The phonetics of function words in German spontaneous speech*

Thomas Wesener

*This research was carried out under the supervision of K.J. Kohler as a further contribution towards the German Research Council (DFG) project "Sound Patterns of German Spontaneous Speech", but outside the funding by grant KO 331/22-1.

1 Introduction

The purpose of this study is to investigate the sentence-level phonetics of function words in German spontaneous speech. Function words are words whose role it is to express grammatical relations rather than lexical meaning. Previous studies on read speech (e.g. Kohler 1979; Kohler 1990b; Kohler 1994 for German) indicate that function words exhibit a great variety of reduction phenomena. They seem to be particularly prone to connected speech processes as they contain information that can often be inferred from the context, and as they are mostly unaccented.

Before the days of spontaneous speech data bases, researchers had to rely mainly on data from read speech, occasional observations in everyday conversation, and introspection in order to explore phonetic aspects of connected speech. Findings based on these methodologies are reviewed and expanded here in the light of a corpus of spontaneous speech. Since many function words occur frequently in the corpus, they are ideal objects for the study of sentence-level phonetics in spontaneous speech.

Traditional word-level accounts of the phonetics of function words prevail in pronunciation dictionaries. For example, the *Duden* pronunciation dictionary's blanket rule assigns a glottal stop to all initial vowels (DUDEN 1990). The *Großes Wörterbuch der deutschen Aussprache* (WDA, Krech et al. 1982) adopts a more differentiated approach, acknowledging that the glottal stop is frequently replaced by a so-called 'soft onset' in unaccented syllables not following a pause; however, it does not specify the phonetic nature of that type of onset.

The WDA as well as other studies in the same tradition (e.g. Meinhold 1973) advanced the phonetic account of German since they adopted a more descriptive approach, observing what speakers actually produced. However, the material investigated were radio broadcasts read by professional speakers. Therefore some of the findings may be due to speech training, rather than reflecting the spontaneous behaviour of phonetically unbiased speakers.

Factors that determine different degrees and forms of reduction include word-external and internal phonetic contexts. They play a major role in reduction and are therefore taken into account when describing reduction phenomena. In section 6 parallel data from a corpus of read speech are explored in order to gain insight into the influence of speaking style. The impact of sentence-accent and word class can only partially be examined since relevant categories often do not occur with sufficient frequency (cf. section 2).

Section 3 presents the concept of canonical labels and discusses the phonetic meaning of labels and their modifications. We then structure the findings according to reduction phenomena concerning consonants and vowels (section 4). The next section deals with the cooccurrences of different reduction phenomena. Appendices A and B give an overview of the absolute frequencies of the items, and of glottal patterns in words starting with a vowel.

2 Method

The paper is based on the *Kiel Corpus of Spontaneous Speech* previously published on CDROM (IPDS 1995; IPDS 1996; IPDS 1997); it examines 112 dialogues between 32 speakers[1]. Kohler, Pätzold, and Simpson (1995) and Kohler, Pätzold, and Simpson (1997) describe the collection and processing of the corpus by means of an appointment-making scenario, in which subjects communicate via headsets to arrange appointments.

Strictly speaking, the material is not as spontaneous as everyday conversation. For example, the technical setup disallowed overlapping speech. Therefore this speaking style is sometimes called non-scripted rather than spontaneous speech. However, we find that 'non-scripted' sounds too technical and does not concisely express the opposition to read speech. Since the material often sounds very natural, we prefer to slightly simplify the terminology and talk about spontaneous speech.

The speakers (18 male, 14 female) are 20 to 60 years old (34 years on average). The bulk of the speakers (27) come from the northern

[1]These dialogues consist of 37119 words, excluding hesitation particles, neologisms, slips of the tongue, and stretches difficult or impossible to identify, but including word-external truncations and false starts as well as technical breaks. This number is lower than in the databases described by Kohler and Rodgers in this volume because they considered five more (isolated) dialogues.

part of Germany (Schleswig-Holstein, Lower Saxony, and Hamburg), and all subjects speak a northern variety of standard German which cannot easily be ascribed to a specific region.

Read speech data derive from the *Kiel Corpus of Read Speech* (IPDS 1994), which features 53 speakers (27 male, 26 female) different from those of the spontaneous corpus, but comparable as to age, language area, and standard variant. Cf. Thon and van Dommelen (1992) and Kohler (1994) for a more detailed description of the composition of the read corpus.

In order to ease access to the data, a Kieldat database was created (cf. Pätzold 1997)[2]. Pattern-matching tools (*awk* scripts) were used to search the database for the items of interest, and gain an overview of the phonetic phenomena captured by the labelling. This was complemented by a narrow phonetic examination of a representative subset of the retrieved tokens in their respective contexts, based on an auditory and instrumental investigation of the speech signal.

For certain questions, the three sessions g07a, g08a, and g09a were chosen for systematic analysis of all relevant items[3]. The items were transcribed manually and classified according to the phonetic characteristics relevant in the respective items. The 21 dialogues were produced by six speakers (four male, two female) who were 25 to 29 years old and came from either Schleswig-Holstein or Lower Saxony (IPDS 1995).

When counting the frequencies of unaccented and accented tokens it turned out that accented function words only occurred in small numbers for most items. Function words were labelled as being unaccented (without a lexical stress mark) by default. If a function word was perceived as accented, a special label (`$''`) was inserted. Although this marker was already included in the process of segmental labelling, the subsequent prosodic labelling was responsible for systematic marking. This means that the number of accented items might increase when the prosodic labelling of all sessions investigated here has been completed. Data from the prosodically labelled data, however, show that even after systematic insertion of the accentuation marker, the number of accented tokens remains small. This

[2]The database represents the status of November 1998.

[3]These sessions contain 7459 words, excluding hesitation particles, neologisms, slips of the tongue, and stretches difficult or impossible to identify.

paper therefore concentrates on the unaccented tokens and discusses accentual aspects only in connection with some items where accented tokens occur with sufficient frequency (cf. e.g. section 4.1.8).

The database was searched for all function words to give an overview of what item occurs how frequently. Appendix A lists 204 orthographic types of function words with 20294 tokens. In some cases, reduction was already indicated in the orthography (e.g. by an apostrophe), which means that some of the orthographic types belong to the same word (e.g. *hab'* and *habe*). These types are combined again for the investigation of reduction phenomena in the respective words.

Not all words were subsequently investigated because of the multitude of different types, of which some are extremely rare. The corpus has some limitations as to the coverage of e.g. personal pronouns. The subjects were focused on interacting with each other within the framework of the appointment-making scenario. They concentrated on their diaries and were generally not talking about other persons. This has led to a shortage of third person pronouns: there is e.g. only one occurrence of *er* in the corpus. On the other hand, first person pronouns are abundant, both *ich* and *wir* occur more than a thousand times[4].

The general criterion for excluding certain words were frequencies below 100 tokens. However, if an item exemplifies an interesting phenomenon and occurs less frequently, it is nonetheless included. On the other hand, there are certain items that occur frequently but do not appear to be of special interest here, and are therefore excluded.

3 Canonical transcription and reduction

Before presenting the findings, some remarks should be made about the labelling of the corpus and its meaning for phonetic analysis. For practical reasons, an automatically generated canonical transcription was the point of departure for labelling. The canonical labels were then modified if necessary.

[4]For *ich*, this amounts to the enormous relative frequency of 0.0366. In other words, one out of 27 words in the corpus is *ich*.

The rules implemented in the automatic generation of labels take into account paradigmatic factors. This means that e.g. the auxiliary verb *haben* is represented with a schwa and an alveolar nasal in the final syllable, the form being placed in a paradigm with *habe* where a schwa can be found. However, all tokens of *haben* are produced without schwa (and with a bilabial nasal, cf. section 4.1.1.3). This means that the generated schwa had to be symbolically deleted in all cases. At this point, the question arises whether one should talk about 'reduction' in the form of schwa 'elision' here. It would not appear to be far-fetched to talk about schwa epenthesis in the canonical form, i.e. the generation of a unit that is not needed to account for the phonetics of the item under discussion.

The above remarks have some impact on the meaning of certain phonetic terms used in this study. An expression like 'connected speech *process*' is to be understood as a convenient shortcut for 'connected speech *phenomenon* involving deviance from a canonical form'. 'Elision' means the absence of a canonically postulated unit. In a similar vein, 'assimilation' does not necessarily imply a complete replacement of one category by another (as implied by the labelling) but also includes cases of incomplete accommodation. All these terms are not understood to express a psychological reality, but are used for the sake of convention and convenience.

One of the most striking characteristics of the material investigated here are a number of function words or collocations of function words which are prone to productions not easily lending themselves to linear segmentation. Non-linear aspects of speech are a challenge to symbol systems based on phonemic analysis. Some non-linear phenomena are therefore marked with special labels. `-q` thus stands for creak in connection with initial vowels, and `-~` for nasalization of vowels instead of contoid nasals. The label `-ma` is employed more generally for cases where a phonological unit is represented by suprasegmental characteristics only, e.g. by a secondary articulation (cf. section 4.2.3.2). However, not all cases of this kind have been marked. This means that the deletion of a label does not necessarily imply the absence of any phonetic correlate.

Finally two points about articulation in an acoustic corpus. Although the trained phonetician can infer a great deal of information on articulation from the acoustic signal and his perception of it, this inference cannot be complete (especially for vowels). Articulatory interpretations are therefore given as tentative reconstructions in the absence of primary articulatory data. The second point is that these articulatory interpretations do not imply a mechanic necessity of certain reduced forms, they only state why its plausible that such forms occur. Articulation is not only determined by mechanical constraints, but also by social factors allowing or disallowing certain types of reduction.

4 Reduction phenomena in function words

4.1 Reduction of consonants

4.1.1 Lenis plosives

In this section, we present different types of realizations of lenis plosives depending on the position in the word and on external contexts. We also consider the influence of place of articulation.

4.1.1.1 Word-initial lenis plosives The initial alveolar plosive in the definite article is produced in three main ways. The most typical case involves reduction from the canonical released plosive via an easily segmentable approximant to a minimal alveolar gesture and ultimately elision. Another direction is the replacement by a nasal after alveolar nasals. The third phenomenon is fricativization after alveolar fricatives.

The most frequent realization is a released plosive (80% of the cases, for the frequencies cf. also table 5 on p. 361). An approximant occurs in every tenth token. The degree of stricture can be reduced even further: figure 1 shows an extreme case of gesture reduction in a production of *noch über den* as [nɔ̟ʊ̟β̞ɵə̻̃], where the correlates of *den* are contained in a nasalized schwa and a drop in F1 between the preceding vowel and the schwa. This drop seems to serve as a cue to the presence of a residual laminal gesture (symbolized here by [◌̻]).

In the case of nasalization of the canonical plosive /d/, the resulting alveolar nasal is part of a longer portion of alveolar nasality, the

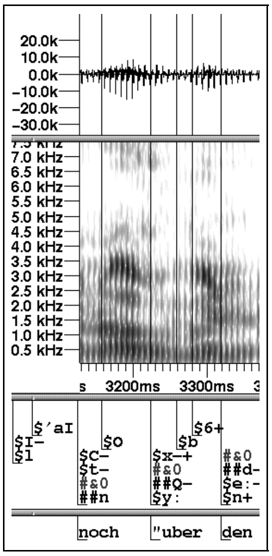


Figure 1: 'Elision' of /d/ in *noch über den* (g071a019).

other part being the final nasal of the preceding word[5]. Replacement of the alveolar plosive by a nasal mainly occurs after prepositions. The high rate of nasalization for *der* results among others from the frequent occurrence of *in der* in the corpus (66 cases). More than half of the occurrences of *in der* are produced with a nasal instead of the oral plosive, which is triggered by the preceding nasal in *in*. However, certain speakers (e.g. KAK) never produce a nasal stop instead of the oral one in the sequence *in der*. Rather than being an automatic process, nasalization can be controlled by the speaker.

The reduction of *in der* to realizations such as [ɪnˑɐ] synchronically mirrors diachronic developments which have led to lexicalized contracted forms like *zum* and *im*; these contractions also involve sequences of prepositions and articles (and semantically differ from the original *zu dem* and *in dem*). The prepositional phrase thus seems to be a laboratory of language change; it provides for frequent repetitions of sequences of the same function words.

Fricativization, where an alveolar fricative appears instead of the plosive, is conditioned by a preceding alveolar fricative. Figure 2 shows a case of fricativization of /d/ in *uns das* produced as [ʊnsˑa̠s]. Fricativization is conditioned here mainly by the preceding alveolar fricative in *uns*. In addition there seems to be an internal factor favouring fricativization since the latter occurs almost exclusively in *das* (19 out of 21 cases of fricativization in the definite article[6]). The alveolar fricative at the end of this item seems to regressively influence the articulation of the beginning of the word, which creates ideal conditions for fricativization when additional alveolar frication is present before the word-initial plosive.

In one case a progressive assimilation of place is observed. The

---

[5]The long nasal does not sound as long as an Italian geminate [nː], but is rather [nˑ].

[6]The orthographic items discussed so far can have grammatical functions besides that of definite article, i.e. demonstrative and relative pronoun. In the absence of a reliable parser to assign grammatical function to a large number of tokens, the assigning was carried out manually for a subset of six speakers. For *der*, *die*, *dem*, and *den*, the vast majority of the tokens are definite articles (96, 92, 100, and 99%, respectively). Extrapolating these results to the larger corpora, it seems justified to talk about definite articles when referring to these items. The picture for *das*, however, is different: in 83% of cases it is used as a demonstrative pronoun (‘that’).

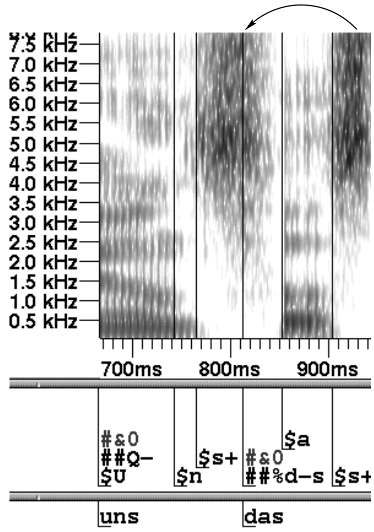


Figure 2: Fricativization of /d/ in *uns das* (g147a003).

initial /d/ is produced as a velar plosive after the velar plosive in *Dienstag, den* (g094a012). In German, progressive assimilation most often occurs in canonical plosive-schwa-nasal sequences where the schwa is not realized and the following nasal adopts the place of articulation of the plosive (cf. p. 333). Whereas these assimilations occur within words, the present case is different in that the assimilation crosses word- and phrase-boundaries. The strong cohesion of an unaccented function word in a collocation of day and date seems to allow for this type of progressive assimilation.

We now turn to the prepositions *bei* and *bis* to investigate the pattern of realization for word-initial /b/. The pattern found resembles that for /d/ in the definite articles. By far the most frequent production is a released plosive (84% for *bei* and 81% for *bis*). It is followed by approximant realizations; but in both items the incidence of approximation (15%) is higher than for /d/ in the definite articles. On the other hand, there are fewer cases involving nasalization and other phenomena.

The realization of word-initial /g/ in function words cannot be investigated properly with the present corpus since /g/ only occurs in the rare items *gegen*, *gewesen*, and *geworden* (together 10 unaccented tokens, of which six are released plosives, and four approximants).

4.1.1.2 Approximation of intervocalic lenis plosives The conjunction *oder* is interesting in that it allows for a comparison between its medial plosive and the initial plosive in *der*. For *oder*, only 54% of the tokens have been labelled as containing a released plosive, the remainder are realized as approximants[7]. For *der*, as much as 80% of the instances are released. Word-internal intervocalic /d/ is thus clearly more prone to reduction than initial /d/.

Data in *aber* show that approximation for /b/ is also significantly more frequent in word-internal than in word-initial position. Intervocalic /b/ was released in only 18% of the cases, which is even less than for /d/ in *oder*. The frequency of intervocalic approximation seems to depend on the place of articulation, bilabial lenis plosives being more prone to reduction than alveolar ones[8].

4.1.1.3 Lenis plosives before word-final nasals Turning now to lenis plosives in word-final plosive-nasal sequences, we first present the phonetics of the final syllables of *haben*, *werden*, and *gegen*. The least reduced form of *haben* is [ha̱b$^{\text{n}}$m̩] (e.g. g093a010 and g105a018). There are also productions where a bilabial approximant ends in a very short stop closure just before the lowering of the velum, resulting in a nasal release: [ɦa̱β̞$^{\text{n}}$m̩] (e.g. g415a004). More frequent, however, is [ɦa̱β̞m̩] without an audible nasal release. The most important realization is with a labial nasal only: around two thirds of the tokens are produced as [ha̱m] or [ɦa̱ɱ] (e.g. g426a015), the latter production being typical for the frequent collocation *haben wir*, which makes up for almost half of all productions of *haben*. Summing up the above productions, we can formulate the line of reduction for the final syllable of *haben* as:

[7] Helgason (1996) discusses examples of intervocalic lenition and temporal compression in *wieder* and *oder*. He also found an instance of *oder* with a tap, which he interpreted as gestural reorganization rather than reduction.

[8] Unfortunately, there is no function word with intervocalic /g/.

(1) [bⁿm̩] → [β̞ⁿm̩] → [β̞m̩] → [m]

It should be noted that in no case of *haben* is there a schwa, not even in the 30 accented productions. Thus, canonical [ha̠bən] does not seem to exist in spontaneous speech. [ha̠bⁿn̩] is not found in the corpus either; progressive assimilation of place of the nasal and the absence of schwa are two sides of the same coin.

However, there is one token of *haben* where the canonical alveolar and not the typically produced bilabial nasal was labelled. Inspection of the token, however, reveals that there rather seems to be a double articulation of a bilabial and an alveolar nasal ([ɦa̠m͡nzi̞]) in the context *haben Sie* (g115a005). The alveolarity seems to derive from the following alveolar articulation rather than from the alveolarity of the canonical /n/, which again has a labial correlate.

The most frequent production of the last syllable of *werden* is [-d̞n̩] with an approximant[9] and a syllabic nasal. There are also productions where the nasal immediately follows the vowel of the first syllable, i.e. where /d/ is dropped, and cases with an alveolar plosive followed by an alveolar nasal. In one case, a flat alveolar fricative seems to have been produced (g424a004).

One production shows creaky voice starting from the first vowel (g421a013). Although it was labelled as involving creak instead of the alveolar plosive, there still seems to be an alveolar approximant. Rather than replacing the plosive, creak is used here to mark the utterance-final position of this item. For the interference of plosive- and utterance-related glottalization cf. the article by Rodgers in this volume.

Inspection of the database thus yields the following line of successive reduction:

(2) [dⁿn̩] → [d̞n̩] → [n]

It should be noted that at the highest level of reduction, the diphthongal vowel [ɛɐ] of *werden* may also be monophthongized so that the word then tends to become homophonous with *wenn*; but there is still a slightly or clearly increased vowel duration as a reflex

[9]The alveolar approximant is symbolized here as [d̞] since [z̞] may be interpreted as involving a grooved tongue shape.

of the diphthong and the alveolar plosive in these cases (e.g. [ʋɛˑn] for *werden* vs [vɛ̠n] for *wenn*).

In all productions of *gegen*, schwa is not present, and a velar nasal is produced instead of an alveolar one. The nasal is mostly preceded by either a voiced velar plosive or a velar approximant. In one case, the whole plosive-schwa-nasal sequence was reduced to [ŋ], which still conveys the exponents of velarity and nasality.

Comparing the findings for *haben*, *werden*, and *gegen*, a discrepancy is obvious as to the frequency of the highest level of reduction with assimilated nasal only. Whereas the second syllable *-ben* in *haben* is predominantly reduced to a labial nasal, a nasal for *-den* in *werden* is less frequent, and a velar nasal for *-gen* in *gegen* occurs only once. The behaviour of the prenasal plosives in *haben* and *werden* resemble the findings for intervocalic plosives in *aber* vs *oder* where [b] also proved to be less robust than [d].

One might argue that the comparative paucity of *werden* and *gegen* does not allow a comparison with *haben*. High incidence might even be a factor for more reduction: the more a word is used, the more frequently simpler patterns for its production might be adopted. This tendency might partially account for the findings in connection with the prenasal plosives. But it seems improbable that frequency is the only factor since it does not play a role for the intervocalic plosives in *aber* and *oder*, which occur in comparable numbers.

To sum up the influence of place, there are more reduced articulations for the bilabial than for the alveolar lenis plosives in both intervocalic and prenasal position. The findings mean that places of articulation show differences in robustness for different reduction phenomena: while alveolar plosives are prone to changes in the *place* of articulation, they seem to be more robust against changes in the *manner* of articulation, or, more precisely, against more open strictures in word-internal intervocalic and prenasal positions.

4.1.2 Correlates of /t/

The preposition *mit* displays interesting phenomena in connection with its final alveolar plosive, e.g. in the sequence *mit mir*. This collocation is notable in that its phonetics does not display the canonical

[tm] in 5 out of 7 cases. There is one example of an alveolar-to-bilabial assimilation for /t/, the result not being a bilabial plosive, but a nasalized bilabial approximant (g194a000).

More important than assimilation of place is creak as a correlate of /t/ which occurs in the remaining 4 cases. The latter can be divided into two categories: one involving an alveolar gesture and creak in connection with /t/, and the other involving creak only. Cases like [ɪ̰t͡ʔm] and [ɪd̰̞̃m] (g081a000 and g086a000, respectively) belong to the first category. The second category contains two cases with creak in the [ɪ] (g085a008 and g413a001) followed by a modal bilabial nasal. In one of these cases (g413a001), an additional drop in F0 and intensity is observed at the beginning of the nasal.

A further point of interest is the production of the alveolar plosive cluster /td/ in *mit dem*, *mit den*, and *mit der*. In 35 out of 56 cases, /t/ has been marked deleted. Nonetheless, even in these renditions a plosive that is different from simple /d/ can be perceived. What is this perception caused by? An obvious guess is increased duration of the closure portion in comparison to the realizations of a canonical /d/. However, the mean duration of the portion is 47 ms, which roughly equals the 49 ms for the same forms of the definite articles not preceded by *mit*; these values clearly contrast with the 79 ms of clusters where /t/ has not been marked deleted. 47 ms could nonetheless be too long a duration in a token where a simple /d/ is perceived; in this case, /d/ would be in an internal position within a bisyllabic unit, and the plosive would have to be shorter than in an initial position. In a similar vein, increased duration of the release could be responsible for the /t/ percept; and indeed, measurements show a slightly longer release of the lenis plosive when /t/ was marked deleted (24 ms, $n = 31$ in comparison to 20 ms, $n = 21$ for the closure portions labelled with a fortis plosive). Other factors might be the intensity of the release and different F0 patterns preceding the closure; but it is hardly feasible to compare these parameters across a variety of microphone settings and prosodic contexts.

/t/ is frequently fricativized or dropped in the context of fricatives. Fricativization occurs e.g. in word-initial alveolar affricates preceded by alveolar frication. The frequent sequence *bis zum* is thus often produced with [ss̞s] at the transition of the two words.

Figure 10 shows that instead of a complete alveolar closure, there is residual alveolar frication with a higher energy cut-off and reduced intensity. A similar reduction of /t/ in the context *nicht so* can be found in figure 14.

The alveolar plosive of *ist* gives further empirical evidence for /t/-elision after fricatives in unaccented syllables (cf. e.g. Kohler 1995). It is absent in 89% of the cases. /t/ in *nicht* displays a similar behaviour, it is not realized in 72% of the productions. In items where /t/ is not preceded by a fricative in the coda, elision is far less frequent. In *mit*, /t/ is 'dropped' in only 23% of the tokens, mostly in the context of a following definite article beginning with /d/ (see above).

4.1.3 Word-final /x/

Doch, *auch*, and *noch* exhibit interesting phenomena in connection with the final /x/. In the pronunciation of isolated words, this corresponds to a voiceless uvular or velar (*auch*) fricative. In the corpus of spontaneous speech, however, the fricative has been labelled as being voiced, replaced by glottal frication (6%), or completely dropped (11%).

The label `x-h` can stand for glottal activity other than that involving abducted vocal folds or breathy voice typically associated with /h/ (cf. section 4.1.4). Figure 3 shows *doch auch noch mal* produced with whispery voice[10] for /x/: [dɔ̟̣ɦ̣a̱ɦ̣ṇnɔ̟̣̃ɦ̣ḥm̥m-]. Whispery voice can also be superimposed on longer stretches between realizations of /x/. *Auch noch mal* was thus produced as [ɑ̰̟ɦ̣ṇɔ̟̣ɦ̥̣ṃm-] (g415a004).

A further possibility for realizing /x/ is shifting voiceless breath, breathy voice or whispery voice onto a following nasal, either partially or completely. Figure 3 shows final voicelessness in *noch* and a stretch of voiceless nasality at the beginning of *mal*. The glottal activity is completely shifted on the nasal in the case of *noch mal* produced as [nɔ̟̣m̤m-] (g105a016).

At this point, findings from spontaneous speech necessitate another rewriting of a chapter of German phonetics and phonology

[10] Adrian Simpson (personal communication), cf. also Laver (1980); whisper is marked here with the ExtIPA diacritic [̣], cf. IPA (1999).

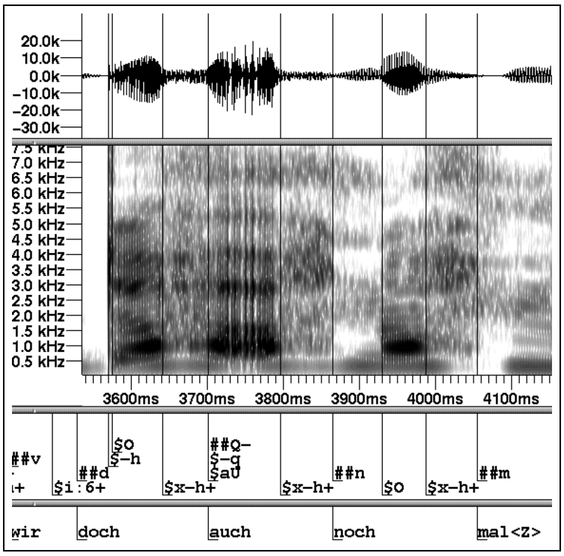


Figure 3: The correlates of /x/ in *doch auch noch m[al]* (g103a005).

which has been the object of numerous discussions in the past, i.e. the dorsal fricatives and their phonological status. Kohler 1990a emphasized that traditional accounts ignored the uvular fricative. In the light of new corpora of spontaneous speech not available at that time, different forms of glottal activity such as breathy and whispery voice have to be added[11].

4.1.4 Word-initial /h/

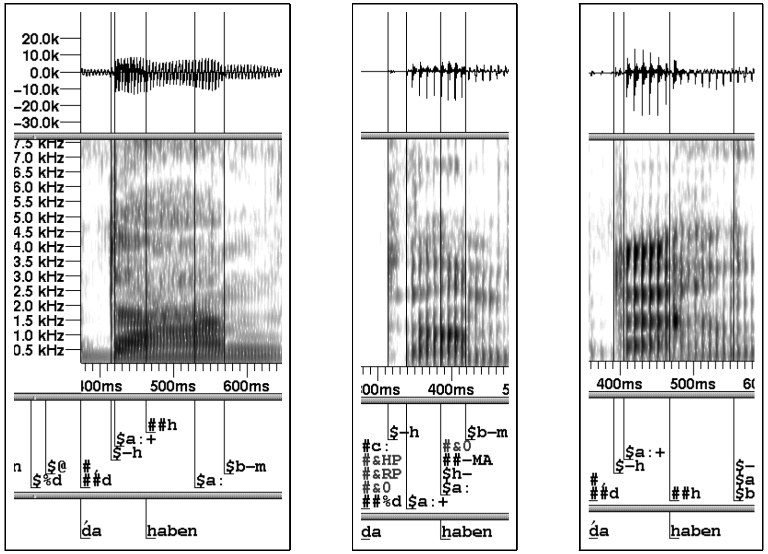
Canonical voiceless glottal frication for /h/ in haben can occur after voiceless plosives, e.g. in *geplant haben* (g105a018), but is not very frequent. The most typical production is a linearly segmentable portion of breathy voice (cf. figure 4(a)). However, in 9% of the cases, the label `-ma` was used to symbolize that there was still an /h/-percept, but that a corresponding portion could not be segmented linearly. An example is *da haben* where breathy voice is superimposed on the first part of the vowel portion, i.e. the vowel of *da*: [d̥ɐ̤ɐm] (cf. figure 4(b)). The correlates of /h/ can also be shifted to the second part of the vowel portion, i.e. the vowel of *haben* (cf. figure 4(c)).

It should be noted that this is just one among a number of examples showing temporal flexibility of glottal activity. The mobility of creak can also be observed in connection with initial vowels where it is not restricted to the initial portion of the vowel but can 'wander' towards the end of the vowel or to preceding sounds (cf. figure 3 for creak shifted to the right in *auch*). Similar shifts can be found for plosive-related creak (cf. figure 13 for creak shifted to the right of the nasal portion in *siebzehnten*, and the article by Kohler in this volume).

4.1.5 Word-final nasals

In this section we present two types of reduction of final nasals, which can be characterized as internal and external. The material investigated are articles ending in a nasal.

[11] All these productions can still be grouped into one phonological unit because they do not contrast with the other dorsal fricatives in the given environments. However, the findings also have an impact on models of phonology and its phonetic implementation, cf. section 7.



(a) medial (b) initial (c) final

Figure 4: Temporal flexibility of breathy voice for /h/ in *da haben*. Left: framed by modal voice (g254a013). Middle: shifted to the left (g071a019). Right: shifted to the right (g425a006).

The internal type occurs in the indefinite articles *einen* and *einem*, where two nasals can mutually influence one another. *Einen* is frequently produced with a single alveolar nasal, i.e. the nasal portion is not perceived as being longer than for *ein*[12] (cf. e.g. figure 7). This contraction is marked in almost half of the cases.

Einem is most frequently realized without schwa in the second syllable, and with a bilabial instead of an alveolar nasal. The absence of the schwa is a prerequisite for the alveolar being assimilated in place to the following bilabial. The resulting nasal portion may sometimes be perceived as a geminate, and sometimes as a single sound. The existence of productions with [nəm] and with [nm] makes it possible heuristically to establish the line of reduction

(3) [nəm] → [nm] → [mm] → [m]

of which the last two steps are the most frequent ones, occurring in 80% of cases in spontaneous speech.

In the external type of phenomena, the following word influences the nasal stretch. A typical example is *ein* being produced with a velar nasal when it precedes a velar, or with a bilabial nasal when it precedes a bilabial. The same articulations are also found for some tokens of *einen* in these contexts. The final syllable of *einem* is in two cases produced as [nŋ] before other velars. These external phenomena occur in 6% of the productions of *ein*, *einen*, and *einem*. Interestingly, they are significantly more frequent in *'n*[13], where they occur in one third of the tokens. The high frequency of bilabial nasals (18) mainly results from the common bilabial context, mostly in *'n bißchen*.

As for the forms of the definite article ending in a nasal, the final nasals in *den* and *dem* can undergo assimilation of place induced by following articulations. 4% of the nasals in *den* are regressively assimilated. In more than half of these tokens there is a velar nasal before words beginning with a velar, and in the other productions the nasal is bilabial in bilabial contexts. In the case of a labial instead

[12]This means that the masculine accusative *einen* can become homophonous with the masculine/neuter nominative and neuter accusative *ein*.

[13]Inspection of the context showed that all tokens of *'n* are indefinite and not definite articles. By contrast, the tokens of *in'n* and *auf'n* refer to combinations of prepositions and definite articles.

of the alveolar nasal, another lexical category might be involved, i.e. *dem* (see below).

For *dem*, 8% of the cases involve assimilation of place. In one token, there is a velar instead of a bilabial nasal, and in the remaining cases an alveolar nasal has been labelled. This, however, is probably due to a transliteration artefact. The transliterator conceived *dem* to be correct in these constructions and therefore transliterated *dem*, irrespective of the actual pronunciation with an alveolar nasal, which resulted in canonical `m` instead of `n`. These cases thus rather involve the word *den* and not an assimilation of place in *dem* (cf. e.g. *Mittwoch, dem neunundzwanzigsten*, which might as well be *Mittwoch, den neunundzwanzigsten* (g072a016); both meaning 'Wednesday, the twenty-ninth').

4.1.6 Nasalization instead of postvocalic /n/

We now turn to nasalization of a vowel as a replacement for /n/ (marked by `-~` in the labelling). The frequently occurring items *uns* and *und* are appropriate to investigate the influence of the articulation following the canonical nasal. Nasalization as the only correlate of /n/ occurs significantly more frequently in *uns* ([ʊ̃s]) than in *und* (6% vs 1%). The fricative thus seems to encourage nasalization.

Differences with regard to articulatory dynamics can account for this finding. Before a homorganic plosive, an oral occlusion has to be established anyway; before a homorganic fricative, however, an occlusion has to be released again into the fricative. Stop formation is simpler, and therefore executed faster, than formation of a fricative stricture (cf. Butcher 1977). Given the fast movement of the tongue tip vs the slow movement of the velum in the sequence before a plosive, the tongue tip has a high probability of reaching alveolar contact, while the velum is still lowered. Before a fricative, the more complex closing-opening gesture is simplified to a movement into a fricative stricture out of the vowel. Since the velum is lowered during this movement, the vowel gets nasalized.

Another reason for this difference may be the robustness of the alveolar fricative, which is present in all tokens of *uns*. This means that the alveolarity of *uns* is preserved even for nasalized renditions since it is still present in the fricative. In *und*, however, the alveolar plosive is not as robust, it has been marked deleted in more than half of the tokens[14]. The pattern for *und* is parallel to the findings for the indefinite article where a complete elision of the nasal does not occur frequently either (1%). In [a̰ĩ], the information on alveolarity would also be lost[15].

Interestingly, the synchronic behaviour of the nasal in *uns* and *und* resembles a diachronic development in other Germanic languages: whereas German has *uns*, the correspondent English form is *us* without nasal (cf. Kohler 1996). Nasalization as a replacement of the contoid nasal may have been the predecessor of this elision. English thus went a step further and dropped the nasal completely, but only before fricatives: it has not given up the nasal in *and* corresponding to German *und*.

4.1.7 Postvocalic /l/

Postvocalic /l/ is frequently not realized as an alveolar lateral involving a central closure. There are two categories of reduction labelled in the corpus: the first is vocalization of the lateral, resulting in a diphthongization of the preceding vowel (labelled as `%l`). The endpoint of the diphthong can have lateral quality, i.e. the sides of the tongue may be lowered although there is no alveolar contact. The second category is elision of the lateral where no traces of a lateral can be perceived (`l-`). The two categories divide a continuum rather than reflecting different types of reduction: there are borderline cases which can be interpreted as involving small quality changes in a diphthong, or as monophthongs with formant transitions induced by following articulations.

The frequent items *also* and *mal* show a high rate of reduction of the lateral. They are produced without a contoid for /l/ in about half of the cases (49% and 47%, respectively). Both words have a pragmatic rather than semantic and syntactic function, which together with their high incidence seems to allow frequent reduction

[14]Although in some of these cases, the deleted stop can still have correlates, e.g. in a slightly increased length of the closure in following words starting with an alveolar plosive.

[15]However, this may happen in certain regional varieties of German where *ein* can be produced as [ə].

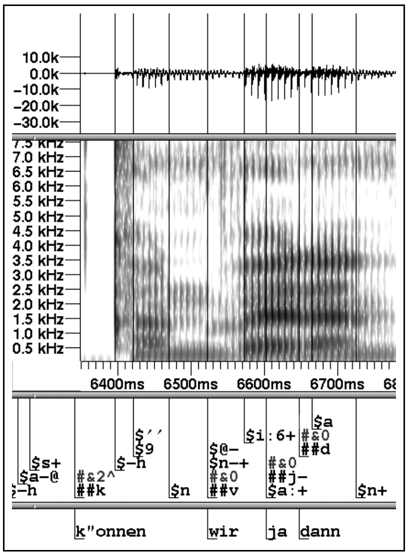


Figure 5: The correlates of *ja* in *können wir ja dann* (g071a010).

of the lateral.

4.1.8 Word-initial /j/

The productions of the modal particle *ja* typically display a rather long transition from the palatal glide to the open vowel, which makes it well-nigh impossible to state where exactly the boundary between the two vocoids lies. In fast speech, the glide can 'disappear' (`j-` in 6% of the tokens), which frequently occurs in the sequence *wir ja*. An example is *können wir ja dann* being produced as [kœ̃ɱṽɛ̠ɛd̥ãn] where [ɛ̠ɛ] is the phonetic exponent of the phonological sequences /ɪr/ and /ja/ (cf. figure 5). Where does the /ja/-percept come from in these cases? It seems that the quality of this phonetic diphthong is raised too much for it to represent only *wir*. In addition, the diphthong appears to be longer than would be necessary for *wir* alone. The correlates of /j/ are thus non-segmentably contained in the quality and the duration of the diphthong.

Ja meaning 'yes' or introductory 'well' is rather unlikely to be produced without a glide (4 out of 807 cases) in comparison to the modal particle *ja* above. This seems to be due to positional and prosodic factors. While the modal particle typically occurs sentence-medially, the affirmative/introductory *ja* is mostly found sentence-initially. In addition, the latter is often accented, while the modal particle only rarely receives sentence-accent (6 cases).

4.2 Reduction of vowels

4.2.1 Glottal activity at the onset of word-initial vowels

Canonically, initial vowels are presumed to be preceded by a glottal stop, represented in the transcription by the label `Q`. Where speakers do not produce a glottal stop, the label is marked deleted (`Q-`). The label `-q` stands for creaky voice. Figure 6 displays four possible combinations of labels: `Q` for a glottal stop only, `Q -q` for a glottal stop with laryngealization of the following vowel, `Q- -q` for creak without a glottal stop, and `Q-` for absence of any glottal reflex.

Although the labels give a good general idea about glottal activity, they do not specify the whole range of phonetic phenomena that can be observed. For example, the label for a glottal stop `Q` does not always represent a 'neat' glottal stop: it may mark a leaky closure, where a slight opening at the posterior part of the vocal folds seems to allow some air to escape, resulting in glottal friction filtered by supraglottal resonances. Speaker THS, for example, occasionally produces a stretch of relatively strong voiceless glottal activity in which F1 is clearly visible (cf. the beginning of *einen* in figure 7). The auditory impression and the spectrogram point to whisper in which there is only an inter-arytenoid opening (Laver 1980). This opening generates more noise than when the vocal folds are abducted as well (e.g. in aspiration). Other options to mark discontinuity at the beginning of the vowel are a harsh voice quality, a drop in F0, or a dip in the energy contour (cf. the article by Rodgers in this volume for more detailed information).

Table 1 shows that a glottal stop on its own (`Q`) is the least frequent realization (that holds for all function words beginning with a vowel, cf. table 7). Creak (`Q- -q`) and absence of a glottal reflex (`Q-`)

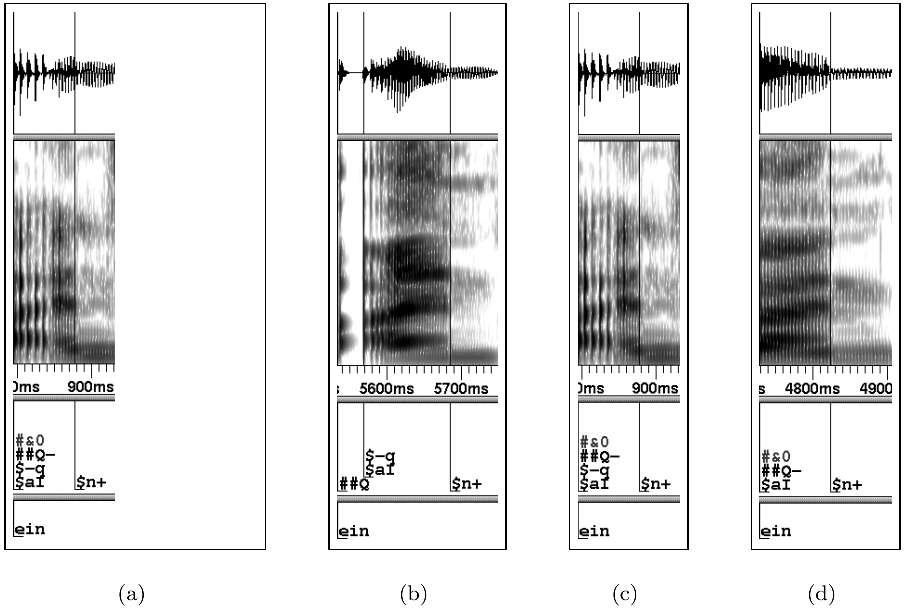


Figure 6: Combinations of glottal stop and creaky voice in *ein* (female speakers). a) glottal stop (g363a010), b) glottal stop plus creak (g107a001), c) creak only (g091a030), and d) absence of a glottal reflex (g194a012).

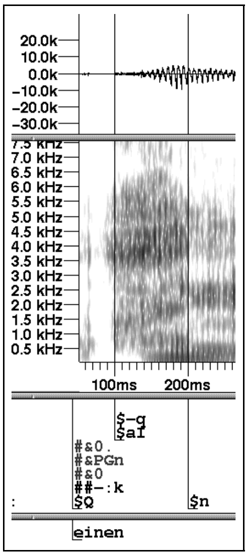


Figure 7: Voiceless glottal activity in *einen* (g081a013).

Table 1: Relative frequencies of the combinations of glottal stop and creak for the items discussed in this section, with n indicating the number of unaccented tokens containing a vowel, and in the next four columns the percentages of the four glottal realizations with reference to n. * stands for all inflections of the item.

item	n	`Q-`	`Q- -q`	`Q`	`Q -q`
*ein**	366	20	57	2	21
es	243	67	27	2	4
ich	1240	62	28	2	8
und	672	28	47	7	17
uns	185	49	44	2	5

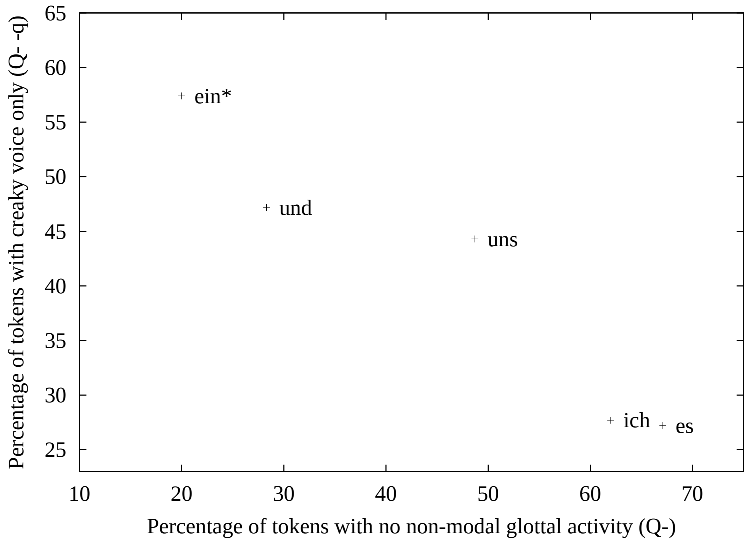


Figure 8: Percentages of the two types of vowel onset without glottal stop for *ein**, *und*, *uns*, *ich*, and *es*.

are the most important combinations, and therefore are visualized in figure 8. Glottal stop plus creak (`Q -q`) is quite common in *ein** and *und*.

Absence of a glottal reflex is more frequent than creak in *ich*, *es*, and *uns*, while it is less frequent in *ein* and *und*. Position in the utterance seems to be one factor for this pattern. *Ich* often occurs enclitically with modal verbs, *es* and *uns* are frequently encliticized in pronominal sequences. The indefinite articles and the conjunction *und*, on the other hand, often occur phrase-initially. Enclitic position seems to favour absence of a glottal reflex.

Why is absence of any glottal reflex much more frequent for *ich* and *es* than for all other function words (cf. figure 8, and also figure 17)? Besides position, another reason seems to lie in the duration of the voiced portion following the glottal activity. The mean duration of that portion (i.e. the vowel) in *ich* is 55 ms, and in *es* 58 ms. The voiced portions relating to the vowels in *in* (58 ms) and *im* (64 ms) are complemented by the voiced portions of the following nasals, which yields a longer stretch of voice than in *ich*. In *auf* (99 ms) and *auch* (109 ms), the vowel is directly followed by voiceless fricatives as it is in *ich* and *es*, but the durations of the diphthongs clearly exceed those of the monophthongs in the latter items. It is only the voiced portions in *ich* and *es* that are as short as 60 ms before voicelessness sets in. Their final voiceless fricatives typically involve an open glottis, and an only slightly advanced opening of the glottis in these short vowels decreases the probability of creak.

Figure 8 shows that the percentage of cases without a glottal reflex is lower for *uns* than for *ich* and *es*, but higher than for the indefinite article. We have argued above that creaky voice is avoided in *ich* and *es* because of the phonatory environment where the opening of the glottis for voicelessness runs contrary to the glottal adjustment needed for creak. The indefinite article, by contrast, does not contain such a glottal abduction. *Uns* occupies an intermediate position in that the voiced stretch is potentially longer than in *ich* and *es*[16], while it occurs mainly enclitically, in contrast to the forms of *ein* which also occur phrase-initially.

When comparing the glottal combinations for *uns* with those for the minimally different *und* (cf. table 7), a shift with regard to the most common combination occurs. Whereas most instances of *uns* have no non-modal activity at all, the most typical production of *und* is with creaky voice. This, together with the fact that the glottal stop occurs more frequently in *und* than in *uns*, probably results from *und* signalling a syntactic boundary, whereas *uns* is often attached to other function words in constructions like *wir uns*. The boundary between two clauses seems to be more important than the transition between two unaccented personal pronouns with strong syntactic cohesion, and so the boundary is made more prominent by glottal activity.

In one token of *und* glottal activity is retained while the vowel itself is not produced. Thus, in the sequence *treffen, und dann* the phonetic correlates of *und* are contained in a laryngealized nasal portion (cf. figure 9). However, this case is ambiguous: creak may

[16]At least in 94% of the tokens where the nasal is not dropped in favour of nasalization of the vowel, cf. section 4.1.6.

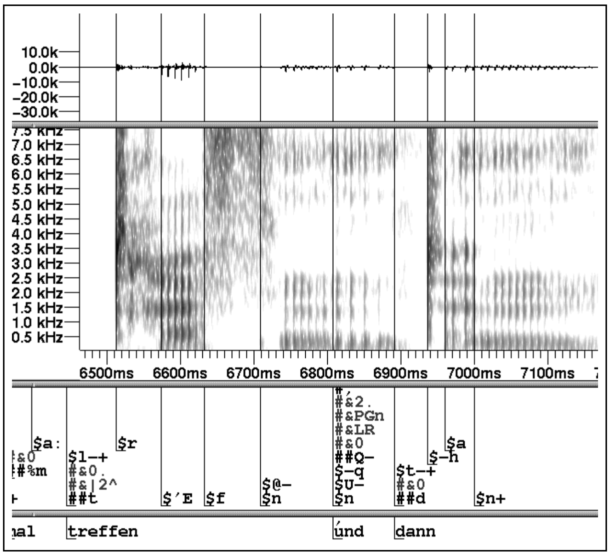


Figure 9: [n̰] for *und* in *treffen, und dann* (g075a004).

be related to both the onset of *und* and to the prosodic boundary between *treffen* and *und*. This is suggested by the continuation of glottalization into the vowel of *dann*, marking a turning point from a falling to a rising pitch across the boundary.

The manifestation of -*und*- in numbers such as *vierundzwanzigster* is completely different in that in 66 out of 68 cases the morphemic boundary before -*und*- is not marked by any non-modal glottal activity. Only in two cases is there a juncture produced by creak. Since there is no glottal stop in the canonic transcription of -*und*-, the label for laryngealization has not consistently been added where necessary, and it is not possible to automatically search the corpus for different combinations of glottal activity. But the results of the manual investigation of six speakers probably hold for the larger corpus as well. The interesting point is that when *und* is not positioned at syntactic boundaries but becomes melted into lexical units, the boundary-marking function of non-modal activity no longer applies, and glottalization is therefore not used.

4.2.2 Monophthongization

Since changes in vowel quality are not consistently indicated in the labelling, tokens of a relevant item were transcribed in a subset of the database (sessions g07a, g08a, and g09a) in order to obtain information on the realizations of diphthongs. Although it is not always easy to differentiate diphthongs from monophthongs plus formant transitions, particularly for short vowel durations, in most cases a clear auditory decision could be taken.

To talk about the indefinite article first, the vowel is reduced to a monophthong in 23 out of 82 tokens (28%). The monophthongs vary in quality between [e̞] and [ɑ̟]; frequently the more central [ɐ] is found. Only in one case is the monophthong as central as [ə]. Monophthongization does not occur in the accented indefinite articles; sentence-accent disallows the rather radical quality changes possible in the less salient unaccented tokens.

The diphthong in the definite article *der* derives from the vocalization of postvocalic /r/[17]. 59 out of 77 tokens (77%) showed monophthongs instead of diphthongs. The quality of the monophthongs lies in the area [ɐ]-[ə]-[ɛ][18].

The percentage of monophthongization is lower in the forms of *ein* than in *der*. The more frequent diphthongal production of the indefinite article appears to be the consequence of more distant starting and ending points of the diphthong in the indefinite articles in comparison to the diphthong in *der*.

4.2.3 Absence of segmentable vowels

4.2.3.1 Vowel elision Elision of word-final schwa is so common that it is already indicated in the orthography by an apostrophe. It

[17]This vocalization is obligatory in standard variants of German, and the canonical labels therefore symbolize diphthongs.

[18]For *die* the vowel quality almost exclusively lies between [i] and [ɪ]. There is only one occurrence of a form with schwa. This seems to differentiate *die*-productions from *der*-productions where schwa is common. Thus, bottom-up strategies on the part of the listener seem to suffice for these items, which facilitates the task of identifying different gender and case relations in the sentence.

Table 2: Vowel elision in unaccented indefinite articles. n_V indicates the number of tokens with a vowel, n_{V-} the number of tokens without a vowel.

item	n_V	n_{V-}	item	n_V	n_{V-}
ein	107	1	*'n*	3	65
eine	68	0	*'ne*	0	39
einem	54	0	*'nem*	1	1
einen	119	0	*'nen*	2	27
einer	11	0	*'ner*	0	5
eines	7	0	*'nes*	0	0
	366	1		6	137

is particularly frequent in the first person singular of *haben*, which has been transliterated as *hab'* (and labelled without schwa) in more than three quarters of the cases. In addition, two cases of *habe* have also been labelled as involving schwa elision, so that 76% of *habe/hab'* do not contain a schwa.

Schwa elision is even more frequent in closed final syllables, especially when schwa is surrounded by sonorants such as /n/. 98% of all tokens of *einen*, *Ihnen*, and *können* are produced without a vowel in the second syllable.

Turning to the elision of vowels other than schwa, the diphthong in the indefinite article is frequently affected. This has led to the orthographic forms *'n*, *'ne*, *'nem*, *'nen*, and *'ner* for *ein*, *eine*, *einem*, *einen*, and *einer*, respectively. The diphthong is elided in 27% of all unaccented tokens of the indefinite article (138 out of 510 cases, cf. table 2). Elision is particularly frequent for *ein/'n*, which often occurs in common sequences such as *'n bißchen*. On the other hand, the shortened genitive form *'nes* does not occur at all.

4.2.3.2 Non-linear correlates of vowels The preposition *zum* is a good example for non-linear aspects of speech. In a number of cases, there is no segmentable vowel portion, but the correlates of the vowel were perceived to still be present. Figure 10 shows a case where /u/ is contained in the velarization of the alveolar fricative and the

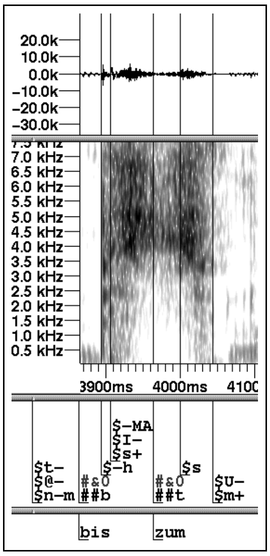


Figure 10: 'Elision' of /u/ in *bis zum [einundzwanzigsten]* (g315a008).

bilabial nasal, and in the labialization of the fricative: [t̪sʷˠmˠ] (for the realization of /t/ in this example cf. p. 336). In two other cases, the token was perceived as containing no exponents of the vowel whatsoever. These cases could be accounted for in linear terms, i.e. by elision of the vowel.

Sich and *mich* are also notable as to the realization of their nucleus. Both items can be produced with contoids only, e.g. [sʲʝ] (figure 11(a)) and [mʝç] (figure 11(b)). In these productions, the palatal fricative — and in the case of *sich* additionally the alveolar fricative — is an ideal environment for the nucleus to be a fricative as well. Since /ɪ/ is always produced as more peripheral and palatal [i] before [ç], the step towards a voiced palatal fricative is just a small one. In the terms of Articulatory Phonology, the fricative can be described as a blend resulting from a complete overlap of the components of palatal frication and vocoidness. These productions have

Table 3: Absolute frequencies of monophthongization and the glottal stop in unaccented forms of the indefinite article in the sessions g07a, g08a, and g09a.

		ʔV	
		+	−
Monophthongization	+	2	21
	−	15	44

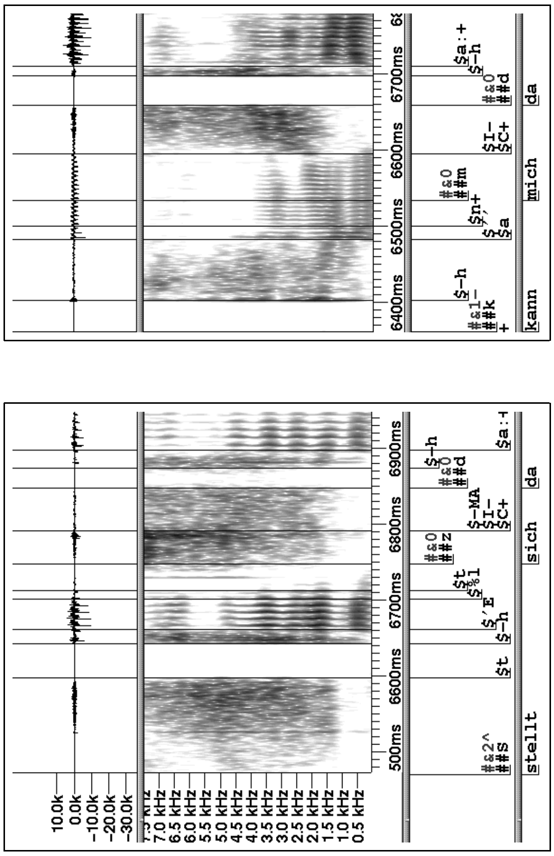
consequences for our view as to how syllables can be organized in German. Whereas syllabic sonorants have always been recognized (cf. section 4.1.1.3), syllabic sounds with frication have so far not been described for German.

5 Cooccurrences of different reduction phenomena

Table 3 displays the relationship between monophthongization and the absence of a glottal stop before the monophthong in unaccented tokens of the indefinite article (cf. p. 351 for more information on monophthongization in the indefinite article). While 25% of the diphthongal productions show a glottal stop, it is only 9% of the monophthongized realizations. There is thus a tendency of monophthongization to imply absence of a glottal stop.

The above relationship can be thought of as an example of a ‘reduction harmony’. It seems reasonable that more reduced forms on one hypo-hyper scale (in this case monophthongized vowels) often cooccur with more reduced forms on another scale (absence of glottal stop). In other words, speakers tend to avoid discrepancies between hypo- and hyper-productions.

We now turn to examples involving the ‘elision’ of vowels and the reduction of bilabials in alveolar contexts. Figure 12 displays a case where *bin* is contained in a strongly palatalized alveolar nasal: *damit bin ich* produced as [d̥a̠mid̞̃n̥nʲj]. It should be noted that this stretch does not sound like *damit nicht*, probably because of the long



(a) sich

(b) mich

Figure 11: Missing vocoid nuclei in *sich* (g071a018) and *mich* (g085a008).

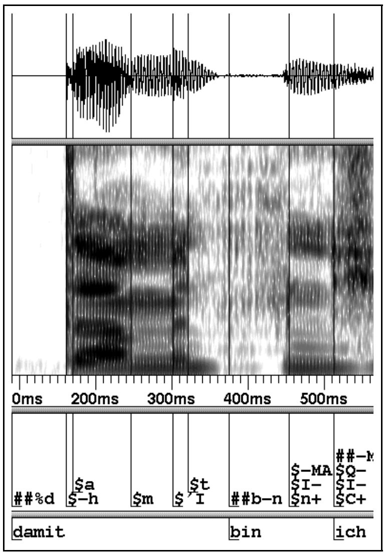


Figure 12: *bin* as a palatalized alveolar nasal in *damit bin ich* (g253a007).

portion of voiceless nasality.

Figure 13 presents a token of *zum* produced with an alveolar instead of a bilabial nasal in *bis zum siebzehnten* ([b̞̃ɪss̞snˠz̥-]). It should be noted that the example involves velarization of the alveolar nasal without a segmentable vowel portion (cf. section 4.2.3.2). Thus, the bilabial-to-alveolar assimilation seems to presuppose the ‘elision’ of the vowel which moves the surrounding alveolar elements closer together.

Another example of a ‘missing’ vowel portion facilitating the influence of an alveolar context on a bilabial articulation is *paßt mir nicht* in figure 14, where *mir* is contained in the bilabial component of a double articulation: [p̞ɑ̟stn͡mˑɪçt̞].

It is possible that the reduction of bilabials in alveolar contexts is restricted to function words in spontaneous speech. Contrary to the better known and more frequent assimilation of alveolars to bilabials,

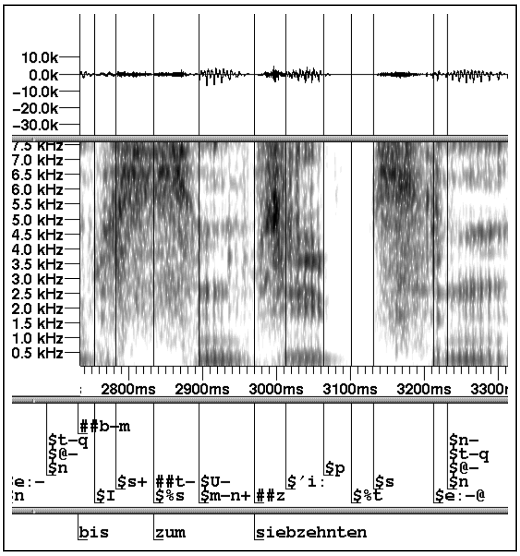


Figure 13: *bis zum siebzehnten* (g212a009).

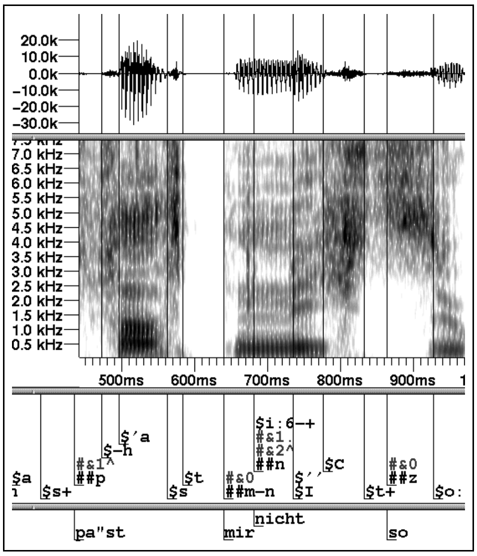


Figure 14: *[das] paßt mir nicht so [gut]* (g143a007).

Table 4: Absolute and relative frequencies (%) of types of vowel onset in unaccented productions of the indefinite articles for both spontaneous and read speech.

	`Q`	`Q -q`	`Q- -q`	`Q-`	total
spontaneous	6 (2)	77 (21)	210 (57)	73 (20)	366
read	31 (5)	94 (16)	286 (49)	176 (31)	587

it tends to require rather strong contractions involving 'elision' of vowels other than /ə/[19].

In both types of cooccurrence discussed above, the involved phenomena cooperate in creating articulatory stretches that are more homogeneous than the canonical productions. In the case of labials in alveolar environments, it seems justified to talk about alveolar prosodies (in the sense of Firth 1948) uniting canonically more diverse stretches.

6 Speaking style and reduction

When thinking of spontaneous speech, a typical association is that it involves more reduction than read speech. To throw more light on the topic of style and reduction, this section investigates realizations at the onset of articles. The first topic is glottal activity at the onset of indefinite articles (cf. section 4.2.1). The second area are the correlates of initial /d/ in definite articles (cf. section 4.1.1.1).

As to the glottal activity in indefinite articles, we expected to find more tokens without a glottal reflex for spontaneous speech.

Table 4 and figure 15 display the frequencies of the four glottal combinations for unaccented productions of the indefinite articles in spontaneous and read speech. In both speaking styles, the most frequent combination is creak (`Q- -q`) rather than the traditionally postulated glottal stop (`Q`; cf. also Kohler 1994). But whereas in read speech absence of any glottal reflex (`Q-`) is the next most frequent,

[19] However, there is also one token with an alveolar nasal for /b/ where *siebenundzwanzigsten bis* was produced as [zimʊ̃nsv̥a̠nn̥sij̥stn̥nə̃s] without elision of the vowel in *bis* (g212a007).

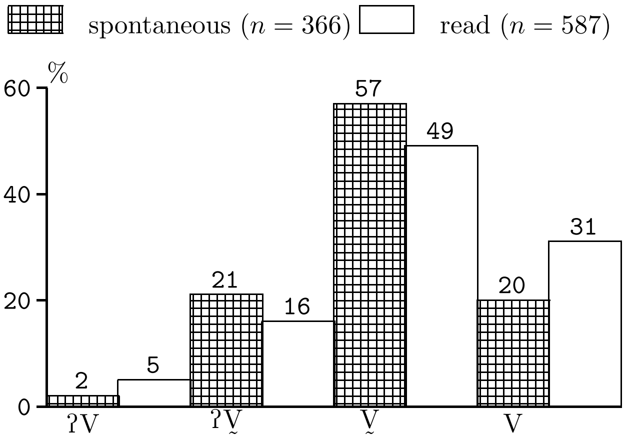


Figure 15: Relative frequencies (%) of types of vowel onset in unaccented productions of the indefinite articles for both spontaneous and read speech.

in spontaneous speech glottal stop with creak (`Q -q`) ranks second, closely followed by absence of any glottal reflex (`Q-`). The main difference between the two speaking styles is in cases without a glottal stop: creak (`Q- -q`) is more common in spontaneous than in read speech where absence of a glottal reflex (`Q-`) is more frequent. Summing the two combinations with `-q`, creak is again more common in spontaneous than in read speech (78 vs 65%). Cases with glottal stop (`Q` and `Q -q`) have roughly the same frequencies in spontaneous and read speech (23 vs 21%), so that for spontaneous speech a glottal reflex is marked in more cases.

A χ^2-test showed the difference to be significant at the 0.1% level. The higher incidence of a glottal reflex in spontaneous speech contradicts our expectation. An explanation lies in the high level of dysfluency phenomena in spontaneous speech: a glottal reflex is more likely after pauses, breathing, and other dysfluencies typical of spontaneous speech (cf. the article by Rodgers in this volume)[20]. By offering an environment full of dysfluencies, spontaneous speech

[20] One could argue that relative to the frequency of unfavourable contexts, the number of items without a glottal reflex is comparatively high in spontaneous speech. But the high frequency of dysfluencies in spontaneous speech and their

Table 5: Absolute and relative frequencies (%) of /d/ realizations in unaccented productions of the definite articles for both spontaneous and read speech.

	D -d	D	D-N	other	total
spontaneous	2104 (80)	269 (10)	146 (6)	96 (4)	2615
read	897 (84)	101 (10)	21 (2)	43 (4)	1063

thus disfavours the absence of a glottal reflex.

Turning to the second topic with regard to style and reduction, our hypothesis predicts that there are fewer released plosives and more reduced realizations in spontaneous than in read speech.

Table 5 and figure 16 display the frequencies of different types of production of *der*, *die*, *das*, *dem*, and *den* for spontaneous and read speech. The first pair of columns stand for the standard case in which there is a released alveolar plosive. The next pair refers to cases where no release has been labelled, i.e. mostly approximants. The frequency of nasalization is indicated in the third pair. Other less frequent productions are grouped together in the last two columns (e.g. deletion and fricativization, for the latter cf. p. 331).

To turn first to spontaneous speech, the production involving a plosive and its release is clearly most common (it occurs on average in 80% of the cases). The most frequent reduced form is approximation (10% of the tokens). An alveolar nasal occurs in 6% of the tokens.

In read speech, there are more standard productions involving a plosive and its release, and fewer nasals instead of alveolar plosives than in spontaneous speech. It should be noted that /d/ is slightly more often preceded by an alveolar nasal in the read corpus than in the spontaneous corpus (13 vs 12% of the tokens) so that more plosive nasalization in spontaneous speech is not due to a more frequent nasal context. The frequency of approximant productions is the same as in spontaneous speech. A χ^2-test proved the differences between spontaneous and read speech realizations of /d/ in the def-

absence in read speech is inherent to the speaking styles and thus of a different nature than statistical differences in the frequency of segmental contexts, e.g. the frequency of nasals preceding initial /d/ (see below).

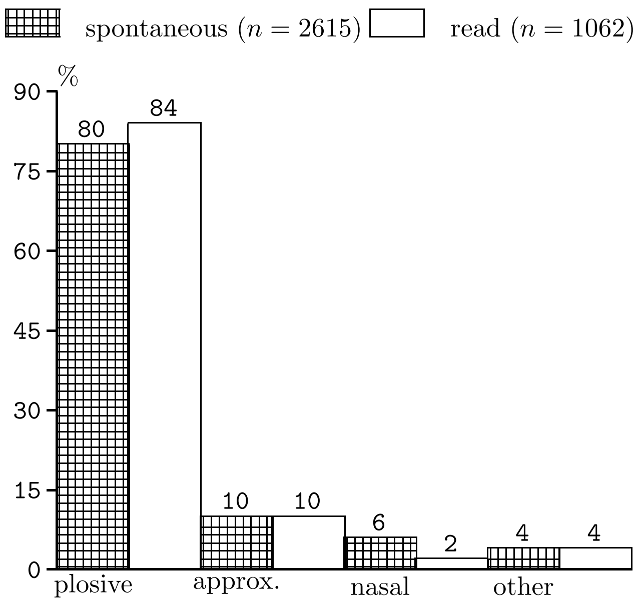


Figure 16: Relative frequencies (%) of different correlates of initial /d/ in unaccented *der*, *die*, *das*, *dem*, *den*.

inite articles to be highly significant ($p < 0.001$ because of large n). Our expectation has thus been confirmed: there are more reduced forms in spontaneous than in read speech.

We have shown that contrary to our expectation, there is more glottalization in spontaneous productions of the indefinite articles than in read productions because of the dysfluencies of spontaneous speech, which disfavour the stronger form of reduction (i.e. absence of glottal reflex) found more frequently in read speech. On the other hand, nasals for /d/ in the definite articles are more frequent in spontaneous than in read speech. There is thus no simple relation between speaking style and the frequency of connected speech phenomena of the kind 'spontaneous speech is necessarily more reduced than read speech'. In general, the range and degree of reduction phenomena is probably greater for spontaneous than for read speech because more information delivered by the situational context allows for less verbal redundancy. But this tendency can be ruled out by intervening variables inherent to the speaking styles.

7 Conclusions

The findings show the importance of corpora of spontaneous speech for a more exhaustive description of the phonetics of connected speech. The investigated function words display previously unknown phonetic exponents of phonological units, e.g. new forms of realization of /x/. Some exponents can have far-reaching consequences, e.g. for possible structures of phonetic syllables to include productions with syllabic fricatives.

Non-linear phenomena illustrate the tolerance of temporarily 'disordered' realizations of phonological units in speech. They point towards a type of phonology that does not stick to the mere concatenation of phonemes, however complex these units might be thought of to be structured internally in terms of features. Although phonetic exponents do not per se change phonological accounts, they always affect the phonology-phonetics interface. It seems more economical to incorporate non-linear aspects in phonological accounts than to exclude them from phonology, because this way the problem is only shifted to the transition from phonology to phonetics.

Many of the phenomena can be accounted for by overlap in terms of Articulatory Phonology (Browman and Goldstein 1992). For example, in [d̥ɐ̤ɐ̤m] for *da haben*, the glottal gesture for /h/ completely overlaps the preceding vowel. However, Articulatory Phonology runs into serious trouble when it comes to the replacement of one gesture by another, as happens in the case of plosive-related glottalization (Kohler 1992). Creating an interarytenoid opening for whispery voice as the correlate of /x/ is a further case of gestural reorganization where a supralaryngeal gesture is replaced by a laryngeal one. Another theory promising insight into the organisation of spontaneous speech is Firthian prosodic analysis. Firth (1948) regards glottal activity as a prosody rather than as a sound, which fits in well with the observed temporal freedom of the correlates of /h/ and /x/. In any case, the phonological account of the phonetics of spontaneous speech needs to be elaborated.

Besides being of theoretical interest for both phonetics and phonology, the results can also be used in teaching German as a foreign language. In an advanced stage, learners should be trained with reference to various degrees of hypo speech (Lindblom 1990) to improve their command of stylistic levels, especially in the perception of spontaneous speech.

A Frequencies of the items

Table 6: Absolute frequencies of the orthographic function word types in the spontaneous speech database (un = unaccented, ac = accented).

item	total	un	ac	item	total	un	ac
ab	118	93	25	*aber*	181	177	4
als	31	30	1	*also*	399	384	15
am	452	446	6	*an*	68	68	0
ans	1	1	0	*auch*	394	244	150
auf	119	118	1	*auf'm*	4	4	0
auf'n	1	1	0	*aufs*	2	2	0
aus	12	9	3	*bei*	300	298	2
beim	15	15	0	*bin*	186	160	26
bis	397	372	25	*da*	508	399	109
daher	4	3	1	*damit*	3	3	0
dann	802	712	90	*darf*	10	7	3
das	1186	1026	160	*daß*	175	170	5
deinen	1	1	0	*dem*	252	248	4
den	660	646	14	*denen*	5	4	1
denn	199	183	16	*der*	411	408	3
des	40	40	0	*dessen*	1	0	1
die	295	287	8	*dies*	1	0	1
diese	31	26	5	*diesem*	11	10	1
diesen	15	11	4	*dieser*	10	7	3
dieses	10	4	6	*doch*	133	108	25
du	2	2	0	*durch*	11	11	0
dürfte	4	3	1	*ein*	121	108	13
eine	80	68	12	*einem*	56	54	2
einen	134	119	15	*einer*	11	11	0
eines	7	7	0	*er*	1	1	0
es	278	278	0	*etwas*	31	26	5
falls	3	2	1	*für*	146	145	1
fürn	1	1	0	*fürs*	2	2	0

item	total	un	ac	item	total	un	ac
gegen	11	8	3	*gewesen*	2	1	1
geworden	1	1	0	*hab'*	159	133	26
habe	48	39	9	*haben*	150	120	30
halt	10	10	0	*hat*	10	9	1
hatt'	2	2	0	*hatte*	7	7	0
hatten	26	24	2	*hätt'*	19	13	6
hätte	27	16	11	*hätten*	37	18	19
hinter	3	2	1	*ich*	1360	1303	57
Ihnen	294	229	65	*Ihr*	2	1	1
Ihre	6	6	0	*Ihrem*	5	5	0
Ihren	6	5	1	*Ihrer*	3	3	0
im	198	198	0	*in*	285	283	2
in'n	5	5	0	*ins*	11	11	0
ist	510	494	16	*ja*	259	253	6
je	1	1	0	*kann*	138	100	38
konnte	1	1	0	*konnten*	2	2	0
können	160	111	49	*könnt'*	14	8	6
könnte	95	55	40	*könnten*	93	53	40
mal	280	277	3	*man*	33	33	0
mein	4	4	0	*meine*	12	11	1
meinem	7	6	1	*meinen*	18	13	5
meiner	1	1	0	*mich*	64	57	7
mir	475	408	67	*mit*	190	188	2
mit'm	1	1	0	*möcht'*	5	4	1
möchte	15	10	5	*möchten*	6	4	2
müssen	77	57	20	*muß*	63	50	13
müßt'	6	4	2	*müßte*	17	9	8
müßten	46	33	13	*'n*	68	68	0
na	18	18	0	*nach*	63	49	14
nachdem	1	1	0	*ne*	23	8	15
'ne	39	39	0	*'nem*	2	2	0
'nen	29	29	0	*'ner*	5	5	0
nicht	317	243	74	*noch*	394	382	12

item	total	un	ac	item	total	un	ac
nun	26	25	1	*ob*	28	25	3
obwohl	2	0	2	*oder*	179	160	19
's	154	154	0	*schon*	167	154	13
sei	1	1	0	*sein*	49	44	5
seinem	1	1	0	*sich*	71	71	0
sie	12	12	0	*Sie*	328	300	28
sind	75	68	7	*so*	166	137	29
soll	23	23	0	*sollen*	9	8	1
sollte	10	9	1	*sollten*	51	28	23
sondern	3	3	0	*sowas*	1	1	0
sowie	2	1	1	*über*	23	23	0
übern	1	1	0	*übers*	2	2	0
um	116	115	1	*und*	720	696	24
uns	203	200	3	*unser*	5	5	0
unsere	13	13	0	*unserem*	1	1	0
unseren	3	3	0	*unserer*	8	7	1
unseres	1	1	0	*unserm*	1	1	0
unsre	5	4	1	*unsrer*	3	3	0
unter	22	20	2	*vom*	158	156	2
von	85	81	4	*vor*	28	21	7
während	14	8	6	*wann*	65	40	25
war	21	20	1	*waren*	2	2	0
wär'	99	79	20	*wäre*	153	109	44
wären	4	4	0	*warum*	1	1	0
was	80	69	11	*weder*	1	1	0
wegen	7	6	1	*weil*	43	42	1
welche	3	3	0	*welchem*	1	1	0
welchen	1	1	0	*welcher*	6	6	0
welches	1	1	0	*wenn*	176	155	21
wer	1	1	0	*werd'*	7	5	2
werde	3	1	2	*werden*	28	25	3
wie	224	211	13	*wieder*	110	86	24
will	3	3	0	*wir*	1020	1017	3

item	total	un	ac	item	total	un	ac
wird	42	38	4	*wo*	16	13	3
wohl	26	26	0	*wollen*	51	38	13
wollt'	10	7	3	*wollte*	13	6	7
wollten	11	10	1	*womit*	1	1	0
würd'	101	89	12	*würde*	222	184	38
würden	23	21	2	*zu*	161	159	2
zum	151	149	2	*zur*	18	18	0
zwar	7	7	0	*zwischen*	41	40	1

In cases where an orthographic form stands for different words or grammatical categories, the following functions are designated:

da: modal particle, not spacial 'there'

der, *die*, *das*, *dem*, *den*: cf. footnote 6

*ein**: indefinite article (accented tokens can be numerals)

ja: modal particle, not 'yes' or introductory 'well'

sein: auxiliary verb, not possessive pronoun

B Glottal patterns for function words beginning with a vowel

Table 7: Relative frequencies of the combinations of glottal stop and creaky voice for unaccented items, with n_u indicating the number of unaccented tokens, n_V the number of unaccented tokens containing a vowel, and in the next four columns the percentages of the four glottal realizations with reference to n_V. * stands for all inflections of the item.

item	n_u	n_V	`Q-`	`Q- -q`	`Q`	`Q -q`
ab	93	93	22	44	14	20
aber	177	177	23	50	6	21
also	384	384	14	60	4	23
am	446	446	22	48	4	25
an	68	68	26	44	6	24

item	n_u	n_V	Q-	Q- -q	Q	Q -q
auch	244	244	20	62	2	16
auf	118	117	25	52	2	21
aus	9	9	0	100	0	0
*ein**	367	366	20	57	2	21
ein	108	107	23	53	4	20
eine	68	68	15	63	3	19
einem	54	54	28	63	0	9
einen	119	119	16	53	0	31
einer	11	11	18	73	0	9
eines	7	7	29	71	0	0
er	1	1	0	100	0	0
es	278	243	67	27	2	4
ich	1303	1240	62	28	2	8
Ihnen	229	229	34	55	0	10
*Ihr**	20	20	25	60	5	10
im	198	184	24	55	3	18
in	283	269	35	50	4	10
ist	494	449	43	48	2	6
oder	160	160	35	50	3	13
über	23	23	26	70	0	4
um	115	112	34	46	6	14
und	696	672	28	47	7	17
uns	200	185	49	44	2	5
*unser**	31	31	19	42	6	3
unter	20	20	10	65	10	15

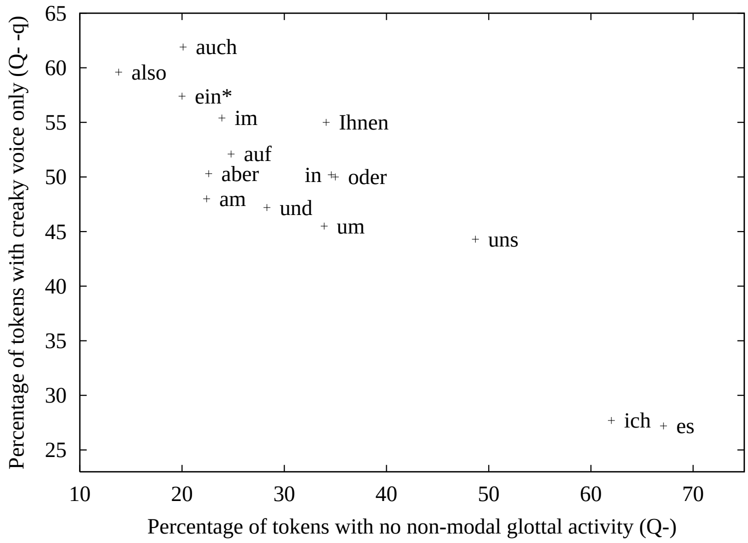


Figure 17: Percentages of the two types of vowel onset without glottal stop for items from table 7 with $n_V > 100$.

References

Browman, C. P. and L. Goldstein (1992). Articulatory phonology: An overview. *Phonetica 49*, 155–180.

Butcher, A. R. (1977). Coarticulation in intervocalic voiceless plosives and fricatives in connected speech. In A. R. Butcher, K. J. Kohler, and H. Künzel (Eds.), *Experimentelle Untersuchungen zur Koartikulation und Steuerung im Deutschen – Experimental Investigation into Coarticulation and Articulatory Control*, AIPUK 8, pp. 154–212.

DUDEN (1990). *Das Aussprachewörterbuch* (3 ed.). Mannheim: Dudenverlag.

Firth, J. R. (1948). Sounds and prosodies. *Transactions of the Philological Society*, 127–152.

Helgason, P. (1996). Lenition in German and Icelandic. In A. P. Simpson and M. Pätzold (Eds.), *Sound Patterns of Connected Speech: Description, Models, and Explanation*, AIPUK 31, pp. 211–218.

IPA (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

IPDS (1994). *The Kiel Corpus of Read Speech*, Volume 1, CD-ROM#1. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

IPDS (1995). *The Kiel Corpus of Spontaneous Speech*, Volume 1, CD-ROM#2. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

IPDS (1996). *The Kiel Corpus of Spontaneous Speech*, Volume 2, CD-ROM#3. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

IPDS (1997). *The Kiel Corpus of Spontaneous Speech*, Volume 3, CD-ROM#4. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

Kohler, K. J. (1979). Kommunikative Aspekte satzphonetischer Prozesse im Deutschen. In H. Vater (Ed.), *Phonologische Probleme des Deutschen*, Number 10 in Studien zur deutschen Grammatik, pp. 13–39. Tübingen: Narr.

Kohler, K. J. (1990a). Comment on German. *Journal of the International Phonetic Association 20*(2), 44–46.

Kohler, K. J. (1990b). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In W. J. Hardcastle and A. Marchal (Eds.), *Speech Production and Speech Modelling*, pp. 69–92. Dordrecht: Kluwer.

Kohler, K. J. (1992). Gestural reorganization in connected speech: A functional viewpoint on 'Articulatory Phonology'. *Phonetica 49*, 205–211. Commentary on Browman and Goldstein's target article in the same volume.

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

Kohler, K. J. (1995). *Einführung in die Phonetik des Deutschen* (2nd ed.). Grundlagen der Germanistik 20. Berlin: Erich Schmidt.

Kohler, K. J. (1996). The phonetic realization of /ə/ syllables in German. In A. P. Simpson and M. Pätzold (Eds.), *Sound Patterns of Connected Speech: Description, Models, and Explanation*, AIPUK 31, pp. 11–14.

Kohler, K. J. (1999). Plosive-related glottalization phenomena in read and spontaneous speech: A stød in german? In K. J. Kohler (Ed.), *Phrase-level Phonetics and Phonology of German*, AIPUK 34.

Kohler, K. J., M. Pätzold, and A. P. Simpson (1995). *From scenario to segment: the controlled elicitation, transcription, segmentation and labelling of spontaneous speech.* AIPUK 29.

Kohler, K. J., M. Pätzold, and A. P. Simpson (1997). From the acoustic data collection to a labelled speech data bank of spoken Standard German. In A. P. Simpson, K. J. Kohler, and T. Rettstadt (Eds.), *The Kiel Corpus of Read/Spontaneous Speech — Acoustic data base, processing tools and analysis results*, AIPUK 32, pp. 1–29.

Krech, E.-M., E. Kurka, H. Stelzig, E. Stock, U. Stötzer, and R. Teske (Eds.) (1982). *Großes Wörterbuch der deutschen Aussprache.* Leipzig: VEB Bibliographisches Institut Leipzig.

Laver, J. (1980). *The phonetic description of voice quality.* Cambridge: Cambridge University Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (Eds.), *Speech Production and Speech Modelling.*, pp. 403–439. Dordrecht: Kluwer Academic Publishers.

Meinhold, G. (1973). *Deutsche Standardaussprache. Lautschwächungen und Formstufen.* Wissenschaftliche Beiträge der Friedrich-Schiller-Universität Jena. Jena: Friedrich-Schiller-Universität.

Pätzold, M. (1997). *KielDat* – data bank utilities for the *Kiel Corpus.* In A. P. Simpson, K. J. Kohler, and T. Rettstadt (Eds.), *The Kiel Corpus of Read/Spontaneous Speech — Acoustic data base, processing tools and analysis results*, AIPUK 32, pp. 117–126.

Rodgers, J. E. J. (1999). Accent- and prosody-dependent glottalization of initial vowels in content and function words in read and spontaneous speech. In K. J. Kohler (Ed.), *Phrase-level Phonetics and Phonology of German*, AIPUK 34.

Thon, W. and W. van Dommelen (1992). PHONDAT 90: Rechnerverarbeitbare Sprachaufnahmen eines umfangreichen Korpus des Deutschen. In K. J. Kohler (Ed.), *Phonetisch-akustische Datenbasis des Hochdeutschen: Kieler Arbeiten zu den PHONDAT-Projekten 1989–1992*, AIPUK 26, pp. 41–79.