Some non-sequential phenomena in German function words

Thomas Wesener
Institute of Phonetics and Digital Speech Processing,
University of Kiel
tw@ipds.uni-kiel.de

Labellers of connected speech corpora encounter a number of cases which are difficult to handle with a linear sequence of segmental labels. Based on material from a corpus of German unscripted speech, this paper presents non-sequential aspects of laryngeal activity for /x/ and /h/, and of supralaryngeal articulation in vocoid exponents of postvocalic /r/. Comparative data for /r/ from a corpus of German read speech indicate that differences across speaking styles may be independent of durational variation. It is discussed how models of the phonology-phonetics interface and Prosodic Analysis can account for the findings.

1 Introduction

Corpora of connected speech contain a number of cases where a sequential ordering of phonological units derived from isolated words cannot be mapped unchanged onto segments of the speech signal. Barry & Fourcin (1992) present a case where plosive-related creaky voice for /t/ in *point zero* occurs at the beginning instead of at the end of the nasal portion. With reference to the phonological sequence /nt/ derived from isolated citation forms, its phonetic correlates can be called non-sequential.

The *Kiel Corpus of Spontaneous Speech* (IPDS 1995–1997) reveals several phenomena involving temporal 'disalignment' in German, among them the temporal flexibility of creaky voice. Kohler (2001) discusses productions of *könnten* 'could' where, as in Barry & Fourcin's example from English, creaky voice frequently is the only exponent of /t/. Besides being located at its sequential place ([-nn̰n]), creak can be shifted to the beginning ([-n̰n]) or to the end of the nasal stretch ([-nn̰]); alternatively, it can cover the whole nasal portion ([-n̰n̰]). Whereas the shifting of creak is still limited to the surrounding nasals in the above examples, more radical deviations from the canonical alignment may also occur. In *achtzehnten elften* 'eighteenth November' (g411a004[1]), creak related to the inter nasal /t/ in *achtzehnten* occurs shortly before a functionally different instance of creak that marks the initial vowel of *elften*. The

[1] References to the *Kiel Corpus of Spontaneous Speech* are structured as follows: g411a is the first dialogue in recording session g41a, 004 the turn number within this dialogue. For each referenced example, speech output and graphic representation can be found at the URL http://www.ipds.uni-kiel.de/publikationen/audiobsp.en.html.

DOI:10.1017/S0025100301001025 Journal of the International Phonetic Association (2001) 31/1

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

plosive-related creak in *achtzehnten* is shifted to the left out of the nasal stretch onto the preceding vowel: [-ḛnn]. At the same time, creak corresponding to the initial vowel in *elften* is produced at the end rather than at the beginning of that vowel and thus shifted to the right. This shifting in opposite directions increases the temporal separation of the two functionally different instances of creak in this example, possibly in order to give each of them more perceptual weight.

This paper looks at a number of other non-sequential phenomena in the *Kiel Corpus of Spontaneous Speech*: laryngeal correlates of German /x/ and /h/ and supralaryngeal activity in connection with /r/. The focus will lie on function words, which are typically unaccented and therefore particularly prone to productions that differ substantially from citation forms. In addition, certain function words occur with great frequency, allowing for quantitative statements even in unscripted speech with its high variability.

Items such as *doch* 'yet' and *auch* 'too' show non-modal glottal activity as the only correlate of /x/, the dorsal stricture just being one of open approximation. These non-modal phonation types again show a considerable degree of temporal flexibility. The frequently occurring phrase *da haben* 'there have' is an ideal combination to observe the temporal flexibility of breathy voice in connection with /h/.

Non-sequentiality is not limited to laryngeal activity. A prominent supralaryngeal issue is the vocoid correlates of postvocalic /r/. The phonological combination of vowel plus /r/ does not always correspond in its phonetics to a diphthong ending in a raised open central vocoid for /r/; /r/ can be coded throughout the vocoid stretch by a monophthongal quality that is more open and central than the corresponding /r/-less vowel (Simpson 1998). In this paper, formant patterns for *wir* 'we' are set against those for monophthongal *wie* 'how'.

In order to shed some light on the topic of speaking style, results for /r/ are compared with findings from the *Kiel Corpus of Read Speech* (IPDS 1994). Finally, the results are discussed in their relation to models of the phonology-phonetics interface (Articulatory Phonology, Window Model of Coarticulation) and Firthian Prosodic Analysis.

2 Method

The data derive from 117 dialogues between 42 speakers that were elicited within an appointment-making scenario and published as the *Kiel Corpus of Spontaneous Speech*. It contains the speech signals and corresponding files with segmental labels, plus suprasegmental labelling for a subset of the signals (e.g. information on sentence accent and intonational patterns). The speakers are 20 to 60 years old, with an average age of 32; the majority are students or academics and come from the north-western part of Germany (Schleswig-Holstein, Lower Saxony and Hamburg). Most speak a variety of German that cannot be easily ascribed to a specific region and that may be called Northern Standard German.

Data from the *Kiel Corpus of Read Speech* allow for a comparison of read and unscripted speech. The corpus features 53 speakers (27 male, 26 female) different from those of the unscripted corpus, but comparable as to age, education, and language variety; for a more detailed description of the composition of the read corpus see Thon & van Dommelen (1992).

The *KielDat* utility (Pätzold 1997) was used to create databases of these corpora (labelling version: January 2000). One database contains all segmentally as well as suprasegmentally labelled dialogues,[2] a second the complete unscripted material,

[2] Complete sessions g07a, g08a, g09a, g14a, g19a, g21a, g25a, and g31a; isolated dialogues g202a, g274a, g287a, g297a, and g306a.

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

comprising in addition the dialogues with segmental, but without suprasegmental labelling,[3] and a third the complete read corpus.

Items of interest were retrieved from the database, and scripts were used to give an overview of phenomena captured by the labelling. Some cases where a phonological unit is only represented by correlates which do not lend themselves to sequential labelling, e.g. by a secondary articulation, are marked with the label -MA followed by the label of the relevant phonological unit which is marked as deleted (examples can be found in figures 2(b) and 2(c) below, where -MA precedes h- and a-, respectively). These cases can easily be retrieved. However, not all cases of this kind have been marked with -MA because the non-sequential correlates were not always detected by the labellers, which means that the deletion of a relevant label does not necessarily imply the absence of any phonetic correlate. In order to capture phenomena which had not been labelled, the automatically retrieved data were complemented by an auditory investigation of the relevant sequences.

An acoustic analysis was carried out for the vocoid correlates of /r/ using the tools *klara* and *ksort* (Willems 1987, Scheffers & Simpson 1995) to estimate formant frequencies (window size 30 ms, Hamming window, pre-emphasis factor 0.95, LPC order 16). To exclude durational artefacts, only unaccented productions in the suprasegmentally labelled database without hesitational lengthening were investigated.

The frequency values were converted to perceptual units since Hz-values do not indicate whether changes are relevant in communication and in the judgment of quality changes by phoneticians. Thus, a value change of 50 Hz in an F2 of 1500 Hz may be perceptually negligible, but the same change for an F1 of 500 Hz is much more prominent due to the non-linear organization of frequency perception; both changes, however, look the same when depicted in a linear graph. The scale of ERB rate according to Moore (1997) was adopted to reflect the perceptual significance of value changes:

$$\mathrm{E} = 21.4\ log_{10}(4.37F + 1).$$

ERBs are the equivalent rectangular bandwidths of auditory filters estimated by using the so-called notched-noise method (Patterson 1976). The scale of ERB rate has a higher resolution than the Bark scale (Zwicker 1961) for frequencies below 500 Hz, which is of relevance mainly to F0 and F1. According to Traunmüller (1990), the Bark scale is a measure of tonotopic position, whereas ERB rate is a measure of frequency resolution and therefore seems to be more appropriate for the present study.

3 Results

3.1 Word-final /x/

There is a long tradition of publications on the phonological status of the dorsal fricatives in German. The literature is dominated by the discussion of whether the observed [ç] and [x] belong to one or to two phonological units. The basis of phonetic observations for these phonological discussions is more or less limited to velar fricatives occurring after open and back vowels, and palatal fricatives occurring after all other vowels, after the sonorants /n/, /l/ and /r/, and morpheme-initially.

Kohler (1990) broadens the basis of phonetic observations for phonological interpretations by pointing out that besides the voiceless palatal and velar fricatives, uvular fricatives also occur, and are indeed more frequent than velar fricatives. They

[3] Complete sessions g10a, g11a, g12a, g36a, g37a, g38a, g41a, and g42a.

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

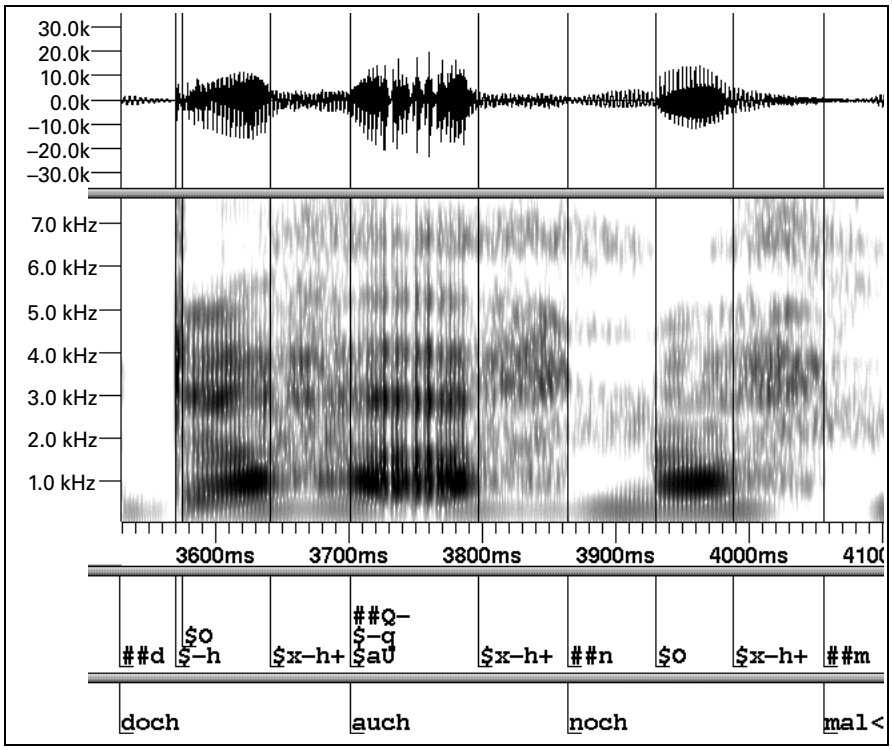


Figure 1 Non-modal phonation types as the only correlate of /x/ in *doch auch noch m*[*al*] 'yet once more' (g103a005).

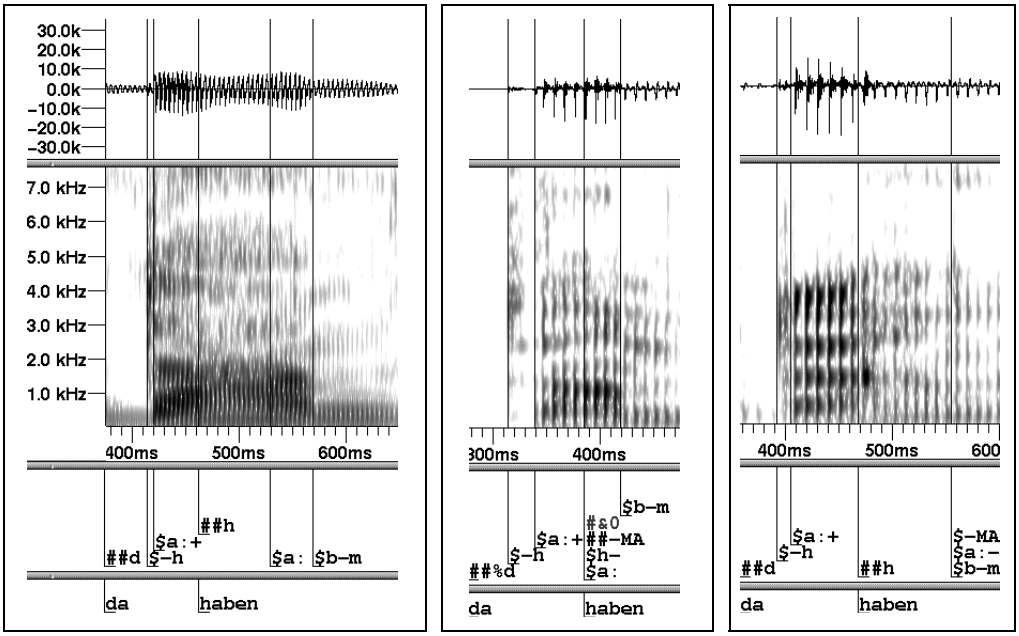
are typical after open and mid-open vowels (e.g. in *nach* 'to', *noch* 'still' and *doch* 'yet'), while velar fricatives can occur after close back vowels or monophthongs ending in that vowel quality (e.g. in *auch* 'also').

Inspection of the *Kiel Corpus of Spontaneous Speech*, however, reveals that /x/ has even more phonetic exponents than the three fricatives described above. In places where a velar or uvular fricative would occur in citation form pronunciations, auditory and visual observation shows that there is no dorsal frication in a number of cases (labelled as x–). In other productions, the labelling indicates that although no dorsal activity is observed, phonatory correlates of /x/ are present (x–h). These correlates are also found in some of the cases where the fricative is marked as deleted.

Voicelessness as the only exponent of the phonological unit /x/ can be thought of as a residual gesture once the dorsal stricture is one of open rather than close approximation. In voiced contexts, the non-modal phonation types of breathy voice and whispery voice are also found as the only correlates of /x/. Figure 1 shows an example with several occurrences of /x/ in *doch auch noch m*[*al*] 'yet once more', none of which involving dorsal frication: [dɔ̤ɦa̤ɦn̤ɦm̥-]. Whispery voice is symbolized here with [ɦ] plus the ExtIPA diacritic for whisper [̥] (cf. IPA 1999). This symbolization is preferred to a vocoid symbol plus [̥] because the quality of the whispery vocoids is difficult to determine. Shifting the diacritic [̥] indicates that the end of the whispery portion is already devoiced.

Although the first two instances of /x/ have sequentially ordered correlates, the first seems to be related to the temporal 'disalignment' of the following vowel-related creak

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press



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

Figure 2 Temporal flexibility of breathy voice for /h/ in *da haben* 'there have'. Left: framed by modal voice (g254a013). Middle: shifted to the left (g071a019). Right: shifted to the right (g425a006).

in *auch*: shifting creak to the right and thus away from the whispery voiced stretch for /x/ may serve to separate the two non-modal phonation types and enhance their perceptibility (cf. the production of *achtzehnten elften* in section 1 above).

The third instance of /x/ (in *noch mal*) involves considerable overlap of non-modal phonation with the following nasal. While the overlap is only partial in this example, non-modal phonation can also be completely shifted onto the following nasal, e.g. [nɔ̰̤̈m-] (g105a016). In addition, non-modal phonation can be superimposed on longer stretches between productions of /x/. *Auch noch mal* was thus produced as [q̰fin̰ɔ̰fiṃ-] (g415a004).

3.2 Word-initial /h/

Whereas voicelessness is an exponent of German /h/ after voiceless articulations and pauses, breathy voice occurs in voiced contexts. (This pattern is similar to other Germanic languages, cf. Engstrand & Nordstrand (1984) for Swedish.) A typical production of *da haben* 'there have' has a linearly segmentable portion of breathy voice (cf. figure 2(a) above). Breathy voice, however, can also be shifted, giving rise to non-sequential productions. The label -MA was used in a number of cases 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*: [d̰ɒ̤̈m] (cf. figure 2(b)). The correlates of /h/ can also be shifted to the second part of the vowel portion, i.e. the vowel of *haben*: [d̰ɒ̤̈m] (cf. figure 2(c)).

Non-sequential productions do not result from shorter absolute durations, but can be as long as sequential ones. The portion of the signal corresponding to /aːhaː/ in the

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

non-sequential example of figure 2(c) measures about 150 ms and thus is as long as the sequential example of figure 2(a).

3.3 /r/-diphthong in *wir*

Syllable-final postvocalic /r/ in Northern Standard German has vocoid rather than contoid correlates. In this section, the acoustic properties of these vocoids are investigated, taking the frequent personal pronoun *wir* 'we' as an example. Regarding a citation form pronunciation [v̥ɪɐ̯] as the point of departure, one would expect to find diphthongal formant movements, with a rising F1 and a falling F2 marking the transition from a close front vowel to an open central vowel.

In a first step, all unaccented tokens of *wir* without hesitational lengthening were grouped together, and F1 and F2 were automatically estimated at five points in time (10, 30, 50, 70 and 90 percent of vowel duration). The trajectories of monophthongal *wie* 'how' were measured in the same way for comparison.

Contrary to expectation, the formants in *wir* are rather flat and almost run in parallel with those in *wie* (cf. figure 3). The main difference is a lower F2 and a higher F1 for *wir*, so that the formants run within those for *wie*.

Although F1 and F2 are rather flat for *wir*, F1 shows a slight upward movement with a maximum at about 70% of vowel duration (cf. figure 3). This movement may represent a trace of diphthongality, and it could be tested in a perception experiment whether a slight upward move in F1 increases the *wir* percept for formant patterns intermediate between those for *wir* and *wie*. Most probably, however, the bulge is the product of averaging out trajectories of different shapes: while the majority of the productions displays a flat F1, some more diphthongal versions provide the slight upward tendency in the averaged curve.

In order to gain insight into the influence of duration on the formant trajectories, the productions of *wir* were subdivided into durational classes with steps of 30 ms. Figure 4(a) shows the trajectories for three durational classes in unscripted speech. It can be seen that the longer the vowel, the greater the difference of initial and final values.

Figure 4(b) displays the same durational classes for read speech. Initial values for F2 and final values for F1 are higher than for the same classes in unscripted speech. In general, read speech shows more 'typical' diphthongal formant movements, also for short durations. Although there are differences in class size and number of speakers, the mean duration of the vowels in the 31–60 ms class in read speech is identical to the one in unscripted speech (45 ms). Different formant patterns across styles do not seem to be due to vocoid duration.

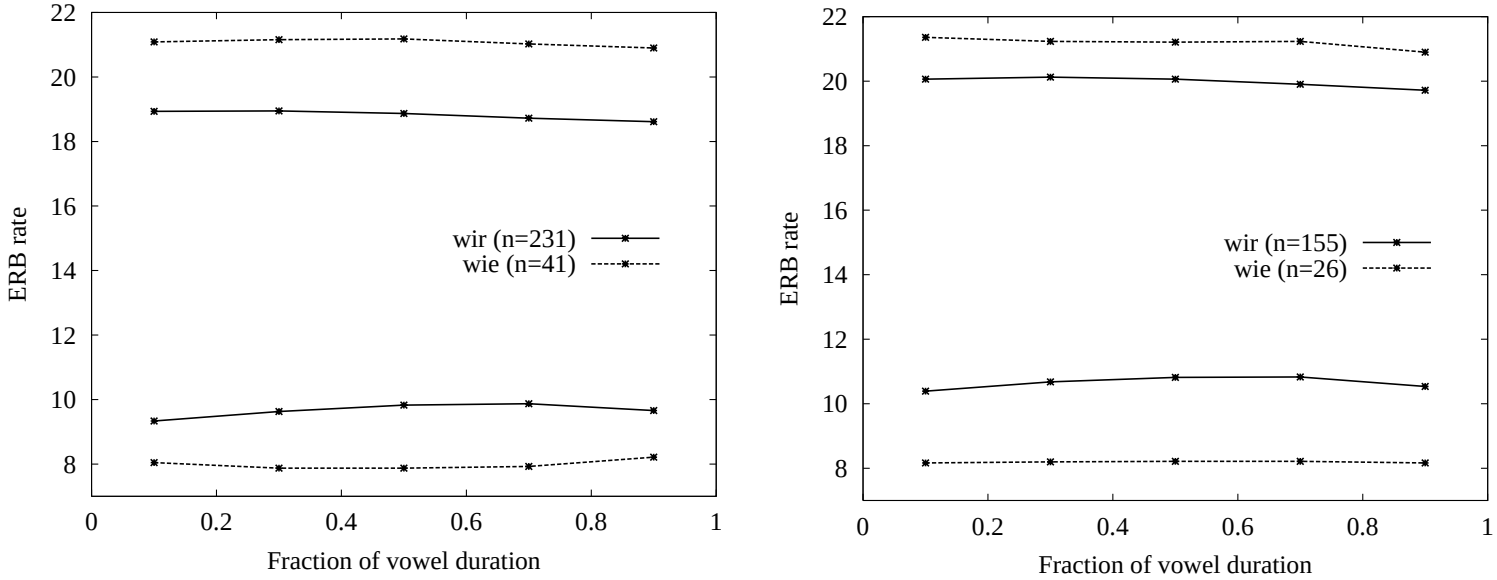
4 Discussion

Work with corpora of connected speech can increase the knowledge of phenomena related to certain phonological units, such as /x/, /h/ and /r/ in this paper. The impact of these findings, however, goes beyond the level of description; they call for accounts from models of the phonology-phonetics interface and challenge the sequential organization of the phonological input to these models.

Componential approaches seem promising when accounting for the phonetic data reported here. Articulatory Phonology (AP, Browman & Goldstein 1992) can adduce the two mechanisms of gestural overlap and reduced gestural magnitude to account for many of the observations in connection with /x/ and /h/. But AP cannot explain 'long' non-sequential productions by referring to increased overlap. Since gestures have

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press



(a) male **(b)** female

Figure 3 Trajectories of F1 and F2 in the vowels of *wir* and *wie* in unscripted speech (male and female speakers).

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

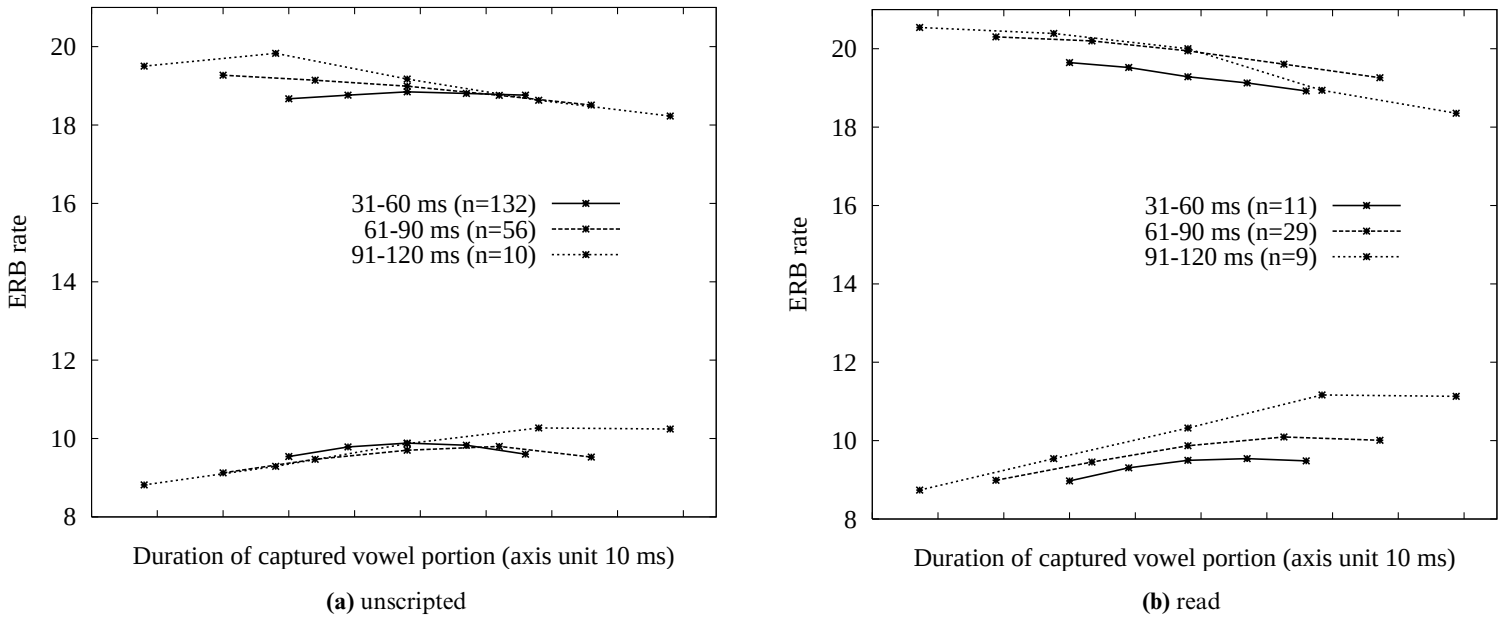


Figure 4 Trajectories of F1 and F2 in the vowel of *wir* for different durational classes in unscripted and read speech (male speakers). The curves are centred around the vowel midpoints.

intrinsic durations in the model, overlap, other variables remaining constant, leads to shorter durations.

Although the sliding of gestures allows for more temporal flexibility than traditional 'beads on a string' phonologies, AP retains a fundamental sequentiality: the computational model that generates the gestural score is fed with sequential phonological strings, and the gestures are phased according to these sequential specifications. The model does not account for the shifting of phonetic correlates beyond immediately neighbouring sounds, as in the example of *achtzehnten*.

The Window Model of Coarticulation (WM, Keating 1990) takes sequential allophonic feature specifications as its input and selects the possible range of physical values, i.e. the window, corresponding to each segmental feature. The behaviour of a given articulator during an utterance is then modelled by the interpolation across the windows of adjacent segments. The model is devised for contextual phenomena in a linear sequence of segments: 'Coarticulation refers to articulatory overlap between neighboring segments, which results in segments generally appearing assimilated to their contexts' (Keating 1990: 452). A window is attributed to each segmental slot, and the temporal alignment of windows seems to be fixed. It is not clear how the model can deal with productions that display all exponents of different phonological units simultaneously, and that do not show a sequential organization of phonetic correlates any longer.

Whereas AP and WM are models of the phonology-phonetics interface ultimately getting their input in the form of sequential phonological specifications, Prosodic Analysis (PA) emphasizes the fundamental phonological relevance of phenomena beyond sequences of contrastive sounds derived at the word level. PA's concept of prosodies as not being attributed to specific slots within sequential phonological strings may be useful for capturing non-sequential aspects of speech.

Firth (1948) regards vowel-initial glottalization in German as a prosody marking the junction between lexical items. In a similar vein, he interprets breathiness in connection with English /h/ as a prosodic signal of initiality. He notes that 'the aitchiness, aitchification, or breathiness of sounds and syllables, and similarly their creakiness or ''glottalization'' are more often than not features of the whole syllable or set of syllables' (Firth 1948: 146).

In the corpus investigated, productions are indeed found where non-modal glottal activity is not limited to one segment or to a segment boundary, but spreads over longer stretches (cf. *auch noch mal* in section 3.1 above). Furthermore, we have seen examples with limited glottal activity that is shifted within or across syllables. In both cases, glottal activity is not localized at a certain point within sequences of phonological units. This temporal flexibility of glottal activity seems to undermine its junctural role at first sight. But although it requires a certain temporal tolerance on the part of the hearer, we have seen cases where the shifting may serve to enhance the perceptibility of two functionally different glottal events by separating them.

Why are non-modal phonation types prominent candidates for non-sequential productions? One reason probably is that laryngeal activity is comparatively independent from supralaryngeal articulations. Articulatory interpretations, however, do not imply a mechanical necessity of certain forms, they only state why it is plausible that such forms occur. Articulation is influenced by mechanical constraints, but it is mainly determined by communicative and social factors that prescribe, allow or prohibit certain phenomena.

In a cross-language perspective, temporal flexibility of glottal correlates for /x/ and /h/ is probably not exclusive to German. The variation of dorsal fricatives with breathy voice and breath in German is found as an alternation in Czech voicing assimilation (Dankovičová 1999), and it seems worthwhile to investigate whether non-sequential productions of the glottal correlates are found in Czech. /h/-correlates in English and

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

Swedish also seem promising areas to extend our knowledge on non-sequential aspects of speech.

Procedural interpretations that derive 'reduced' productions from underlying phonological forms by manipulating segmental strings via a set of rules bear the danger of missing important aspects of the material investigated (cf. Simpson 1992). From a procedural point of view, one might postulate that the vowel preceding /r/ is 'elided' in monophthongal productions of *wir*: [ɪɐ] → [ɐ]. Both phonological units, however, are phonetically coded throughout the monophthong: F1 is intermediate between the endpoints of the diphthongal productions (cf. figure 4(a)). A declarative understanding of concrete phonetic properties as direct exponents of abstract phonological units seems more appropriate to account for the data.

Vowel–/r/ combinations are often said to involve /r/-'vocalization', again implying a procedural derivation. In items like *wir*, however, only vocoid productions of /r/ are found. Even if there is an inflectional paradigm with an alternation between vocoid and contoid correlates of /r/, as e.g. in *ihr* 'her (m./n. sg.)' vs *ihre* 'her (f. sg.)', one has to adduce language history, reflected in the orthography, to justify regarding the contoid as primary and the vocoid as derived. It is doubtful, however, that this diachronic argument is relevant in a synchronic account of speech patterns. Rather than being a 'vocalized' realization of consonantal /r/, the vocoid broadly transcribed as [ɐ] is the exponent of /r/, the latter being an abstract phonological unit and neither a contoid nor a vocoid.

A common belief is that 'reduction', and style in general, is the result of increased speech rate. This assumption led to terms such as 'lento' vs 'allegro style' and 'fast speech rules' (Dressler et al. 1972), and it is also implicitly adopted in Articulatory Phonology. Since unscripted speech is often supposed to be faster than read speech, it is expected to display more radical deviations from isolated word forms. The findings for /r/, however, show that it is important to disentangle style and speech rate. Similar durational patterns across speaking styles can be connected with different productions. In these cases, the phonetic coding is determined by style, not by duration, i.e. by communicative function and not by the mechanics of the vocal tract.

Acknowledgements

Part of the work reported here was carried out within the project 'Sound Patterns of German Spontaneous Speech' funded by the German Research Council DFG (grant Ko 331/22–1). I would like to thank the project supervisor Klaus Kohler for the discussion of a number of topics raised here. I am also grateful to an anonymous reviewer, Adrian Simpson and Michel Scheffers for valuable comments on earlier versions of this paper. All responsibility lies of course with me.

References

Barry, W. J. & Fourcin, A. J. (1992). Levels of labelling. *Computer Speech and Language* **6**, 1–14.

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

Dankovičová, J. (1999). Illustrations of the IPA: Czech. In IPA, 70–73.

Dressler, W., Fasching, P., Chromec, E., Wintersberger, W., Leodolter, R., Stark, H., Groll, G., Reinhart, J. & Pohl, H. D. (1972). Phonologische Schnellsprechregeln in der Wiener Umgangssprache. *Wiener Linguistische Gazette* **1**, 1–29.

Engstrand, O. & Nordstrand, L. (1984). A new phonetic interpretation of /h/ in Swedish. *Journal of Phonetics* **12**, 195–206.

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

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

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press

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

IPDS (1995–1997). *The Kiel Corpus of Spontaneous Speech*, vols. 1–3, CD-ROM#2–4. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

KEATING, P. A. (1990). The window model of coarticulation: articulatory evidence. In Kingston, J. & Beckman, M. E. (eds.), *Papers in Laboratory Phonology* I: *Between the Grammar and Physics of Speech*, 451–470. Cambridge: Cambridge University Press.

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

KOHLER, K. J. (2001). Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German? In Grønnum, N. & Rischel, J. (eds.), *To Honour Eli Fischer-Jørgensen. Festschrift on the Occasion of her 90th Birthday, February 11th, 2001* (Travaux du Cercle Linguistique de Copenhague **31**), 174–211. Copenhagen: C. A. Reitzel.

MOORE, B. C. J. (1997). Aspects of auditory processing related to speech perception. In Hardcastle, W. J. & Laver, J. (eds.), *The Handbook of Phonetic Sciences*, 539–565. Oxford & Cambridge, MA: Blackwell.

PATTERSON, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America* **59**, 640–654.

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

SCHEFFERS, M. T. M. & SIMPSON, A. P. (1995). LACS: Label assisted copy synthesis. In *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden, 13–19 August, 1995*, vol. 2, 346–349.

SIMPSON, A. P. (1992). Casual speech rules and what the phonology of connected speech might really be like. *Linguistics* **30**, 535–548.

SIMPSON, A. P. (1998). Accounting for the phonetics of German *r* without processes. In Pompino-Marschall, B. & Mooshammer, C. (eds.), *Papers of the Conference on 'The Word as a Phonetic Unit', ZAS, Berlin, October 22–23, 1997* (ZAS Working Papers in Linguistics **11**), 91–104.

THON, W. & VAN DOMMELEN, W. (1992). PHONDAT 90: Rechnerverarbeitbare Sprachaufnahmen eines umfangreichen Korpus des Deutschen. In Kohler, K. J. (ed.), *Phonetisch-akustische Datenbasis des Hochdeutschen: Kieler Arbeiten zu den PHONDAT-Projekten 1989–1992*, *AIPUK* **26**, 41–79.

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

WILLEMS, L. F. (1987). Robust formant analysis for speech synthesis applications. *Proceedings of the European Conference on Speech Technology, Edinburgh, September 1987*, vol. 1, 250–253.

ZWICKER, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America* **33**, 248.

https://doi.org/10.1017/S0025100301001025 Published online by Cambridge University Press