distinction and determine its sources, and (3) to verify the acoustic parameter of VOT as an appropriate measure of the fortis/lenis distinction. I will review each of these below.

In §2.3 and §2.4 I outlined the High German consonant shift and analyzed the *binnenhochdeutsche Konsonantenschwächung* 'internal High German consonant lenition' as a byproduct of this shift. After the fortis stops /p t k/ had shifted, the fortis/lenis distinction was essentially lost in High German dialects. Loan words with fortis stops entering these dialects after the shift were then deaspirated, with the exception of a few words such as *packen* 'to pack, leave'.

In §5.1 and §5.2 I discussed the various sources of variation in VOT found in the study. A small amount of the variation was found to stem from social, linguistic, and experimental factors, but the majority of the variation was found to be lexically specific and/or free variation. I propose that the major source of this variation is influence from SG. Whereas older loan words such as *packen* [$\mathrm{p^h}$akə] may have been exceptions to the lenition found in other words such as *Post* [poʃt], the increasing number of loan words with fortis stops such as *tanken* [$\mathrm{t^h}$ankə]and *Toast* [$\mathrm{t^h}$oʃt] may be introducing a general fortis/lenis distinction in this dialect.

In §5.3 I compared the acoustic results to other similar studies. In contrast to Willi (1996), closure duration was not found to be an acoustic correlate of the fortis/lenis distinction, but VOT was. The total mean VOTs for fortis and lenis stops were comparable to Jessen's study of Standard German.

7. Conclusion

This study has improved our understanding of the ongoing merger of stops in the Swabian dialect of Isny im Allgäu. Three key factors in the study led to these results: (1) including a modest number of speakers, (2) using multiple repetitions of each word, and (3) performing acoustic analysis. These three factors led to the observation of both inter- and intra-speaker variation, which would have been very difficult to observe without any one of these factors. This is in contrast to many dialectological studies (especially in the German tradition), which have rarely used acoustic analysis, and often rely on single occurrences of words. These advantages have come at some cost however. Using multiple repetitions of words limits the total number of words. Another limitation of this study is that the speakers were from a fairly small age range. The slight correlation with age found might be more evident if a wider range of ages were chosen.

This study also raises questions about speech perception. How might the extreme variation found in this study affect perception? Would these speakers perceive the fortis/lenis distinction as well as other speakers without such variation? This study has taken the first step in addressing such questions, by examing the two-way effects of loan words on the dialect. Not only are some loan words adapted to fit the phonology of the dialect, but some words retain their original pronunciation, and as the number of these words grow, they can have a reverse effect on the loan words which had formerly been adapted.

8. REFERENCES

Allen, J. Sean, Joanne L. Miller, and David DeSteno. Individual talker differences in voice-onset-time. Journal of the Acoustical Society of America 113/1.544-552.

Bachmann, Armin R. 2000. Die Mundart von Eslarn in der Oberpfalz. Stuttgart: Steiner.

Barbour Stephen, and Patrick Stevenson. 1998. Variation im Deutschen: soziolinguistische Perspektiven. Berlin: De Gruyter

Berroth, Daniela. 2001. Altersbedingter Mundartgebrauch: Wandel und Kontinuität in einem mittelschwäbischen Dialekt. Stuttgart: Steiner.

Braune, W. 1987. Althochdeutsche Grammatik, 14th edn. by Hans Eggers. Tübingen: Niemeyer.

Bybee, Joan. 1985. Morphology: a study of the relation between meaning and form. Amsterdam: John Benjamins.

Bybee, Joan. 2000. Lexicalization of sound change and alternating environments. Papers in Laboratory Phonology 5, ed. by Michael Broe and Janet Pierrehumbert. 250-268. Cambridge: Cambridge University Press.

Chambers, J.K. 2002. Dynamics of dialect convergence. Journal of Sociolinguistics, 6/1.117-130.

Cho, Taehong, and Peter Ladefoged. 1999. Variation and universals in VOT: evidence from 18 languages. Journal of Phonetics 27.207-229.

Chomsky, Noam, and Morris Halle. 1968. The sound pattern of English. New York: Harper.

Clausing, Stephen. 1979. Notker's Anlautgesetz. Journal of English and Germanic Philology 78.358-373.

Crosswhite, Katherine. 1999. Intra-paradigmatic homophony avoidance in two dialects of Slavic. UCLA Working Papers in Linguistics.1-18.

Davis, Garry, Gregory Iverson, and Joseph Salmons. 1999. Peripherality and markedness. Beiträge zur Geschichte der deutschen Sprache und Literatur 121/2.177-200.

Docherty, Gerry. 1992. The timing of voicing in English obstruents. Berlin: Foris Publications.

Fischer, Hermann. 1904-1936. Schwäbisches Wörterbuch. Tübingen: Verlag der H. Laupp´schen Buchhandlung.

Flemming, Edward. to appear. Contrast and Perceptual Distinctiveness. The Phonetic Bases of Markedness, ed. by B. Hayes, R. Kirchner, and D. Steriade. Cambridge: Cambridge University Press.

Fourcin, A. J., and E. Abberton. 1971. First applications of a new laryngograph: medical and biological illustration, 21, 172–182; reprinted in Volta Rev. 74, 161–176.

Francis, Alexander L., Valter Ciocca, and Jojo Man Ching Yu. 2003. Accuracy and variability of acoustic measures of voicing onset. Journal of the Acoustical Society of America 113/2.1025-1032.

Goblirsch, Kurt Gustav. 1994. Consonant strength in Upper German dialects. North-Western European Language Evolution, Supplement 10. Denmark: Odense University Press.

Goblirsch, Kurt Gustav. 1997. Notker's Law and Consonant Strength. North-Western European Language Evolution 31-32.135-144.

Iverson, Gregory K. and Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. Phonology 12.369-397.

Jessen, Michael. 1997. Obstruents in German. Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung. Stuttgart: Universität Stuttgart.

Keller, R.E. 1961. German dialects: phonology and morphology. Manchester: Manchester University Press.

Kingston, John and Randy L. Diehl. 1994. Phonetic Knowledge. Language 70.419-454.

Kluge, Friedrich. 2002. Etymologisches Wörterbuch der deutschen Sprache, 24 ed. bearbeitet von Elmar Seebold. Berlin: Walter de Gruyter.

Kucera, H., and W.N. Francis. 1967. A computational analysis of present-day American English. Providence, RI: Brown University Press.

Liberman, A.M, K.S. Harris, H.S. Hoffman, and B. C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology 54.358-368.

Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H&H theory. Speech production and speech modeling, ed. by William J. Hardcastle and Alain Marchal. Dordrecht: Kluwer.

Lisker, Leigh. 1975. Is it VOT or a first-formant transition detector? Journal of the Acoustical Society of America 57/6.1547-1551.

Lisker, Leigh, and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. Word 20.384-422.

Lisker, Leigh, and Arthur S. Abramson. 1970. The voicing dimension: some experiments in comparative phonetics. Proceedings of the Sixth International Congress of Phonetic Sciences, 563-567. Prague: Academia.

Milroy, Lesley, and Matthew Gordon. 2003. Sociolinguistics: method and interpretation. Oxford: Blackwell.

Ohala, John J. 1974. Phonetic explanation in phonology. Proceedings of the parasession on natural phonology, 251-274. Chicago: Chicago Linguistic Society.

Page, Richard B. 1999. On Notker's Anlautgesetz. Interdigitations: essays for Irmengard Rauch, ed. by Gerald F. Carr, Wayne Harbert and Lihua Zhang. New York: Peter Lang Publishing.

Ruoff, Arno. 1981. Häufigkeitswörterbuch gesprochener Sprache. Tübingen: Niemeyer.

Russ, Charles V. J. (ed.) 1990. The dialects of Modern German. London: Routledge.

Sonderegger, Stefan. 1987. Althochdeutsche Sprache und Literatur. Berlin: de Gruyter.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. Papers in Laboratory Phonology 5, ed. by Michael Broe and Janet Pierrehumbert. Cambridge: Cambridge University Press.

Stevens, Kenneth N. 1998. Acoustic phonetics. Cambridge, MA: MIT Press.

Stevens, Kenneth N., and Dennis H. Klatt. 1974. Role of formant transition in the voiced-voiceless distinction for stops. Journal of the Acoustical Society of America 55.653-659.

Vennemann, Theo. 1988. Preference laws for syllable structure: and the explanation of sound change. Berlin: Mouton de Gruyter.

Wells, C.J. 1987. German: A linguistic history to 1945. Oxford: Clarendon Press.

Winteler, J. 1876. Die Kerenzer Mundart des Kantons Glarus in ihren Grunzügen dargestellt. Leipzig & Heidelberg: Winter.

Willi, Urs. 1996. Die segmentale Dauer als phonetischer Parameter von 'fortis' und 'lenis' bei Plosiven im Zürichdeutschen. Stuttgart: Steiner.

Wright, Richard. 2001 Perceptual cues in contrast maintenance. The role of speech perception phenomena in phonology, ed. by Elizabeth Hume and Keith Johnson, 251-277. San Diego: Academic Press.

8. Appendices

Appendix A. Wordlist items

SG spelling	Pronunciation according to SW	gloss	aspirated	attested first in century x (Kluge 2002)	lexical frequency - # per million (Ruoff 1981)	log frequency
backen	[pakʰə]	'to bake'		9th	88	4.48
Bahre	[parə]	'stretcher'		8th	2	0.69
beten	[pɛtə]	'to pray'		8th	26	3.26
Boot	[poːt]	'boat'		15th	16	2.77
Dach	[taχ]	'roof'		9th	56	4.03
danken	[tankʰə]	'to thank'		8th	4	1.39
Delle	[tɛlə]	'dent'		13th	0	0.00
Gang	[kaŋ]	'hallway, path'		8th	22	3.09
kann	[kʰaː]	'can.3.sg'	✓	(können) 8th	4262	8.36
Paare/ Pärlein	[par] / [pɛrlə]	'several, couples'		13th	20/ 10	2.71
packen	[pʰakʰə]	'to pack, leave'	✓	16th	38	3.64
Pech	[pɛç]	'bad luck'		9th	12	2.48
Pelz	[pɛlts]	'fir, pelt'		10th	0	0.00
Peter	[pʰetɐ]	personal name	✓	4th-8th	-	1.61
Post	[poːʃt]	'mail'		16th	24	3.18
Tag	[tak]	'day'		8th	1722	7.45
tanken	[tʰankʰə]	'to fill up car with gasoline'	✓	18th-20th	2	0.69
Tanne	[tanə]	'fir tree'		9th	20	3.00
Tasse	[tʰasə]	'cup'		16th	8	2.08
Teller	[tɛlɐ]	'plate'		13th	6	1.79
Toast	[tʰoːʃt]	'toasted bread'	✓	19th	0	0.00
Tod	[toːt]	'death'		8th	8	2.08
überbacken	[ibɐpakə]	'to brown, broil'		8th	0	0.00
Verdacht	[fɐtaχt]	'suspicion'		16th	2	0.69
vergangen	[fɐkaŋə]	'former'		(gehen) 8th	0	0.00
verkanntet	[fɐkʰantet]	'jammed'	✓	not listed	2	0.69
verpacken	[fɐpʰakə]	'to wrap up, pack'	✓	16th	6	1.79
vertauschen	[fɐtaʊʃə]	'to exchange'		16th	0	0.00

APPENDIX B – WORDLIST

<u>Wortliste</u>

Bitte sprechen Sie jedes Wort in der folgenden Liste auf Schwäbisch, wie Sie normalerweise Schwäbisch sprechen. Die Wörter sind auf Hochdeutsch geschrieben, und in dem Satz "Sagen Sie ____ lauter" eingebettet.

1. Sagen Sie 'Boot' lauter
2. Sagen Sie 'auf' lauter
3. Sagen Sie 'Gang' lauter
4. Sagen Sie 'Pech' lauter
5. Sagen Sie 'Bahre' lauter
6. Sagen Sie 'packen' lauter
7. Sagen Sie 'Teller' lauter
8. Sagen Sie 'vertauschen' lauter
9. Sagen Sie 'Bahre' lauter
10. Sagen Sie 'Peter' lauter
11. Sagen Sie 'vertauschen' lauter
12. Sagen Sie 'Delle' lauter
13. Sagen Sie 'Post' lauter
14. Sagen Sie 'Pfarrer' lauter
15. Sagen Sie 'kann' lauter
16. Sagen Sie 'Toast' lauter
17. Sagen Sie 'überbacken' lauter
18. Sagen Sie 'beten' lauter
19. Sagen Sie 'Stein' lauter
20. Sagen Sie 'Dach' lauter
21. Sagen Sie 'Tasse' lauter
22. Sagen Sie 'Pelz' lauter
23. Sagen Sie 'Verdacht' lauter
24. Sagen Sie 'schön' lauter
25. Sagen Sie 'Tanne' lauter
26. Sagen Sie 'Tod' lauter
27. Sagen Sie 'Bluse' lauter
28. Sagen Sie 'verkanntet' lauter
29. Sagen Sie 'Toast' lauter
30. Sagen Sie 'faul' lauter
31. Sagen Sie 'Ehre' lauter
32. Sagen Sie 'Tag' lauter
33. Sagen Sie 'beten' lauter
34. Sagen Sie 'Gruß' lauter
35. Sagen Sie 'Post' lauter
36. Sagen Sie 'Gang' lauter
37. Sagen Sie 'Pelz' lauter
38. Sagen Sie 'Samstag' lauter
39. Sagen Sie 'Häuschen' lauter
40. Sagen Sie 'Tasse' lauter
41. Sagen Sie 'Paare' lauter
42. Sagen Sie 'Wurst' lauter
43. Sagen Sie 'Tanne' lauter
44. Sagen Sie 'schwätzen' lauter
45. Sagen Sie 'Teller' lauter
46. Sagen Sie 'vergangen' lauter
47. Sagen Sie 'Tod' lauter
48. Sagen Sie 'faul' lauter
49. Sagen Sie 'Peter' lauter
50. Sagen Sie 'Stein' lauter
51. Sagen Sie 'schaffen' lauter
52. Sagen Sie 'überbacken' lauter
53. Sagen Sie 'Delle' lauter
54. Sagen Sie 'Pfarrer' lauter
55. Sagen Sie 'tanken' lauter
56. Sagen Sie 'Pech' lauter
57. Sagen Sie 'Tod' lauter
58. Sagen Sie 'schön' lauter
59. Sagen Sie 'kann' lauter
60. Sagen Sie 'Boot' lauter
61. Sagen Sie 'Ehre' lauter
62. Sagen Sie 'Samstag' lauter
63. Sagen Sie 'verpacken' lauter
64. Sagen Sie 'Tanne' lauter
65. Sagen Sie 'Gruß' lauter
66. Sagen Sie 'Bluse' lauter
67. Sagen Sie 'danken' lauter
68. Sagen Sie 'Toast' lauter
69. Sagen Sie 'Wurst' lauter
70. Sagen Sie 'schaffen' lauter

71. Sagen Sie ‘Tag’ lauter
72. Sagen Sie ‘Post’ lauter
73. Sagen Sie ‘auf’ lauter
74. Sagen Sie ‘Häuschen’ lauter
75. Sagen Sie ‘Verdacht’ lauter
76. Sagen Sie ‘Pech’ lauter
77. Sagen Sie ‘schwätzen’ lauter
78. Sagen Sie ‘lieben’ lauter
79. Sagen Sie ‘Peter’ lauter
80. Sagen Sie ‘Wurst’ lauter
81. Sagen Sie ‘vergangen’ lauter
82. Sagen Sie ‘Stein’ lauter
83. Sagen Sie ‘Bahre’ lauter
84. Sagen Sie ‘Tasse’ lauter
85. Sagen Sie ‘faul’ lauter
86. Sagen Sie ‘verpacken’ lauter
87. Sagen Sie ‘Delle’ lauter
88. Sagen Sie ‘schön’ lauter
89. Sagen Sie ‘Dach’ lauter
90. Sagen Sie ‘Bluse’ lauter
91. Sagen Sie ‘vertauschen’ lauter
92. Sagen Sie ‘beten’ lauter
93. Sagen Sie ‘Samstag’ lauter
94. Sagen Sie ‘verpacken’ lauter
95. Sagen Sie ‘Boot’ lauter
96. Sagen Sie ‘lieben’ lauter
97. Sagen Sie ‘tanken’ lauter
98. Sagen Sie ‘kann’ lauter
99. Sagen Sie ‘Pfarrer’ lauter
100. Sagen Sie ‘Tod’ lauter
101. Sagen Sie ‘Ehre’ lauter
102. Sagen Sie ‘überbacken’ lauter
103. Sagen Sie ‘schwätzen’ lauter
104. Sagen Sie ‘Toast’ lauter
105. Sagen Sie ‘Tanne’ lauter
106. Sagen Sie ‘backen’ lauter
107. Sagen Sie ‘Pelz’ lauter
108. Sagen Sie ‘schaffen’ lauter
109. Sagen Sie ‘faul’ lauter
110. Sagen Sie ‘danken’ lauter
111. Sagen Sie ‘Post’ lauter
112. Sagen Sie ‘Gruß’ lauter
113. Sagen Sie ‘Paare’ lauter
114. Sagen Sie ‘Teller’ lauter
115. Sagen Sie ‘Bluse’ lauter
116. Sagen Sie ‘Gang’ lauter
117. Sagen Sie ‘lieben’ lauter
118. Sagen Sie ‘beten’ lauter
119. Sagen Sie ‘schön’ lauter
120. Sagen Sie ‘auf’ lauter
121. Sagen Sie ‘packen’ lauter
122. Sagen Sie ‘Tasse’ lauter
123. Sagen Sie ‘Wurst’ lauter
124. Sagen Sie ‘verkanntet’ lauter
125. Sagen Sie ‘Pelz’ lauter
126. Sagen Sie ‘schaffen’ lauter
127. Sagen Sie ‘tanken’ lauter
128. Sagen Sie ‘Delle’ lauter
129. Sagen Sie ‘Pfarrer’ lauter
130. Sagen Sie ‘Verdacht’ lauter
131. Sagen Sie ‘Teller’ lauter
132. Sagen Sie ‘Stein’ lauter
133. Sagen Sie ‘vergangen’ lauter
134. Sagen Sie ‘Tod’ lauter
135. Sagen Sie ‘Ehre’ lauter
136. Sagen Sie ‘backen’ lauter
137. Sagen Sie ‘Dach’ lauter
138. Sagen Sie ‘Gang’ lauter
139. Sagen Sie ‘danken’ lauter
140. Sagen Sie ‘Peter’ lauter
141. Sagen Sie ‘lieben’ lauter
142 Sagen Sie ‘packen’ lauter
143. Sagen Sie ‘Samstag’ lauter
144. Sagen Sie ‘tanken’ lauter
145. Sagen Sie ‘Tanne’ lauter
146. Sagen Sie ‘auf’ lauter
147. Sagen Sie ‘Bahre’ lauter
148. Sagen Sie ‘Boot’ lauter
149. Sagen Sie ‘kann’ lauter
150. Sagen Sie ‘Häuschen’ lauter

151. Sagen Sie ‘vertauschen’ lauter
152. Sagen Sie ‘Post’ lauter
153. Sagen Sie ‘Bluse’ lauter
154. Sagen Sie ‘Paare’ lauter
155. Sagen Sie Ehre lauter
156. Sagen Sie ‘Tasse’ lauter
157. Sagen Sie ‘verkanntet’ lauter
158. Sagen Sie ‘schaffen’ lauter
159. Sagen Sie ‘Gang’ lauter
160. Sagen Sie ‘Delle’ lauter
161. Sagen Sie ‘Gruß’ lauter
162. Sagen Sie ‘Verdacht’ lauter
163. Sagen Sie ‘schwätzen’ lauter
164. Sagen Sie ‘vergangen’ lauter
165. Sagen Sie ‘Teller’ lauter
166. Sagen Sie ‘überbacken’ lauter
167. Sagen Sie ‘Tag’ lauter
168. Sagen Sie ‘Samstag’ lauter
169. Sagen Sie ‘packen’ lauter
170. Sagen Sie ‘beten’ lauter
171. Sagen Sie ‘vertauschen’ lauter
172. Sagen Sie ‘auf’ lauter
173. Sagen Sie ‘Bahre’ lauter
174. Sagen Sie ‘Pech’ lauter
175. Sagen Sie ‘danken’ lauter
176. Sagen Sie ‘faul’ lauter
177. Sagen Sie ‘verpacken’ lauter
178. Sagen Sie ‘Stein’ lauter
179. Sagen Sie ‘Paare’ lauter
180. Sagen Sie ‘Pelz’ lauter
181. Sagen Sie ‘backen’ lauter
182. Sagen Sie ‘Wurst’ lauter
183. Sagen Sie ‘kann’ lauter
184. Sagen Sie ‘Peter’ lauter
185. Sagen Sie ‘verkanntet’ lauter
186. Sagen Sie ‘schön’ lauter
187. Sagen Sie ‘Dach’ lauter
188. Sagen Sie ‘Tag’ lauter
189. Sagen Sie ‘Gruß’ lauter
190. Sagen Sie ‘verkanntet’ lauter
191. Sagen Sie ‘Häuschen’ lauter
192. Sagen Sie ‘danken’ lauter
193. Sagen Sie ‘Boot’ lauter
194. Sagen Sie ‘verpacken’ lauter
195. Sagen Sie ‘backen’ lauter
196. Sagen Sie ‘Dach’ lauter
197. Sagen Sie ‘packen’ lauter
198. Sagen Sie ‘Paare’ lauter
199. Sagen Sie ‘Pfarrer’ lauter
200. Sagen Sie ‘überbacken’ lauter
201. Sagen Sie ‘schwätzen’ lauter
202. Sagen Sie ‘backen’ lauter
203. Sagen Sie ‘vergangen’ lauter
204. Sagen Sie ‘Pech’ lauter
205. Sagen Sie ‘Verdacht’ lauter
206. Sagen Sie ‘tanken’ lauter
207. Sagen Sie ‘Tag’ lauter
208. Sagen Sie ‘Häuschen’ lauter
209. Sagen Sie ‘Toast’ lauter
210. Sagen Sie ‘lieben’ laut