

# Prosodically-conditioned variation in the realization of domain-final stops and voicing assimilation of domain-initial fricatives in German

Claudia Kuzla

Max Planck Institute for Psycholinguistics  
Postbus 310, 6500 AH Nijmegen, The Netherlands  
E-mail: Claudia.Kuzla@mpi.nl

## ABSTRACT

This study investigates the realization of stop-fricative sequences across prosodic boundaries in German. According to phonological descriptions of German, voiced fricatives following voiceless obstruents undergo assimilatory devoicing. Results from acoustic analysis show that this process is gradient rather than categorical, and sensitive to prosodic structure, showing a larger extent of assimilation across smaller boundaries. In addition, also in the temporal patterns of these sequences, evidence of prosodic structure was found, with both the domain-final stop and the domain-initial fricatives being longer at larger boundaries.

## 1. INTRODUCTION

Prosody is often used to refer to the hierarchically organized structure of speech that constrains various phonological and phonetic processes. For example, some phonological and phonetic rules (e.g., *liaison* in French, *raddoppiamento sintattico* in Italian, *linking-r* in British English, *Lenis Stop Voicing Rule* in Korean) are known to apply across a lower prosodic boundary (e.g., the Prosodic Word boundary), but to be blocked by a higher prosodic boundary (e.g., Intermediate Phrase or Accentual Phrase boundaries) [1, 2].

In German, assimilatory devoicing of fricatives following voiceless obstruents (e.g., [z] → [s]/[t]\_) is generally taken to be a phonological process that occurs across lexical word boundaries, as in *setzt [z]ich* [3]. However, no studies have systematically examined how the German voicing assimilatory process is bounded by prosodic structure. Thus, the present study examines voicing assimilation with prosodic domain limitations taken into account, by investigating detailed acoustic characteristics of domain-initial fricatives (/z, v/).

In order to examine the effect of prosodic structure on voicing assimilation in German, several factors must be taken into account.

First, prosodic structure itself has a systematic influence on the detailed phonetic realization of segments. One of its correlates has been found to be domain-initial consonant

strengthening [6, 7, 8]. Currently, the influence of this process on the realization of fricatives is not yet well established. The question thus arises whether, for example, a durational increase of initial fricatives would interact with the voicing assimilation process.

However, there is no consensus about the higher-level prosodic constituents in German. It is widely accepted that there is the Prosodic Word (Wd), and the Intonational Phrase (IP), but the existence of an intermediate phrase is still a controversial issue of discussion, as Grabe has pointed out [4]. The question is whether the detailed acoustic aspects of domain-final and domain-initial elements can provide further evidence for potential multiple levels of prosodic structure.

Second, some studies have suggested that what has traditionally been considered as a “phonological process” (such as place assimilation) is phonetically gradient rather than categorical [5]. Thus, this study will also explore whether voicing assimilation in German is gradient, or categorical.

Finally, in German, there is a phonological “voicing” contrast in fricative pairs such as /v, f/ and /z, s/. However, there is a distributional asymmetry between /v, f/ and /z, s/ pairs, in that both /v/ and /f/, but only /z/ occur word-initially (the “Initial-s-Constraint” [6]). One might expect that the categorical distinctiveness between /v/ and /f/ affects assimilatory devoicing, predicting less devoicing for /v/ than for /z/, because the latter has no voiceless (or tense) counterpart to contrast with in word-initial position. However, it has been argued that the /v/-/f/ distinction also involves the fortis-lenis dimension [3]. It is thus possible that assimilatory devoicing of /v/ may not create a more /f/-like sound. Whatever an appropriately grounded description of the contrast in /v, f/ might be, the crucial point is that the assimilation of the voiced member of a pair may be limited by the existence of a voiceless counterpart.

## 2. METHODS

Eight native speakers of Northern Standard German, four male and four female, read sentences of four different syntactic types, which were constructed to elicit various prosodic boundaries in /t#C[fric]/ and /ə#C[fric]/ sequences (Table 1).



- a. SENTENCE: Er mag, was Clara gemalt hat. Senken und **Huegel** sind auf dem Bild.  
(*He likes what Clara has drawn. Hollows and **hills** are shown on the picture.*)
- b. LIST: Weil sie Berge gemalt hat, Senken und **Huegel** gemalt hat, und Duenen gemalt hat.  
(*Because she has drawn mountains, hollows and **hills**, and dunes.*)
- c. COMPLEMENT CLAUSE: Weil sie vorhat, Senken und **Huegel** zu malen, faehrt sie zum Aschberg.  
(*Since she wants to draw hollows and **hills**, she is going to the Aschberg.*)
- d. WORD: **Clara** hat Senken und Huegel gemalt.  
(***Clara** has drawn hollows and hills.*)

**Table 1.** Example of sentence set for sequence /t#z/

The target fricatives were /f, v, z/, followed by /ε/ in the nouns *Felder* ('fields'), *Waelder* ('forests') and *Senken* ('hollows'). The preboundary context was varied to be /ə/ or /t/ in the verb forms *hat* ('has') vs. *hatte* ('had'). In order to control for pitch accent, a primary sentence accent was indicated on another word in the utterance by bold printing, leading to (ideally complete) deaccentuation of the target words. Speakers were instructed to imagine themselves uttering the target sentences in an appropriate discourse context in a natural conversation, for example utterance 1b being the answer to the question *Why did he like Clara's picture?* Therefore they were familiarized with the speech materials and practiced the desired sentence accent placement for about 10 minutes before data collection. Each sentence set was repeated five times in a pseudo-randomized order. In total, 960 sentence tokens were collected (3 fricatives x 4 sentence types x 2 preboundary contexts x 8 speakers x 5 repetitions). Acoustic measurements included durations of the preboundary syllable (/t/-closure +aspiration +/ə/ in *hatte*, /a/+t/-closure+aspiration in *hat*), duration of /t/-closure, duration of voicing during /t/-closure (to calculate %voicing of /t/), pause duration (if any), duration of postboundary fricative, and duration of voicing during the fricative (to calculate %voicing of /z, f, v/).

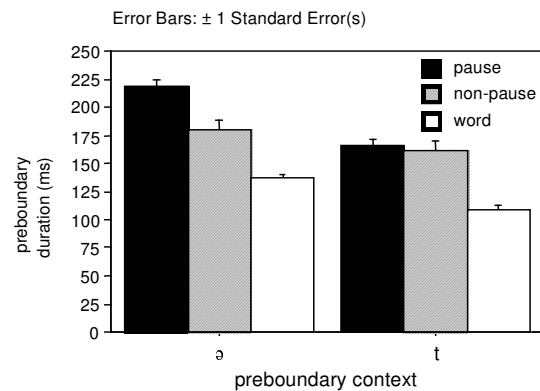
There was considerable variation across speakers in the realization of sentence types b and c (see Table 1) in terms of pausing. According to the model of perceptual strength of prosodic boundaries by de Pijper and Sanderman [9], a pause is a salient cue for a strong prosodic boundary. A weak boundary on the other hand is characterized by absence of pause and absence of melodic discontinuity, which corresponds to the word condition in the present data (i.e., utterance type d). Given the lack of agreed-upon differential criteria for higher-level prosodic units, all utterances other than "word" type were subdivided into "pause" and "non-pause" boundary types, assuming "non-pause" (np) to be a weaker boundary than "pause" (p), but stronger than "word" (wd). A pause was defined to be any portion of silence in the signal longer than 20 ms which

separated either the offset of vowel formants from frication onset (if the preboundary segment was /ə/), or preboundary aspiration from postboundary onset of frication (if the preboundary segment was /t/). In the latter context, absence of pause implied also absence of a distinguishable aspiration phase, because the stop closure was released into the following fricative.

### 3. RESULTS

The results regarding /z/ are based on eight speakers, while only preliminary results will be provided for /v, f/ in this paper.

**3.1 Preboundary lengthening.** Since preboundary lengthening is an important correlate of prosodic boundaries, we examine whether durational differences of the domain-final syllable support the grouping obtained by boundary strength as described in 2. Repeated measures ANOVAs with Huynh-Feldt corrections show main effects of boundary type,  $F(1.8,12)=22.2$ ,  $p<0.01$ , and of preboundary context,  $F(1,6)=36.8$ ,  $p<0.01$ . Posthoc comparisons show a three-way distinction  $p>np>wd$  in the /ə/-final context, but only a two-way distinction  $\{p=np\}>wd$  in the /t/-context (see Figure 1).



**Figure 1.** Preboundary syllable duration

However, if we include pause duration in the analysis, there is again a main effect of boundary,  $F(2,12)=188$ ,  $p<0.01$ , but also a clear three-way distinction in both the /ə/- and the /t/-context (see Figure 2). The assumption that not exclusively the duration of the preboundary segments, but also overall temporal characteristics including pauses are important correlates of prosodic structure, is in line with the model of perceptual boundary strength [9]. Therefore, the use of this split in the following analyses is justified.

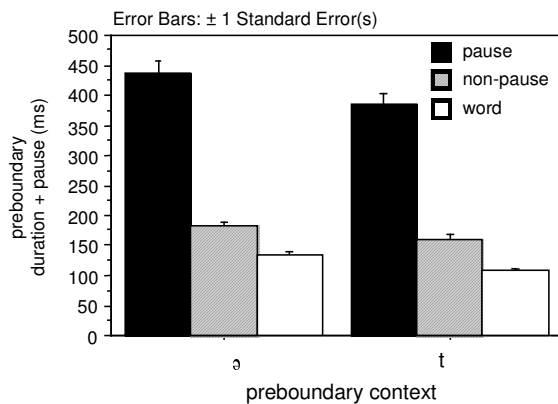
#### 3.2. /z/

**Preboundary /t/-closure duration.** There were main effects of boundary ( $F(1.5, 12)=7.3$ ,  $p<0.05$ ) and context ( $F(1,6)=19.2$ ,  $p<0.01$ ), and a significant Context x Boundary interaction ( $F(2,12)=59$ ,  $p<0.01$ ). Posthoc



comparisons showed a two-way distinction ( $\{p=np\}>wd$ ) in the /ə/-context: When /t/ was the onset of the preboundary syllable (in *hatte#*), a slightly but significantly longer /t/-closure before stronger boundaries was found (a difference of about 10 ms). If /t/ was the domain-final segment (in *hat#*), the closure was lengthened in the non-pause condition, but not in the pause condition, as compared to the word condition, showing  $np>(p=wd)$ .

The reason why the pause condition did not show a lengthening effect on the /t/-closure might be that the pause itself and the /t/ aspiration contribute to the extended temporal pattern marking a large boundary. On the other hand, in the non-pause condition, speakers' strategy might be to lengthen domain-final elements more than in the pause condition, presumably to compensate for the lack of additional cues such as pause and aspiration (the /t/-closure is released into the following homorganic fricative).

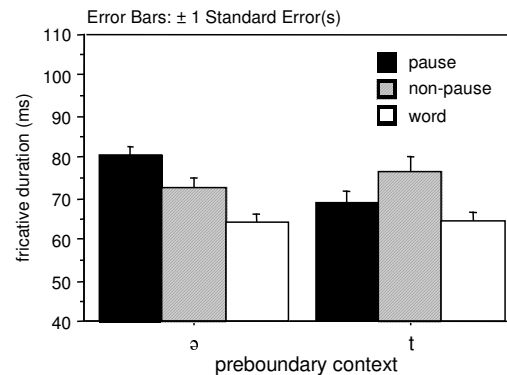


**Figure 2.** Preboundary syllable + pause duration

**%Voicing during /t/-closure.** Neither the context nor the boundary effect on the percentage of voicing during the /t/-closure reached significance. However, there was a trend that in both context conditions, there was less voicing as the boundary became larger. [Note that a comparison between %voicing and /t/-closure duration indicated that the tendency of /t/ to be more voiced before smaller boundaries is not entirely attributable to a decrease in /t/-closure duration. It might thus also reflect the effect of prosodic boundaries of different size.]

**Post-boundary fricative duration.** There was a main effect of boundary ( $F(1.5,12)=7.3$ ,  $p<0.05$ ), showing a longer duration after a higher prosodic boundary. However, posthoc comparisons showed a significant two-way distinction for the /ə/- context ( $p>wd$ ), but not for the /t/-context (see Figure 3). Instead, for the /t/-context, the fricative in the non-pause condition was significantly longer than in the word condition ( $np>wd$ ), thus showing the boundary effect. [Note that the shortened fricative for the pause condition as compared to the non-pause condition may again be accounted for by the compensatory temporal

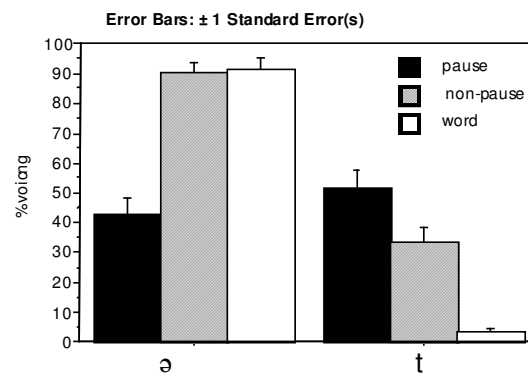
function of a pause, as discussed above with respect to preboundary /t/-closure duration.]



**Figure 3.** Fricative duration for /z/

**%Voicing during fricative.** There was a strong Context x Boundary interaction ( $F(2,12)=59$ ,  $p<0.01$ ), showing the expected assimilatory devoicing in the /t/-context (see Figure 4). Considering the pause condition first, voicing of /z/ was reduced to roughly 50% not only in the /t/-context, but also in the /ə/-context. Therefore, any reduced amount of voicing associated with the pause condition may be simply due to the fact that it is difficult (i.e. requires much articulatory effort) to produce fully voiced obstruents at the beginning of utterances [10].

Turning to the non-pause and word conditions, as shown in Figure 4, /z/ was fully voiced in intervocalic position (the /ə/-context), with no difference between the non-pause and word boundary types ( $p<\{np=wd\}$ ). However, after /t/, /z/ was fully devoiced in the word condition, but only partially in the np condition ( $\{p=np\}>wd$ ). This distinction indicates that assimilatory devoicing is gradient and sensitive to prosodic boundary strength.



**Figure 4.** %Voicing for /z/



### 3.3. Preliminary results for /f, v/.

Based on the data of four speakers, we can report the following trends in the durational patterns, and in the voicing assimilation of /v/:

First, contrary to the findings for /z/, both /f, v/ were shorter after a pause than in the other two prosodic conditions. But, as for /z/, both /f, v/ were longer in duration in the non-pause condition than in the word condition. This latter result again suggests that domain-initial strengthening, in terms of increased duration, is used to compensate for the absence of the strong cue “pause” in signaling prosodic structure.

Second, as shown in Figure 5, the overall pattern in %voicing for /v/ was similar to that for /z/ (cf. Figure 4). However, unlike /z/, in the (assimilatory) context of /t/, /v/ did not undergo complete devoicing in the word condition, with a substantial amount of voicing being maintained. As introduced at the outset of the paper, this is presumably driven by the maintenance of the phonological contrast between /v/ and its voiceless (or tense) counterpart /f/.

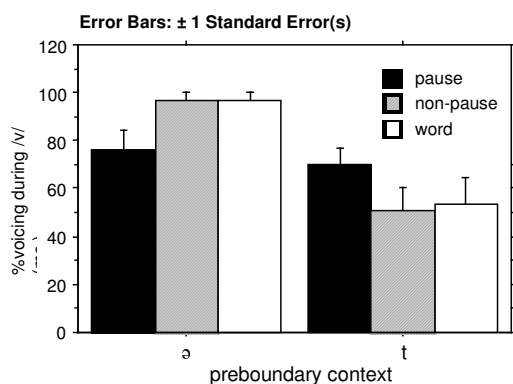


Figure 5. %Voicing for /v/

### 4. CONCLUSION

The results of the present study can be summarized as follows. First, there was a three-way distinction in preboundary lengthening, being correlated with the three-way grouping based on a model of perceptual boundary strength [9]. When pauses were included in the durational analysis, the three-way distinction became even more robust. Second, the duration of fricatives was generally longer after a larger prosodic boundary, showing some evidence of domain-initial strengthening. Third, the assimilatory devoicing of /z/ (to a preceding voiceless /t/) varied as a function of prosodic boundary: the degree of assimilation increased as the size of prosodic boundaries became smaller. This indicates the gradient aspect of voicing assimilation. However, /v/ did not show such an effect, which is presumably conditioned by the need to

maintain the phonological contrast with its voiceless (or tense) counterpart.

Overall, the results of the current study add to a growing body of evidence that phonological processes such as voicing assimilation in German are bounded by high-level prosodic structure. Furthermore, the prosodically-conditioned phonetic variation can be seen as reflecting a multiple level distinction of prosodic constituents larger than a prosodic word.

### REFERENCES

- [1] M. Nespor and I. Vogel, *Prosodic phonology*. Dordrecht: Foris, 1986.
- [2] S.-A. Jun, “The accentual phrase in the Korean prosodic hierarchy.” *Phonology*, vol.15 (2), pp.189-226, 1998.
- [3] K. J. Kohler, *Einfuehrung in die Phonetik des Deutschen*. Berlin: Erich Schmidt, 1995.
- [4] E. Grabe, *Comparative intonational phonology: English and German*. (MPI Series in Psycholinguistics 7). Wageningen: Ponsen and Looijen, 1998.
- [5] M. Jessen, *Phonetics and phonology of tense and lax obstruents in German*. (Studies in functional and structural linguistics 44). Amsterdam/Philadelphia: Benjamins, 1998.
- [6] F. Nolan, “The descriptive role of segments: evidence from assimilation.” In G. Docherty and D. R. Ladd (eds), *Papers in Laboratory Phonology II*, Cambridge, UK: Cambridge University Press, pp261-280, 1992.
- [7] P. A. Keating, T. Cho, C. Fougerson, and C. Hsu, “Domain-initial strengthening in four languages.” In J. Local, R. Ogden, and R. Temple (eds), *Papers in Laboratory Phonology VI*, Cambridge, UK: Cambridge University Press, in press.
- [8] T. Cho and P. A. Keating, “Articulatory and acoustic studies of domain-initial strengthening in Korean.” *Journal of Phonetics*, vol.29, pp.155-190, 2001.
- [9] R. J. de Pijper and A. A. Sanderma, “On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues.” *Journal of the Acoustical Society of America*, vol. 96 (4), pp.2037-2047, 1994.
- [10] P. A. Keating, “Physiological effects on stop consonant voicing.” *UCLA working papers in phonetics* 59, pp.29-34, 1984.