Phonology **19** (2002) 189–218.
DOI: 10.1017/S0952675702004311 Printed in the United Kingdom

*Laryngeal features in German**

Michael Jessen
Bundeskriminalamt, Wiesbaden

Catherine Ringen
University of Iowa

It is well known that initially and when preceded by a word that ends with a voiceless sound, German so-called 'voiced' stops are usually voiceless, that intervocalically both voiced and voiceless stops occur and that syllable-final (obstruent) stops are voiceless. Such a distribution is consistent with an analysis in which the contrast is one of [voice] and syllable-final stops are devoiced. It is also consistent with the view that in German the contrast is between stops that are [spread glottis] and those that are not. On such a view, the intervocalic voiced stops arise because of passive voicing of the non-[spread glottis] stops. The purpose of this paper is to present experimental results that support the view that German has underlying [spread glottis] stops, not [voice] stops.

---

1 Introduction

In spite of the fact that voiced (obstruent) stops in German (and many other Germanic languages) are markedly different from voiced stops in languages like Spanish, Russian and Hungarian, all of these languages are usually claimed to have stops that contrast in voicing. For example, Wurzel (1970), Rubach (1990), Hall (1993) and Wiese (1996) assume that German has underlying voiced stops in their different accounts of German syllable-final devoicing in various rule-based frameworks. Similarly, Lombardi (1999) assumes that German has underlying voiced obstruents in her optimality-theoretic (OT) account of syllable-final laryngeal neutralisation and assimilation in obstruent clusters. Noske (1999) also

* We wish to acknowledge our debt to Jill Beckman, Stuart Davis, Grzegorz Dogil, Paula Fikkert, Gunnar Hansson, Martin Krämer, Donka Minkova, Olga Petrova, Rosemary Plapp, Jurek Rubach, Dan Silverman, Otto Sohn, Szilárd Szentgyörgyi, members of the audiences at *WCCFL* 20 and *HILP* 5, reviewers and an associate editor of *Phonology* for valuable comments and suggestions, Vannakhane Daongam for assistance with recording subjects and performing measurements in Experiment II, and Tomomasa Sasa and Chungyuan Jing for help with manuscript preparation. We are especially grateful to George Woodworth for his generous assistance and advice on the statistical analysis. We are responsible for any errors. The research reported in this paper was supported, in part, by a grant from the Arts and Humanities Initiative, University of Iowa (1999–2000) to Catherine Ringen. Authors' names are listed alphabetically.

presents an OT account of German syllable-final neutralisation in which she assumes that the feature [voice] is distinctive.

Other researchers, investigating rather different sets of questions, have suggested that the relevant feature is not [voice], but rather [tense] or [spread glottis]. These include Kloeke (1982), Meinhold & Stock (1982), Anderson & Ewen (1987), Jessen (1989, 1996, 1998), Iverson & Salmons (1995) and Petrova *et al.* (2000). Recent support for the claim that [spread glottis] and not [voice] is the relevant feature in German comes from modern acoustic-phonetic studies of German dialects. According to the review of VOT studies by Braun (1996), the vast majority of German dialects show long-lag VOT in orthographic *p*, *t*, *k*, which are interpreted as aspirated, and short-lag VOT in *b*, *d*, *g*, interpreted as voiceless unaspirated. Prevoicing (negative VOT) in *b*, *d*, *g*, i.e. full voicing, is very uncommon in word/utterance-initial position. This shows that most German dialects implement the distinction with presence *vs*. absence of aspiration word/utterance-initially, so that both stop series are voiceless in that position. A recent VOT study on the Palatinate dialect of German by Scharf & Masur (2002) shows the same pattern.

German stops are voiceless utterance-initially (and word-initially unless preceded by a voiced sound), both voiced and voiceless stops occur intervocalically and stops are voiceless syllable-finally. This distribution is consistent with an analysis in which the contrast is one of [voice], and syllable-final stops are devoiced. It is also consistent with the view that in German the contrast is between stops that are [spread glottis] and those that are not. On such a view, the intervocalic voiced stops arise because of passive voicing of the non-[spread glottis] stops. Passive voicing means that without the devoicing effect of glottal spreading, stops can be voiced during most or all of closure if they occur in the context of sonorant sounds, even if there are no active voicing gestures (vocal fold slacking, tongue-root advancement, etc.) on the part of the speaker (Westbury 1983, Westbury & Keating 1986). The [spread glottis] analysis is slightly more attractive, because it more easily accounts for the initial voicelessness of 'voiced' stops in German. Specifically, initial position is a context where stop voicing can only occur by active means (Westbury & Keating 1986). Such active voicing in initial context is usually expected for languages that employ [voice], but not if [spread glottis] is the relevant feature.

However, this is not the only difference between the two views. If there are underlying voiced stops in German, then these stops should be voiced in word-medial clusters when preceded by a voiceless stop. For example, in *Jagden* 'hunting PL', the underlying /g/ of *Jag* is claimed to surface as [k], because of syllable-final devoicing, but the underlying /d/ is supposed to surface as voiced (<*Ja*/g+d/+*en*) (Rubach 1990, Lombardi 1991, 1999, Noske 1999). On the other hand, if the only voicing of German stops comes about by passive voicing between sonorants, as claimed by Jessen (1998) and by Petrova *et al.* (2000), then both stops in a word like *Jagden* should be voiceless, because neither stop is flanked by sonorants.[1]

Petrova *et al.* (2000) present an OT-based typology which builds on the insights of Lombardi's (1999) OT account of syllable-final laryngeal neutralisation and assimilation in obstruent clusters, but which avoids the empirical inadequacies of her analysis.[2] They claim that German has underlying [spread glottis] stops, not underlying voiced stops, as claimed by Lombardi. According to this view, the voiced stops that do occur in German are the result of passive voicing.[3]

The purpose of this paper is to present experimental results that bear on the issue of whether German has underlying voiced stops or [spread glottis] stops. In our studies, we investigated the voicing of so-called 'voiced' stops in medial clusters. As noted above, it is well known that German stops are voiceless (aspirated or unaspirated) in utterance-initial position (see Braun 1996, Jessen 1998 for discussion), but data on voicing and aspiration in medial clusters have not been available. The structure of this paper is as follows. First, we will briefly sketch the differences between Lombardi's [voice] account of German and the [spread glottis] account of Petrova *et al.* (2000). Second, we present results from our first experiment which support the [spread glottis] account. Third, we report on another experiment which provides additional support for the [spread glottis] account and give an OT (McCarthy & Prince 1993, 1995, Prince & Smolensky 1993) account for these new data.

We assume that phonological features should be grounded in phonetic reality and that phonological analyses (including OT accounts) primarily model the categorical aspects of linguistic sounds and sound structure, while modern phonetics focuses on more gradient or variable aspects of speech production and perception (cf. Cohn 1993, Clements & Hertz 1996, Keating 1996, among others). We proceed with a few preliminaries on each of these two aspects of the relation between phonology and phonetics.

The two phonological features assumed in this paper are [spread glottis] and [voice], commonly referred to and formally treated as 'laryngeal features' (Clements 1985). The phonetic grounding of the feature [spread glottis], which was first proposed by Halle & Stevens (1971), is most straightforwardly stated in articulatory terms. We take sounds

[1] In Jessen (1998), and the studies reviewed therein, closure voicing in German occurs occasionally in word-initial orthographic *b*, *d*, *g* preceded by pause or by a voiceless obstruent (hence in a position where the context to the left is not a sonorant). But this occurrence of voicing is rare and unsystematic. Many speakers never produce voiced stops in these contexts, while others often produce only some tokens of *b*, *d*, *g* with voicing and the majority of tokens voiceless. We interpret this unsystematic occurrence of voice production as hypercorrect pronunciation, which should not be modelled in the phonology.

[2] Petrova *et al.* (2000) note that Lombardi's analysis does not accurately describe voice assimilation in Russian and Hungarian, both [voice] languages.

[3] Actually, morpheme-internally between sonorants, by Lexicon Optimisation (Prince & Smolensky 1993) there would be underlying voiced stops in words such as *Magen* 'stomach' in the [spread glottis] account of Petrova *et al.* (2000).

specified as [spread glottis] to be implemented with an active glottal opening gesture. The most reliable, but also methodologically the most difficult, way of demonstrating an active glottal opening gesture experimentally is by showing activity of the posterior cricoarytenoid muscle, which spreads the glottis (Löfqvist & Yoshioka 1980).[4] On the acoustic/auditory level, [spread glottis] in stop production is implemented primarily by aspiration (often above 50 ms, but with much variation according to boundaries, stress, place of articulation and other factors). If the stop occurs before a sonorant consonant, [spread glottis] might also be manifested as partial or full devoicing of that sonorant. Both of these criteria for the specification of [spread glottis] will be employed in this paper and both can be unified by the use of long-lag VOT as the measurement criterion (Lisker & Abramson 1964).[5]

We take the feature [voice] in stop production to be prototypically implemented with voicing during closure. Stop voicing during closure can be achieved by the active enlargement of the oral cavity mentioned above.[6] Stop voicing can also occur in a language if the stop appears in contexts and conditions favourable for the occurrence of voicing. We use the term 'passive voice' or 'passive voicing' to refer to this possibility. The fact that oral gestures enlarging the oral cavity – as well as pulmonic gestures ensuring sufficient subglottal pressure – are required in the reliable production of stop voicing shows that, in terms of its phonetic implementation, [voice] is not just a laryngeal feature and that, unlike

[4] Slighter amounts of glottal opening can also be created passively, i.e. due to biomechanical-aerodynamic factors without muscular activity. Passive glottal opening in non-[spread glottis] stops has been explained on the theoretical level for English by Stevens (1998) and demonstrated empirically for German by Jessen (1998). Consequently, the categorical difference between [spread glottis] and non-[spread glottis] stops in English and German is not so much in terms of presence *vs.* absence of glottal opening as in terms of active (and large) glottal opening *vs.* passive (and slight) glottal opening. These aspects and further phonetic criteria for active *vs.* passive glottal opening are discussed in Jessen (1998).

[5] Occasionally, stops specified as [spread glottis] are not implemented with aspiration or sonorant devoicing. One context where this can occur is in front of a second stop (see the word *Reitgerte* to be analysed in (18)). In this situation it is possible that the [spread glottis] feature of the first stop cannot be manifested as aspiration because the stop is either not released at all or because its release is immediately followed by the formation of the closure associated with the second stop. A second context where [spread glottis] is usually not implemented with aspiration is in clusters with a preceding fricative. This case is explained in more detail by Iverson & Salmons (1995). A third situation (style-dependent) where a stop specified as [spread glottis] does not show aspiration or sonorant devoicing occurs if in a word like *leiten* 'lead' the canonical schwa of the second syllable is not produced and the stop is instead released directly into the syllabic production of the final nasal (Kohler 1979). The term 'nasal release' used in Table I refers to this possibility.

[6] According to Jessen (2001), stops specified as [voice] might, in specific contexts and languages, be implemented with 'low frequency properties' (Kingston & Diehl 1994, 1995) other than closure voicing, in particular, low fundamental frequency, low first formant frequency and probably breathy/slack voice in adjacent vowels.

[spread glottis], it does not have a simple, single-gesture articulatory definition. Instead, with its voicing-during-closure correlate, it is more straightforwardly defined in acoustic/auditory terms (Stevens & Keyser 1989, Ladefoged 1997).

In our experiments we found variation in (passive) voice in intervocalic position. We consider this voicing to be phonetically conditioned by the voiced context. The detail of the variation has a phonetic explanation. In particular, the facts that females are more likely to exhibit variation in stop voicing and that velar stops are more likely to be voiceless than alveolar stops have phonetic explanations, and hence are not appropriately analysed phonologically. Similarly, the *degrees* of aspiration (variation in VOT) which depend on, for example, stress or place of articulation have phonetic explanations and hence, we believe, should be treated as part of phonetics, not phonology.

2 Voice analysis

Lombardi (1999) presents an analysis of voice assimilation in obstruent clusters which is intended to account for voicing assimilation patterns found in natural languages. Assuming that voice is privative, she adopts two identity constraints. One is a positional faithfulness constraint of the type proposed by Beckman (1997), IDENTONSET(lar), which requires that segments in an onset in a presonorant position retain their underlying specifications for voice. The other is the general constraint IDENT(lar), which requires that corresponding input and output segments have the same laryngeal specification. In addition, she adopts a markedness constraint *LAR, which prohibits laryngeal specifications ([voice] as well as [spread glottis]), and a constraint AGREE, which requires that obstruents in a cluster agree in voicing.

(1) a. IDENTONSET(lar)
A presonorant consonant in an onset should be faithful to underlying laryngeal specification.

b. IDENT(lar)
Consonants should be faithful to underlying laryngeal specification.

c. AGREE
Obstruent clusters should agree in voicing.

d. *LAR
Do not have laryngeal features.

Lombardi claims that in German the constraints are ranked IDENTONSET (lar) ≫ *LAR ≫ AGREE, IDENT(lar).

We give relevant German data in orthographic form in (2), because the two analyses to be contrasted are based on different claims about what the phonetic facts are. Lombardi assumes that orthographic *b*, *d*, *g* are underlyingly voiced and that orthographic *p*, *t*, *k* are underlyingly unmarked for a laryngeal feature. Data cited in Lombardi (1999) are indicated by [L].

(2)

Lob	'praise NOM' [L]	Lobes	'praise GEN' [L]
Tag	'day SG'	Tage	'day PL'
Tier	'animal'		
jagen	'to hunt' [L]	(sie) jagten	'(they) hunted'
Jagd	'hunting SG' [L]	(die) Jagden	'(the) hunting PL' [L]
(es) deckte	'(it) covered'		
reiten	'to ride'	Reitgerte	'riding crop'

So-called syllable-final devoicing is accounted for as in (3):

(3)

	IdOns(lar)	*Lar	Agree	Id(lar)
a. /tɑg/				
☞ i. [tɑk]				*
ii. [tɑg]		*!		
b. Ja/g+d/+en 'hunting PL'				
☞ i. Ja[k.d]en		*	*	*
ii. Ja[k.t]en	*!			**

An underlying voiced obstruent retains its voice feature when it is in an onset before a sonorant, as illustrated by the second example in (3) and by the examples in the tableaux in (4):

(4)

	IdOns(lar)	*Lar	Agree	Id(lar)
a. ja/g/+en				
☞ i. ja.[g]en		*		
ii. ja.[k]en	*!			*
b. /d/ec/k+t/+e				
☞ i. [d]e[k.t]e		*		
ii. [t]e[k.t]e	*!			*

An intervocalic voiceless stop is faithful to its input laryngeal specification because it is in an onset, as indicated in (5):

(5)

rei/t/+en	IdOns(lar)	*Lar	Agree	Id(lar)
☞ a. rei.[t]en				
b. rei.[d]en	*!	*		*

The optimal candidate for an input cluster as in (6) with a voiceless stop followed by a voiced stop has an identical voiceless–voiced cluster.

(6)

Rei/t+g/erte	IdOns(lar)	*Lar	Agree	Id(lar)
☞ a. Rei[t.g]erte		*	*	
b. Rei[d.g]erte		**!		*
c. Rei[t.k]erte	*!			*

The optimal output for an input such as *jagten* '(they) hunted' in (7), with a voiced stop followed by a voiceless stop, has a voiceless cluster:

(7)

ja/g+t/+en	IdOns(lar)	*Lar	Agree	Id(lar)
☞ a. ja[k.t]en				*
b. ja[g.t]en		*!	*	

3 Spread glottis account

The OT account of German presented in Petrova *et al.* (2000) has the constraints in (8)–(12).[7] The first is a high-ranked constraint against voiced [spread glottis] stops. Although Lombardi does not discuss this constraint, we assume that she would also need such a constraint to prevent an input with a voiced [spread glottis] stop from being designated as optimal in presonorant position in an onset.

(8) *[voi/sg]

Voiced spread glottis stops are prohibited (Davis 1998, Petrova *et al.* 2000).[8]

The second is a faithfulness constraint for the feature [spread glottis]:

(9) Ident-IO[sg]

Correspondent input and output segments have the same specification for [spread glottis] (McCarthy & Prince 1995, Petrova *et al.* 2000).

Additional constraints are PassiveVoice and markedness constraints against voiced obstruents and [spread glottis] segments:

(10) PassiveVoice

Obstruents are voiced between sonorants (Petrova *et al.* 2000).

A violation is assessed for an obstruent that is between sonorants but not voiced.

(11) *[voice]

Voiced obstruents are prohibited (Alderete 1997).

A violation is assessed for any obstruent that is voiced, whether the voice is 'passive' or underlying.

(12) *[sg]

[spread glottis] segments are prohibited.

[7] In the analysis of Petrova *et al.* (2000), privative voice is assumed for easier comparison with Lombardi's analysis, but nothing in that analysis or in the extensions proposed in this paper depend on this assumption. If it turns out that voice is binary, as recently argued by Rubach (1997) and Mascaró & Wetzels (2001), among others, nothing significant in our analysis would change.

[8] In languages such as Hindi, which do have voiced spread glottis stops, the constraint *[voi/sg] must be lower-ranked than it is in German.

According to this account, all underlying stops in German are voiceless, either with or without the feature [spread glottis], so there is no syllable-final devoicing of stops. Forms such as *Tag* and *Jagden* are analysed as in (13) and (14). Following Iverson & Salmons (1995), Petrova *et al.* (2000) assume that a [spread glottis] stop with a singly linked [spread glottis] feature is realised phonetically as an aspirated stop when it precedes a sonorant, and that a stop with a multiply linked [spread glottis] stop is unaspirated. The optimal output is correctly predicted to have an aspirated stop. Lombardi does not give an account of the aspiration in *Tag*:

(13)

/t$^{[sg]}$ɑk/	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
☞ a. [t$^{[sg]}$ɑk][9]					*
b. [tɑk]		*!			
c. [d$^{[sg]}$ɑk]	*!			*	*

Id-IO[sg] ≫ *[sg]

In (14) the input for *Jagden* 'hunting pl' has two voiceless non-[spread glottis] stops. Note that on this account, the second stop is voiceless (unaspirated) in the output, whereas it is voiced in (3b.i) on Lombardi's account.

(14)

Ja/k+t/+en	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
☞ a. Ja[kt]en					
b. Ja[kd]en				*!	

In (15) the input has an intervocalic non-[spread glottis] stop and the optimal candidate has a voiced stop. On Lombardi's account, the same surface form would be optimal, but the voicing of the intervocalic stop is assumed to be underlying.

(15)

ja/k/+en	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
☞ a. ja[g]en				*	
b. ja[k]en			*!		
c. ja[k$^{[sg]}$]en		*!	*		*

PassVoi ≫ *[voice]; Id-IO[sg] ≫ *[voice]

As noted above, it is well known that in utterance-initial position and after voiceless sounds, orthographic *b*, *d*, *g* are pronounced as voiceless unaspirated stops. On the [spread glottis] analysis, an initial non-[spread glottis] stop is predicted to be voiceless (unaspirated), as illustrated in

[9] The actual output, given a constraint proposed later in this paper, is [t$^{[sg]}$ak$^{[sg]}$]. See note 19 for further discussion.

(16). Note that here too, the optimal candidate is different from that in Lombardi's analysis, where the initial stop is voiced.

(16)

/t/eckte	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
☞ a. [t]eckte					
b. [d]eckte				*!	
c. [t[sg]]eckte		*!			*
d. [d[sg]]eckte	*!	*		*	*

In (17), where the input has an intervocalic [spread glottis] stop, the optimal output will violate PassiveVoice, but the alternative candidates violate higher-ranked constraints.

(17)

rei/t[sg]/+en	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
☞ a. rei[t[sg]]en			*		*
b. rei[t]en		*!	*		
c. rei[d]en		*!		*	
d. rei[d[sg]]en	*!			*	*

*[voi/sg] ≫ PassVoi; Id-IO[sg] ≫ PassVoi

On Lombardi's account, the optimal form is voiceless as well, but no account of the aspiration is given.

When the first stop in a cluster is underlyingly [spread glottis], as in (18), this feature is preserved in the output, but since the stop is not in presonorant position it is not aspirated. Hence, this analysis predicts that the entire cluster should be voiceless, with no aspiration on the second stop.

(18)

Rei/t[sg]+k/erte	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
a. Rei[tk]erte		*!			
☞ b. Rei[t[sg]k]erte					*

When the second stop in a cluster is specified as [spread glottis], it is predicted to be aspirated, as illustrated in (19) for *jagten* '(they) hunted'.

(19)

ja/k+t[sg]/+en	*[voi/sg]	Id-IO[sg]	PassVoi	*[voice]	*[sg]
a. ja[kt]en		*!			
☞ b. ja[kt[sg]]en					*

The differences between the [voice] analysis and the [spread glottis] analysis are summarised in (20):

(20) *Predictions*: [voice] *vs.* [spread glottis]

	[voice]	[spread glottis]
deckte	[d]eckte	[t]eckte
Tier	[t]ier	[t^h]ier
Jagden	Ja[kd]en	Ja[kt]en
jagten	ja[kt]en	ja[kt^h]en
Reitgerte	Rei[tg]erte	Rei[tk]erte

4 Experiment I

Six native speakers of German were recorded digitally. The recordings were transferred to a computer (SiliconGraphics Indy Workstation) and analysed acoustically with a speech analysis software package (ESPS/*waves*+). Three speakers were female and three male. Four (two males and two females) were from the North of Germany and two were from the South. Subjects' ages ranged from 25 to 40.

Each subject read a list of words twice. The words of interest for this study were randomised with fillers of different segmental make-up, thus making it unlikely that the subjects would guess the purpose of the experiment; when asked, all confirmed that they had not.

The list contained stimuli with word-medial stops contrasting for the voicing or glottal spreading feature that is at issue here. In some cases the contrasting stops occurred intervocalically (e.g. *Hecke* 'hedge' *vs. Egge* 'harrow'); in others they were preceded by a stop across a syllable boundary (e.g. *Sie jagten* 'they hunted' *vs. die Jagden* 'the huntings'). In a word-medial stop cluster, we did not do measurements for VOT or voice duration for the first stop (e.g. *die Ja**g**den*), since there is no question that this stop is voiceless.

For word-initial position the laryngeal contrast was not investigated here since there already is a broad range of literature available on this position (see Jessen 1998). But we did include an example of a word beginning with orthographic *d*.

Positive VOT was measured for intervocalic stops, for the second (post-voiceless) stop in a word-medial cluster and for initial stops in voiceless contexts.[10] Voicing during closure was also measured.[11] The

[10] Positive VOT, as measured in this experiment and in the experiment in §5, comprises the duration of the release burst plus any subsequent aspiration if it exists. Orthographic *b*, *d*, *g* in intervocalic position (e.g. in words like *Egge*) can be fully voiced during closure. Even in those fully voiced stops, voicing amplitude usually reduces gradually during closure and reaches a low value at the end of stop closure. Due to this reduction of voicing amplitude the stop burst visually interrupts the voicing pattern until voicing begins again right after the burst. Due to this voice-interruption pattern it is justified to report positive VOTs (corresponding to burst duration) even in fully voiced intervocalic stops. Often a different situation occurs in initial stops with prevoicing, where the burst can be very weak, and can be superimposed on a continuing voicing pattern. However such cases did not occur in our data (see note 11).

[11] Voicing during closure is expressed as a categorical judgment of the percentage of closure during which voicing occurs. Negative VOT, prevoicing or voicing lead

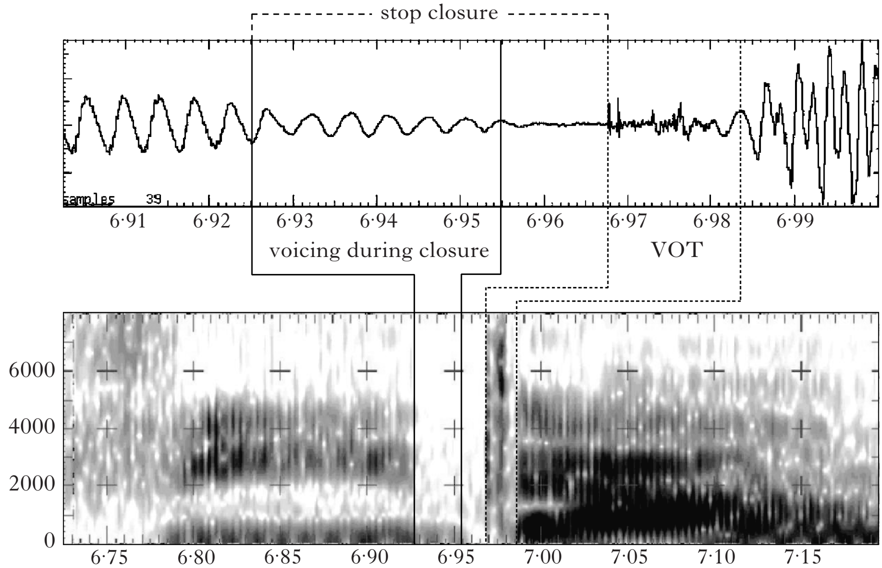

Figure 1

Measurement criteria for Experiment I, on the basis of waveform (upper) and spectrogram (lower) (from a token of *wieder*).

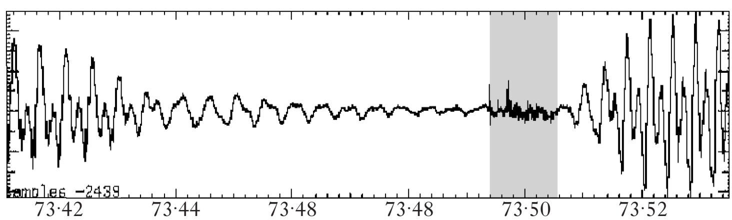

Figure 2

First special case (see note 10): fully voiced closure with discernible positive VOT (grey) (from a token of *Widder*).

labelling conventions are illustrated in Figs 1–3 (see §5 for more details on the measurement criteria employed for both experiments presented in this paper).

Positive VOT in the long-lag range was taken as evidence for aspiration, which in turn was interpreted as implementation of [spread glottis]. No absolute threshold for the difference between short and long-lag

(which are synonymous), as introduced by Lisker & Abramson (1964), is a special case of voicing during closure that applies most straightforwardly to word-initial stops preceded by silence or voiceless sounds. Our word-initial example (the first item in Table I) lacked any prevoicing consistently across speakers.

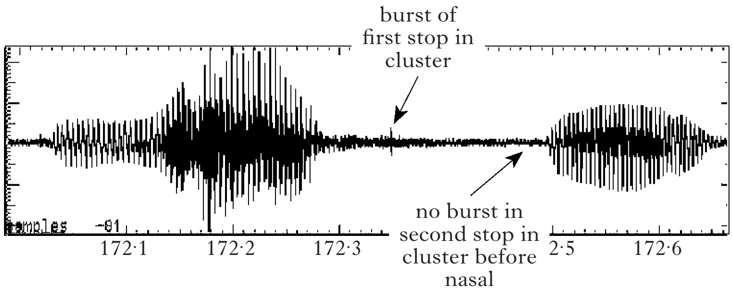

Figure 3

Second special case (see note 14): second stop in cluster is released directly into the nasal, i.e. it has no discernible burst (from a token of *die Jagden*).

VOT was used, one reason being that established boundary values are usually obtained from word-initial stops within a stressed syllable (see Lisker & Abramson 1964), whereas the stops we tested occurred word-medially in an unstressed syllable (with schwa as the syllable nucleus). Instead, the interpretation of positive VOT values was carried out in a more relational manner. According to this approach, the VOT values of voiceless stops in intervocalic word-medial position (which occur in the words *Hecke* to *Kette* in Table I) were taken as the baseline upon which the VOT values of the second stop in a word-medial cluster (see rows indicated with § in Table I) were interpreted. This approach is justified by the fact that the VOT values for intervocalic word-medial unstressed position reported in Table I are consistent with those reported in the German phonetics literature, where the majority of the evidence indicates presence of aspiration (Jessen 1998 for an overview).[12] That aspiration in German can occur even in word-medial unstressed syllables is noteworthy in its own right. It underlines the persistence of aspiration in German and argues against the view – frequently encountered in the literature – that aspiration or the occurrence of [spread glottis] is predictable from prosodic context, in particular foot-initial position (see for example Yu 1992).

The results of this study are given in Table I.[13] 'VOT_1' is the VOT (in ms) on the first reading for the stop given in bold. 'VOT_2' is the VOT on

[12] In some of the studies reported in Jessen (1998), including his own investigation, the values reported for intervocalic word-medial orthographic *p*, *t*, *k* are slightly longer than in the corresponding stops in Table I. This is at least partially due to the fact that criteria for the measurement of aspiration duration can differ slightly (e.g. voice onset *vs.* F2 onset, voice onset *vs.* aspiration offset, etc.). Positive VOT was the preferred method in this study because it is the most straightforward way of unifying aspiration and sonorant devoicing methodologically (cf. §1).

[13] Words not glossed in the text are *rege* 'lively', *wieder* 'again'.

Nasal release as indicated in Table I occurs when a stop is released directly into a syllabic nasal (Kohler 1979). Acoustically, tokens labelled as 'nasal release' have no visible burst and the nasal starts right after the end of closure. For this reason, standard VOT-measurement criteria could not be applied to these tokens.

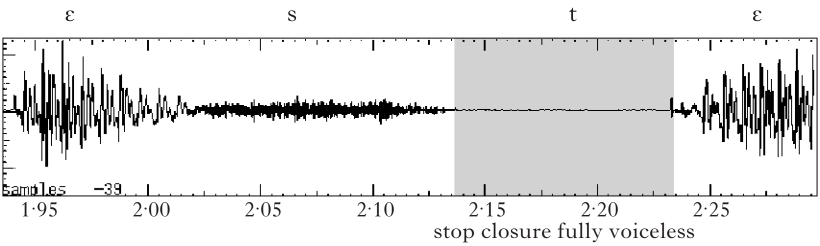

Figure 4

Initial stops are voiceless (from a token of *es deckte*, spoken by M-2).

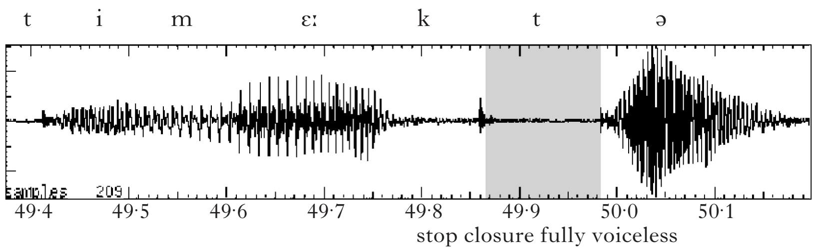

Figure 5

Second stop in word-medial cluster is voiceless (from a token of *die Mägde*, spoken by M-2).

the second reading. 'Voicing$_1$' indicates the degree of voicing during closure for the stop in bold on the first reading and 'Voicing$_2$' indicates the degree of voicing during closure for the stop in bold on the second reading. Information on voicing is only provided for orthographic *b*, *d*, *g*. Orthographic *p*, *t*, *k* are all voiceless (except for a very short tail of voicing into closure, which is probably universal and not perceivable, and hence can be ignored for present purposes). All test words are given in their orthographic form since what is at issue is whether the stops are voiced, aspirated, or voiceless and unaspirated. (Note that all the target sounds under the heading 'intervocalic' are single intervocalic stops.)

The results of this experiment support the analysis in which [spread glottis], not [voice], is the contrastive feature in German stops. All speakers had voiceless stops in initial position in *es deckte*, as illustrated in Fig. 4. This can be seen in the first row for each speaker in Table I. As noted above, this is consistent with results of earlier investigations of word-initial stops preceded by a pause or a voiceless sound in German (see Jessen 1998, which includes a literature review). It is the expected result, given the [spread glottis] analysis, but not the [voice] analysis. Second, none of the speakers had any voicing during closure of (orthographic) *d* in *Jagden* and (*die*) *Mägde* 'the maids' (see the last two rows in the table for each speaker, which are preceded by •). That is, the medial

M-1			VOT$_1$	VOT$_2$	Voicing$_1$	Voicing$_2$
word-initial		es **d**eckte	14	13	voiceless	voiceless
intervocalic		Hec**k**e	48	43		
		stec**k**e	45	46		
		hä**tt**e	47	42		
		Ke**tt**e	61	29		
		re**g**e	19	no release	fully voiced	fully voiced
		E**gg**e	32	19	fully voiced	fully voiced
		wie**d**er	9	9	fully voiced	fully voiced
		Wi**dd**er	12	8	fully voiced	fully voiced
clusters	§	es deck**t**e	52	46		
	§	Sek**t**e	52	39		
	§	sie jag**t**en	62	56		
	•	die Jag**d**en	18	20	voiceless	voiceless
	•	die Mäg**d**e	41	15	voiceless	voiceless
M-2						
word-initial		es **d**eckte	14	14	voiceless	voiceless
intervocalic		Hec**k**e	51	47		
		stec**k**e	50	44		
		hä**tt**e	45	51		
		Ke**tt**e	41	36		
		re**g**e	18	23	fully voiced	fully voiced
		E**gg**e	16	21	fully voiced	fully voiced
		wie**d**er	13	18	fully voiced	fully voiced
		Wi**dd**er	16	16	fully voiced	fully voiced
clusters	§	es deck**t**e	67	49		
	§	Sek**t**e	52	53		
	§	sie jag**t**en	55	60		
	•	die Jag**d**en	26	26	voiceless	voiceless
	•	die Mäg**d**e	28	24	voiceless	voiceless
M-3						
word-initial		es **d**eckte	13	18	voiceless	voiceless
intervocalic		Hec**k**e	41	36		
		stec**k**e	39	26		
		hä**tt**e	26	32		
		Ke**tt**e	29	31		
		re**g**e	12	no release	fully voiced	fully voiced
		E**gg**e	no release	6	fully voiced	fully voiced
		wie**d**er	8	9	fully voiced	fully voiced
		Wi**dd**er	11	9	mostly voiced	fully voiced
clusters	§	es deck**t**e	41	23		
	§	Sek**t**e	22	19		
	§	sie jag**t**en	nasal release	nasal release		
	•	die Jag**d**en	nasal release	nasal release	voiceless	voiceless
	•	die Mäg**d**e	17	19	voiceless	voiceless

F-1			VOT$_1$	VOT$_2$	Voicing$_1$	Voicing$_2$
word-initial		es **d**eckte	15	12	voiceless	voiceless
intervocalic		Hec**k**e	41	36		
		stec**k**e	39	40		
		hä**tt**e	46	36		
		Ke**tt**e	39	42		
		re**g**e	23	10	fully voiced	fully voiced
		E**gg**e	25	15	mostly voiced	mostly voiced
		wie**d**er	16	12	fully voiced	fully voiced
		Wi**dd**er	13	21	fully voiced	mostly voiced
clusters	§	es deck**t**e	36	35		
	§	Sek**t**e	28	24		
	§	sie jag**t**en	nasal release	nasal release		
	•	die Jag**d**en	nasal release	nasal release	voiceless	voiceless
	•	die Mäg**d**e	21	20	voiceless	voiceless
F-2						
word-initial		es **d**eckte	3	12	voiceless	voiceless
intervocalic		Hec**k**e	32	33		
		stec**k**e	27	28		
		hä**tt**e	37	46		
		Ke**tt**e	45	46		
	*	re**g**e	no release	29	mostly voiceless	mostly voiced
		E**gg**e	18	no release	mostly voiced	fully voiced
		wie**d**er	8	18	fully voiced	mostly voiced
		Wi**dd**er	12	14	mostly voiced	mostly voiced
clusters	§	es deck**t**e	51	50		
	§	Sek**t**e	error	71		
	§	sie jag**t**en	59	29		
	•	die Jag**d**en	22	38	voiceless	voiceless
	•	die Mäg**d**e	23	30	voiceless	voiceless
F-3						
word-initial		es **d**eckte	11	18	voiceless	voiceless
intervocalic		Hec**k**e	43	48		
		stec**k**e	32	36		
		hä**tt**e	43	39		
		Ke**tt**e	35	49		
	*	re**g**e	16	19	mostly voiceless	mostly voiced
	*	E**gg**e	21	20	voiceless	voiceless
		wie**d**er	15	14	mostly voiced	fully voiced
		Wi**dd**er	20	16	mostly voiced	mostly voiced
clusters	§	es deck**t**e	18	16		
	§	Sek**t**e	26	23		
	§	sie jag**t**en	18	19		
	•	die Jag**d**en	19	18	voiceless	voiceless
	•	die Mäg**d**e	25	17	voiceless	voiceless

Table I

Results of Experiment I. VOT$_1$ and VOT$_2$ = VOT (in ms) for stops in bold on first and second readings. Voicing$_1$ and Voicing$_2$ = degree of voicing during closure for stops in bold on first and second readings.

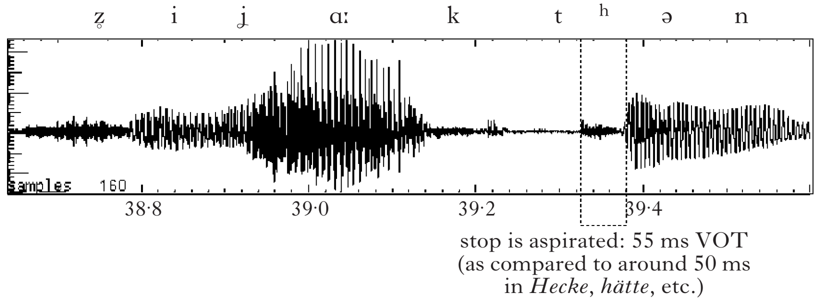

Figure 6

Usually aspiration on second orthographic *p*, *t*, *k* in cluster (from a token of *sie jagten*, spoken by M-2).

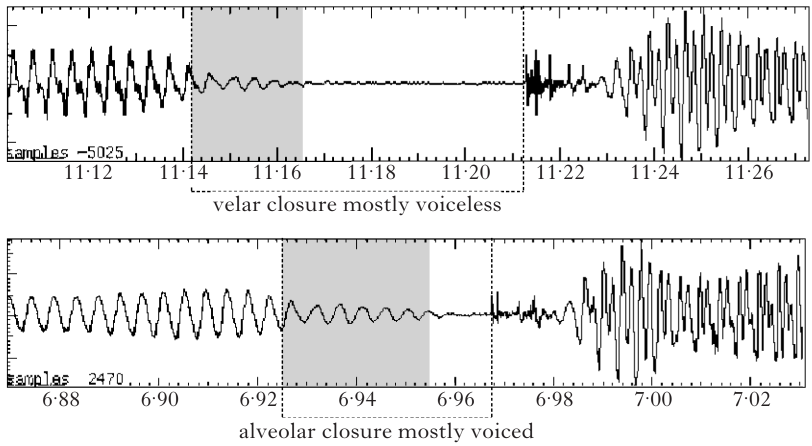

Figure 7

Orthographic *g* more (often) voiceless than orthographic *d* (from tokens of *rege* (upper) and *wieder* (lower), spoken by F-3).

cluster was always [kt], not [kd], as can be seen in the representative example in Fig. 5. These results are not surprising if the underlying stops are all voiceless, but it is difficult to explain if the underlying cluster is /gd/ or /kd/.

In words like *deckte* and *jagten* (see the forms preceded by § in Table I), all but one subject produced some orthographic *t*'s with aspiration, and of these, most subjects produced all of them with aspiration comparable to that found with single intervocalic [spread glottis] stops. For example, subject M-1 has VOTs for the (orthographic) *t* in *deckte*, *Sekte* 'sect' and *jagten* which are comparable to those for intervocalic aspirated stops in *Hecke*, *stecke* 'stick 1SG PRES', *hätte* 'have 1,3SG SUBJ II' and *Kette* 'chain'. This is the distribution of aspiration predicted by the [spread glottis] analysis, on the assumption the [spread glottis] stops before vowels are aspirated (see Fig. 6 for an example). The one anomaly is subject F-3, whose VOTs indicate that these *t*'s lack aspiration.

In intervocalic position, four speakers produced orthographic *d* and *g* as voiced (often fully voiced, i.e. without any interruption of voicing before release, otherwise with termination of voicing very shortly before release). The remaining two speakers (F-2, F-3) produced orthographic *d* as voiced or mostly voiced as well, but some orthographic *g*'s were voiceless between vowels (see Fig. 7). These (fully or mostly) voiceless productions are preceded by asterisks in Table I.

This variability suggests that the intervocalic voicing is phonetically conditioned by the surrounding voiced segments. This variability is difficult to understand in the [voice] analysis, since it would mean that underlying voiced stops would be optionally *devoiced* intervocalically, the environment most conducive to voicing. In this respect, our account differs slightly from that of Petrova *et al.* (2000) and Jessen & Ringen (2001a, b), in that we assume that the output of the phonology has only voiceless unaspirated stops for underlying non-[spread glottis] stops, as in (21) for *Egge* 'harrow', and that passive voicing is phonetic.

(21)

E/k/e	*[voi/sg]	Id-IO[sg]	*[voice]	*[sg]
a. E[g]e			*!	
☞ b. E[k]e				
c. E[k$^{[sg]}$]e		*!		*

The fact we find both voiced and voiceless stops intervocalically has a phonetic explanation. Voicing of non-[spread glottis] stops is conditioned by the flanking voiced segments. The fact that we find more voiceless velars than alveolars also has a phonetic explanation. There is a universal tendency for percentage of voicing and voice duration to proceed in the order [b] > [d] > [g]. This universal tendency can be explained by the aerodynamic principle that equalisation of the transglottal pressure difference, which is fatal to voicing, occurs more easily when the volume of the oral cavity behind the occlusion is small (as in [g]) than if it is large (as in [b]). Even more important is the fact that with occlusions further forward in the oral cavity there is more compliant tissue available that can passively yield to the build-up of intraoral air pressure, thereby creating additional expansion of the oral cavity volume and consequently delaying equalisation of the transglottal pressure drop (see Ohala 1983, Keating 1984 for details). Evidence for the order [b] > [d] > [g] in terms of voice duration has been found for German by Jessen (1998: 322). This is only a recent example of a number of studies on German with the same general result (Jessen 1998: 57ff for a review). Specifically, the reduced voicing of [g] as compared to [d] which has been found here has been reported before (see in particular Stock 1971).

The fact that we found voicing to be less likely with females than males also can be explained phonetically. On average, women have smaller vocal tracts than males, making passive voicing less likely to occur since equalisation of the transglottal pressure difference occurs more easily when the volume of the oral cavity behind the occlusion is small.[14]

5 Experiment II

Many phonologists, including Vennemann (1972) and Rubach (1990), have claimed that there is a difference between *Handlung* 'action' and *handlich* 'handy': the stop in *Handlung* is voiced, but the stop in *handlich* is voiceless. Following Vennemann (1972) and others, Lombardi suggests that forms such as *Handlung* are syllabified with [dl] in the onset, in spite of the fact that, in general, German does not allow such clusters in onsets.[15] On this analysis, since [d] is in an onset it is not expected to undergo devoicing. Rubach (1990), on the other hand, rejects the suggestion that the syllabification violates the general principles of German syllable structure and argues instead for a cyclic account in which the input is *Han*/dl/+*ung*. On cycle 1 the /d/ escapes syllable-final devoicing because the /d/ is syllabified as an onset of the syllable *dl* (with syllabic *l*), and on the second cycle, where syllable-final devoicing is no longer active, /d/ is resyllabified into the coda and /l/ is resyllabified as the onset of the syllable *lung*.

If, as assumed by Lombardi, [voice] is the relevant feature in German and obstruents are voiceless in codas, *Handlung* requires some additional assumptions, such as the cycle or unmotivated syllable structure, whereas *handlich* requires no such special treatment. On the [spread glottis] assumption, *Handlung* requires no special treatment, but the fact that /t/ in *handlich* is not subject to (phonetically conditioned) passive voice needs to be explained.

We designed a second experiment, which included words like *Handlung* (*Handl+ung*) and *handlich* (*hand+lich*). We also included words like *neblig* (*nebl+ig*) and *erheblich* (*erheb+lich*), in which the syllabification principles of German are not violated by the onset clusters *bl* and *gl*. Finally, we included words such as *Ummantl-ung* and *bekannt-lich*, which

[14] An alternative to assuming that passive voice is phonetically conditioned is to assume that it is phonological, as in Petrova *et al*. (2000), and that the variation in voicing is accounted for by phonetic constraints on the implementation of the feature [voice].

[15] As Wagner (2000) points out, the only reason for assuming the exceptional onsets is to account for the lack of devoicing.

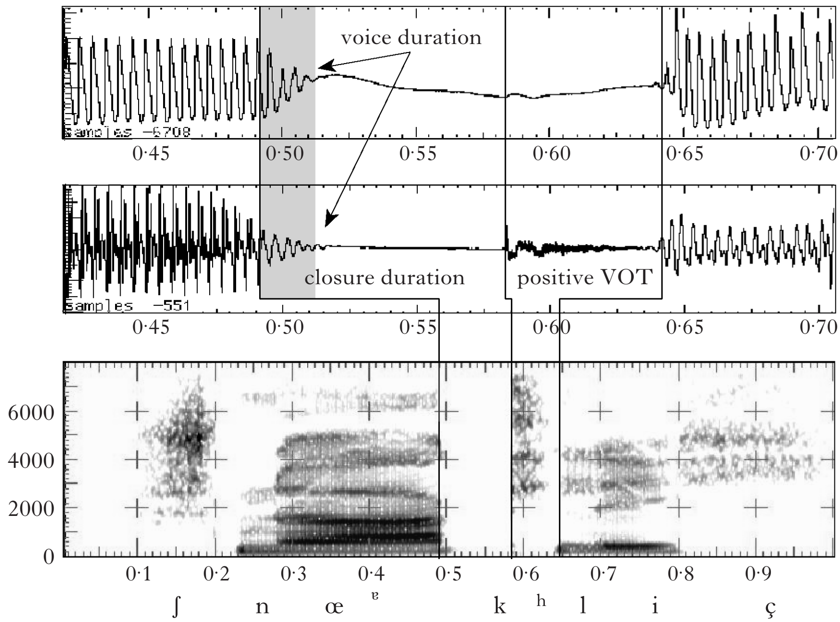

Figure 8

Measurement criteria for Experiment II, on the basis of EGG signal (upper), waveform (middle) and spectrogram (lower) (from a token of *schnörklig*).

are structurally identical to *Handlung* and *handlich* but differ in having underlying [spread glottis] stops instead of non-[spread glottis] stops on the [spread glottis] account. A complete list of the target words, which were randomised with fillers of different segmental make-up, is given in (22) (hyphens are included to indicate morpheme structure). The words were presented to subjects in regular German orthography, i.e. without hyphens.

(22)	a.	Handl-ung	'action'	b. hand-lich	'handy'
		nörgl-ig	'cranky'	kärg-lich	'sparse'
		nebl-ig	'foggy'	erheb-lich	'considerable'
	c.	Ummantl-ung	'coating'	d. bekannt-lich	'well-known'
		schnörkl-ig	'ornate'	merk-lich	'remarkable'
		popl-ig	'lousy'	prinzip-lich	'principled'

Ten native speakers from the North of Germany, seven males and three females ranging in age from 23 to 32, read the twelve words in (22), giving 120 tokens. The subjects were recorded digitally in a soundproof booth with two channel recordings (channel 1 for speech and channel 2 for EGG). The recordings were transferred to a computer (Silicon

Graphics Indy Workstation) and analysed with a speech analysis software package (ESPS/*waves+*). The stop in the stop–sonorant cluster was measured for closure duration, voice duration, voice percentage and positive VOT. Closure duration is the distance from closure onset to the onset of stop release. Voice duration is the distance from closure onset up to the offset of voicing during closure (which is maximally up to the end of closure) and voice percentage the percentage of closure during which voicing occurs. Positive VOT ranges from the onset of stop release to the beginning of voicing associated with the following sonorant (or with the following vowel in the case of the first experiment, where the same VOT-measurement criterion was used). The beginning of stop closure was taken as the beginning of a rapid drop in signal amplitude of the preceding vowel or sonorant, as visible in the waveform and the time-aligned spectrogram. Voicing offset during closure or voicing onset after release was taken to be the last or first voicing period, respectively (possibly with low amplitude), that was clearly identifiable. All measurements were made on the basis of the speech signal, but the electroglottographic signal was used to provide further verification of the amount of voicing during closure and the beginning of voicing after release.[16] The measurement procedures are illustrated in Fig. 8.

The results, pooled across the ten subjects and the three places of articulation, labial, alveolar and velar, are given in Table II. The table shows the results for all the words of type (a), represented by *Handl-ung*, for type (b) words, represented by *hand-lich*, for type (c) words, represented by *Ummantl-ung*, and for type (d) words, represented by *bekannt-lich*.

	closure duration		voice duration		% voice during closure		positive VOT	
	mean	sd	mean	sd	mean	sd	mean	sd
Handl-ung	72	(30)	37	(26)	58	(37)	25	(27)
hand-lich	82	(27)	17	(7)	23	(13)	45	(23)
Ummantl-ung	78	(22)	21	(13)	29	(20)	49	(25)
bekannt-lich	75	(26)	18	(7)	27	(12)	50	(25)

Table II
Results of Experiment II (duration in ms).

16 Of the 120 tokens, five had to be excluded from analysis because the measurement criteria could not be applied consistently, due to the presence of glottalisation. These glottalisation patterns were limited exclusively to words in which the target stop is preceded by a nasal consonant (this was the case in certain tokens of *Handlung*, *handlich*, *bekanntlich*). This pattern is consistent with Kohler (1994), who reports cases of glottalisation of stops in the vicinity of nasals in German.

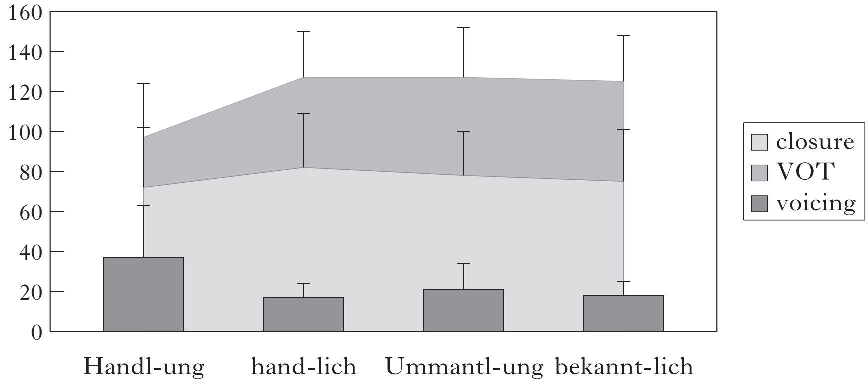

Figure 9
Graphical representation of results reported in Table II.

Most striking from these results is the distinction between *Handl-ung*-type words and the remaining three word types in terms of average voice duration, voice percentage and VOT: voice duration and percentage is much greater and VOT much smaller in *Handl-ung*-type words (a) than in any of the other three types. Short VOT (indicating lack of aspiration) and large voice duration/percentage mean that speakers generally have [d] in *Handlung* (as well as [b] in *neblig* and [g] in *nörglig*). The long VOTs in *handlich*, *Ummantlung* and *bekanntlich*-type words indicate stop aspiration and (partial) devoicing of the following sonorant consonant.

The results in Table II are represented graphically in Fig. 9. Duration values (in ms) are presented on the *y*-axis, the four different word types on the *x*-axis. Mean VOT and closure duration are represented as planes and voice duration as columns against the background of the closure interval. Standard deviations are also given. Voice percentage is not represented directly but can be inferred from the degree to which the voicing column overlaps with the closure plane. The graph illustrates that *Handlung*-type words differ from the other word types in having the highest voice duration and percentage and the lowest VOT values, while closure durations differ little between the four types.

The data were analysed by repeated measures ANOVA. The fixed factors were boundary, L-feature and place. The factor 'boundary' refers to the distinction between presence and absence of a phonological word boundary between the target stop and the following sonorant (to be explained below); 'L-feature' refers to the distinction between underlying presence and absence of the [spread glottis] feature associated with the relevant stop; 'place' designates the distinction between labial, alveolar and velar places of articulation of the stop. The assignment of the twelve words in (22) to the three factors is shown in (23).

(23)

	boundary	*L-feature*	*place*
Handl-ung	PW-medial	non-[sg]	alveolar
nörgl-ig	PW-medial	non-[sg]	velar
nebl-ig	PW-medial	non-[sg]	labial
hand-lich	PW-final	non-[sg]	alveolar
kärg-lich	PW-final	non-[sg]	velar
erheb-lich	PW-final	non-[sg]	labial
Ummantl-ung	PW-medial	[sg]	alveolar
schnörkl-ig	PW-medial	[sg]	velar
popl-ig	PW-medial	[sg]	labial
bekannt-lich	PW-final	[sg]	alveolar
merk-lich	PW-final	[sg]	velar
prinzip-lich	PW-final	[sg]	labial

Speakers ($n = 10$) were treated as a random effect. Fixed effects (assignable causes of variation) were tested for significance via F-tests. Significant effects were followed up with pairwise comparisons for which we analysed adjusted and unadjusted p-values, confidence intervals and effect sizes (Cohen 1988). The four different acoustic parameters, closure duration, voice duration, voice percentage and positive VOT, serve as the dependent variables in this experiment. Results are discussed separately for each of these four parameters in (24).

(24) a. *Closure duration*
Of the assignable causes of variation, only place was statistically significant ($F(2,103) = 17{\cdot}43$, $p < {\cdot}0001$). Mean closure duration for labial (92 ms) was significantly different from its value for alveolar (64 ms) and velar (72 ms) place of articulation (Tukey p-values $\leqslant {\cdot}0001$). The effects are large (0·80 standard deviation or greater).

b. *Voice duration*
Boundary, L-feature and place have highly significant effects on voice duration ($F(1,95{\cdot}8) = 18{\cdot}38$, $p < {\cdot}0001$; $F(1,94{\cdot}7) = 8{\cdot}61$, $p = {\cdot}0042$; $F(2,95{\cdot}5) = 14{\cdot}15$, $p < {\cdot}0001$, respectively). There is a highly significant boundary by L-feature interaction ($F(1,94{\cdot}7) = 10{\cdot}47$, $p = {\cdot}0017$). Velar (18 ms) *vs.* alveolar (19 ms) place is insignificant; all other place of articulation contrasts are significant at $p < {\cdot}0001$ (labial = 33 ms). Non-[spread glottis] is significantly different from [spread glottis] only for PW-medial position ($p = {\cdot}0001$). PW-final is significantly different from PW-medial position only for non-[spread glottis] stops ($p < {\cdot}0001$). Examination of the means (see Table II and Fig. 9) indicates that this is due to the fact that the PW-medial non-[spread glottis] combination is associated with voice duration almost twice as long as any of the other combinations. Main effects of boundary and L-feature are uninterpretable because of the interaction. All other significant effects are large (0·80 standard deviation or greater).

c. *Voice percentage*

Boundary and L-feature have highly significant effects on voice percentage ($F(1,95{\cdot}3) = 18{\cdot}39, p < {\cdot}0001$; $F(1,94{\cdot}4) = 9{\cdot}24, p = {\cdot}0031$, respectively). There is a highly significant boundary by L-feature interaction ($F(1,94{\cdot}4) = 15{\cdot}06$, $p = {\cdot}0002$). Non-[spread glottis] is significantly different from [spread glottis] only for PW-medial position ($p < {\cdot}0001$). PW-final is significantly different from PW-medial position only for non-[spread glottis] stops ($p < {\cdot}0001$). The interpretation of this effect and further findings are identical to the results for voice duration, given above.

d. *Positive VOT*

Boundary, L-feature and place have highly significant effects on VOT ($F(1,94{\cdot}4) = 8{\cdot}05$, $p = {\cdot}0056$; $F(1,94{\cdot}1) = 15{\cdot}87$, $p = {\cdot}0001$; $F(2,94{\cdot}3) = 16{\cdot}64$, $p < {\cdot}0001$, respectively). There is a highly significant boundary by L-feature interaction ($F(1,94{\cdot}1) = 7{\cdot}72$, $p = {\cdot}0066$). Non-[spread glottis] is significantly different from [spread glottis] only for PW-medial position ($p < {\cdot}0001$); PW-final is significantly different from PW-medial position only for non-[spread glottis] stops ($p = {\cdot}0009$). This pattern appears to be accounted for by the fact that the combination PW-medial/non-[spread glottis] has about half the VOT (25 ms) compared to the other combinations of L-feature and boundary, which have roughly equal onset times (45–50 ms) (see also Table II and Fig. 9). There is a significant interaction of place and boundary ($F(2,94{\cdot}4) = 4{\cdot}28$, $p = {\cdot}0166$) which involves two 'outliers' – final/velar (67 ms onset) and medial/labial (25 ms). All other combinations of place and boundary have roughly equal onset times (36–44 ms). Final/velar is significantly different from final/alveolar, final/labial and medial/velar. Medial/labial is significantly different from medial/velar. Presumably, final/velar is significantly different from medial/labial; however, we do not report confounded contrasts since they can be viewed as a chain of interpretable contrasts; *viz.* final/velar *vs.* medial/velar and medial/velar *vs.* medial/labial. Main effects are uninterpretable since each factor is involved in at least one significant interaction. With the exception of labial/medial *vs.* velar/medial (effect = 0·79), all significant effects are large (0·80 or larger).

The statistical results from (24) that are most important for the present discussion are the significant interactions between boundary and L-feature for VOT as well as for voice duration and percentage. As noted in (24), the significant interactions are explainable by the fact that it is only when medial position and underlying non-[spread glottis] status are *combined* that strikingly low VOT and high voice duration/percentage values occur. This statistical result supports our conclusion, expressed above, that only the words *neblig*, *Handlung* and *nörglig* were generally voiced unaspirated while all the other nine words in the experiment were generally voiceless aspirated.[17]

The [spread glottis] account sketched so far would predict that *Handlung* has a voiced stop because of passive voicing. However, it provides no explanation of why *handlich* does not have a voiced stop as well.

There are, however, reasons to conclude that passive voicing is blocked in *handlich*. In German, although an initial stop in a word is usually voiceless in a voiceless environment (*es deckte*) it may be voiced in a voiced environment as in *sage* [b]*ank* 'say bench' (Künzel 1977). In word-final position, however, no passive voicing occurs, *Ta*[k] *und Nacht* (**Ta*[g] *und Nacht*) 'day and night'. This suggests that there is a constraint requiring that word-final obstruents are [spread glottis] in German. Such a constraint is also proposed in Holsinger (2000).

A high-ranked constraint requiring that stops at the end of the prosodic word be [spread glottis] as in (25) can explain why there is no passive voicing in *Tag und Nacht* and why *Han*[d]*lung* has [d] but *han*[t]*lich* has [t] (with aspiration/sonorant devoicing), without cycles or levels or otherwise unmotivated syllabification.

(25) PW-R[sg]

A prosodic word-final obstruent is [spread glottis].

This account requires that we consider *handlich* (but not *Handlung*) to be two prosodic words, an assumption that is independently motivated (Wiese 1996).[18] Prosodic word boundaries in (26) and (27) are indicated with { }.

(26)

Han/t/l+ung	PW-R[sg]	Id-IO[sg]	*[voice]	*[sg]
a. {Han[d]lung}			*!	
☞ b. {Han[t]lung}				

The surface form is *Han*[d]*lung* because of the phonetically conditioned voicing of non-[spread glottis] stops flanked by voiced segments.

[17] It has been shown for English, German and many other languages that closure duration in labial stops is generally longer than in alveolar and velar stops, while the opposite relation occurs for VOT; the difference between alveolar and velar place is often subject to language-specific and context-specific differences (cf. Docherty 1992, Jessen 1998). Consistently, in Experiment II average closure duration across L-feature and other factors is 92 ms for labial, 72 ms for velar and 64 ms for alveolar stops, while average VOT is 30 ms for labial, 41 ms for alveolar and 55 ms for velar stops. Furthermore, the result that labial stops have significantly longer voice duration than alveolar and velar stops is consistent with the discussion of phonetic voicing in §4. Notice in (24) that place has no significant effect on voice *percentage*. This has to do with the fact that although labials have the largest voice durations they also have the largest closure durations compared to other places of articulation.

[18] For example, Wiese (1996) points out that suffixes that are prosodic words can be 'gapped': *mütter- und väter**lich*** 'motherly and fatherly', but suffixes that are not prosodic words cannot: **winz- oder ries**ig*** 'tiny or huge'. According to Wiese, all consonant-initial suffixes are separate prosodic words.

(27)

han/t/+lich	PW-R[sg]	ID-IO[sg]	*[voice]	*[sg]
a. {han[d]}{lich}	*!		*	
☞ b. {han[t^[sg]]}{lich}		*		*
c. {han[t]}{lich}	*!			

The surface form is *han*[t^[sg]]*lich*. Phonetically conditioned voicing is relevant only for non-[spread glottis] stops and hence is irrelevant for *han*[t^[sg]]*lich*.

Assuming that *Ummantlung* and *bekanntlich* have underlying [spread glottis] stops, the correct forms are also designated as optimal in the [spread glottis] analysis, as illustrated in the tableaux in (28).[19] Since the stops in these words are [spread glottis], passive voicing does not occur. In words like *Umman*[t^[sg]]*lung* and *bekann*[t^[sg]]*lich* we understand [t^[sg]] to represent stop aspiration and partial devoicing of the following sonorant.

(28)

		PW-R[sg]	ID-IO[sg]
a.	Umman/t^[sg]/l+ung	PW-R[sg]	ID-IO[sg]
	☞ i. {Umman[t^[sg]]lung}		
	ii. {Umman[d]lung}		*!
b.	bekann/t^[sg]/+lich		
	☞ i. {bekann[t^[sg]]}{lich}		
	ii. {bekann[t]}{lich}	*!	*
	iii. {bekann[d]}{lich}	*!	*

Although it appears that *Handlung*-type words were produced with a voiced stop from the pooled results in Table II, examination of productions of individual words indicates that some *Handlung*-type words were actually produced with voiceless stops. This variation between voiced and voiceless tokens is also evident from the standard deviations for voice duration and voice percentage shown in Table II and Fig. 9,

[19] Note that according to the analysis of Petrova *et al*. (2000), underlying [spread glottis] stops will retain their [spread glottis] feature in word-final position. Given the constraint PW-R[sg] assumed here, *all* word-final stops will be [spread glottis]. It has been claimed that all word-final stops may be optionally aspirated (Lotzmann 1975, Knetschke & Sperlbaum 1987). Although in our account all word-final stops will be [spread glottis], they will not be predicted to be aspirated, since we assume that aspiration only occurs when a [spread glottis] stop occurs before a sonorant. It is possible that what has been claimed to be optional aspiration is actually optional release. It has also been claimed that aspiration is quite common at the end of an utterance/before pause (Lotzmann 1975, Knetschke & Sperlbaum 1987, Hall 1992). This might be the result of a more general tendency of a large utterance-final/prepausal glottal opening gesture, which occurs after all sounds, not just stops (see Lisker *et al*. 1969 and Jessen 1998: 208–211 for illustrations of utterance-final glottal opening using transillumination, and Engstrand & Nordstrand 1984 for discussion of the perceptual effect that a word like *Anna* played backwards is heard as *Hanna*). These are questions that require further research.

which are much higher in *Handlung*-type words (a) than in words of the other three types (b)–(d).

The failure of passive voicing to occur in some *Handlung*-type words is, we assume, another instance of the variation of the phonetically conditioned passive voicing,[20] as was also found for some intervocalic non-[spread glottis] stops.[21]

Since *dl* is not a possible onset in German, while *gl* and *bl* are, we might have expected that the *d* would be more likely to be syllabified in the coda than the *g* and *b* if syllabification were relevant to the distribution of voiced and voiceless stops, and hence that the stop in *Handlung* would be more frequently voiceless during closure than the stops in *nörglig* and *neblig*. This is not the case. Six speakers had voiceless stops in *nörglig*, five in *Handlung* and three in *neblig*.

Vennemann's (1972) account of the variation between a voiced and voiceless stop in *Handlung*-type words is different syllabification. He suggests that there are two groups of speakers: those who have a voiced stop in all *Handlung*-type words syllabify these words as *Han.dlung*, *nör.glig* and *ne.blig*. He claims that those who have [t] in words like *Handlung* do so because [dl] is an impermissible cluster. Hence he claims that the other group of speakers have [t] in *Handlung*, but [g] and [b] in *nörglig* and *neblig*. That is, he suggests that the second group of speakers syllabify these words as *Hand.lung*, *nör.glig* and *ne.blig*. Our data show that Vennemann's account cannot be correct. There are speakers with voiceless stops in *nörglig* and *neblig*, which Vennemann's account denies, and those with [t] in *Handlung* do not necessarily have [g] and [b] in *nörglig* and *neblig*.

Another account of the variation is that there are two dialects, one in which the *Handlung*-type words are pronounced with a voiceless stop and another in which they are pronounced with a voiced stop. Our data are inconsistent with this claim as well. Our speakers, who are homogeneous with respect to dialect, use more voiced stops than voiceless stops in *Handlung*-type words, but most speakers have voiced stops in some

[20] In earlier work (Jessen & Ringen 2001b) we accounted for the variation by crucially unranked *[voice] and PASSIVEVOICE.

[21] In previous work (Jessen & Ringen 2001b), we assumed a constraint requiring that stops in codas be [spread glottis] so that the voiceless tokens of the stops in *Handlung*-type words would be [spread glottis]. We made this assumption because *average* VOTs for the voiceless stop tokens in *Handlung*-type words were similar to the average VOTs for *handlich*, *Ummantlung* and *bekanntlich*. Actually, however, when we looked at the individual tokens, we found both long-VOT and short-VOT versions of voiceless stops in *Handlung*-type words, leaving open the question of whether these stops should be specified as [spread glottis] or not. However, in a new recording in which words of the type in (22) were read by another group of North German speakers, all *Handlung*-type words (those with and those without voicing during closure) were consistently produced with short positive VOT. These data support the analysis presented here, as shown in (26), and suggest that our earlier coda [spread glottis] analysis was incorrect. The slight discrepancy between the experiment presented in §5 and the new study mentioned here could be due to the different set-ups of the stimuli (isolated words in the present experiment *vs*. words embedded in sentences in the new study).

Handlung-type words and voiceless stops in others.[22] Finally, as noted in Brockhaus (1995), others have suggested that the variation is style-dependent ([t] being less formal). Our data show that variation occurs even in a single (relatively formal) situation, since most of our speakers use both voiced and voiceless stops.[23]

6 Conclusion

In conclusion, we have presented new experimental data which support an account of German stops that involves an underlying feature of [spread glottis], suggesting that the claims of researchers such as Kloeke (1982), Meinhold & Stock (1982), Anderson & Ewen (1987), Jessen (1989, 1996, 1998), Iverson & Salmons (1995) and Petrova *et al.* (2000) are correct. These facts are difficult to account for if the feature [voice] is assumed. We found that speakers consistently have only voiceless (obstruent) stops in word-medial clusters. We also found that even in intervocalic position some speakers do not have consistent voicing of non-[spread glottis] stops. These results, in conjunction with the well-known fact that speakers consistently have only voiceless (obstruent) stops in

[22] See Brockhaus (1995: 37–88) for an extensive review of the analyses of *Handlung*-type words in the literature.

[23] Cases similar to the ones addressed in the *Handlung/handlich* experiment were investigated by Bonnin (1964). Bonnin presents an auditory analysis of a large number of sequences of word-medial stop+sonorant in German *Umgangssprache* (spontaneous, relatively informal speech). Unfortunately, in his analysis he does not distinguish between the four different linguistic types shown as (a)–(d) in (22). But it is nevertheless remarkable that the percentage of stop tokens perceived by him as voiced during closure (his term for this is *media*) is close to zero for his North German speakers. If it is assumed that roughly 25% of these stop–sonorant sequences were of the *Handlung*-type – and some clearly were, since he explicitly mentions some word examples of this type – then this would mean that the percentage of *Handlung*-type words that are voiceless during closure is very high. This is in contrast to the present study, where enough *Handlung*-type tokens are voiced so that there is a significant voice-duration difference between them and the other three stop types. Another difference lies in aspiration/VOT. Bonnin perceived only about 10% of all tokens as aspirated (his term for this is *tenuis aspirata*); the others were unaspirated. This is in contrast to the large number of tokens with relatively long VOT (most of them among the types (b)–(d)). Part of the discrepancy between Bonnin (1964) and the present study might be due to the stylistic difference between casual speech in the former as opposed to more formal speech in the latter case. Notice that lack of voicing in *Handlung*-type words is consistent with the assumptions about casual speech mentioned in connection to the overview of Brockhaus (1995). And furthermore one would expect aspiration to reduce with a decrease in formality (and presumably an increase in speech rate, which would imply a reduction of many duration variables, most probably including aspiration duration). But another reason for the discrepancy might lie in the difference between the production and the perception level. That closure voicing is of very little perceptual value in the perception of word-medial *b, d, g vs. p, t, k* in German has been shown by Kohler (1979) and Jessen (1998); i.e. full voicing during closure might be present acoustically but still not be perceived. (This low perceptual value of closure voicing is another argument against the importance of the feature [voice] in German.)

word-initial position (unless preceded by sonorant segments), are consistent with the [spread glottis] account, but are difficult to explain if [voice] is the active feature.

We found that speakers are not consistent in their pronunciation of *Handlung*-type words: sometimes the stop is voiced and sometimes it is voiceless. This variation also can be accounted for if [spread glottis] is the underlying feature, but is difficult to explain if [voice] is the underlying feature. In order to account for the data, we have adopted the constraints for German assumed in Petrova *et al.* (2000) and a constraint which requires that stops in prosodic word-final position be [spread glottis]. One difference between our account and that of Petrova *et al.* is that we assume that the variable passive voicing is phonetic, not phonological.

These results are important for a number of reasons. In addition to the obvious implications for German phonology, they call into question any theoretical conclusions based on analyses of German in which [voice] is assumed to be the feature of contrast. These results also raise questions about the correct analysis of other languages which have fully voiced stops (in some positions) and aspirated stops (e.g. English, Swedish, Turkish), and which have traditionally been analysed as having a [voice] contrast. And finally, they show how the results of phonetic studies can bear on questions of phonological analysis.

REFERENCES

Alderete, John (1997). Dissimilation as local conjunction. *NELS* **27**. 17–32.

Anderson, John M. & Colin J. Ewen (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

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

Bonnin, Gunther M. (1964). Some acoustic aspects of final stop allophones in contemporary German. *Phonetica* **11**. 65–100.

Braun, Angelika (1996). Zur regionalen Distribution von VOT im Deutschen. In Angelika Braun (ed.) *Untersuchungen zu Stimme und Sprache*. Stuttgart: Steiner. 19–32.

Brockhaus, Wiebke (1995). *Final devoicing in the phonology of German*. Tübingen: Niemeyer.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* **2**. 225–252.

Clements, G. N. & Susan R. Hertz (1996). An integrated approach to phonology and phonetics. In Jacques Durand & Bernard Laks (eds.) *Current trends in phonology: models and methods*. Salford: ESRI. 143–173.

Cohen, Jacob (1988). *Statistical power analysis for the behavioral sciences* (2nd edn). Hillsdale, N.J.: Erlbaum.

Cohn, Abigail C. (1993). Nasalisation in English: phonology or phonetics? *Phonology* **10**. 43–81.

Davis, Stuart (1998). An Optimality-Theoretic analysis of the distribution of [h] and aspirated stops in American English. Paper presented at the 4th Midcontinental Workshop on Phonology, University of Michigan, Ann Arbor.

Docherty, Gerard J. (1992). *The timing of voicing in British English obstruents*. Berlin & New York: Foris.

Engstrand, Olle & Lennart Nordstrand (1984). A new phonetic interpretation of /h/ in Swedish. *JPh* **12**. 195–206.

Hall, Tracy A. (1992). *Syllable structure and syllable-related processes in German*. Tübingen: Niemeyer.

Hall, Tracy A. (1993). The phonology of German /ʀ/. *Phonology* **10**. 83–105.

Halle, Morris & Kenneth N. Stevens (1971). A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report* **101**. 198–213.

Holsinger, David (2000). *Lenition in Germanic*. PhD dissertation, University of Wisconsin, Madison.

Iverson, Gregory K. & Joseph C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* **12**. 369–396.

Jessen, Michael (1989). *Laryngale Mechanismen in der Phonetik und Phonologie des Deutschen. Auslautverhärtung, Stimmassimilation und Vokaleinsätze aus universalgrammatischer Sicht*. Ms. Revised version of 1989 MA thesis, University of Bielefeld.

Jessen, Michael (1996). The relevance of phonetic reality for underlying phonological representation: the case of tense versus lax obstruents in German. In Kleinhenz (1996). 294–328.

Jessen, Michael (1998). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Jessen, Michael (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T. Alan Hall (ed.) *Distinctive feature theory*. Berlin & New York: Mouton de Gruyter. 237–294.

Jessen, Michael & Catherine Ringen (2001a). Are there any underlying voiced stops in German? In Caroline Féry, Antony Dubach Green & Ruben van de Vijver (eds.) *Proceedings of HILP 5*. Potsdam: University of Potsdam. 66–85.

Jessen, Michael & Catherine Ringen (2001b). On the status of [voice] in German. *WCCFL* **20**. 304–317.

Keating, Patricia A. (1984). Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics* **59**. 29–34.

Keating, Patricia A. (1996). The phonology–phonetics interface. In Kleinhenz (1996). 262–278.

Kingston, John & Randy L. Diehl (1994). Phonetic knowledge. *Lg* **70**. 419–454.

Kingston, John & Randy L. Diehl (1995). Intermediate properties in the perception of distinctive feature values. In Bruce Connell & Amalia Arvaniti (eds.) *Phonology and phonetic evidence: papers in laboratory phonology IV*. Cambridge: Cambridge University Press. 7–27.

Kleinhenz, Ursula (ed.) (1996). *Interfaces in phonology*. Berlin: Akademie-Verlag.

Kloeke, W. U. S. van Lessen (1982). *Deutsche Phonologie und Morphologie*. Tübingen: Niemeyer.

Knetschke, Edeltraud & Margret Sperlbaum (1987). *Zur Orthoepie der Plosiva in der deutschen Hochsprache: eine auditiv-komparative Untersuchung*. Tübingen: Niemeyer.

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

Kohler, Klaus J. (1994). Glottal stops and glottalization in German: data and theory of connected speech processes. *Phonetica* **51**. 38–51.

Künzel, Hermann J. (1977). *Signalphonetische Untersuchung Deutsch–Französischer Interferenzen im Bereich der Okklusive*. Frankfurt: Lang.

Ladefoged, Peter (1997). Linguistic phonetic descriptions. In William J. Hardcastle & John Laver (eds.) *The handbook of phonetic sciences*. Oxford: Blackwell. 589–618.

Lisker, Leigh & Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20**. 384–422.

Lisker, Leigh, Arthur S. Abramson, Franklin S. Cooper & Malcolm H. Schvey (1969). Transillumination of the larynx in running speech. *JASA* **45**. 1544–1546.

Löfqvist, Anders & Hirohide Yoshioka (1980). Laryngeal activity in Swedish obstruent clusters. *JASA* **68**. 792–801.

Lombardi, Linda (1991). *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts, Amherst.

Lombardi, Linda (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *NLLT* **17**. 267–302.

Lotzmann, Geert (1975). *Zur Aspiration der Explosivae im Deutschen : ein sprechwissenschaftlich-phonetischer Beitrag zur deutschen Hochlautung*. Göppingen: Kümmerle.

McCarthy, John & Alan Prince (1993). *Prosodic morphology I : constraint interaction and satisfaction*. Ms, University of Massachusetts, Amherst & Rutgers University.

McCarthy, John & Alan Prince (1995). Faithfulness and reduplicative identity. In Jill Beckman, Laura W. Dickey & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst: GLSA. 249–384.

Mascaró, Joan & Leo Wetzels (2001). The typology of voicing and devoicing. *Lg* **77**. 207–244.

Meinhold, Gottfried & Eberhard Stock (1982). *Phonologie der deutschen Gegenwartssprache*. Leipzig: VEB Bibliographisches Institut.

Noske, Manuela (1999). Deriving cyclicity: syllabification and final devoicing in German. *The Linguistic Review* **16**. 227–252.

Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.) *The production of speech*. New York: Springer. 189–216.

Petrova, Olga, Rosemary Plapp, Catherine Ringen & Szilárd Szentgyörgyi (2000). Constraints on voice: an OT typology. Paper presented at the 74th Annual Meeting of the Linguistic Society of America, Chicago.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory : constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder.

Rubach, Jerzy (1990). Final devoicing and cyclic syllabification in German. *LI* **21**. 79–94.

Rubach, Jerzy (1997). Polish voice assimilation in Optimality Theory. *Rivista di Linguistica* **9**. 291–342.

Scharf, Gabriele & Harald Masur (2002). Voice onset time in normal speakers of a German dialect: effects of age, gender and verbal material. In Fay Windsor, M. Louise Kelly & Nigel Hewlett (eds.) *Investigations in clinical phonetics and linguistics*. Mahwah, N.J.: Erlbaum. 327–339.

Stevens, Kenneth N. (1998). *Acoustic phonetics*. Cambridge, Mass: MIT Press.

Stevens, Kenneth N. & Samuel J. Keyser (1989). Primary features and their enhancement in consonants. *Lg* **65**. 81–106.

Stock, Dieter (1971). *Untersuchungen zur Stimmhaftigkeit hochdeutscher Phonemrealisationen*. Hamburg: Buske.

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

Wagner, Michael (2000). *Laryngeal features in West Germanic*. MA thesis, Humboldt University, Berlin.

Westbury, John R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *JASA* **73**. 1322–1336.

Westbury, John R. & Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *JL* **22**. 145–166.

Wiese, Richard (1996). *The phonology of German*. Oxford: Clarendon.

Wurzel, Wolfgang U. (1970). *Studien zur deutschen Lautstruktur*. Berlin: Akademie-Verlag.

Yu, Si-Taek (1992). *Unterspezifikation in der Phonologie des Deutschen*. Tübingen: Niemeyer.