Three influences on glottalization in read and spontaneous German speech*

Jonathan Rodgers

*This research was carried out with funding by German Research Council (DFG) grant KO331/22–1 under the supervision of K.J. Kohler.
1 Introduction

1.1 Overview

Glottalization phenomena is an umbrella term used to describe glottal stop and any deviation from canonical modal voice. A recent symposium on non-modal vocal fold vibration and voice quality (cf. Kohler 1999a, Rodgers 1999) highlighted *inter alia* the interest in the linguistic and paralinguistic functions of a scale of glottal activity from glottal stop to low frequency irregular glottal pulsing to breathiness and breathy voice.

Four main environments are associated with glottalization phenomena in German:

1. Plosive-related glottalization
2. Onset-related glottalization
3. Utterance-final glottalization
4. Truncation glottalization.

These four influences suggest glottalization has both a linguistic and paralinguistic function. In plosive-related glottalization, a plosive can be reinforced or replaced by glottal stop or other phenomena, while in onset-related glottalization, such phenomena serve as a boundary signal of a word of morpheme beginning with a vowel, and may also mark sentence accent. These are linguistic functions. Paralinguistically, prosodic phrase final relaxation of phonation can lead to glottalization alternating with breathy voice and breathiness, but not glottal stop; and in truncation glottalization, glottal stop and/or glottalization result from a tensing of phonation when utterances are interrupted.

The paper by Kohler (1999b) in this volume (p.281ff.) addresses the issue of plosive-related glottalization, and is referred to here. The other three influences are examined initially in isolation from each other, and then the interplay of the factors is addressed.
1.2 Onset-related glottalization

Initial vowels in German are canonically realized with a glottal stop. A blanket rule in the authoritative *Duden* pronunciation dictionary assigns a glottal stop to all initial vowels (Duden 1990). In the *Großes Wörterbuch der deutschen Aussprache* Krech et al. (1982) acknowledge the possibility of a "soft onset" in unstressed syllables that do not follow a pause, without expanding on the phonetic properties of this soft onset. Prior to this simplistic statement, Krech (1968) had examined in greater detail the incidence of the glottal stop in German, but for a variety of reasons her approach is flawed. Krech’s research into glottal activity was in connection with the *Großes Wörterbuch der deutschen Aussprache*, and its prescriptive approach is clear: it focuses on read speech from largely classical texts (with a preponderance of Goethe) produced by a small number of professional speakers. The most extreme evidence of this prescriptivism is in her condemnation of creaky voice — which this study will show to be the most common phonatory correlate of juncture — as pathological:

“Eine andere pathologische Variante des Glottisschlages bildet der geknarrte Einsatz, bei dem mehrere aufeinanderfolgende Glottisschläge zu hören sind”.

Here, by contrast, the approach is descriptive: spontaneous as well as read speech is examined, and data from a large number of non-professional speakers are investigated. The genuine speech data described here suggest that at a phonological level at least four different realizations are possible, and this paper examines the factors that condition the distribution of these realizations.

This paper extends to spontaneous speech the analysis of read speech by Kohler (1994), which examined glottal phenomena in the Kiel Corpus of Read Speech published in the same year (IPDS 1994). Since then, three volumes of the Kiel Corpus of Spontaneous Speech have been published (IPDS 1995, 1996, 1997a), and prosodic labelling has been made available for all read speech and the first volume of spontaneous speech. As well as comparing the distribution of realizations as a function of factors marked in the prosodically labelled read and spontaneous corpora, this research extends to data
which is not prosodically labelled.

The minimal four phonological categories are those captured by the labelling of the Kiel Corpora of Read and Spontaneous Speech using the conventions in Table 1 (p.176). The symbols used in this table are not intended as phonetic, which is why they appear in sans serif font and without the phonetic or phonemic brackets that would imply the status of a transcription at some level. These four categories are far from exhaustive; the labelling does not aim to capture all details in a narrow phonetic transcription, but rather, broadly to classify phonetic phenomena to orientate the researcher in the database. Further realizations subsumed under these four labellings are also possible, and are discussed in Section 6 (p.244).

Figures 1(a)–1(d) (p.177) show wide-band spectrograms and label windows for examples of these four possibilities, in unaccented realizations of the word *ich*.

At least four factors are assumed to condition these realizations, and are the object of this study:

1. speech style
2. word class
3. position in the utterance
4. sentence accent.

<table>
<thead>
<tr>
<th>Realization</th>
<th>Label</th>
<th>Symbolization</th>
</tr>
</thead>
<tbody>
<tr>
<td>glottal stop</td>
<td>Q</td>
<td>?V</td>
</tr>
<tr>
<td>glottal stop and glottalization</td>
<td>Q -q</td>
<td>?V</td>
</tr>
<tr>
<td>glottalization without glottal stop</td>
<td>Q- -q</td>
<td>V</td>
</tr>
<tr>
<td>absence of glottal stop and glottalization</td>
<td>Q-</td>
<td>V</td>
</tr>
</tbody>
</table>

Table 1: Labelling conventions: four realizations of canonical glottal stop in column I are labelled in the Kiel corpora using IPDS modified SAMPA as in column II, and symbolized in tables and figures in this paper as in column III.
Figure 1: Wide-band spectrograms and label windows for four realizations of canonical glottal stop in unaccented *ich* in files g091a000, g096a000, g077a003 and g074a000, from left to right. Figure 1(a) shows glottal stop, marked as $Q$, symbolized as $\hat{V}$; Figure 1(b) shows glottal stop and glottalization, marked as $Q\rightarrow q$, symbolized as $\hat{V}^3$; Figure 1(c) shows glottalization without glottal stop, marked as $Q\rightarrow -q$, symbolized as $\hat{V}$; Figure 1(d) shows absence of glottal stop and glottalization, marked as $Q\rightarrow$, symbolized as $\hat{V}$. 

177
In demarcating the influences as four separate factors, the aim is to simplify hypotheses. It is acknowledged that these factors may interact, and these interactions are borne in mind in the statistical analysis of findings.

For example, an interaction may occur between the factors of sentence accent and word class: function words are typically unaccented, and unaccented words may be less clearly articulated than accented; if the vowel that begins a function word is not realized with a glottal stop, the unanswerable question arises of whether this is because it is unaccented or because it is a function word.

Speech style and position in the utterance may interact similarly. On the one hand, read speech may contain more carefully articulated items, including glottal stops — assuming that neat glottal stops are a property of careful speech. On the other, glottal stops may be expected sentence-initially, and the sentence structure of read speech is clearly marked by the punctuation of the text: the content of the utterance is given, and the reader knows from punctuation markers — .!? — where the sentence starts and ends. The majority of the sentences in read speech are short or very short, and are single sentences. This means that if a glottal stop occurs in a phrase-initial vowel in read speech, it may be due either to the greater precision in read speech, or to the fact that the speaker is aiming to bring out the sentence structure of a given text, which clearly cannot be the case for spontaneous speech, which is ipso facto unscripted.

1.2.1 Speech style

Patterns of glottalization may be different in spontaneous and read speech. Read speech is likely to be more carefully spoken, being elicited in a formal context, yet it may also be more fluent and lack hesitations and disjuncture phenomena. As mentioned above, sentence structure is also clear in the text of read speech, so may be expected to be more clearly reflected by phonetic markers corresponding to prosodic structure in the speech signal. There are both read and spontaneous corpora (IPDS 1994, 1995, 1996, 1997a) for comparing the influence of the different speech styles on glottalization.
1.2.2 Word class

The patterns for glottalization may be different in content and function words. Matthews (1997) defines content words as words with lexical meaning, and function words as words with grammatical meaning. The difference in patterns may also be partly due to expected interactions of sentence accent and word class. Function words are typically unaccented, and for this reason may be less likely to be marked by a glottal reflex. The Kiel corpora mark whether lexical items are content or function words (Kohler et al. 1995:27). A detailed investigation of the realization of function words in spontaneous and read speech is offered by Wesener (1999) in this volume (p.323ff.). In tables and figures these terms are abbreviated to cw (content words) and fw (function words).

1.2.3 Position in the utterance

A glottal stop may be more common phrase-initially than -medially, as well as after a phrase boundary marked by pause, hesitation, false starts, interruption or breath. Ceteris paribus, the vocal folds are held apart for respiration (Laver 1994, Lieberman and Blumstein 1988), but in speech, they are typically slightly closed, such that phonation can begin: a vocalic utterance made from this starting point is more likely to begin with a glottal stop with or without glottalization, or simple glottalization, before modal voicing is reached (Lofqvist and Oshima 1993). As well as a reduced inventory of punctuation marks (., ?) to mark sentence structure in the orthographic transliteration (Kohler et al. 1995:15), the corpora use other symbols to indicate the other relevant prosodic factors (phrase boundaries, breath, pause, false starts) (Kohler et al. 1995:16–17;77), so that the influence of these factors can also be examined.

1.2.4 Sentence accent

Kohler et al. (1995) maintain a distinction between stress and accent. Lexical stress is an aspect of word phonology; primary lexical stress is marked by a ‘ before the vowel of the stressed syllable, and secondarily stressed syllables of non-initial elements of compounds
Table 2: Conventions for labelling four levels of sentence accent in the Kiel corpora.

<table>
<thead>
<tr>
<th>unaccented</th>
<th>&amp;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>partially accented</td>
<td>&amp;1</td>
</tr>
<tr>
<td>accented</td>
<td>&amp;2</td>
</tr>
<tr>
<td>reinforced</td>
<td>&amp;3</td>
</tr>
</tbody>
</table>

are marked by " before the vowel of the relevant syllable. By default function words have no lexical stress marked in the canonical transcription. Sentence accent is an aspect of sentence prosody. It is an attribute of the word in an utterance, and is indicated in the corpus by attaching to the word on which it is phonetically manifested the appropriate symbol: four levels of sentence accent are recognized, symbolized as in Table 2 (p.180).

The primary lexical stress marks indicate potential locations where sentence accent manifests itself phonetically. The symbol ‘’ is also used to mark cases where sentence accent is attached to a location other than that marked by the lexical stress marker (e.g. Augenblick accented on the final syllable), or if the sentence accent falls on lexical items which have been given no lexical stress: where function words receive sentence accent, this is marked by ‘’; since they otherwise have no lexical stress mark to which to attach a sentence accent label. The labels ‘’ (primary and secondary lexical stress), ‘’ (shifted or added accent) and 0, 1, 2 and 3 (four levels of sentence accent) provide a means of access to prosodic phenomena.

In the first instance, vowels are separated into “accented” and “unaccented”, terms which correspond to conflations of the parameters ±accent and ±stress: +accent is marked by \&1 \&2 \&3, −accent by \&0; +stress by ‘’ ‘’ , −stress by ” or no marking. In “accented” vowels (+accent and +stress), sentence accent is realized on the lexically primarily stressed vowel. In “unaccented” vowels (−accent and/or −stress), the vowel is lexically either unstressed or secondarily stressed, or the vowel may be lexically primarily stressed but sentence accent is not realized on it. Accented vowels are expected to be associated with a glottal reflex, that is, simple glottalization,
or a glottal stop with or without glottalization.

1.2.5 Hypotheses

The following hypotheses have been set up for the examination of onset-related glottalization in this paper.

Speech style

- Glottal stop, with or without glottalization, is more frequent in read than spontaneous speech
- Absence of glottal reflex is more frequent in spontaneous than read speech
- Glottalization without a glottal stop is more frequent in spontaneous than read speech

Word class

- Glottal reflex is more frequent in content than function words

Position in the utterance

- Glottal stop is more frequent phrase-initially than medi ally
- Glottalization without a glottal stop, or absence of any glottal reflex whatsoever, is most frequent phrase-medially

Sentence accent

- Glottal reflex is more frequent in accented vowels
- Unaccented vowels are realized simply with glottalization or with no glottal reflex whatsoever
1.2.6 Heuristic procedure

The examination of onset-related phenomena has a four-part structure:

1. formulation of hypotheses on the basis of previous research and expectations about likely factors in glottalization;

2. automated examination of label files for prosodically labelled data to establish distribution of minimal four glottal realizations according to four putative factors, and statistical evaluation of findings;

3. examination of signal files in order to:
   (a) evaluate consistency and accuracy of all labelling;
   (b) gain insights into items that contradict hypotheses, and as necessary expand the four glottal categories;

4. extrapolation from findings from statistically supported label files and phonetically analysed signal files, to corpus items that have not been prosodically labelled.

1.3 Utterance-final glottalization in spontaneous speech

In many languages a low falling intonation is often terminated by creak or creaky voice (Laver 1994:196), indeed in English, termination of this sort is sometimes used for a regulative function, as speakers use creak as a signal of yielding the floor to another speaker at the end of a turn.\(^1\)

Low \(f_0\) and utterance-final relaxation are believed to be the prerequisites for laryngealization of this kind: the glottis ceases modal vibration as it prepares for non-speech function, e.g. inhalation. Such opening is the opposite of glottal stop, which is not expected in this

---

\(^1\)Indeed, Laver points out that for speakers who adopt this convention, a low falling intonation before a silent pause, but without a creaky termination, and particularly with simultaneous avoidance of eye-contact, signifies that the speaker has not yet reached the end of his or her turn, and is resisting the possibility of take-over of the speaker-role by the other participant.
Glottalization in read and spontaneous German speech

1.4 Truncation glottalization

Truncation glottalization is the opposite of glottalization associated with $f_0$-declination and utterance-finality. It describes instances where, utterance-internally, articulation is abruptly halted, due to error or replanning of the utterance. Such truncation is most effectively achieved by cutting off the air stream at the glottal valve, by tensing the vocal folds for a glottal stop (Nakatani and Hirschberg 1994, Local and Kelly 1986). Truncations are marked in the label files of the corpora, so can be isolated and analysed.

2 Method

2.1 Materials

This study examines read and spontaneous speech corpora which have been published as signal and label files (IPDS 1994, 1995, 1996, 1997a), and for much of which prosodic labelling is available. The appointment-making paradigm used to elicit the data, as well as further recording technicalities, are described in Kohler, Pätzold and Simpson (1995, 1997). The dialogues which have been prosodically as well as segmentally labelled allow insights into the effect of prosodic structures on Connected Speech Processes. For read speech, prosodic labelling is available for all published material, comprising 31,382 labelled words from 53 speakers (26 female, 27 male). For details of the read corpus see Thon and van Dommelen (1992), Kohler (1994). For spontaneous speech by contrast, prosodic labelling is available for only a subset of the entire spontaneous speech data published to date. Labelling is an ongoing process, however, such that the amount of prosodically labelled data available is continually increasing. The prosodically labelled dialogues available so far comprise 16,599 word entries from 26 speakers (9 female, 17 male) from eight complete sessions and five single dialogues (i.e. 61 dialogues in
total) shown in Table 3(a) (p.185). For practical reasons addressed in Section 3.1.5 (p.196), inferential statistics are carried out on a portion of the prosodically labelled data (referred to as $n^{\text{spontaneous}}_p$ for partial), whereas descriptive statistics refer to the full set (referred to as $n^{\text{spontaneous}}_t$ for total). The spontaneous speech data for which there is as yet no prosodic labelling comprise 20,730 labelled words from 16 speakers (eight female, eight male) from eight complete sessions (i.e. 56 dialogues in total) shown in Table 3(b) (p.185). To the 37,329 lexical words from the 117 spontaneous dialogues are added 599 truncations and false starts (which do not occur in the read database); the database excludes hesitation particles, neologisms, slips of the tongue and stretches difficult or impossible to identify.

Figures 2–3(b) (p.186 and 187) show the relative distribution of speakers in the three databases examined, by plotting the percentage of items with a glottal stop in the canonical transcription that were produced in each database by each speaker. Especially in the read database, certain speakers are responsible for a high proportion of the items, since it was intended that one speaker of each gender would record all materials: female speaker k62 (PHONDAT90) and rtd (PHONDAT92) are identical, as are male speaker k61 (PHONDAT90) and kko (PHONDAT92). In the spontaneous speech data, no such control was exerted, and the amount each speaker utters is largely a result of his/her own loquaciousness. Overall, the balance of speakers is good, with no one speaker contributing fewer than 15 items, which is enough for a representative sample. Table 4 (p.186) shows the proportion of items spoken by each gender. Krech’s speaker set featured an imbalance of genders (65% male, 35% female). Table 4 (p.186) shows that only for prosodically labelled data are the Kiel data imbalanced in this way, and for read and prosodically unlabelled speech — as well as overall — the balance is closer to even. These figures refer to words containing a glottal stop in the canonical transcription: for all lexical words, however, of which there are 68,711 in the entire corpora, 38,204 (56%) are produced by male speakers, and 30,507 (44%) by female.
Table 3: Dialogues constituting spontaneous speech database. Table 3(a) shows prosodically labelled data, Table 3(b) shows prosodically unlabelled data.

(a) Labelled

<table>
<thead>
<tr>
<th>Session</th>
<th>No. of dialogues</th>
<th>No. of turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>g07a</td>
<td>7</td>
<td>115</td>
</tr>
<tr>
<td>g08a</td>
<td>7</td>
<td>109</td>
</tr>
<tr>
<td>g09a</td>
<td>7</td>
<td>159</td>
</tr>
<tr>
<td>g14a</td>
<td>7</td>
<td>77</td>
</tr>
<tr>
<td>g19a</td>
<td>7</td>
<td>151</td>
</tr>
<tr>
<td>g202a</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>g21a</td>
<td>7</td>
<td>87</td>
</tr>
<tr>
<td>g25a</td>
<td>7</td>
<td>119</td>
</tr>
<tr>
<td>g274a</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>g287a</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>g297a</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>g306a</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>g31a</td>
<td>7</td>
<td>102</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>61</strong></td>
<td><strong>993</strong></td>
</tr>
</tbody>
</table>

(b) Unlabelled

<table>
<thead>
<tr>
<th>Session</th>
<th>No. of dialogues</th>
<th>No. of turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>g10a</td>
<td>7</td>
<td>141</td>
</tr>
<tr>
<td>g11a</td>
<td>7</td>
<td>116</td>
</tr>
<tr>
<td>g12a</td>
<td>7</td>
<td>149</td>
</tr>
<tr>
<td>g36a</td>
<td>7</td>
<td>119</td>
</tr>
<tr>
<td>g37a</td>
<td>7</td>
<td>140</td>
</tr>
<tr>
<td>g38a</td>
<td>7</td>
<td>110</td>
</tr>
<tr>
<td>g41a</td>
<td>7</td>
<td>112</td>
</tr>
<tr>
<td>g42a</td>
<td>7</td>
<td>116</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>56</strong></td>
<td><strong>1003</strong></td>
</tr>
</tbody>
</table>
Figure 2: Relative distribution of speakers and genders in read database: bars show proportion of items with canonical glottal stop produced by each speaker, with female speakers at left (k62–k14) and male speakers at right (k61–k13) (n = 7334). k62=k62+rtd, k61=k61+kko (see text).

<table>
<thead>
<tr>
<th>Database</th>
<th>% male</th>
<th>% female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read speech</td>
<td>54</td>
<td>46</td>
</tr>
<tr>
<td>Labelled spontaneous speech</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>Unlabelled spontaneous speech</td>
<td>56</td>
<td>44</td>
</tr>
<tr>
<td>Entire database</td>
<td>58</td>
<td>42</td>
</tr>
</tbody>
</table>

Table 4: Gender profile in data shown in Figures 2–3(b), p.186 (n_{read} = 7334, n_{spontaneous labelled} = 4602, n_{spontaneous unlabelled} = 5582, n_{entire} = 17518).

2.2 Procedure

Two different approaches are used to search for items of interest, depending on the question at issue. For the examination of utterance-final glottalization, it is sufficient to search a database for prosodic labels representing a falling intonation (for 2. see Kohler et al. 1995:83), while labels for truncations can be similarly sought (for
Figure 3: Relative distribution of speakers and genders in spontaneous databases: bars show proportion of items with canonical glottal stop produced by each speaker, with female speakers at left (ANS–SOV and HEE–FRA) and male speakers at right (TIS–BLA and MAW–ARK). Figure 3(a) shows distribution in labelled spontaneous speech ($n = 4602$), Figure 3(b) shows unlabelled spontaneous speech ($n = 5582$).
The interplay of hypotheses makes the approach of a database more hazardous for the issue of vowel onset-related glottalization. Pätzold (1997) proposes generating a KielDat database and searching it using *awk* scripts that exploit the KielDat utilities, to gain an overview of the data in the corpus. Although the factors corresponding to the hypotheses raised in Section 1.2.5, p.181 can be examined through the labels in the database, their interaction is complex, and searching the database directly can make for correspondingly complex — and erroneous — *awk* scripts. To overcome the intractability of the database, a lexicon is generated to facilitate the examination of different combinations of proposed influences. In principle such a lexicon could be made for any examination, with the lexicon containing a subset of the information in the database, pared down to correspond to the issues under investigation.

For this examination, two databases were created, one of spontaneous speech representing the prosodically labelled dialogues in Table 3(a) (p.185), and one of read speech from materials published in IPDS (1994), and from each a lexicon was generated. For each vowel whose canonical realization features a glottal stop, the lexicon contains a field that contains the following information about that vowel: is it in a content or function word; is it initial or medial in the word; does it occur after a phrase boundary; with which of the four glottal realizations shown in Table 1 (p.176) is it realized; is the word in which it appears accented, and if so with what level of sentence accent; is the vowel primarily or secondarily lexically stressed or not. Further fields contain the orthographic and canonical representations, the filename and path of the file, and optionally durational information.

Once the criteria have been established by which the levels of the factors will be sorted, this search is entirely automatic, and is supplemented by narrow phonetic investigation using waveforms, spectrograms, and spectra in *xassp* (IPDS 1997b) as well as auditory analysis.

There are two aims to the detailed phonetic examination of the signal files. The first is to evaluate the accuracy and consistency of the segmental and prosodic labelling of the entire corpus, and to this end, a representative proportion (10%, or where the cell con-
Vowels can belong to one of the following classes, where the vowel of interest is underlined in the orthography:

1. lexically stressed word-initial vowels
   e.g. Arbeit, canonically Qa6baIt

2. lexically unstressed word-initial vowels
   e.g. August, canonically QaUg’Ust

3. compound-internal vowels
   e.g. Wochenende, canonically v’Ox@n#Q“End@

4. vowels in words with shifted or added stress, marked by ’’
   e.g. Augenblick, canonically Q’aUg@n#bl’’Ik
     but realized as Q’aUg@n#bl’’’Ik
   or ich, canonically QIC+
     but where accented realized as Q’’IC+, + marking a function word

5. word-internal vowels
   e.g. geir@t, canonically g@Q’I6t

There is a complex mapping of these five classes onto the four categories of glottalization /C8V, /C8, V and V shown in Table 1 (p.176) and the four levels of sentence accent 0, 1, 2 and 3 shown in Table 2 (p.180), hence the simplification of the terms “accented” and “unaccented” in Section 1.2.4 (p.179). For function words the five classes are simplified: lexical stress is not canonically marked in function words, so the distinction of the first two classes is lost; where they are accented, the accent is always realized on the relevant (i.e. initial) vowel (marked by ’’), except in one word, obwohl, which occurs only once; function words cannot be compounds, nor are there any word-internal vowels in function words.
2.3 Inferential and descriptive statistics

The parametric tests (e.g. analysis of variance) that are often used for the statistical analysis of speech elicited under laboratory conditions are not appropriate for the evaluation of the data in this study. As here, the data in a corpus study typically comprise absolute frequencies rather than means, with no measure of dispersion, and the conditions assumed by tests like ANOVA, namely homogeneity of variance (Levene test) and normality of distribution (Kolmogorov-Smirnov test) are not always met. Furthermore, because of the uncontrolled nature of the spontaneous data elicitation, different speakers contribute different numbers of items, and for different items there may be wildly differing n. Since these criterial conditions are not met, non-parametric analogues to parametric tests are used (cf. Siegel and Castellan 1988). As the non-parametric analogue to a multivariate analysis of variance with planned comparisons Lienert (1978) proposes a multifactorial contingency structure analysis with $\chi^2$ partition according to Lancaster (1949, 1950a, 1950b, 1960, 1967, 1969) which has been implemented in a program described in Fillbrandt (1986), and is a complex implementation of the $\chi^2$ test. The model of the analysis is speech style (2) $\times$ word class (2) $\times$ position in utterance (2) $\times$ sentence accent (2) $\times$ glottal category (4).

This analysis could only be applied to the subset of the prosodically labelled spontaneous data shown in Table 5 (p.191). These were the data available when the VAX cluster at the Computing Centre of the University of Kiel, on which the program was installed, was decommissioned; the program has yet to be ported to a more modern computing system, and was no longer running when new prosodically labelled data became available. These new data will be analysed when the program becomes available.

There is an overwhelming similarity between the patterns in the original and expanded database, documented in figures in Appendix C (p.268), so that findings reported here are held to be valid for the data not yet analysed by K.S.A., which will be analysed once the program becomes available again.

By contrast with inferential statistics, the descriptive statistics in Section 3 (p.192) refer to the full database detailed in Table 3(a) (p.185).
<table>
<thead>
<tr>
<th>Session</th>
<th>No. of dialogues</th>
<th>No. of turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>g07a</td>
<td>7</td>
<td>115</td>
</tr>
<tr>
<td>g08a</td>
<td>7</td>
<td>109</td>
</tr>
<tr>
<td>g09a</td>
<td>7</td>
<td>159</td>
</tr>
<tr>
<td>g141a</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>g192a</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>g202a</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>g213a</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>g256a</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>g274a</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>g287a</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>g297a</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>g306a</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>g315a</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>31</strong></td>
<td><strong>525</strong></td>
</tr>
</tbody>
</table>

Table 5: Prosodically labelled dialogues constituting spontaneous speech database for which K.S.A. was carried out.
3 Findings: onset-related glottalization

3.1 Label files

11,936 vowels from the spontaneous and read corpora are examined. The examination of the distribution of glottalization focuses on regular words, so excludes from a total of 4,775\(^2\) vowels in the spontaneous speech database 171 entries comprising 151 hesitations (see Appendix B, p.267) (<"ah">, <"ahm"), 12 technical breaks (<;T>) and eight false starts (/+), leaving 4,602 items. No items of this kind are found in the read speech database, which contains 7,334 vowels suitable for examination.

Table 6 (p.193) offers an overview of the distribution of the four glottal categories in Table 1 (p.176) (cf. Figures 1(a)–1(d), p.177) according to the four factors addressed in Sections 1.2.1–1.2.4 (pp.178–179), on the basis of the automated label search. In Figures 4–7 (pp.194–197), the absolute values in Table 6 (p.193) are represented as relative values: the four glottal categories add up to 100\%. For each group of hypotheses in Section 1.2.5 (p.181) each figure in turn shows the relative distribution of the four glottal realizations according to the two levels of each of the four corresponding factors, pooling the remaining three factors.

An overview of the findings of the contingency structure analysis is shown in Table 7 (p.194); all main effects and significant interactions are shown, and addressed in detail in Sections 3.1.1–3.1.5 (pp.192–196).

3.1.1 Speech style

Figure 4 (p.194) shows the relative distribution of glottal categories by speech style, pooling word class, position in the utterance, and sentence accent. As predicted, a glottal stop, with or without glottalization, is more frequent in read than spontaneous speech, while

\(^2\)Other users searching for Q in the canonical transcription and/or one of the four realizations in Table 1 (p.176) will find only 4,773 items. The two missing files are g095a007 and g086a005: file g095a007.mix has an initial vowel which lacks Q in the canonical transcription, while file g086a005.mix contains an illegitimate label combination (see Section 6, p.244). In this research, as in future updates to the database, both errors are corrected.
Table 6: Distribution of glottal categories in spontaneous and read speech according to four factors of speech style, word class, position in the utterance, and sentence accent (for “accented” and “unaccented” see Section 2.2, p.186), \( n = 11936 \) (\( n_{spontaneous} = 4602 \), \( n_{read} = 7334 \)).
Table 7: Overview of main effects and significant interactions in contingency structure analysis ($n = 9896$ ($n_{spontaneous} = 2562$, $n_{read} = 7334$)).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech style</td>
<td>3</td>
<td>250.52</td>
<td>0.0005</td>
</tr>
<tr>
<td>Word class</td>
<td>3</td>
<td>990.55</td>
<td>0.0005</td>
</tr>
<tr>
<td>Position in the utterance</td>
<td>3</td>
<td>842.15</td>
<td>0.0005</td>
</tr>
<tr>
<td>Sentence accent</td>
<td>3</td>
<td>1095.45</td>
<td>0.0005</td>
</tr>
<tr>
<td>Speech style $\times$ word class</td>
<td>3</td>
<td>17.87</td>
<td>0.0005</td>
</tr>
<tr>
<td>Speech style $\times$ position in the utterance</td>
<td>3</td>
<td>65.05</td>
<td>0.0005</td>
</tr>
<tr>
<td>Speech style $\times$ sentence accent</td>
<td>3</td>
<td>369.72</td>
<td>0.0005</td>
</tr>
<tr>
<td>Position in the utterance $\times$ sentence accent</td>
<td>3</td>
<td>272.94</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

Figure 4: Relative distribution of glottal realizations in spontaneous and read speech, pooling word class, position in the utterance, and sentence accent ($n_{spontaneous} = 4602$, $n_{read} = 7334$).

A realization with no glottal reflex is more frequent in spontaneous than read speech. Glottalization without a glottal stop is more frequent in spontaneous than read speech. The main effect of speech style is highly significant ($\chi^2 = 366.24$, $df = 3$, $p < 0.0005$).
3.1.2 Word class

Figure 5 (p.195) shows the relative distribution of glottal categories by word class, pooling speech style, position in the utterance, and sentence accent. As predicted, a glottal reflex is more frequent in content than function words: a glottal reflex of some kind is absent in only 10% of all content words. In function words, by contrast, 40% are realized without any glottal reflex. The main effect of word class is highly significant ($\chi^2 = 1179.03, df = 3, p < 0.0005$).

3.1.3 Position in the utterance

Figure 6 (p.196) shows the relative distribution of glottal categories by position in the utterance, pooling speech style, word class and sentence accent. As predicted, a glottal stop is more frequent phrase-initially than -medially, whereas glottalization without a glottal stop, or absence of any glottal reflex whatsoever, is most frequent phrase-medially. The main effect of position in the utterance is highly significant ($\chi^2 = 1118.05, df = 3, p < 0.0005$).
3.1.4 Sentence accent

Figure 7 (p.197) shows the relative distribution of glottal categories by sentence accent, pooling speech style, word class and position in the utterance. As predicted, a glottal reflex is more frequent in vowels that are accented than in those that are not: only 6% of accented vowels are realized with no glottal reflex whatsoever. By contrast, over 75% of unaccented vowels are realized simply with glottalization or with no glottal reflex whatsoever. The main effect of sentence accent is highly significant ($\chi^2 = 1377.88, df = 3, p < 0.0005$).

3.1.5 Interactions

Sections 3.1.1–3.1.4 (pp.192–196) have shown statistically significant support for the hypotheses advanced in Section 1.2.5 (p.181). This section examines in more detail the distribution of these effects in the two speech styles examined, and in different places in the utterance, with a view to underlining the significance of the effects, and
clarifying the importance of certain of the factors; in this paper interactions are offered for a subset of the data, although the similarity of the distributions in the data analysable before and after K.S.A. was unavailable suggest that the findings are valid for all spontaneous data (cf. Section 2.3, p.190).

**Interaction between speech style and word class**  
Figure 5 (p.195) showed the relative distribution of data for word class, in which speech style was pooled; Figures 8(a) and 8(b) (p.199) show the distribution of glottal categories by word class in spontaneous and read speech respectively. Similar overall patterns are seen for both spontaneous and read corpora, albeit with interesting differences, and in both corpora the main effect of word class is statistically highly significant: in the spontaneous data $\chi^2 = 294.79, df = 3, p < 0.0005$, in the read data $\chi^2 = 668.92, df = 3, p < 0.0005$. In both read and spontaneous speech a glottal reflex is more frequent in content than function words, with only about 10% of content words realized with no glottal reflex. However, the distribution of the glottal reflexes within the content words is different between speech
styles, a glottal stop being more frequent in read than spontaneous speech: 46% in read speech, and 28% in spontaneous. The corollary to this difference is seen in the distribution of function words between speech styles, where the absence of any glottal reflex is more frequent in spontaneous than read speech; 45% are unmarked in spontaneous speech, and 36% in read. The interaction of speech style and word class is highly significant ($\chi^2 = 17.87, df = 3, p < 0.0005$).

Interaction between speech style and position in the utterance Figure 6 (p.196) showed the relative distribution of data for position in the utterance, in which speech style was pooled; Figures 9(a) and 9(b) (p.200) show the distribution of glottal categories by position in the utterance in spontaneous and read speech respectively. As was the case for the factor of word class, similar overall patterns are seen for both spontaneous and read corpora, again with equally interesting differences. In both spontaneous and read corpora, the main effect of position in the utterance is statistically highly significant: for spontaneous speech $\chi^2 = 202.83, df = 3, p < 0.0005$, and for read speech $\chi^2 = 672.79, df = 3, p < 0.0005$. In both read and spontaneous speech a glottal reflex is more frequent phrase-initially than -medially, but the proportion of glottal stops is higher in read than spontaneous speech: 24% and 34% for glottal stops without and with glottalization in read speech, against 13% and 26% in spontaneous. The proportion of phrase-initial vowels with no glottal reflex is identical for spontaneous and read speech, at 21%, and the difference between the speech style lies in the category of simple glottalization, which stands at 41% for spontaneous and 21% for read. For phrase-medial vowels the proportion of simple glottalization is similar across speech styles; realizations with a glottal stop are more frequent in read than spontaneous speech, and realizations without any glottal reflex are more frequent in spontaneous than read speech, both as may be expected. The interaction of speech style and word class is highly significant ($\chi^2 = 65.05, df = 3, p < 0.0005$).

Interaction between speech style and sentence accent Figure 7, p.197 showed the relative distribution of data for sentence accent, in which speech style was pooled; Figures 10(a) and 10(b)
Glottalization in read and spontaneous German speech

![Graph](image)

(a) Spontaneous speech

![Graph](image)

(b) Read speech

Figure 8: Relative distribution of glottal realizations in content and function words in spontaneous and read speech, pooling position in the utterance and sentence accent ($n_{\text{spontaneous}} = 4602$ ($n_{\text{spont-cw}} = 1377$, $n_{\text{spont-fw}} = 3225$), $n_{\text{read}} = 7334$ ($n_{\text{read-cw}} = 2874$, $n_{\text{read-fw}} = 4460$)).
Figure 9: Relative distribution of glottal realizations in phrase-initial and -medial vowels in spontaneous and read speech, pooling word class and sentence accent ($n_{\text{spontaneous}}^t = 4602$ ($n_{\text{spont-initial}}^t = 1195$, $n_{\text{spont-medial}}^t = 3407$), $n_{\text{read}} = 7334$ ($n_{\text{read-initial}} = 2018$, $n_{\text{read-medial}} = 5316$).
show the distribution of glottal categories by sentence accent in spontaneous and read speech respectively. As was the case for the factors of word class and position in the utterance, similar overall patterns are seen for both spontaneous and read corpora, again with equally interesting differences. Both the spontaneous and read speech corpora show a highly significant main effect of sentence accent: for spontaneous speech \( \chi^2 = 334.71, df = 3, p < 0.0005 \), for read speech \( \chi^2 = 756.08, df = 3, p < 0.0005 \). A glottal reflex is more frequent in accented than unaccented vowels in both speech styles, with a similar, low, proportion of vowels being realized with no glottal reflex whatsoever (7% in spontaneous speech, 5% in read). Realizations with a glottal stop, without or with glottalization, are more frequent in read than spontaneous speech — 17% and 36% respectively for read speech, against 8% and 27% for spontaneous; the difference lies in the category of simple glottalization, which marks onset in 59% of accented vowels in spontaneous speech, as opposed to 42% of those in read. Similarly, glottal stops are more frequent in unaccented vowels in read speech than spontaneous, whereas simple glottalization or absence of any glottal reflex is more frequent in spontaneous than read speech. The interaction of speech style and word class is highly significant (\( \chi^2 = 369.72, df = 3, p < 0.0005 \)).

**Interaction between position in the utterance and sentence accent** The interaction between position in the utterance and sentence accent is statistically highly significant, \( \chi^2 = 272.94, df = 3, p < 0.0005 \). Figures 11(a) and 11(b) (p.205) show the relative distribution of the four glottal realizations in phrase-initial and -medial vowels respectively, by sentence accent. The main effect of sentence accent is statistically highly significant in both phrase-initial and -medial vowels: for phrase-initial realizations \( \chi^2 = 58.69, df = 3, p < 0.0005 \), for phrase-medial realizations \( \chi^2 = 1379.94, df = 3, p < 0.0005 \).

Phrase-medially the distribution of glottal realizations by sentence accent is very similar to that in Figure 7 (p.197), where position in the utterance is one of the pooled factors: very few accented vowels are realized with no glottal reflex whatsoever (4% in phrase-medial vowels in Figure 11(b), p.205, and 6% overall, in Figure 7, p.197), and, in order, simple glottalization, glottalization with a glottal stop,
Figure 10: Relative distribution of glottal realizations in accented and unaccented vowels in spontaneous and read speech, pooling word class and position in the utterance ($n_{\text{spontaneous}} = 4602$ ($n_{\text{spont}}-\text{accented}\cdot = 1247$, $n_{\text{spont}}-\text{unaccented}\cdot = 3355$), $n_{\text{read}} = 7334$ ($n_{\text{read}}-\text{accented}\cdot = 2191$, $n_{\text{read}}-\text{unaccented}\cdot = 5143$)).
and a simple glottal stop mark accented vowels. Phrase-medially as well as overall, unaccented vowels are most frequently marked by simple glottalization or the absence of any glottal reflex (43% for both in Figure 11(b), p.205, 38% and 37% in Figure 7, p.197).

Phrase-initiality alters the pattern observed. Glottal stops, with or without glottalization, are far more frequent in both accented and unaccented vowels, and the frequency of cases marked by simple glottalization is far lower. The number of vowels marked with no glottal reflex is higher for accented vowels phrase-initially than medially, whereas for unaccented vowels the opposite is true: even unaccented vowels are likely to be marked with some glottal reflex when phrase-initial.
Table 8: Distribution of realizations with and without a glottal reflex in content words in spontaneous speech, separated according to sentence accent. Figures which contradict hypotheses are shown in bold. \( n_{spon-cw} = 1377 \) \( (n_{spon-cw-‘accented’} = 890, n_{spon-cw-‘unaccented’} = 487) \)

<table>
<thead>
<tr>
<th></th>
<th>&amp;0</th>
<th>&amp;1</th>
<th>&amp;2</th>
<th>&amp;3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ?Vcw )</td>
<td>25</td>
<td>8</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>( Vcw )</td>
<td>57</td>
<td>38</td>
<td>186</td>
<td>12</td>
</tr>
<tr>
<td>( V_{cw} )</td>
<td>304</td>
<td>127</td>
<td>407</td>
<td>10</td>
</tr>
<tr>
<td>( V_{cw} )</td>
<td>101</td>
<td>12</td>
<td>36</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2 Signal files

The automated search of the label files is one step in the examination of the data, albeit one offering an effective way of gaining insight into the distribution of given phenomena, because of the quality of the labelling, in terms of its consistency and relative level of phonetic detail. The next stage is a detailed analysis of the signal files, most specifically those signal files containing items that run counter to the hypotheses in Section 1.2.5 (p.181), which are otherwise well supported by the data. The two criterial groups of items for examination of the signal file, regardless of speech style, are unaccented vowels realized with a glottal stop, and accented vowels realized without a glottal reflex. The following sections examine signal files corresponding to anomalous items in Tables 8, 9, 10 and 11 (pp.204, 210, 217 and 229).

3.2.1 Spontaneous speech

Anomalies in content words in spontaneous speech Table 8 (p.204) offers a detailed breakdown of realizations separated according to the levels of sentence accent listed in Table 1 (p.176).

Given the two assumptions that unaccented words are unlikely to be realized with a glottal stop, and that, conversely, accented words are unlikely to be realized without a glottal reflex, the first candidates for examination of the speech signal are the 25 unaccented
Figure 11: Relative distribution of glottal realizations in accented and unaccented vowels phrase-initially and -medially, pooling speech style and word class in ($n_{initial} = 3213$ ($n_{init-"accented"} = 657$, $n_{init-"unaccented"} = 2556$), $n_{medial} = 8723$ ($n_{init-"accented"} = 2781$, $n_{med-"unaccented"} = 5942$)).
words realized with a simple glottal stop, and the 49 accented words realized without a glottal reflex. These categories are shown in bold in Table 8 (p.204).

**Unaccented vowels realized with a glottal stop** Unaccented vowels are not expected to be realized with a full glottal stop. The simple glottal stop in almost half of the unaccented cases (12) in Table 8 (p.204) can be attributed to a preceding phrase boundary. Of the remaining thirteen, two are mislabelled (see Appendix A, p.264), while the other eleven feature a breathy onset which is marked as a glottal stop. The two speakers in these cases — KAK and THS — are both male, yet typically speak with a breathy voice quality; this may mean that an onset that in another speaker may not be perceived as a glottal stop is perceived as such in these speakers, even though lacking an archetypically hard onset.

**Accented vowels realized without a glottal reflex** Accented vowels are expected to show a glottal reflex of some kind. Of the 49 accented vowels which are nevertheless unmarked by a glottal reflex, two are mislabelled (see Appendix A, p.264). In 19 of the remaining cases, the onset of the vowel is marked by a strategy other than glottalization, but typically involving a change in voice quality. For example, the realization of *allerdings* in Figure 12 (p.207) has an unusual voice-quality which seems to serve as marking the onset.

Similarly, in the realization of *Ostermontag* in Figure 13 (p.208), the onset seems to be marked by a change in $f_0$: a dip in amplitude and $f_0$ is visible at 1450ms. This case also shows that between sonorants the glottal stop may not in any case be realized, especially in connected speech. The same phenomenon is shown in the realization of *einverstanden* in the phrase *[mit] einverstanden*[.] in Figure 14 (p.209), which is spoken so rapidly that there is no closure, and no perceived glottal stop.

These *allegro* realizations are different from the remaining 28 cases of accented vowels unmarked by a glottal reflex: in cases like that in Figure 14 (p.209), articulations are temporally compressed. In the cases below, however, some form of reorganization seems to take place.
Figure 12: Change in voice quality used to mark onset in [.] Aller-
dings [bin] in file g087a003.
Figure 13: Change in $f_0$ and amplitude used to mark onset in 
[...obwohl Ostermonta] in file g315a010.
Figure 14: Absence of glottal stop in accented word, concomitant with fast rate of articulation in [...]mit einverstanden [...] in file g076a004.
Table 9: Distribution of realizations with and without a glottal reflex in function words in spontaneous speech, separated according to sentence accent. Figures which contradict hypotheses are shown in bold.  

\[
\begin{array}{l|cccc}
\text{fw} & \&0 & \&1 & \&2 & \&3 \\
\hline
\text{?fw} & 102 & 11 & 19 & 0 \\
\text{fw} & 335 & 47 & 45 & 8 \\
\text{f} & 1055 & 76 & 102 & 5 \\
\text{fw} & 1376 & 14 & 30 & 0 \\
\end{array}
\]

\[
\begin{align*}
(n_{\text{spon-fw}} & = 3225 \ (n_{\text{spon-fw-"accented"}} = 357, \\
& n_{\text{spon-fw-"unaccented"}} = 2868))
\end{align*}
\]

In eight cases, the coda of the preceding syllable becomes the onset of the vowel under examination, e.g. [n] in \textit{achtzehn Uhr}, [s] in \textit{ins Auge fassen}. And in seven cases, where the vowel of interest is word-internal, it is merged with the preceding syllable, as in \textit{geirrt} and \textit{veranstaltten} in Figures 15(a) and 15(b) (p.211). In the realization of \textit{geirrt} glottalization in the preceding vowel marks a discontinuity in the signal related to the vowel onset, but temporally dissociated from it. The voicing in the \textit{irr} segment itself is modal, with glottalization in the preceding \textit{ge} syllable. In the latter case, \textit{veran} is reduced to a voiceless phase with formant structure visible in it, but the transition from breath to modal voice represents a comparable discontinuity.

**Anomalies in function words in spontaneous speech** A more detailed breakdown of realizations, separated according to the levels of sentence accent listed in Table 1 (p.176), is offered in Table 9 (p.210).

Given the two assumptions that unaccented words are unlikely to be realized with a glottal stop, and that, conversely, accented words are unlikely to be realized without a glottal reflex, the first candidates for examination of the speech signal are the 102 unaccented words realized with a simple glottal stop, and the 44 accented words realized without any glottal reflex. These are shown in bold in Table 9 (p.210).
Figure 15: Left: reduction of *geirrt* in *[Monat] geirrt [, es]*. in file g086a012; realization features glottalization of preceding vowel. Right: syllable reorganization in *[*<A*> veranstalten [könnten]*] in file g086a004; realization features between change in voice quality (breath→voice) generating discontinuity.
Unaccented vowels realized with a glottal stop  Of the 102 unaccented vowels realized with a simple glottal stop, 74 are phrase-initial, where a glottal stop may be expected, and 13 are mislabelled, leaving 15 to be explained. Of the 13 mislabellings, seven realizations have breathy onset or are voiceless, so are not in this sense archetypical glottal stops, although they were placed in that category by the labelers, suggesting this is a difficult labeling decision: the four glottal categories identified by the labels are underspecified, and this issue is addressed in Section 6 (p.244).

The remaining 15 cases are, similarly, different from canonical realizations of a glottal stop, and may be classified as allegro realizations similar to those discussed for content words. For example, in realizations of und and oder after -tag, there is coarticulation of the /k/ and the glottal onset of the vowel (double articulation of the plosive), as in Figures 16(a) and 16(b) (p.213); in these cases the plosive release is not as heavily aspirated as in isolation, and the onset to the vowel may be slightly breathy.

Accented vowels realized without a glottal reflex  Of the 44 accented vowels realized without glottal reflex, five are mislabelled (see Appendix A, p.264). Of the remaining 39 vowels, one — ist in file g082a004 shown in Figure 17 (p.214) — is phrase-initial, follows breath, and has breathy onset. The marking of such cases appears to be inconsistent, in separate but comparable cases the same phenomenon is marked by the presence or the absence of a glottal stop: this discrepancy is discussed in Section 6 (p.244).

The other 38 cases suggest the existence of other strategies beside glottalization for marking vowel onset. For example, the glottal stop is frequently lost where the vowel follows another sonorant, as is the case for auch in Sie auch in Figure 18(a) and mir auch in Figure 18(b) (p.215). In Figure 18(a) the whole phrase da haben Sie auch is radically reduced, and a dip in $f_0$ generates a discontinuity that serves as an index to the sentence accent, whereas in Figure 18(b) a similar dip in $f_0$ is reinforced by one in amplitude.

Coarticulation between sonorants across a boundary between function words can also include a following nasal, as bei Ihnen is reduced to /hǐn/ in Figures 19(a) and 19(b) (p.216); the creak in the /n/ of the latter example is probably due to utterance-finality.
Figure 16: Coarticulation of /k/ of -tag with onset of vowel, marked as Q albeit unaccented. 16(a) shows /Diens(tag und [der] in file g085a002, 16(b) shows /Diens(tag oder [Mittwoch] in file g086a004.
Figure 17: Phrase-initial, post breath breathy onset in [...]<A> Ist [gebongt.] in file g082a004 at 500ms.
Figure 18: Absence of glottal stop at transition between sonorants in function words. 18(a) shows /[...]/da haben Sie auch wieder [recht] in file g071a019; dip in \( f_\theta \) serves as index of vowel onset. 18(b) shows /[...]/mir auch [passen] in file g274a001; dip in amplitude and \( f_\theta \) serves as index of vowel onset.
Figure 19: Reduction and reorganization of /CJ/CQ/CP /A3 /C1 /D2

in [b] and g096a012 in 19(b); creak in /n/ is due to utterance-finality.

in 19(a) and g096a011
Glottalization in read and spontaneous German speech

3.2.2 Read speech

Anomalies in content words in read speech

Table 10 (p.217) offers a detailed breakdown of realizations separated according to the levels of sentence accent listed in Table 1 (p.176).

Given the two assumptions that unaccented words are unlikely to be realized with a glottal stop, and that, conversely, accented words are unlikely to be realized without a glottal reflex, the first candidates for examination of the speech signal are the 87 unaccented words realized with a simple glottal stop, and the 98 accented words realized without a glottal reflex. These categories are shown in bold in Table 10 (p.217).

Unaccented vowels realized with a glottal stop

Of the 87 unaccented vowels marked by glottal stop 25 are phrase-initial and 16 mislabelled (see Appendix A, p.264), leaving 46 to be explained. Of these, 34 may simply be categorized as clearly articulated speech. Such a category is not found in the spontaneous speech data addressed above, but in read speech many speakers are fluent yet precise in their articulation, occasionally bordering on hyperarticulation (Lindblom 1990). Such realizations may also feature a slower speech rate than is found in spontaneous speech, although, equally, certain speakers (e.g. dlm) are able to maintain a fast speech rate while articulating clearly. In slow speech, there is greater time for articulators to reach their trajectories, and “breaks” between words

<table>
<thead>
<tr>
<th></th>
<th>k0</th>
<th>k1</th>
<th>k2</th>
<th>k3</th>
</tr>
</thead>
<tbody>
<tr>
<td>?Vcw</td>
<td>87</td>
<td>19</td>
<td>310</td>
<td>21</td>
</tr>
<tr>
<td>?Vcw</td>
<td>162</td>
<td>61</td>
<td>648</td>
<td>33</td>
</tr>
<tr>
<td>Vcw</td>
<td>405</td>
<td>89</td>
<td>734</td>
<td>32</td>
</tr>
<tr>
<td>Vcw</td>
<td>175</td>
<td>11</td>
<td>83</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 10: Distribution of realizations with and without a glottal reflex in content words in read speech, separated according to sentence accent. Figures which contradict hypotheses are shown in bold. \((n_{\text{read-cw}} = 2874, n_{\text{read-cw-"accented"}} = 2045, n_{\text{read-cw-"unaccented"}} = 829))\)
may be reinforced, irrespective of accent. Slower and more careful speech is found in the isolated sentences than in connected texts. The influence of speech rate is addressed in Section 6 (p.244).

Similar explanations can be invoked for the remaining 12 items as were used to account for similar items in spontaneous speech. For example, the vowel onset may coarticulate with a preceding voiceless segment to produce a breathy or voiceless onset to the vowel, as shown in ausführen in Figure 20 (p.218), where the preceding word is Druckschrift.

A similar but subtly different case is shown in Figure 21 (p.219). Following the heavily accented nach (marked by k3) there is a drop in amplitude visible in the waveform and in the energy plot, and we may assume that the vocal folds fail to vibrate because there is
Figure 21: Voiceless onset to vowel following voiceless segment and heavily accented syllable. Phrase is /[...]/nach acht [/] in file dlme002.

...insufficient transglottal pressure drop after such a heavily accented syllable (cf. Ohala 1975), and because the vocal folds will also be widely abducted for production of a fricative (Löfqvist and Oshima 1993).

The same speaker (dlm) provides two further realizations where speech rate and offset of the preceding syllable determine the onset of the relevant vowel. The realization of aus in fast speech in file dlme070 shown in Figure 22(a) (p.221) is comparable to the phenomenon noted in spontaneous speech in Figures 16(a) and 16(b) (p.213), where the /k/ of -tag becomes the onset of the following vowel; here the /k/ of Regensburg becomes the onset of aus. Fast speech rate also contributes to the realization of Intercity in file...
dlme003 shown in Figure 22(b) (p.221). The /t/ of mit (in which the /m/ has the quality of the preceding vowel) is glottally reinforced, and is perceived as almost tap-like by dint of ballistic tongue movement (cf. Rodgers et al. 1997:136; Abercrombie 1967).

In the remaining cases there is evidence of temporal indeterminacy, where the glottal gesture that may be expected to mark the onset of the vowel can be dissociated from that location: provided such discontinuity is present in the neighbourhood of the stop, however, it seems robustly to cue vowel onset. In Figures 23–24(b) (pp.222 and 223) (erhalten in file k29but1, endlich in file k62tk021 and eingetroffen in file k29but1) there is creak in the offset of the preceding word, whilst the onset of the relevant vowel is modal. From the cases it is clear that the temporal alignment of glottalization can be highly variable. In Figures 24(a) and 24(b) (p.223) it is arguable that the realization should have been labelled as Q−q.

**Accented vowels realized without a glottal reflex** Of the 98 accented vowels realized without a glottal reflex, 46 are mislabelled (see Appendix A, p.264), leaving 52 to be explained. In the absence of a glottal stop, other strategies are used in these items to mark a discontinuity which elicits a percept of vowel onset. Perceptually, such strategies appear robust, in that they need not be located exactly at the point where a glottal stop would be expected, but, in the case of glottalization, or of dips in \( f_0 \) or amplitude, are subject to considerable temporal variability.

The commonest marker of vowel-onset related break where a glottal reflex is absent is a dip in \( f_0 \) and amplitude, which to an extent are interdependent. In Figures 25(a) and 25(b) (p.224), which show the words Abfahrtszeit and ankome in the files rtde006 and hpts097 respectively, the discontinuity is clear.

Although they are most often found together, as above, \( f_0 \) and amplitude can each be independently manipulated to effect a discontinuity. In Figures 26(a) and 26(b) (p.225) there is little dip in amplitude, and it is rather the lowering of \( f_0 \) that causes a discontinuity signalling sentence accent.

In Figures 27(a) and 27(b) (p.226) by contrast, \( f_0 \) remains relatively constant, whilst it is a dip in amplitude that signals the discontinuity. The figures show Ärzte in file k61be038, and beirren in
Figure 22: Offset of preceding syllable becomes onset of vowel in fast speech. 22(a) shows *aus* in *[Re-gens]h burg aus* in file *dlme070*. 22(b) shows *Intercity* in *[z\\text{w}a]r mit *Intercity* in file *dlme003*.
Figure 23: Variable temporal alignment of glottalization as index of vowel onset in *erhalten* in [...]zu erhalten [...] in file k29butt1.
Figure 24: Variable temporal alignment of glottalization as index of vowel onset. 24(a) shows *endlich* in *[...]Wann endlich* in file k62tk021. 24(b) shows *eingetroffen* in *[Butter eingetroffen sei]* in file k29butt1.
Figure 25: Discontinuity in the speech signal marked by dip in \( f_\theta \) and amplitude. 25(a) shows Abfahrtszeit in [...]oder Abfahrtszeit] in file rtdc006. 25(b) shows ankomme in [...]Uhr ankommme .] in file hpts097.
Figure 26: Discontinuity in the speech signal marked by dip in $f_0$. 26(a) shows Auskunft in [. . . ]eine Aus[kunft]/ in file kkos065. 26(b) shows endlich in [. . . ]kam en[dlich]/ in file k28but2.
Figure 27: Discontinuity in the speech signal marked by dip in amplitude. 27(a) shows Ärzte in [. . . /Die Ärzte] in file k61be038. 27(b) shows beirren in [nicht/ beirren /] in file k28butt2.

Furthermore, the choice of which strategy to use may depend on the speaker. Two separate speakers’ (k27 and k28) realizations of auf in the phrase forderten ihn auf are shown in Figures 28(a) and 28(b) (p.228) with $f_0$ and energy analyses respectively. In the first, ampli-
Glottalization in read and spontaneous German speech

227

tude rises rapidly after the change from nasal to oral segment, and it is a slight dip in $f_0$ that marks the vowel-onset related break; in the latter by contrast, $f_0$ is little altered, rather it is a discontinuity in the signal seen most clearly in the waveform and energy window at the transition from nasal to oral that marks the accented vowel: rather than “ramping” gradually, the amplitude remains relatively constant throughout the vowel (cf. Leshite (1960)’s observations on internal open juncture in pairs such as a nice man — an iceman, home-acre — hoe-maker, see the meat — see them eat, seem able — see Mabel).

A further possibility for marking a discontinuity is glottalization that is temporally dissociated from the relevant vowel. Whereas changes in $f_0$ and amplitude occur in the speech signal where the glottal reflex would be expected, glottalization may shift from the vowel onset, to a preceding or following segment. Where the following segment is glottalized this is captured by the labelling as -q and features in the previous tables in the categories /CV/ and /V/. Figures 29(a) and 29(b) (p.229) show examples of discontinuity marked by creak before the accented syllable, in the words beirren and Uhr in files k23butt2 and d1ms058 (cf. realizations of beirren in Figures 29(a), p.229 and 27(b), p.226).

Equally, glottalization of offset of the preceding syllable may co-occur with a dip in amplitude and $f_0$ to mark a discontinuity, as in the realization of andere in file ugae053 in Figure 30 (p.230).

Anomalies in function words in read speech

Table 11 (p.229) offers a more detailed breakdown of realizations separated according to the levels of sentence accent listed in Table 1 (p.176).

Given the two assumptions that unaccented words are unlikely to be realized with a glottal stop, and that, conversely, accented words are unlikely to be realized without a glottal reflex, the first candidates for examination of the speech signal are the 423 unaccented words realized with a simple glottal stop, and the 17 accented words realized without any glottal reflex. These are shown in bold in Table 11 (p.229).
Figure 28: Two different strategies used to mark sentence accent in the same phrase. In file k27butt2 in 28(a) $f_0$ (second pane from bottom) is cue to vowel onset, in file k28butt2 in 28(b) energy (bottom pane) is used. Phrase in both cases is /forderten/ ihn auf [,].
Figure 29: Creak in offset of preceding syllable used to mark discontinuity concomitant with sentence accent. 29(a) shows beirren in [nicht] beirren [,] in file k23butt2. 29(b) shows Uhr in [sech]zehn Uhr [Züge] in file dlms058.

Table 11: Distribution of realizations with and without a glottal reflex in function words in read speech, separated according to sentence-accent. Figures which contradict hypotheses are shown in bold. ($n_{\text{read-fw}} = 4460$ ($n_{\text{read-fw-"accented"}} = 146$, $n_{\text{read-fw-"unaccented"}} = 4314$))
Figure 30: Discontinuity in signal concomitant with sentence accent marked by (1) creak in offset of preceding syllable and dip in (2) $f_0$ and (3) amplitude. Figure shows *andere* in [einzige an[dere Möglichkeit] in file ugae035.
Unaccented vowels realized with a glottal stop \ Of the 423 unaccented vowels realized with a glottal stop 329 are phrase-initial and 18 are mislabelled, leaving 76 to be explained. Of these 77, 52 can simply be categorized as clearly articulated speech, within which the rate may be fast, normal, or slow (cf. p.218). Certain of the remaining 25 items cannot be regarded as canonical glottal stops, sometimes having a breathy onset, a breathy quality during the closure, or being voiceless. Such voiceless onsets typically follow voiceless consonants, e.g. ist after Mitternacht in file k11mr037 in Figure 31(a) and und after Wald in file k65be057 in Figure 31(b) (both p.232) As suggested in Section 3.2.1 (p.212), these cases with breathy onset appear to represent a difficult labelling decision.

Accented vowels realized without a glottal reflex \ Of the 17 accented vowels realized without a glottal reflex, 7 are mislabelled (see Appendix A, p.264), leaving ten to be explained. As is the case for comparable content words, i.e. accented without a glottal reflex, we find discontinuity signalled by changes in $f_0$ and amplitude, and by temporally dissociated creak. In the case of function words the speech rate is typically faster, in which light two examples are particularly noteworthy. Figures 32(a) and 32(b) (p.233) both show drastic articulatory reduction, and in both cases the syllables are not heavily accented ($\&1$). Both contain the phrase das ist (in the phrase ja das ist zu früh in Figure 32(b), in the phrase nein , das ist mir dann zu spät in Figure 32(a)), and in both the quality of the vowel is smeared onto the following /s/, which can be heard as /s/ in isolation.

3.3 Extrapolation

Figures 33(a) and 33(b) (p.234) and Figures 34(a) and 34(b) (p.234) show the absolute and relative distribution of the four glottal categories in the database for which prosodic labelling is not available.

As for data for which prosodic labelling is available, the first cases examined are the anomalous categories, that is, the 157 content words realized without a glottal reflex (cwQ00) and the 135 function words realized with a glottal reflex (fwQ10). Furthermore,
Figure 31: Voiceless onset to vowel marked as Q after voiceless consonant. 31(a) shows *ist* in [Mitternacht ist] in file k11mr037. 31(b) shows *und* in [Wald und Feld] in file k65be057.
Figure 32: Smearing of vowel quality of [i] onto following \( /\mathbf{CJ}/ \) shows ist in [\ldots] ja das ist zu früh in file hpte042. (a) shows ist in [\ldots] das ist mir dann zu spät in file hpte041.
Figure 33: Distribution of four glottal categories in content and function words in prosodically unlabelled data ($n_{content} = 1758, n_{function} = 3824$).

Figure 34: Relative distribution of four glottal categories in content and function words in prosodically unlabelled data ($n_{content} = 1758, n_{function} = 3824$).
a representative subset (9–10%) of other categories was examined, viz:

- 10/100 of cwQ10
- 40/393 of cwQ11
- 100/1108 of cwQ01
- 50/550 of fwQ11
- 150/1749 of fwQ01
- 140/1392 of fwQ00

In further support of the hypotheses, it was found that:

- 157 content words without a glottal reflex are unaccented or mislabelled
- 135 function words with a plain glottal stop are phrase-initial and/or accented
- Examination of randomized subset of other categories shows similar patterns for unlabelled as for labelled speech, i.e. glottal reflex associated with phrase-initiality, sentence accent, content word; absence of glottal stop in function words unless initial, or accented.

4 Findings: utterance-final glottalization

The spontaneous speech database contains 759 items marked with #&2.#&pg, that is, a phrase-final fall to the bottom of the speaker’s range. Table 12 (p.236) shows the distribution of these items according to whether any glottalization present may be attributable to other documented causes, viz plosive- or onset-related glottalization.

Clearly, phrase-finality favours glottalization. Where other properties already dispose the speaker to the production of glottalization, a strongly glottalized item will be realized: every one of the 22 cases where there is potential glottalization at vowel-onset and in a plosive
<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>Glottalized</th>
<th>% glottalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>onset-related</td>
<td>156</td>
<td>135</td>
<td>87</td>
</tr>
<tr>
<td>onset- and plosive-related</td>
<td>22</td>
<td>22</td>
<td>100</td>
</tr>
<tr>
<td>plosive-related</td>
<td>59</td>
<td>50</td>
<td>85</td>
</tr>
<tr>
<td>others</td>
<td>522</td>
<td>493</td>
<td>94</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>759</td>
<td>700</td>
<td>92</td>
</tr>
</tbody>
</table>

Table 12: Distribution of phrase-final items realized on falling intonation in spontaneous speech.

context contains a glottal reflex. Furthermore 85% of cases where plosive-related glottalization may be present, and 87% where onset-related glottalization may occur, are marked by a glottal reflex. It seems that in these cases where a fall in $f_\theta$ over a longer stretch of the utterance is already causing relaxation of the voicing gesture, a secondary influence such as plosive- or onset-related glottalization will only be enhanced by the phrase-final environment.

Phrase-final laryngealization affects 94% of the items where no other influence may be responsible for glottalization. Within these phrase-final cases, however, a range of possibilities is evident, summarized in the four utterances shown in Figures 35(a)–35(d), 36(a)–36(d), 37(a)–37(c) and 38(a)–38(d) (pp.239, 240, 241 and 242).

In Figures 35(a)–35(d) (p.239) the utterance is:

\[ Tja, \, da \, bin \, ich<Z> \, Ihrer \, Meinung. \, Wunderbar. \, <P> \]
\[ Vielen \, Dank \, <A>. \]

with phrase-final words underlined. One aim of this example is to show that glottalization can occur regardless of the length of the phrasal unit: for two of the four words here — tja and wunderbar — a word comprises the entirety of the phrase. The $f_\theta$ analysis shows the falling fundamental frequency, and in tja and wunderbar the alternation with breath is clear, while the entire realization of Meinung is with breathy voice.

Figures 36(a)–36(d) (p.240) shows phrase-final items from the following utterance:
Glottalization in read and spontaneous German speech

Here the alternation between glottalization and breathiness is clear: the $f_0$ analysis fails to find periodicity in the signal in two of the sections, and from the spectrogram it is clear that the periods are widely spaced, and at the end of the utterance also irregularly spaced. In many cases a difference is discernible between items that are turn-final, and those that are turn-internal, as the turn-internal items are less strongly glottalized than the turn-final (cf. Figures 38(a)–38(d), p.242).

For certain speakers, however, the alternation between glottalization and breath is less clear. Figures 37(a)–37(c) (p.241) show sections from the following utterance:

$<A>$ <Schmatzen> $f_0$, da wür’ ich auch $<Z>$ voll mit einverstanden. Also das würde mir grade gut in den Terminkalender passen. $<A>$ Dann $<Z>$ denk’ ich, halten wir das fest, und dann fahren wir da und $<P>$ denn $<Z>$ ja, könnten wir noch mal gucken, ob wir uns nachher vielleicht noch mal irgendwann treffen, um $<Z>$ $<A>$ Nachbereitung zu machen und $<Z>$ ei $<Z>$ nisch mal auf ’n Glas Wein, oder so.

Of the three phrase-final items, the turn-final *fassen* shows irregular periods in both vowels, then breath; the $f_0$ analysis fails to find periodicity in the signal. In *passen* and *frei*, however there is a section of glottalization (particularly strong in *frei*) followed by further modal voice, and only after this comes a breathy section. In this utterance the phrase-final items are not realized on a particularly low $f_0$, such that the speaker may be using creak to signal phrase-finality, and the breath that follows is not because the speaker’s vocal folds are relaxing, but in preparation for the next speech event. This possibility is interesting because it has been assumed that breathiness arises because the vocal folds are relaxing after creak, but here there is a phase where the vocal folds clearly tense again to resume modal voice after glottalization, and before breath.
The following utterance is shown in Figures 38(a)–38(d) (p.242):

\[ \text{Da muß ich zu einem Besuch nach Leipzig, } +/\text{das}<Z>/+ \]
\[ \text{das ist leider nichts zu machen. Ich könnte eher vorschlagen, } +/A/+ \]
\[ \text{direkt nach Muttertag, am Dienstag, dem zehnten. Das wäre dann } +/P/+ \text{ eine Woche später. Quasi.} \]

This series is chosen as showing several of the features discussed above within one utterance. The \( f_0 \) analysis in each case shows a fall in fundamental frequency. Glottalization is clear in Leipzig and \( \text{später} \), and breathiness also in \text{machen} and \text{quasi}. This final \text{quasi} comes as something of an afterthought, and is both breathy and glottalized. As mentioned in connection with Figures 36(a)–36(d), there is stronger glottalization in turn-final than turn-internal items.

5 Findings: truncation glottalization

Spontaneous speech is rich in phenomena that are regarded in laboratory speech as noise: misarticulations, errors, interruptions. The truncation phenomena examined here offer valuable insights into speaker’s articulatory planning, much as pathology can inform us about healthy bodies. Of 599 cases in the spontaneous speech corpus marked by \$/+\$, \$/+/, \$/+/ or \$/-\ 142 are of specific interest in that the truncation is within a sonorant, such that glottal activity can be reliably inferred from analysis of the speech signal. In all cases, an attempt to make an abrupt glottal closure is evident. What follows this truncation depends on how rapidly the speaker resumes his or her turn, which is itself an indicator of how great a disruption to the articulatory plan the truncation represents, and how rapidly the plan can be repaired.

Where the speaker misarticulates, glottal stricture is accompanied by oral closure as well, so that the speech signal reveals an abrupt glottal closure, and a strong oral stop. In Figure 39(a) (p.245), the particle \text{ke} can be interpreted as a misarticulated onset to \text{kommt} in part of a hesitant utterance \text{das}<Z> <\dd>+/+ \text{ke}=+/+ \text{kommt}: the initial /k/ is heavily fricated, then follow two periods of glottalization, and a clear glottal stop, which in turn has
Glottalization in read and spontaneous German speech

Figure 35: Utterance-final glottalization in [...] Tja, d[a bin ich<Z> Ihrer] Meinung[/] wunderbar. <P> [Vielen] Dank <A>[.] in file g076a014.
Figure 36: Utterance-final glottalization in [\textit{Schmatzen}] Ja, da wäre ich auch voll mit. Also das würde mir gerade gut in den Terminkalender passen. Dann halten wir das fest, und dann fahren wir da und denken an noch mal gucken, ob wir uns nächstes Mal noch mal intensiver treffen. Und dann, ja, könnten wir noch mal gleich, einfach mal auf 'n Glas Wein, oder so.
Figure 37: Utterance-final glottalization in “Schmatzen” <A>. Das würde mir auch sehr <Z> gut <Z> passen. Doch bin ich frei <A>. Dann wär’ ich sagen, <A> dann werden wir dieses Wochenende stark ins Auge <A> fassen in file g095a014.
Figure 38: Utterance-final glottalization in [Da muß ich zu einem Besuch nach Leipzig, +/das <Z>/+ das ist leider nichts zu machen, +/das <P>/+ die Woche nach der Mutertag am Dienstag, +/das <Z>/+ das wäre dann +/das <Z>/+ Quasi in der >. Quelle: manhänd ergebene Merkmale von file g072a008.
Glottalization in read and spontaneous German speech

243

a clear release. Similarly in Figure 39(b) (p.245), there is a misarticulation in a hesitant utterance: <äh> <A> vielleicht <P> <häs> +/ku=/+ die Woche<Z> <P>. Here, however, the role of the ku is unclear, as the next word is die. There is the same heavily fricated /k/ release, and several periods of creaky voice, before a silence for closure followed by a strong alveolar burst.

The truncation in some items is necessitated by a post hoc reorganization of the utterance, where the speaker realized that what has been planned contains an error. Figure 40 (p.246) shows the utterance +/zum Vorberei=/+ zur Vorbereitung, where the speaker realizes that the agreement of the condensed form of zu + definite article is wrong, and changes from masculine or neuter to feminine. The glottal stop at 4610ms is clear, with some low frequency vibration persisting.

The planning error can occur over a wider domain than a couple of words. In file g274a008 (for figure see Kohler 1999b:287) the speaker means to say in Hannover auf der Messe, but elliptically realizes this as in der Messe, and corrects this after a pause, such that the entire utterance is: in 'ner Messe , <P> +/a=/+ in Hannover auf der Messe. Glottalization at 2050ms marks the truncation, then as the speaker reorders her thoughts, at 2320ms in the abortive onset to +/a=/+ (perhaps an attempted auf), and the onset to in at 2410ms.

In file g315a009 (for figure see Kohler 1999b:288) the speaker misarticulates danach, producing a voiceless velar stop /k/ instead of the fricative /x/; he continues am O(stersonntag), which is then truncated, because the speaker has become aware of the error, and he repairs the faulty second syllable of danach. The utterance is thus da . +/*nak , am<Z> O=/+ _nach am<Z> Ostersonntag, with glottalization discernible in the truncated O=/+ at 4950ms, and the onset to the repaired _nach at 5130ms.

Figures 41(a) and 41(b) (p.247) show the use of repeated glottal stops and glottalization to break off an utterance that needs repeated repair. The speaker means to say ich freue mich sehr, daß Sie mich eingeladen haben, Sie zu besuchen but instead realizes this as . . . +/daß ich S<Z>/+ daß Sie mich eingeladen haben , +/sich zu<Z>/+ <P> <Lachen> Sie zu besuchen. In the first figure, Sie is prolonged as the error becomes clear then ended at 4215ms,
with some low frequency energy still visible after the closure for the /d/. The articulatory plan may be reorganized during this prolonged vowel, as the closure is short before a strong alveolar burst in \textit{daß} at 4290ms. In the second figure the \textit{zu} is similarly prolonged, but here the closure is at the glottis, with glottalization visible from 5890ms, before a long pause and before the repaired utterance resumes.

In some cases, glottalization is present when speech is resumed, before vocal fold vibration becomes modal. Figure 42 (p.248) shows a similar grammatical error to that mentioned in connection with Figure 40 (p.246), as the speaker uses a simple preposition instead of a contracted form with the indefinite article in +/im+/ <P> im Dezember. Glottalization is clear from 8320ms, but there is also glottalization in the resumption of speech, at 8410ms. A similar but more extreme case is shown in Figure 43 (p.249). The similarity is that a grammatical error — choosing das for die — causes a reorganization. In this case, the glottalization at 1475ms is onset-related, but the glottalization in the start of einlädig at 2000ms starts in die, and is probably also due to the resumption of the utterance. Nasal exhalation is clearly audible at 1820ms, perhaps because the glottal closure is reinforced by an oral closure, and the velic closure, perhaps less directly controlled than the oral closure, is released.

6 Discussion

Three out of four putative factors affecting glottalization have been examined in this paper: onset-related, utterance-final, and truncation glottalization. The fourth factor, plosive-related glottalization, is addressed in detail by Kohler (1999b) in this volume (p.281ff.).

This study has aimed to establish the distribution in spontaneous and read speech databases of four combinations of labels that correspond to four broad possibilities of glottalization, and then to see how these combinations co-occur with further labels that indicate word class, position in the utterance, and sentence accent. Statistical analysis shows that the hypotheses from Section 1.2.5 (p.181) are extremely strongly supported: the different glottal realizations are consistently associated with combinations of levels of the four factors of speech style, word class, position in the utterance, and
Figure 39: Truncation glottalization following misarticulation. 39(a) shows 
[kw=+/kommt+]/ in file g075a008. 39(b) shows 
[Woche<Z>P<häsci>]+/ku=/+ die 
in file g141a001.
Figure 40: Truncation glottalization following grammatical error in 
[+\textit{zum}] \textit{Vorberei}=+ zur \textit{Vorbereitung} in file g094a012.
Figure 41: Truncation glottalization in repeatedly repaired utterance in 
[...]/Sie<Z>/+- daß [Sie mich eingeladen haben , +/sich] zu<Z>/+ <P> <<Lachen> Sie [zu besuchen] in file g092a000.
Figure 42: Glottalization in truncation and resumption of utterance in [...] +/m/+ <P> im [Dezember] in file g081a014.
Figure 43: Glottalization in truncation and resumption of utterance in [...] +/das ein/+ die ein\textsuperscript{tägige} Arbeitssitzung/ in file g141a000.
sentence accent. These factors are all in binary opposition, so that with some accuracy it may be possible to say how their different levels map on to the different realizations, to the point of saying where along a scale of glottalization a combination of factors may place an item. In its crudest conception that scale may run from a full glottal stop, corresponding to an accented phrase-initial content word in read speech to an unaccented phrase-medial function word in spontaneous speech.

Prior to this research the main study into the incidence of glottalization was that of (Krech 1968). The deficiencies of that research have been pointed out, and bear reiteration here. Krech’s approach was prescriptive — destined for inclusion in the *Großes Wörterbuch der deutschen Aussprache* — whereas this research is descriptive in its outlook. Her study used a small number of professional speakers, predominantly male, whereas those examined here were typically phonetically naïve. The number and distribution of speakers shown in Figures 2–3(b) (p.186 and 187) and Table 4 (p.186) shows that the speaker sample examined here more accurately represent the speech community. This study has also been larger and broader in scope, largely thanks to advances in computing not available to Krech: the CD-ROMs of the Kiel Corpora have allowed an unprecedentedly large amount of data — 17,518 vowels — to be examined: Krech examined 9,886. This study has also investigated other potential influences on glottalization — plosive-related, utterance-final and truncation glottalization — and examined the interactions of these various environments. Finally, Krech’s study focused exclusively on read speech, so in this respect was untypical of genuine speaker behaviour. Corpora of spontaneous speech data offer excellent opportunities for speech research. Similarly to the research on German in this volume (Wesener 1999) (p.323ff.) is that of Ogden (1998), who offers a Declarative Phonology (Scobbie et al. 1996) account of destructive processes, such as deletion of function words, in English phonetics and phonology, using data drawn from MARSEC, a Machine Readable Spoken English Corpus (Roach et al. 1993).

The disadvantage of a corpus is that there is no control over the data. The advantages of corpus research, however, are potentially greater. Generalizations made about the material are legitimate generalizations about spoken language, precisely because observations
are based on heterogenous material. The database also provides natural rather than idealized material. Ogden cites Rischel (1992):

“Phonology has so far been based on very exaggerated idealizations of speech and exaggerated expectations about the power of rule machinery as the format in which to take care of variation”.

Indeed, phonetics can benefit even more than phonology from investigating genuinely natural speech, and corpus materials offer a rich source of data for examining the kind of Connected Speech Processes addressed by Nolan (1996) and Abramson and Lisker (1996).

Establishing that the data — as represented by the labels — broadly support the hypotheses represents one step in the study. Inter alia these findings confirm and extend those of Dilley et al. (1996), who examined the influence of pitch accent and phrase boundaries. The labels serve as an initial point of orientation, and the insights their distribution gives into phonetic phenomena are a stepping-off point for further examination. In this sense, the underspecification of phonetic phenomena through labels is not a weakness of the corpora, but rather an intentional feature of them. Detailed phonetic analysis of high quality speech data is the end point of corpus research, rather than counting labels.

The four categories in Table 1 (p.176) are phonetically underspecified: clearly there are more possibilities at vowel onsets than those that they outline. Two particular phenomena are brought to light by detailed phonetic analysis of the categories: phrase-initial onsets and an expanded definition of glottalization.

In phrase-initial realizations there is often inhalation then exhalation, and in the exhalation phase the vowel may be realized, resulting in a breathy onset. This phenomenon is diversely marked as Q (see Figure 44(a), p.252), Q- (see Figure 17, p.214), or in one case as Q- -h (see Figure 44(b), p.252). Clearly three characterizations of one phenomenon must be unified. The labelling with Q- -h is rejected as an inaccurate representation of the phonetic phenomenon at hand: in the entire database this combination of symbols is used only once, and for a case where the glottal friction it refers to is not comparable with that in the first syllable of Hammer, for example, which the labels suggest it should be. The use of Q is proposed for cases like
Figure 44: Phrase-initial post breath onsets marked by Q in \textit{ich} in file g096a001 in 44(a) and by Q-\textit{h} in \textit{am} in g082a009 in 44(b).
that in Figure 44(a) (p.252), where there is a hard onset, and Q− for 
those like in Figure 17 (p.214) where the hard onset is lacking.

To some extent this phenomenon overlaps with other realizations, 
mentioned in Section 3.2.1 (p.212): these are vowels that have a 
breathy onset or are voiceless, but have been marked as Q, although 
they are not archetypical glottal stops. Perceptually, they function 
as such, however, so are admitted to that category.

The body of realizations marked by −q is also expanded to in- 
clude realizations that may not be archetypically creaky. The label 
−q is originally used to describe creaky voice defined as slow and 
irregular vibration of the vocal folds. There is evidence to suggest 
that the label should be extended to describe cases of irregularity in 
the amplitudinal as well as temporal domain. Figure 45 (p.254) has 
a realization of auch in das paßt mir da auch nicht marked as Q− 
albeit with level 2 sentence accent. Here, a change in the amplitude 
envelope signals the onset. A change in amplitude envelope (and $f_0$) 
marking onset is also shown in Figure 13 (p.208).

A similar phenomenon occurs in the realization of Ihnen in Fig- 
ure 19(a) (p.216), which is also marked with a level 2 sentence accent 
but has Q− to show absence of a glottal stop. As in the previous case, 
there are two function words involved, suggesting a word class effect 
(cf. Wesener 1999, p.323ff.), and the lowering in the amplitude en- 
envelope (as also in $f_0$) serves to mark the onset. In the same way as 
breathy onsets perceived as glottal stops are admitted to the category 
Q, creak represents the appropriate category for such amplitudinal ir-
regularity. Where the original four categories in Table 1 (p.176) can 
be expanded — consistently — to embrace such phenomena, their 
underspecified nature becomes an economical but powerful method 
of describing a corpus.

Read speech provides further support for such an expansion of 
the category q to embrace non-standard realizations of glottalization. 
Section 3.2.2 (p.217) suggests that the key to the category is discon-
tinuity, be it marked by a dip in amplitude (Figures 27(a) and 27(b), 
p.226) or in $f_0$ (Figures 26(a) and 26(b), p.225), or by both together 
(Figures 25(a) and 25(b), p.224), or by creak, temporally dissociated 
from the site of the expected glottal reflex (Figures 29(a) and 29(b), 
p.229), or by a combination of all three (Figure 30, p.230). While 
the dip in amplitude or $f_0$ must be at the site of the expected glot-
Figure 45: Extending the category of creak: glottal reflex marked by amplitude envelope in […] mir da auch [nicht] in file g074a012.
tal reflex for the discontinuity concomitant with the initial accented vowel to be perceived, glottalization shows considerable temporal indeterminacy, and is in this sense a very robust cue.

Detailed phonetic analysis of read speech signal files (p.218) made clear that glottal stops occurred naturally in careful read speech, without the necessity for the vowel in question to be phrase-initial or accented. Careful speech typically features fewer reduction and assimilation processes, such that stops in general, including glottal stops, tend to be realized fully. Speech rate and precision of articulation are in principle independent (Laver 1994:158); however, although it is possible to speak slowly but sloppily, and rapidly but carefully (e.g. speaker dlm), much of the read speech analysed here is both careful and produced at a slow to normal rate. The issue here is production rather than perception, but different sources suggest casual speech can be perceived at a rate of 10–15 phonemes per second (Miller 1967, 1981a, 1981b) and rapid speech at a rate of 20–30 per second (Liberman et al. 1967): the careful speech examined here is at the lower end of this scale.

It was also noted that slower and more careful articulation was observed in isolated sentences (e.g. Berlin, Marburg sentences) than in passages (e.g. Der Nordwind und die Sonne, Die Buttergeschichte). Clearly, even a passage in turn comprises mere sentences, but the greater informational coherence in a narrative structure, and the reflection of natural speech rhythms in the texts in question appear to elicit readings of greater fluency. Even under laboratory conditions, well rehearsed narratively coherent passages can be used to elicit speech that contains Connected Speech Processes comparable to those in spontaneous speech, whereas isolated sentences tend to be less fluent (Rodgers 1998, Tunley 1999).

An economical summary of the position of the four underspecified categories, and how they might usefully be expanded, is provided by the phrase I.I.A., an abbreviation for Internationale Automobilausstellung, one of the events for which speakers in the spontaneous speech database were required to make an appointment. The I. and last A. are typically accented, with the middle A. unaccented. These

---

\(^3\)Cole and Jakimik (1980) suggest artificially sped-up speech can be perceived at a rate as high as 40–50 phonemes per second.
<table>
<thead>
<tr>
<th>Label</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>glottal stop</td>
</tr>
<tr>
<td>Q -q</td>
<td>glottal stop and glottalization</td>
</tr>
<tr>
<td></td>
<td>breathy onset to vowel</td>
</tr>
<tr>
<td></td>
<td>temporally dissociated glottalization</td>
</tr>
<tr>
<td>Q-</td>
<td>glottalization without glottal stop</td>
</tr>
<tr>
<td>Q-</td>
<td>absence of glottal stop and glottalization</td>
</tr>
<tr>
<td></td>
<td>dip in $f_\emptyset$ and/or</td>
</tr>
<tr>
<td></td>
<td>dip in amplitude</td>
</tr>
<tr>
<td></td>
<td>glottal stop coarticulated with preceding plosive</td>
</tr>
</tbody>
</table>

Table 13: Expansion of four underspecified categories in Table 1 (p.176).

three vowels cover the whole variety of glottal activity.

Figure 46(a) (p.257) shows an utterance-initial realization of the item, with the vowels marked as realized with glottal stop, creak, and absence of a glottal reflex, respectively, although glottalization is apparent in the final A. The combination of glottal stop and creak is shown in Figure 46(b) (p.257). These two files alone show the four minimal categories outlined in Table 1 (p.176), including the possibility of a breathy onset after a glottal stop.

The following figures show how the category of creak can be expanded. Figure 47(a) (p.258) shows the I. marked by a dip in $f_\emptyset$, the first A. by a dip in amplitude, and the final, accented A. by creak as well as a dip in $f_\emptyset$. Figure 47(b) (p.258) shows the first A. marked by temporally dissociated creak, and in a very heavily reduced realization, which sounds more like ja, Figure 47(c) (p.258) shows the I. marked by dip in $f_\emptyset$ and amplitude, and temporally dissociated glottalization, suggesting the Q- is a mislabelling.

Table 13 (p.256) offers a summary of the phonetic phenomena that might be captured by each label.

The extrapolation from prosodically labelled to unlabelled data in Section 3.3 (p.231) showed the robustness of the phenomena described, since it is possible with a high degree of accuracy to predict the accent of the item, given its glottal category and word class.
Figure 46: Examples of the four minimal categories from Table 1 (p.176) in the item I.A.A. 46(a) shows glottal stop Q, glottalization without a glottal stop Q- -q and mislabelled absence of any glottal reflex Q- (instead of Q- -q) in file g373a000. 46(b) shows glottal stop and glottalization Q -q, and breathy onsets after glottal stop in file g083a000.
Figure 47: Expansion of the category of creak in the item I.A.

(a) shows use of /∅/ in the I. and second A., and a dip in amplitude in file g123a009.

(b) shows temporally dissociated glottalization before the first A. from 813ms in file g413a000. An (c) shows temporally dissociated second A. and a dip in amplitude in the first A. in file g123a009.

(c) shows a dip in /∅/ and amplitude and temporally dissociated glottalization at the onset to the I. in file g123a000.
The economy with which a wide variety of glottal phenomena can be accounted for, using the interplay of relatively few influences, is attractive.

The interplay of further glottalization phenomena has been addressed both here and by reference to Kohler’s article in this volume. In particular it was established (Section 5, p.238) that truncation glottalization and utterance-final glottalization are in contrastive opposition, truncation being a constriction gesture alternating with glottal stop, and utterance-final glottalization being a relaxing gesture, alternating with breathy voice and breath. Unfortunately the corpus contains no example of a truncation glottalization that is utterance-final: in such an instance the glottis would be facing competing demands.

It was also pointed out (Section 4, p.235) that utterance-final and other forms of glottalization — plosive- and onset-related — may interact, as the glottal relaxation concomitant with phrase-final $f_0$-declination may enhance a predisposition to produce a glottal reflex in these two environments. Table 12 (p.236) showed an overwhelming correlation between utterance-finality and glottalization; furthermore utterance-final items where onset- or plosive-related glottalization could occur were typically glottalized (indeed 100% of items where onset- or plosive-related glottalization was possible were glottalized when phrase-final). In such cases glottalization due to separate factors can sometimes be separated, and this issue is addressed here and by Kohler (1999b) (p.281ff.).

Figures 48, 49 and 50 (pp.260, 261 and 262) show examples of phrase-final words where onset- and/or plosive-related glottalization is also present and each source of glottalization is separately discernible. In the realization of ersetzen shown in Figure 48 glottalization affects the entire word, and is discernible in the preceding den as well as ersetzen, but is of a different quality from that phrase-finally where the periods are widely spaced and breathy. The onset-related glottalization in ein- and Auch in Figure 49 is slightly breathy in quality, whereas that related to the /d/ is slightly damped, perhaps due to the nasal. The $f_0$-analysis in the bottom window shows that fundamental frequency is falling throughout the word. In the realization of anbieten in Figure 50 the quality of the onset-related glottalization is harsh, with high-amplitude energy at all frequencies, and the
striations dark and narrow, whereas the phrase-final glottalization declines into breath.

The corpus offers at least two examples where the influences on glottalization can be separated. The utterance *achtzehnten Oktober* in file g083a003 (for figures see Kohler 1999b:291) shows onset-related glottalization for *achtzehnten* and *Oktober* at 10350ms and 10920ms respectively, while utterance-final glottalization can be seen from 11380ms. Plosive-related glottalization is visible from 10800ms, but the quality of this glottalization is clearly different from that marking the onset of *Oktober*: the transition between the two is abrupt and clear at 10920ms.
Figure 49: Three types of glottalization in phrase-final [voll mit] ein-verstanden. Also [...] in file g074a005. Onset-related breathiness and glottalization affect ein- at 1460ms and Also- at 2050ms, plosive-related glottalization affects -standen at 1980ms, and phrase-final glottalization affects entire word: note $f_0$ declination throughout.
Figure 50: Three types of glottalization in phrase-final *anbieten* in file g081a015. Onset-related glottalization affects *an-* at 3850ms, plosive-related glottalization affects *-ten* at 4400ms, and phrase-final glottalization declines to breath at 4430ms.
In the phrase-final realization of *achtzehnten elften* in file g411a004 (for figures see Kohler 1999b:290), onset-related glottalization is visible in *elf* at 2800ms, plosive-related glottalization in *-zehnten* at 2750ms, and phrase-final glottalization leading to breathiness in *-ten* at 3160ms. The onset-related glottalization is, however, qualitatively different from the plosive-related, which in turn is different from the phrase-final glottalization.

Such cases, even though rare, offer important insights into the interplay of different influences of connected speech phenomena, in this case glottalization, and indicate that, much as the patterns of onset-related glottalization can be accounted for by the interaction of relatively few factors — sentence accent, word class, speech style, position in the utterance — the patterns of glottalization in connected speech as a whole can be explained by the interaction of smaller subsets of glottalization environments — onset-, plosive- and phrase-final-related glottalization.
A  Mislabellings

A.1  Spontaneous speech

A.1.1  Content words

<table>
<thead>
<tr>
<th>SS₃Vcw_unaccented</th>
<th>aus</th>
<th>g086a016</th>
<th>14.40s</th>
<th>&amp;0 Q10 → &amp;0 Q11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osterdienstag</td>
<td>g095a012</td>
<td>5.95s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SS₃Vcw_accented</th>
<th>auszumachen</th>
<th>g074a000</th>
<th>3s</th>
<th>&amp;1 Q00 → &amp;1 Q10</th>
</tr>
</thead>
<tbody>
<tr>
<td>einmal</td>
<td>g072a000</td>
<td>2s</td>
<td>&amp;1 Q00 → &amp;0 Q00</td>
<td></td>
</tr>
</tbody>
</table>

A.1.2  Function words

<table>
<thead>
<tr>
<th>SS₃Vfw_unaccented</th>
<th>aber</th>
<th>g086a011</th>
<th>1.1s</th>
<th>&amp;0 Q10 → &amp;0 Q11</th>
</tr>
</thead>
<tbody>
<tr>
<td>aber&lt;Z&gt;</td>
<td>g315a005</td>
<td>4.2s</td>
<td>&amp;0 Q10 → &amp;0 Q00</td>
<td></td>
</tr>
<tr>
<td>also</td>
<td>g096a005</td>
<td>2s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>am</td>
<td>g081a007</td>
<td>1.4s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>am</td>
<td>g084a001</td>
<td>3.4s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>auch</td>
<td>g096a015</td>
<td>2.4s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>eine</td>
<td>g085a001</td>
<td>1.55s</td>
<td>&amp;0 Q10 → &amp;0 Q00</td>
<td></td>
</tr>
<tr>
<td>ob</td>
<td>g085a003</td>
<td>2.5s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>und</td>
<td>g077a005</td>
<td>5.3s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>und</td>
<td>g085a004</td>
<td>3.6s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>uns</td>
<td>g097a000</td>
<td>11.5s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>und</td>
<td>g097a007</td>
<td>3.4s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
<tr>
<td>und</td>
<td>g315a009</td>
<td>0.9s</td>
<td>&amp;0 Q10 → &amp;0 Q11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SS₃Vfw_accented</th>
<th>also</th>
<th>g077a006</th>
<th>0.7s</th>
<th>&amp;1 Q00 → &amp;0 Q00</th>
</tr>
</thead>
<tbody>
<tr>
<td>auch&lt;Z&gt;</td>
<td>g073a001</td>
<td>1.3s</td>
<td>&amp;1 Q00 → &amp;0 Q00</td>
<td></td>
</tr>
<tr>
<td>auch</td>
<td>g192a018</td>
<td>0.8s</td>
<td>&amp;1 Q00 → &amp;0 Q00</td>
<td></td>
</tr>
<tr>
<td>ich</td>
<td>g081a009</td>
<td>0.3s</td>
<td>&amp;1 Q00 → &amp;1 Q10</td>
<td></td>
</tr>
<tr>
<td>und&lt;Z&gt;</td>
<td>g086a007</td>
<td>0s</td>
<td>&amp;1 Q00 → &amp;1 Q10</td>
<td></td>
</tr>
</tbody>
</table>

A.2  Read speech

A.2.1  Content words
Glottalization in read and spontaneous German speech

LS_V_cw_unaccented

<table>
<thead>
<tr>
<th>Word</th>
<th>Start Time (s)</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab</td>
<td>2.03500</td>
<td>0 Q10</td>
</tr>
<tr>
<td>ab</td>
<td>1.97338</td>
<td>0 Q10</td>
</tr>
<tr>
<td>allein</td>
<td>0.79744</td>
<td>0 Q10</td>
</tr>
<tr>
<td>aller</td>
<td>2.47250</td>
<td>0 Q11</td>
</tr>
<tr>
<td>anschließen</td>
<td>28.62006</td>
<td>0 Q10</td>
</tr>
<tr>
<td>aus</td>
<td>2.11606</td>
<td>0 Q10</td>
</tr>
<tr>
<td>ein</td>
<td>1.39400</td>
<td>0 Q10</td>
</tr>
<tr>
<td>Elefant</td>
<td>1.08356</td>
<td>0 Q10</td>
</tr>
<tr>
<td>empfahl</td>
<td>1.58575</td>
<td>0 Q10</td>
</tr>
<tr>
<td>erfolgt</td>
<td>0.92194</td>
<td>0 Q10</td>
</tr>
<tr>
<td>essen</td>
<td>0.94275</td>
<td>0 Q10</td>
</tr>
<tr>
<td>Uhr</td>
<td>8.05594</td>
<td>0 Q11</td>
</tr>
<tr>
<td>Verkehrsampeln</td>
<td>0.62156</td>
<td>0 Q11</td>
</tr>
<tr>
<td>Verkehrsampeln</td>
<td>0.61238</td>
<td>0 Q11</td>
</tr>
<tr>
<td>Verkehrsampeln</td>
<td>1.03487</td>
<td>0 Q11</td>
</tr>
</tbody>
</table>

LS_V_cw_accented

<table>
<thead>
<tr>
<th>Word</th>
<th>Start Time (s)</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abends</td>
<td>0.79731</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Abends</td>
<td>0.49956</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Abends</td>
<td>0.67094</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Achte</td>
<td>0.52838</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Adler</td>
<td>0.53981</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Adler</td>
<td>1.16250</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Adler</td>
<td>0.79338</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Adolf</td>
<td>0.73600</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Alle</td>
<td>0.57769</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Alle</td>
<td>1.05150</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Alle</td>
<td>1.05200</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Allzu</td>
<td>0.66150</td>
<td>2 Q11</td>
</tr>
<tr>
<td>Allzu</td>
<td>0.78625</td>
<td>2 Q10</td>
</tr>
<tr>
<td>Allzu</td>
<td>0.40450</td>
<td>2 Q10</td>
</tr>
<tr>
<td>ankommende</td>
<td>2.36356</td>
<td>2 Q10</td>
</tr>
</tbody>
</table>
A.2.2 Function words

LS.?V_fw_unaccented

ab rtds059 5.23769s &0 Q10 → &0 Q11
als  k67mr029  1.81619s &0 Q10 → &0 Q01
ein  k62mr084  1.57719s &0 Q10 → &0 Q11
ein  k68mr084  1.34587s &0 Q10 → &0 Q11
eine k62ko040  1.26319s &0 Q10 → &0 Q11
einem k62tk030  0.79331s &0 Q10 → &0 Q11
einen k62ko055  0.88287s &0 Q10 → &0 Q11
ihm  k62ko062  0.86769s &0 Q10 → &0 Q01
ihre k12mr056  0.79450s &0 Q10 → &0 Q01
in   k62e2022  1.92906s &0 Q10 → &0 Q11
in   kkos018  6.90375s &0 Q10 → &0 Q11
ist  k61ko025  0.59569s &0 Q10 → &0 Q01
um   k62sr009  2.24063s &0 Q10 → &0 Q01
ums k68mr098  1.26869s &0 Q10 → &0 Q11
und k68mr022  1.29056s &0 Q10 → &0 Q11
und k68mr031  1.78456s &0 Q10 → &0 Q01
und k08mr078  2.03031s &0 Q10 → &0 Q11
und k62mr095  1.90419s &0 Q10 → &0 Q01
also hpte062  1.17500s &2 Q00 → &2 Q10
einem k78butt1  7.35206s &2 Q00 → &2 Q01
Er   k24butt3  7.61694s &2 Q00 → &2 Q10
Er   k61mr008  0.60000s &2 Q00 → &2 Q10
Er   k62mr008  0.62756s &2 Q00 → &2 Q10
und dlme098  4.08506s &1 Q00 → &0 Q10
Unser k67mr070  0.94263s &1 Q00 → &1 Q10

B Hesitations

The 151 hesitation particles found in the spontaneous speech data are examined separately from other lexical items, partly because they cannot be compared with such items in the read speech database, where they are entirely absent, and partly because there are good reasons to believe such items are not even directly comparable with the 2562 items in spontaneous speech that are examined. Pätzold and Simpson (1995) examined a subset (three speakers) of the spontaneous speech data investigated here, establishing that the phonetic quality of the vocalic portions of hesitation particles is significantly different from that of those of comparable lexical items, \textit{inter alia}
contradicting Levelt’s (1983, 1989) assertion that hesitation particles represent the neutral position of the vocal tract.

Table 14 shows the distribution of the four glottal categories according to sentence accent phrase-initially; there are only four phrase-medial hesitation particles. Of these two are unaccented, and realized with glottal stop and glottalization, the other two are accented, and realized with glottalization without glottal stop. The remaining 147 items seem to conform to the pattern observed for other items in spontaneous speech, with the majority featuring a glottal reflex of some kind, typically including creak. The high proportion (7.5%) of accented items realized with no glottal reflex whatsoever may be because, albeit labelled as content words, hesitation particles are not lexical items as such, whose perception relies on decoding of specific acoustic and auditory properties, and may not have the same requirement for marking by a glottal stop as vowels in content and function words.

### C Interactions and their analysis by K.S.A.

Section 3.1.5 (p.196) indicated that interactions could only be examined by K.S.A. in a subset of the prosodically labelled spontaneous corpus. The following figures compare the distributions of the four glottal categories in the data already examined by K.S.A. shown in Table 5 (p.191) and in the full database of prosodically labelled spontaneous speech shown in Table 3(a) (p.185). This full database includes the data that has already been analysed by K.S.A., and
naturally reflects the distribution of that dataset to a certain extent. Indeed, the aim of these figures is to show that, because the distributions in the lesser and greater datasets are so similar, it is legitimate to assume that the findings of the K.S.A. for the lesser dataset also hold true for the greater one. The legitimacy of this assumption will be tested when the K.S.A. software is available again.
Figure 51: Relative distribution of glottal realizations in content and function words in spontaneous speech, pooling position in the utterance and sentence accent in databases shown in Tables 5 (p.191) and 3(a) (p.185) respectively. (In 51(a) $n_{\text{spon-cw}} = 2562$, $n_{\text{spon-fw}} = 777$), in 51(b) $n_{\text{spon-cw}} = 1377$, $n_{\text{spon-fw}} = 3225$).
Figure 52: Relative distribution of glottal realizations in phrase-initial and -medial vowels in spontaneous speech, pooling word class and sentence accent in databases shown in Tables 5 (p.191) and 3(a) (p.185) respectively (In 52(a) \(n^{p}_{\text{spon-initial}} = 2562\) \(n^{p}_{\text{spon-medial}} = 1938\), in 52(b) \(n^{p}_{\text{spon-initial}} = 1195\), \(n^{p}_{\text{spon-medial}} = 3407\)).
Figure 53: Relative distribution of glottal realizations in accented and unaccented vowels in spontaneous speech, pooling word class and position in the utterance in databases shown in Tables 5 (p.191) and 3(a) (p.185) respectively (In 53(a) \( n_{\text{spon}}^{p} = 2562 \) \( (n_{\text{spon}}^{p} - \text{"accented\"} ^{p} = 663, n_{\text{spon}}^{p} - \text{"unaccented\"} ^{p} = 1899) \), in 53(b) \( n_{\text{spon}}^{p} = 4602 \) \( (n_{\text{spon}}^{p} - \text{"accented\"} ^{p} = 1247, n_{\text{spon}}^{p} - \text{"unaccented\"} ^{p} = 3355) \).
Figure 54: Relative distribution of glottal realizations in phrase-initial accented and unaccented vowels, pooling speech style and word class in databases shown in Tables 5 (p.191) and 3(a) (p.185) respectively. In 54(a) $n_{\text{initial}} = 2635$ ($n_{\text{init}-\text{"accented"}} = 521$, $n_{\text{init}-\text{"unaccented"}} = 2114$), in 54(b) $n_{\text{initial}} = 3213$ ($n_{\text{init}-\text{"accented"}} = 657$, $n_{\text{init}-\text{"unaccented"}} = 2556$).
Figure 55: Relative distribution of phrase-medial glottal realizations in accented and unaccented vowels, pooling speech style and word class in databases shown in Tables 5 (p.191) and 3(a) (p.185) respectively. In 55(a) \( n_{\text{medial}} = 7261 \) \( (n_{\text{med} - \text{accented}} = 2334, n_{\text{med} - \text{unaccented}} = 4927) \), in 55(b) \( n_{\text{medial}} = 8723 \) \( (n_{\text{med} - \text{accented}} = 2781, n_{\text{med} - \text{unaccented}} = 5942) \).
References


