Word Boundary Marking at the Glottal Level in the Production of German Obstruents*

Michael Jessen
University of Stuttgart

1. Introduction

Despite the impression – nourished by the regular use of written language among the literate community – that an utterance can easily be segmented into words, phonetic research has shown that on the signal-phonetic level of reality it is very difficult to observe and understand how the unity and demarcation of words is manifested. When, for example, the spectrogram of a portion of naturally produced speech is evaluated in a naïve way by expecting pauses between words, it is quite likely that the portions of silence or near-silence created by the closure phases of stops (especially when voiceless) are more salient than whatever might exist in terms of word demarcation, and that consequently stop closures would be mistaken for word boundaries (thanks to Anne Cutler for a demonstration of this effect). In the search for cues to word segmentation in the phonetic signal it is not known a priori on which location in the syntagmatic dimension, on which level of the speech chain, or with which choice of signal-processing algorithms this search is most likely to succeed. Most of it is a matter of systematic empirical phonetic research on all possible levels.

The present contribution attempts to provide another piece in the puzzle of approaching the word as a phonetic unit. The phonetic domain that will be investigated is speech production at the level of gestures. Based on the results of a transillumination study, we will evaluate patterns of the glottal opening gesture in the voiceless obstruents of German under the production of sequences with and without an intervening word boundary. After reviewing the literature about the potential of glottal opening gestures for the coding of word segmentation, this study will show that the facts provided here are better explained with an acoustic-based than a gesture-based interpretation of word segmentation.

2. Results and perspectives from the literature

Transillumination provides a valid and reliable means of visualizing and quantifying the glottal opening movements that are associated with voiceless stops and fricatives or with /h/ (see Hoole, to appear, for overview). Among the different possible experimental phonetic agendas that can be pursued with transillumination, one major subject of interest has been the glottal opening patterns in sequences of voiceless obstruents and the influence that a word boundary or related morpho-syntactic boundary within these obstruent sequences has on glottal opening behavior.

Frøkjær-Jensen, Ludvigsen & Rischel (1971) found that if a sequence of /s/ followed by /p/ in Danish is separated by a word boundary and if it is produced relatively slowly, the transillumination signal indicates that instead of the single glottal opening gesture present in individual voiceless obstruents a compound glottal opening occurs that is suggestive of two underlying/intended gestures, one for /s/ and one for /p/ (which is aspirated word-initially in Danish). Frøkjær-Jensen, Ludvigsen & Rischel mention the possibility that this, what they call two-peakedness, is a boundary/junctural phenomenon. They also show, however, that as soon as the same sequence is produced less slowly with what they refer to as normal pronunciation the effect disappears and only a one-peak glottal gesture occurs. For the remainder of this contribution a one-peak glottal opening pattern will be referred to as "monomodal", and a two-peak pattern as "bimodal".

In contrast to the Danish case of /s#p/ (# symbolizing a word boundary), in which a compound glottal opening is at least potentially possible, Lindqvist (1972) found for Swedish that

(Löfqvist & Yoshioka 1981), and Dutch (Yoshioka, Löfqvist & Collier 1982). Most of the discoveries and conclusions relevant for the word boundary marking issue were already made in the Löfqvist (1978) and Löfqvist & Yoshioka (1980) studies on Swedish. They were later confirmed for the other Germanic languages as well (cf. also Fukui & Hirose 1983 on Danish and Hoole, Pompino-Marschall & Dames 1984 on German for further consistent evidence). Among all the interesting details reported in the work of Löfqvist and colleagues we concentrate on those cases that fall into the falsification patterns mentioned in (2).

One regular exception to (1) is the case of adjacent identical obstruents.[1] Even if separated by a word boundary only a monomodal glottal opening is found over adjacent identical obstruents. For example, the sequences /s#s/ in *My ace sales* and /k#k/ in *I make cave* are produced with a single glottal opening (Yoshioka, Löfqvist & Hirose 1981). This effect also occurs if certain other obstruents precede or follow the sequence of adjacent obstruents. This situation applies to the sequence /ks#sp/ of Swedish ***Eks spelar*** 'the Eks play', which is produced with a single glottal opening (Löfqvist & Yoshioka 1980: 798). The case of monomodal glottal opening gestures in identical obstruents that are separated by a word boundary constitutes a type (2a) falsification of the Word Boundary Marking Hypothesis in (1).

The data provided by Löfqvist and colleagues also contain cases that constitute type (2b) violations of the Word Boundary Marking Hypothesis. Löfqvist & Yoshioka (1980: 797) show for Swedish that the sequence /sts#/ in *Kvists ilar* 'the Kvists hurry' is produced with a bimodal glottal opening. This is the case although the obstruent cluster is not interrupted by a word boundary, but only followed by it. Apparently, the explanation for the compound glottal opening in this case cannot be found in terms of word segmentation, but it must be a matter of the specific obstruents involved and their sequencing pattern. Another example that illustrates the same phenomenon is the trimodal glottal opening pattern found in the sequence /sts#p/ of Swedish ***Kvists pilar*** 'the Kvists throw'. In this case the word boundary can explain at most two of the three peaks in the sequence, whereas the two peaks associated with the /sts/-part of the sequence must be due to other factors, as in the previous case.

To explain the entire range of glottal opening patterns Löfqvist & Yoshioka (1980: 800) write: "The observed variations in glottal area are obviously related to the segmental properties of an utterance. The most apparent relation is that sounds requiring a high rate of air flow are produced with a separate glottal opening gesture". Analogous statements are made in Löfqvist (1978) and various of the work with his colleagues, mentioned above. In (3) this generalization is repeated with a slightly different wording.

(3) *Löfqvist's rule*:

> A separate glottal opening peak is associated with every aspirated stop and every voiceless fricative, in order to fulfill the necessary aerodynamic conditions for the production of these sounds.

If (3) turns out to be correct (it will be argued that it does), word boundary marking is no longer an immediate explanation for the articulatory patterns of laryngeal gesture organization. Word boundaries, of course, have to be manifested somehow. It seems that as far as voiceless stops are involved this phonetic coding of word boundaries is most straightforwardly provided by the presence or absence (or the absolute degree of) *aspiration*. Consider again the case investigated by Pétursson (1977), in which the sequence /st/ was produced with a single glottal opening word-internally, but a bimodal glottal opening if separated by a word boundary (analogous cases exist for other Germanic languages). The crucial point is that in the Germanic languages investigated by Pétursson, as well as Löfqvist and colleagues, voiceless stops are aspirated when they occur word-initially without further obstruents, but that they are unaspirated when occurring with a preceding voiceless fricative in word-initial position. In the pattern /s#t/ word-initial /t/ is aspirated and hence receives a glottal opening peak of its own, while the word-final voiceless fricative /s/ also receives its own glottal peak according to (3). In the pattern /#st/, on the other hand, the /t/ is unaspirated and receives no glottal peak of its own. Instead, it shares one single glottal opening with the one that already exists due to the voiceless fricative /s/.

Beyond the case of /s#t/ versus /#st/, which could also be explained by the Word Boundary Marking Hypothesis in (1), Löfqvist's generalization in (3) can also capture the problematic cases of the types (2a) and (2b). That with faster speech rates and less formal speech a bimodal gesture in a structure like /s#t/ can be transformed into a monomodal pattern might be explained by the reduction of aspiration duration that is expected to occur with these stylistic changes (cf. Löfqvist 1992: 19). And the bimodal gesture in the word-internal sequence /sts/ found by Löfqvist & Yoshioka (1980) can be explained by the fact that this sequence contains two voiceless fricatives and one unaspirated stop, which according to (3) amounts to a bimodal pattern. In the sequence /sts#p/ with the trimodal pattern found by Löfqvist & Yoshioka the third peak is due to the free-standing word-initial aspirated /p/. The only case that remains problematic even under the account in (3) is the monomodal pattern of adjacent identical obstruents, in particular fricatives. In the sequence /s#s/ two glottal opening peaks are expected, although as a rule only one peak is found (Yoshioka, Löfqvist & Hirose 1981). Contrary to the fricative sequence, a sequence of identical stops across a word boundary is expected to be realized with a monomodal glottal opening if the word-final stop is unaspirated (which is often the case in English). This prediction is borne out for the sequence /k#k/ shown by Yoshioka, Löfqvist & Hirose (1981), as well as the sequence /p#p/ shown by Lisker & Baer (1984: 167). Despite the unexpected monomodal pattern in the fricative sequence there is still evidence in the EMG data of Yoshioka, Löfqvist & Hirose (1981) for a compound structure in /s#s/ as opposed to a truly plain structure in /k#k/ (the same asymmetry between fricative geminates and stop geminates is reported for Japanese by Yoshioka, Löfqvist & Hirose 1982). The merging of glottal gestures in sequences of identical obstruents could be due to a general degemination tendency, that is known for many languages, especially in systems like English, German etc., which do not tolerate true geminates. Also problematic is the finding mentioned by Yoshioka, Löfqvist & Collier (1982) that a monomodal glottal opening pattern is associated with word-initial /sx/ in Dutch, although these two voiceless fricatives are not identical.

This discussion of the literature has shown that the Word Boundary Marking Hypothesis in (1) has been met with skepticism by most of the phoneticians that performed the relevant trans-illumination (and related) experiments. One could argue that due to this skepticism and the success of the alternative explanation in (3) it would be of no current relevance to pursue this issue any longer. However, it seems that among a number of phonologists (1) is still assumed in one or the other fashion. We will concentrate on two proposals along these lines, one made by Browman & Goldstein (1986) and one by Iverson & Salmons (1995).

We saw that in the word-initial voiceless fricative-stop clusters of languages like English and Swedish the voiceless stop is unaspirated (e.g. in a word like *spill*). In order to provide a plausible phonological explanation of this fact, Browman & Goldstein (1986: 227) propose that in the case of clusters such as word-initial /sp/ a single glottal opening gesture coincides with several supraglottal gestures, in this case one alveolar fricative gesture for /s/ and one bilabial closure gesture for /p/. Since the single glottal gesture is approximately coordinated with the fricative, both if occurring individually or if followed by a stop (this is shown in more detail in the literature of Löfqvist and colleagues), there is not sufficient time left after the release of a stop in clusters like /sp/ during which the glottis would be open and thus in an appropriate condition for the production of aspiration. Browman & Goldstein claim that this suprasegmental status of glottal opening, by which the supraglottal gestures of more than one segment are associated with one shared glottal opening gesture, is a typical feature of word-initial consonant clusters in English and other Germanic languages. Browman & Goldstein's position is expressed in (4).

(4) *Browman & Goldstein's rule*:

Words in English and other Germanic languages begin with at most a single glottal gesture.

While the generalization in (4) has no implications for the (2a) problem, is makes the prediction that the false alarm problem in (2b) should not occur, at least not word-initially in Germanic languages. The bimodal glottal opening observed by Löfqvist & Yoshioka (1980) for the word-final sequence /sts/ in Swedish, although being a (2b) type counterexample of (1), is not strictly a counterexample of (4), simply because it occurs word-finally (in the Swedish name *Kvists*) and not at the beginning of the word. If, on the other hand a case should be found in which any sequence of consonants is produced with more than one single glottal opening gesture *word-initially*, Browman & Goldstein's generalization in (4) would be violated (two cases of this type will be shown for German in § 4).

Iverson & Salmons (1995) discuss the same facts that were also brought up by Browman & Goldstein (1986). They provide similar arguments, but arrive at a proposal that differs from the use of articulatory gestures, that is characteristic of Browman & Goldstein's approach to phonology as a whole (see Browman & Goldstein 1992). Working broadly within the tradition of nonlinear phonology (see Kenstowicz 1994), Iverson & Salmons represent glottal opening in the production of voiceless obstruents with the phonological feature [spread glottis], which goes back to a proposal by Halle & Stevens (1971). The case of word-initial clusters of the type /sp/ is represented by Iverson & Salmons with a single instance of the feature [spread glottis] that is associated with the feature matrices of both the fricative and the following stop. With this representation they intend to capture essentially the same as Browman & Goldstein did, namely that only a single glottal opening is associated with the entire /sp/-like cluster and that this prevents the stop from being aspirated. But aside from the broader theoretical implications there is one important difference between the two approaches that has empirical consequences. By considering the multiple association of [spread glottis] as a special case of the Obligatory Contour Principle (see McCarthy 1988) Iverson & Salmons (1995) claim that this multiple association pattern occurs in any word-internal (or, more accurately, morpheme-internal) position, including word-initial, word-medial, and word-final position. In a format similar to (4) this claim is stated in (5).

(5) *Iverson & Salmon's rule*:

At most a single glottal gesture is found for any word-internal consonant sequence in English and other Germanic languages.

The statement in (5) makes the empirical prediction that more-than-single-peak glottal openings should not occur for any word-internal consonant sequence. But this prediction is too strong if we again think about the case of word-final /sts/ in Swedish with its bimodal glottal opening. It might be that with further formal assumptions this case might be accommodated to the analysis of Iverson & Salmons (1995), but as it stands now, evidence of the type present in the Swedish /sts/ cluster is problematic to their account. In order to provide more data that are relevant to the issues discussed here the results of a transillumination study of German will be illustrated and discussed for the remainder of this contribution.

3. Methods

In the transillumination procedure used in this study the glottis is illuminated with the help of a fiberscope that is inserted through the nasal cavity. The light source is attached from outside and feeds cold light of sufficient brightness into the fiberscope. The tip of the fiberscope is positioned in the hypopharynx in a way which provides an unobstructed view of the glottis. The amount of light that passes first through the glottis and then through the tissue of the neck skin is measured with a phototransistor that is attached to the neck and held in place with the help of a neckband. The phototransistor was held at the level of the cricothyroid membrane. The subject (the author) was seated in a dental chair. The fiberscope was inserted through the nasal cavity by an otolaryngologist (Kiyoshi Oshima) under the application of some local anesthesia. Insertion and positioning of the fiberscope was controlled by the physician with the help of the image from a standard video system. The stimulus words were printed in large letters, randomized, and placed on a stand that was adjusted in height to ensure that the list was visible to the subject. Recordings were made both of the transillumination signal and of the audio signal. The audio signal was recorded with a Sennheiser direction-sensitive microphone that was positioned in front of the subject. A number of test tokens were produced by the subject, during which the levels of the transillumination and the audio signal were controlled by the director of the experiment (Anders Löfqvist) with the help of an oscilloscope. The input volume was adjusted such that the audio signal was unclipped for the relevant portions associated with obstruent production. The recording took place at Haskins Laboratories. Both the audio and the transillumination signal were directly digitized into the computer with a sampling rate of 10 kHz, each using preemphasis. The transillumination signal was smoothed using a triangular window of size 35.1 ms. This eliminates the voicing patterns of the signal and facilitates the evaluation of the glottal opening patterns associated with obstruent production.

The primary purpose of the experiment was an investigation of the difference between voiced and voiceless obstruents in German, as documented in Jessen (to appear). Due to this emphasis most of the material read by the speaker contains obstruents in contexts in which they contrast in voicing. This is the case for the data discussed in § 5, where the obstruents occur word-initially and are preceded by a voiceless palatoalveolar fricative across a word boundary. Relevant for the present study are only the voiceless target obstruents. The different patterns investigated for the discussion in § 5 are /ʃ#p/, /ʃ#t/, /ʃ#k/, /ʃ#f/, and /ʃ#s/ (due to the primary interest in the voicing opposition no fricatives at other than labial and alveolar place of articulation were included, since others do not engage in the voiced/voiceless opposition). The patterns were created by combining the carrier word ***rasch*** /ʃ/ 'quickly' with the words ***Pier*** /p/ 'pier', ***Tier*** /t/ 'animal', ***Kir*** /k/ 'kir', ***vier*** /f/ 'four', and ***Sir*** /s/ 'sir', respectively. Some of these examples are somewhat marginal, but they have the advantage of constituting minimal pairs and of containing a following vowel /i/, which is the preferred vowel for transillumination studies. The carrier word *rasch* was preferred over other possible carriers ending in one of the target obstruents, in order to avoid artefacts that might occur with identical adjacent obstruents (see § 2). Each target pattern was repeated approximately twenty times in blocks of approximately ten. The same number of tokens were produced in a second session three weeks later. The exact positioning of the fiberscope in a recording session, as well as the exact placement of the phototransistor, have effects on the intensity level and other details of the transillumination trace. For this reason the results for the two sessions are kept apart in the evaluation of the results. Contrary to § 5, the material discussed in § 4 was less substantial. It contains voiceless obstruents in word-initial clusters such as /ʃp/, /sk/, /ʃpl/ etc. (in this position no voiced/voiceless opposition is possible in German). The material was only repeated a few times and only recorded in one of the two sessions for exploratory purposes. Yet, this material should still be important and interesting enough for some basic conclusions and ideas about what might be expected in a more in-depth investigation of these patterns. Further information on the methods used here is found in Jessen (to appear).

4. Multiple word-internal glottal openings

In this section some of the glottal opening patterns will be discussed that occur in word-initial obstruent clusters in German. Special emphasis will be given to those cases that are immediately relevant for the different hypothesis in the literature that were discussed in § 2. Figure 1 provides an illustration of four subsequent repetitions of the sequence /i#ʃpi/ in the utterance *nie Spier* 'never spar'. The top part of Figure 1 shows the audio waveform, the lower part the smoothed trans-illumination signal. Both representations are temporally aligned and labeled at the bottom for the occurring sequences of sounds. The transillumination curve shows that the word-initial sequence /ʃp/ is produced with a bimodal glottal opening in the first, third, and fourth repetition. No clear two peaks are observable in the second repetition, but the glottal opening gesture still looks complex. This bimodal pattern for word-initial /ʃp/ is in violation of the claims of both Browman &

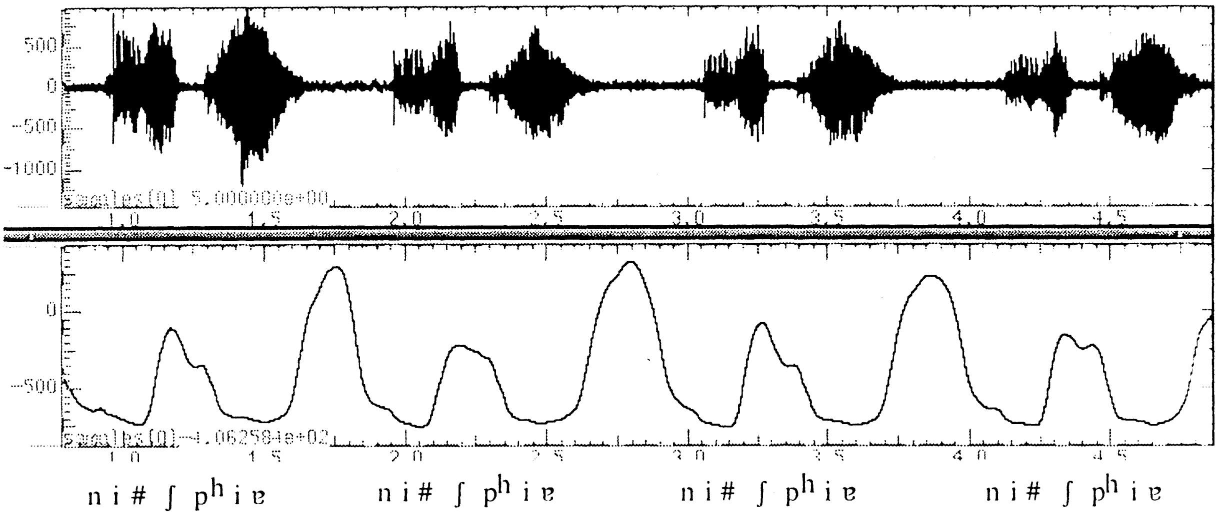

Figure 1: Bimodal glottal opening in /#ʃp/.

Goldstein (1986) in (4) and Iverson & Salmons (1995) in (5). Is it also in violation of Löfqvist's generalization in (3)? Measurements of aspiration duration that were carried out on the basis of spectrograms derived from the acoustic signal indicate values that on an average almost reach 50 ms.[2] Although it is difficult and controversial to divide a continuum of aspiration duration values into the categories aspirated vs. unaspirated, values of almost 50 ms allow the conclusion that at least some aspiration was present in these stops. Apparently the common assumption that stops after word-initial fricatives are necessarily unaspirated in Germanic languages is too strong. Lotzmann (1975) in his transcription-based study of aspiration in German provides a discussion of the literature, including sources that address dialectal differences. He shows that in North German pronunciation aspiration of /p,t,k/ is more common and stronger than in the pronunciation of South German speakers (pp. 9-31). This is consistent with the fact that the speaker in the present study was raised in North Germany (close to Flensburg). He also shows that in the corpus of speech that he investigated stops in the word-initial clusters /ʃp/ and /ʃt/ can be aspirated, even if this is a less preferred option. He found the stop in /ʃp/ to be 34 times aspirated and 101 times unaspirated and the one in /ʃt/ to be 39 times aspirated as against 213 times unaspirated (p. 135). Knetschke & Sperlbaum (1987: 149) in their large transcription study mention that stops in fricative-stop clusters like /ʃp/ and /ʃt/ were aspirated "relatively often" in the speech of German newscasters. These results indicate that aspiration in word-initial fricative-stop clusters is an option German, which had been implemented in the tokens shown in Figure 1. Given that the stops in Figure 1 are aspirated, the finding of a bimodal glottal opening is not in violation, but in confirmation of Löfqvist's generalization in (3).

Another set of examples relevant for the present discussion is shown in Figure 2. The transillumination signal above shows four subsequent repetitions of the sequence /i#ʃpli/ in the utterance *nie Splier* and the one below shows four repetitions of the sequence /i#ʃpri/ in the utterance *nie Sprier* (*Sprier* and *Splier* are nonsense words, but phonotactically normal in German). It can be seen in Figure 2 that across repetitions a single glottal opening occurs for the sequence /ʃpl/, but that a bimodal glottal opening occurs for /ʃpr/. To understand these patterns it is important to point out that /r/ was produced as a voiceless uvular fricative [ʁ̥] in this context. The realization of German /r/ as an uvular fricative is the preferred realization of this phoneme in the syllable onset, and its manifestation as fully voiceless after other voiceless obstruents in the syllable onset is common for German as well (Kohler 1995: 166). If we take into account that the

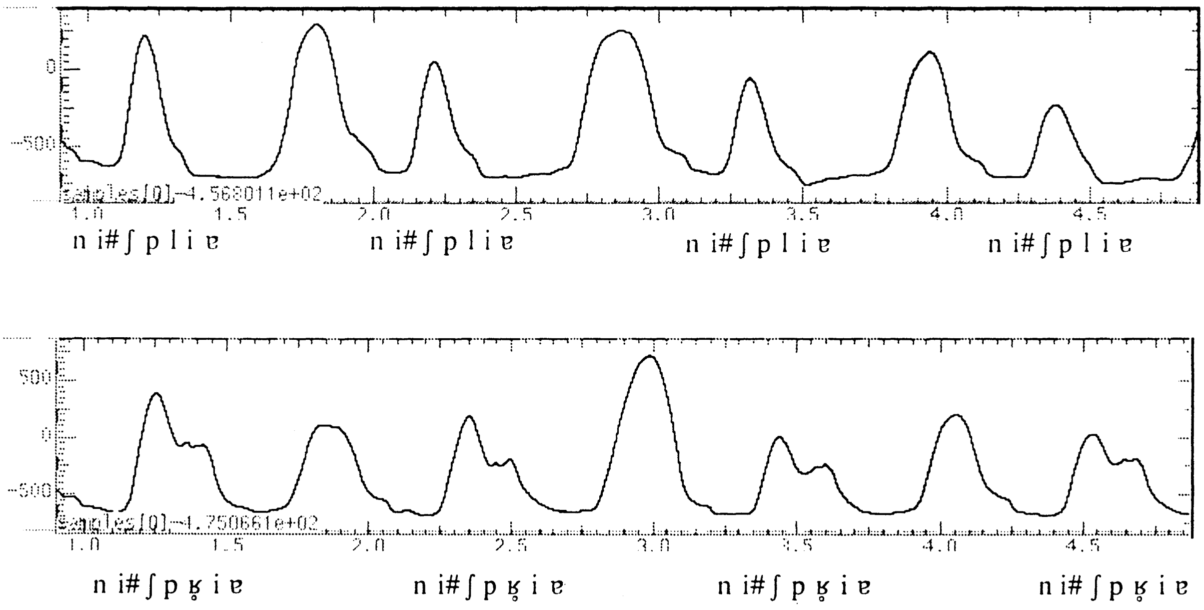

Figure 2: Monomodal glottal opening in /#ʃpl/, bimodal in /#ʃpr/.

sound /r/ in /ʃpr/ is realized as a voiceless fricative, the bimodal glottal opening follows from Löfqvist's generalization in (3). Since the sequence /ʃpr/ contains two voiceless fricatives and one unaspirated stop, two glottal opening peaks should occur, which is borne out in Figure 2. In contrast, the /l/ in /ʃpl/ bears no resemblance to a voiceless fricative and the stop is unaspirated; hence only the single glottal opening associated with /ʃ/ should be found, which is again borne out. The lack of aspiration of the stops here (as measurable in the audio signal, which is not shown here) as opposed to the aspiration of the stops in Figure 1 might be partially explained by the universal phonetic tendency that the duration of segments shortens with the number of segments in a cluster (see Klatt 1976).

Taken together, Figures 1 and 2 have shown us two cases of glottal opening behavior that are explainable by the generalization of Löfqvist in (3), but that present direct counterevidence to the phonological hypotheses in (4) and (5). These two cases show additional instances of the false alarm problem of the Word Boundary Marking Hypothesis in (1), which add to the bimodal pattern found by Löfqvist & Yoshioka (1980) in the word-final sequence /sts/ of Swedish.

5. Gestural aggregation

We now turn to cases that challenge the Word Boundary Marking Hypothesis from the direction of the problem of misses in (2a). It will be shown that a monomodal glottal opening is possible despite the presence of an intervening word boundary. The most systematic research along these lines has been performed by Munhall & Löfqvist (1992). Munhall & Löfqvist show that the sequence /s#t/ is produced with two separate glottal openings in slow speech. With a gradual increase in speech rate the two separate glottal openings turn into a compound bimodal structure, until they finally merge into a plain monomodal glottal opening. This merging of glottal openings with increasing rate of speech is referred to as "gestural aggregation" by Munhall & Löfqvist. The German data to be discussed in this section show that a substantial degree of gestural aggregation occurs even in the relatively slow rate of lab speech that was used in the present study, and that factors other than rate of speech must be held responsible as well.

A corpus of about 200 tokens (ca. 20 tokens • 2 sessions • 5 word initial voiceless obstruents; cf. § 3) was evaluated qualitatively. This was done by printing out each one of the tokens, arranging them according to session number as well as place and manner of articulation of the second obstruent, and trying to find generalizations by "eye balling" the material. One striking observation was the considerable token-to-token variability of the same sequence repeated successively within the same session. It was not at all uncommon to find a very clear bimodal structure being followed in the next repetition by a glottal opening with only a minimal trace of compoundedness left. Fortunately, aside from this high degree of pure free variation it was also possible to observe some linguistically systematic tendencies. As will be shown and explained for the remainder of this section, it was found that gestural aggregation depends to a certain degree on the place and manner of the word-initial obstruent in the position after word-final /ʃ/. As a tendency, the degree of gestural aggregation occurs in the order labial > alveolar > velar and in the order fricative > stop. The representative examples in Figure 3 illustrate these tendencies. The transillumination signal at the top of Figure 3 shows an example of the sequence /ʃ#f/ in *rasch vier* 'quickly four'. The vertical scale of glottal opening intensity is in arbitrary units and is adjusted to fit the entire range of the window, which is the reason why different ranges of values occur in the vertical axes of the three examples (the same holds for any other transillumination signal shown in the figures of this paper). It appears that the glottal gesture visible here is a bit more complex than would be expected from a plain single glottal opening, but it is impossible to see more than one peak. The signal in the middle of Figure 3, which shows an example of the sequence /ʃ#p/ in *rasch*

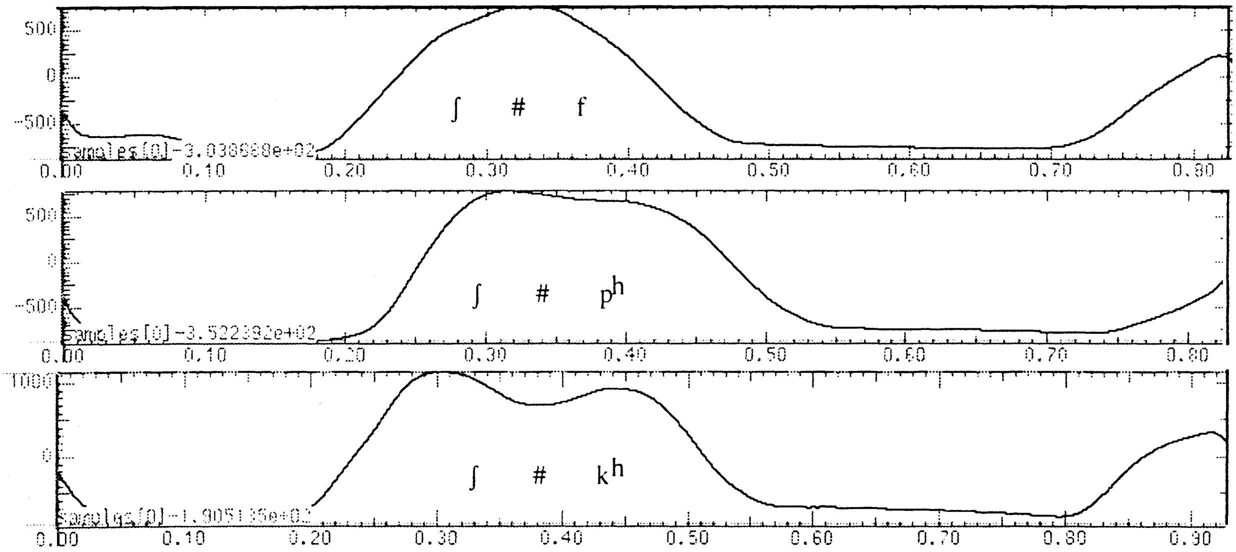

Figure 3: Place- and manner-related differences in gestural aggregation.

Pier emerges as clearly complex, but it is difficult to tell apart the dimension of the first from that of the second component. Finally, at the bottom of Figure 3 it is shown that in the sequence /ʃ#k/ in *rasch Kir* a bimodal glottal opening can be determined with the two peaks clearly distinguishable. From the top to the middle part the influence of manner of articulation can be seen (fricative vs. stop, respectively), while from the middle to the bottom part the influence of place of articulation (here: labial vs. velar, respectively) is identifiable. In each of these steps the compound nature and bimodal appearance becomes clearer.

An explanation for the situation illustrated in Figure 3 can be sought in the different oral-laryngeal timing patterns of voiceless obstruents that differ with respect to place of articulation and the stop-fricative distinction. In the investigation described more explicitly in Jessen (1995) and Jessen (to appear) the dimension and oral-laryngeal coordination of glottal opening has been measured in word-initial obstruents preceded by a vowel. The target words were the same as the ones described in this section (*Pier, Tier, Kir, vier, Sir*), with the only difference that the preceding carrier word ended in a vowel (in the word *nie* /ni/ 'never', as in *nie Pier, nie Tier* etc.) instead of the fricative /ʃ/ used in the present study. Since in those stimuli the target obstruents were surrounded by vowels it was much easier to quantify the entire glottal gesture, whereas in the present setting full quantification was inhibited due to the influence of the glottal gesture associated with /ʃ/. Table 1a,b shows a selection of those results in the investigation of Jessen (1995) and Jessen (to appear) that reveal differences in oral-laryngeal timing that are due to place (Table 1a) or manner (Table 1b) of articulation. The particular parameters addressed in the table are the interval between the onset of the target obstruent and the onset of its glottal opening (OG-OC) and the interval between the onset of the target obstruent and the moment of peak glottal opening associated with that obstruent (P-OC). Table 1a shows the means, as well as the standard deviations (in parentheses), of the parameters OG-OC and P-OC (in ms) for labial, alveolar, and velar place of articulation and the P-values that emerge in one-factor ANOVAs with place of articulation as the independent variable and OG-OC, as well as P-OC as the dependent variables (see different columns). In the case of stops, in which three places exist, the probabilities for each pairwise comparison (Tukey HSD multiple comparisons) is provided separately (with l = labial, a = alveolar, and v = velar). Separate analyses were carried out for stops in the first recording session, stops in the second session, fricatives in the first, and fricatives in the second session (see different rows). Each of the

Table 1a: Oral-laryngeal coordination for different places of articulation.

		OG-OC				P-OC			
Manner	Sess	labial	alveolar	velar	P	labial	alveolar	velar	P
stops	1	5 (7)	5 (3)	6 (11)	l/a: .96 l/v: .90 a/v: .98	125(10)	124 (7)	130(11)	l/a: .86 l/v: .28 a/v: .11
	2	7(5)	16 (10)	17 (19)	l/a: .10 l/v: .07 a/v: .98	111 (7)	115 (7)	128(14)	l/a: .55 l/v: .00 a/v: .00
fricatives	1	-9 (4)	-14 (9)	-	.07	104 (7)	111(15)	-	.1
	2	-14 (8)	-6 (6)	-	.001	95 (9)	105(13)	-	.007

mean values are based on approximately 20 repetitions. Table 1a shows that as a tendency OG-OC and P-OC increase as place of articulation proceeds backwards in the oral cavity. This tendency holds for both stops and fricatives. Although consistent in most cases, this effect is only significant in some cases, i.e. the order velar > labial and velar > alveolar in the parameter P-OC for stops produced in the second recording and the order alveolar > labial in both parameters for fricatives of the second recording.[3] Table 1b shows the results of a comparison between stops and fricatives (columns), separately for labial place of articulation in the first and second, and alveolar place of articulation in the first and second recording (rows). Beyond this difference, the conventions used for Table 1b are identical to Table 1a. Table 1b shows quite a robust pattern according to which higher values of OG-OC and P-OC are found for stops than fricatives. This effect holds across the two different places of articulation involved, and it is statistically significant in each case.

Table 1b: Oral-laryngeal coordination for different manners of articulation.

		OG-OC			P-OC		
Place	Sess	stop	fricative	P	stop	fricative	P
labial	1	5 (7)	-9 (4)	.000	125 (10)	104 (7)	.000
	2	7 (5)	-14 (8)	.000	111 (7)	95 (9)	.000
alveolar	1	5 (3)	-14 (9)	.000	124 (7)	111 (15)	.002
	2	16 (10)	-6 (6)	.000	115 (7)	105 (13)	.005

In summary, Table 1a,b shows the tendency that the time between consonant onset and glottal opening onset as well as between consonant onset and peak glottal opening is in the order velar > alveolar > labial across stops and fricatives (except that the voiceless velar fricative was not investigated) and in the order stops > fricatives across places of articulation. In a schematical format these general relations are illustrated in Figure 4. Figure 4 shows the supralaryngeal configuration (called oral gesture) of an (in this case word-initial) intervocalic obstruent, with supralaryngeal opening for the first vowel, closure (or constriction in the case of fricatives) for the obstruent, and opening again for the second vowel. Above the oral gesture we find a glottal opening gesture that is coordinated relatively late with respect to the oral gesture, and below we find a glottal gesture with relatively early timing. Although Figure 4 does not represent the facts in detail, it suffices to illustrate the difference between the late timing of the glottal gesture in back places of articulation as opposed to front ones and the late timing of stops as opposed to fricatives (for reasons of space only the P-OC parameter is included in Figure 4; it also turned out to be slightly more robust statistically than OG-OC). For the issue at hand, i.e. the gestural organization

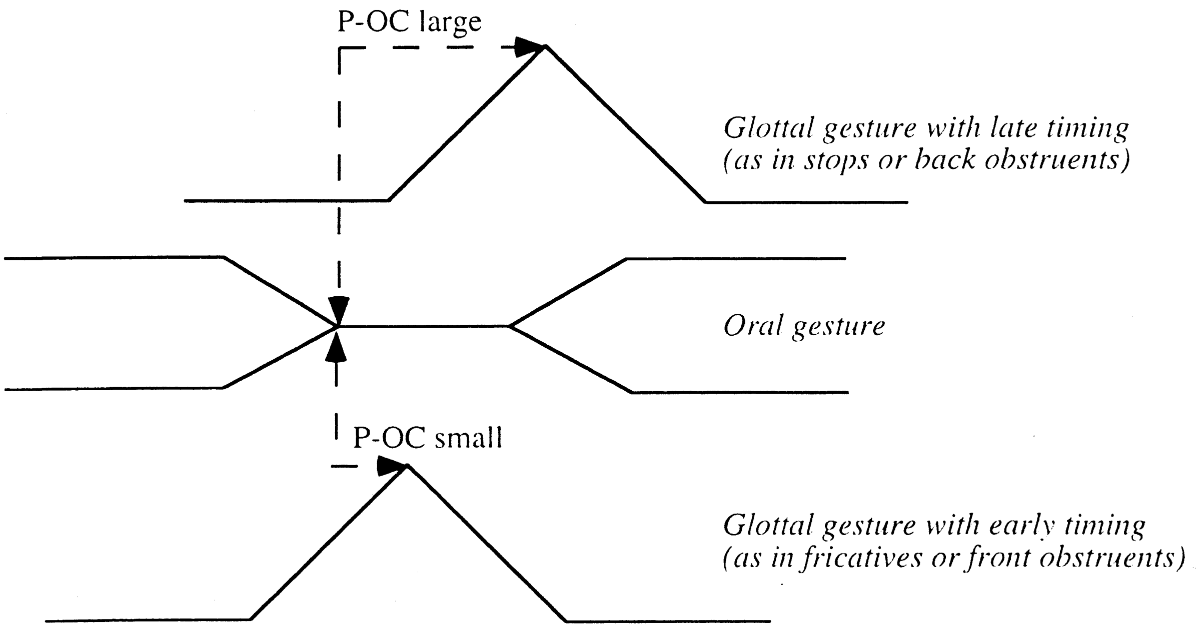

Figure 4: Different patterns of oral-laryngeal coordination.

of obstruent sequences across a word boundary, the different coordination patterns found in Table 1 and illustrated schematically in Figure 4 have the following consequence. Assuming that the sound that precedes the obstruent in question is not a vowel, as in Table 1 and Figure 4, but a voiceless fricative like /ʃ/ (as it is the case in the present investigation) and assuming the constancy of the glottal opening gesture of /ʃ/ across the specifics of the following sound, two different glottal opening patterns are expected, depending on whether the obstruent after /ʃ/ has a relatively early or late timing of glottal opening (Figure 5). Figure 5 shows two hypothetical glottal opening patterns.

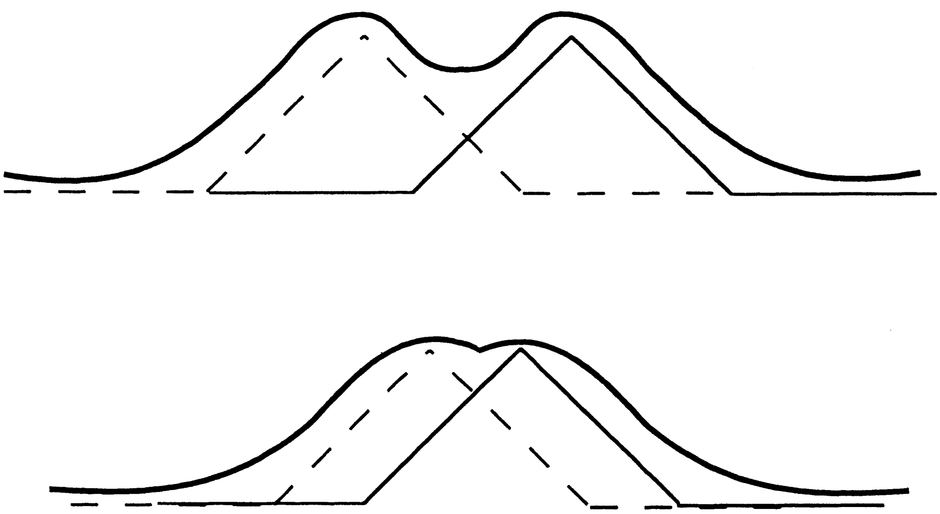

Figure 5: Glottal opening patterns with different timing of the second obstruent.

In the illustration above, the underlying glottal opening of a voiceless fricative like /ʃ/ is combined with the underlying glottal opening of a voiceless obstruent that has a late glottal opening pattern, as in stops and obstruents backwards in the oral cavity. If the two glottal gestures are combined a clear bimodal gesture is expected. In the illustration below, the gesture of the fricative is combined with an early-timed glottal opening found in fricatives or obstruents produced forward in the oral cavity. After combination of the underlying gestures a result is expected in which the bimodal structure is greatly reduced or no longer present. As mentioned earlier in this section and illustrated with Figure 3, the expectations expressed in Figure 5 are borne out as a tendency in the "eye-balling" method pursued here.[4] One possibility of approaching a more quantitative index of gestural aggregation is the use of a velocity calculation of the transillumination signal. If the compound glottal gestures show a clear bimodal pattern two additional zero crossings should occur in the velocity curve (cf. Munhall & Löfqvist 1992 for this criterion). Figure 6 provides an illustration of this procedure. Figure 6 shows the smoothed transillumination signal (above) time-aligned with a velocity calculation (below) of an example of the utterance *rasch Kir*, with the sequence /ʃ#k/. The first zero crossing in the velocity curve is aligned with the first peak, which belongs to /ʃ/. The second zero crossing is aligned with the valley in-between the peaks, and the third one is aligned with the peak that is associated with /k/. Instead of the three zero crossings only one occurs if no two clear peaks exist. Going through the material it also turned out that in a number of cases in which two peaks were identifiable still only a single zero crossing appeared. This was found in examples were the intensity of one of the peaks was considerably lower than that of the other, in which case no direction change occurs in the transillumination signal and subsequently no additional zero crossing in the velocity curve. The number of productions with additional zero crossings found in the corpus is listed in Table 2. Table 2 shows the number of voiceless obstruents occurring after /ʃ/ across a word boundary that were produced with an additional pair of zero crossings in the velocity curve. It is subdivided according to the obstruents involved (columns) and the two different recording sessions (rows). The total number of tokens for each slot is approximately twenty. Thus, for each obstruent and both recordings cases with additional zero crossings turned out to be in the minority. Partially this means that the presence of

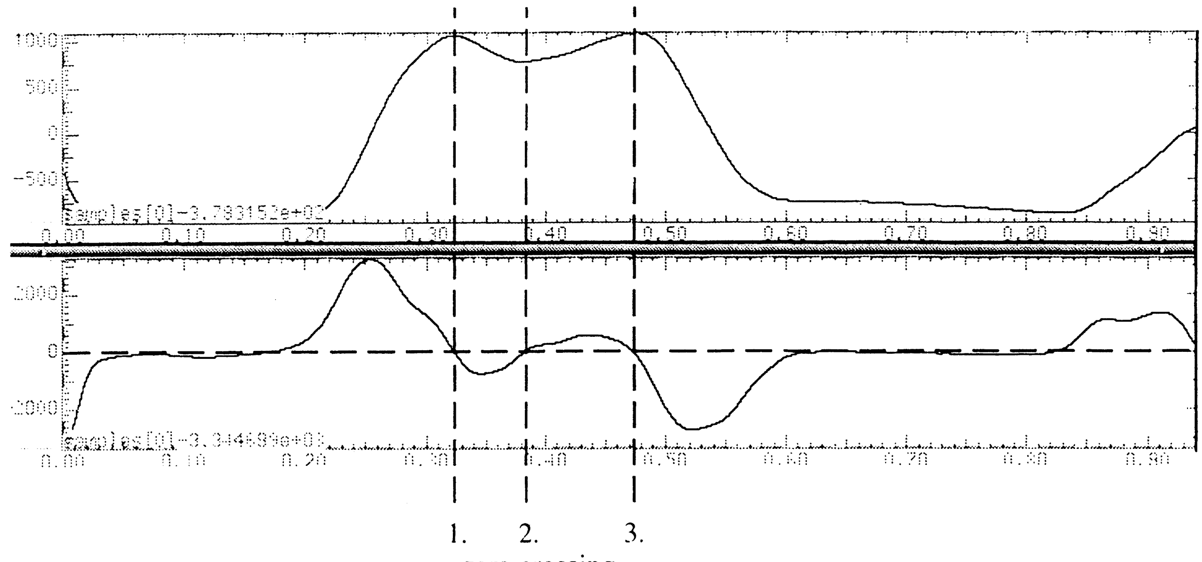

Figure 6: Additional zero crossings in bimodal glottal openings.

additional zero crossings, as objective and methodologically simple it is, might be a criterion that is too conservative for the type of investigation that is pursued here. But this result also indicates that the level of gestural aggregation in this study is relatively high despite the fact that the material was spoken in a relatively slow speech rate and nonspontaneous "lab speech" style (cf. Pétursson 1977 for variation of gestural aggregation in lab speech and Fukui & Hirose 1983 for differences between speakers). Thus, the rate of misses, i.e. cases where despite an intervening word boundary no clear two glottal opening peaks can be observed, is relatively high. But despite an overall small number of additional zero crossings, the absolute values shown in Table 2 are actually consistent with the patterns that are expected from the different configurations of oral-laryngeal timing illustrated in Figures 3 and 5. As a tendency, higher numbers of tokens with additional zero crossings are found in stops as compared to fricatives and in consonants produced further back as compared to consonants further to the front of the oral cavity.

Table 2: Number of tokens produced with additional zero crossings.

	f	s	p	t	k
Session 1	4	5	5	6	6
Session 2	0	0	1	1	3

6. Further aspects of laryngeal word boundary marking

In this contribution the phonetics of word boundary marking was investigated with respect to a quite specific topic, namely the number of glottal opening peaks in sequences of voiceless obstruents. Even if we restrict ourselves to the role of the larynx in the production of cues to word segmentation there are a number of other potential topics that deserve further attention. However, only a few remarks along these lines can be made here.

Glottal opening behavior has been addressed in this study from a largely qualitative and *categorical* point of view, when we were looking at the number of glottal opening peaks occurring in obstruent sequences with and without word boundaries. But it is also possible to look for cues to word boundary marking in the *quantitative* and *gradient* properties of glottal opening gestures.

An interesting design for a transillumination study that proceeds along these lines is found in the work of Cooper (1991) for English. By distinguishing between word-initial stressed, word-medial stressed, word-initial unstressed, and word-medial unstressed syllables in his stimuli, Cooper was able to differentiate the effects of word stress from the effects of word segmentation on glottal opening. The variables that were measured by Cooper include the dimensions of the glottal opening gesture alone, such as glottal opening duration and peak glottal opening, and patterns of oral-laryngeal coordination similar to the ones discussed in connection with Table 1.

As another area of laryngeal word boundary marking we may want to look beyond obstruent production and consider the glottal opening behavior of the sound /h/. One aspect that makes /h/ interesting for the topic of word boundary marking is its phonotactic status. The position in which /h/ is found most commonly in Modern Standard German is at the beginning of the word (e.g. *Hafen* 'harbor', *halten* 'hold', *Heirat* 'marriage'). In this position it occurs in isolation before the vowel and may not combine with other consonants. The restriction against a combination of /h/ with other consonants holds for any position within a (monomorphemic) word. The only position in which /h/ can be found word-internally in German is between two vowels, provided that the preceding vowel is tense (or a diphthong) and the following vowel is not [ə] or [ɐ]. The number of words with word-medial /h/ is quite small and many of them are loans (e.g. *Ahorn* 'maple', *Alkohol* 'alcohol', *Uhu* 'eagle owl', *Vehikel* 'vehicle'). It has been pointed out already by Trubetzkoy (1939: 247) that /h/ is an effective boundary marker in German. If a sequence of consonant and /h/ occurs, it can be inferred that a word or morpheme boundary falls in-between the two sounds.[5]

On the phonetic side it is interesting to notice that /h/ differs from voiceless fricatives and aspirated stops by the fact that glottal opening can coincide with continuing vibrations of the vocal folds. In obstruents glottal opening acts as a devoicing mechanism, but in /h/, where no supra-glottal obstruction occurs, voicing can carry through part or all of its glottal opening. This can be seen in the illustrations in Sawashima & Hirose (1983: 17) for Japanese and Löfqvist & McGowan (1992: 98) for a Swedish speaker. Figure 7 shows the same for German (see also Hartmann 1963

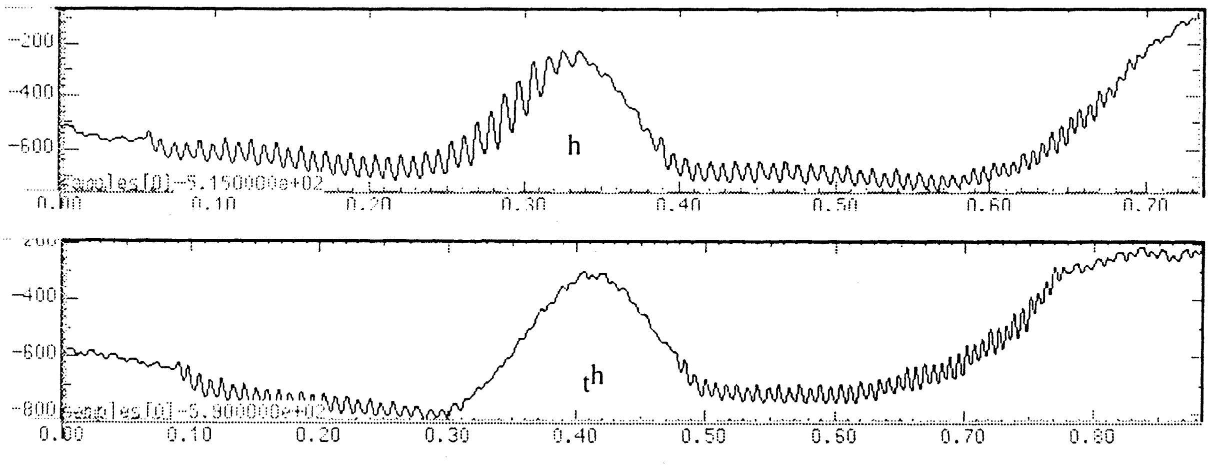

Figure 7: Glottal opening gestures in [h] and [t^h].

and Stock 1971: 100ff. for German). Figure 7 shows the transillumination signal of /h/ in the utterance *nie hier* 'never here' (above) in comparison with the transillumination signal of aspirated /t/ in the utterance *nie Tier* 'never animal' (below). Contrary to the other transillumination examples presented so far, no smoothing was applied to these cases, so that voicing information was preserved. The glottal opening associated with the aspirated stop is almost entirely voiceless, whereas more than half of the glottal opening of /h/ is voiced. In the corpus of /h/ productions there were also tokens with voicing throughout and tokens with less voicing than shown here, but always with more than in aspirated stops. These fact show that glottal opening gestures with substantial voicing are always indicative of the segment /h/ in German and subsequently, for the phonotactic reasons mentioned above, indicative of word-initial position.[6]

Finally, is should be discussed whether the opposite of glottal opening, namely *glottal constriction*, has potential as an index of word segmentation. One case that has been interpreted frequently as a boundary marker in German phonetics and phonology is the glottal stop (Trubetzkoy 1939: 244f., Moulton 1947, among others). The glottal stop has a distribution in German that is not much different from the occurrence of /h/, discussed above, except that it is probably more sensitive to the degree of stress in the syllable that it initiates (see Moulton 1947, Kohler 1995: 100ff., 168f., Wiese 1996: 58ff.). According to the pronouncing dictionary *Großes Wörterbuch der deutschen Aussprache* (Krech et al. 1982) the glottal stop can occur not only word-initially, but also word-internally between two vowels, if the second vowel is stressed (e.g. *The[ʔ]ater* 'theater', *Di[ʔ]ode* 'diode'), though other pronouncing dictionaries do not agree. Most of these items are loans, analogously to the situation for /h/, though with more examples, provided the presence of glottal stop in this context is in fact a stable phenomenon (which is doubtful in light of the fact that Kohler 1994 found no glottal stops in cases of this type). Contrary to /h/, which is commonly classified as phonemic in German, the glottal stop has a wider range of realizations and more variability in its occurrence than /h/. This is one of the major reasons why the glottal stop is usually not considered phonemic in German. The occurrence and realization of what is understood and transcribed as the glottal stop in German has been investigated in detail by Krech (1968) and Kohler (1994). As one of their results, both found that glottal stops are produced more frequently after voiceless obstruents than after vowels (and sonorants). Notice that, like /h/, the glottal stop does not usually combine with preceding obstruents word-internally (Kohler 1995: 101). This implies that post-voiceless position, which turned out to be particularly favorable for the expression of the glottal stop, is a context in which the glottal stop occurs word-initially (e.g. in *das [ʔ]Ohr* 'the ear').[7]

Glottal stop and the laryngealization that accompanies or replaces it is visible in the transillumination signal as a gesture-like lowering of intensity beyond the baseline that is associated with the adjacent vowels. An illustration of the glottal stop in word-initial position surrounded by vowels is provided in Figure 8 (cf. a similar view of the glottal stop in Löfqvist & McGowan 1992: 98 for a Swedish speaker). Figure 8 shows the audio signal (above) and unsmoothed transillumination signal (below) of a glottal stop found in the utterance *nie ihr* 'never you PL', representative of several repetitions. It can be inferred from Figure 8 that the glottal constriction produced in

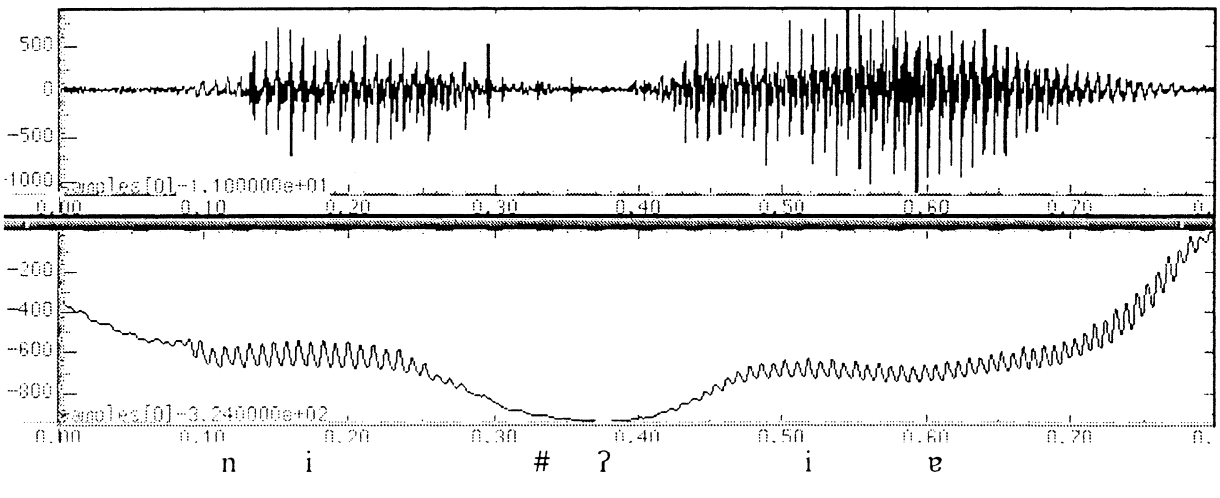

Figure 8: Glottal constriction in [ʔ].

a glottal stop leads to a compression of the vocal fold tissues that reduces the conduction of light through the glottis beyond the level found in the closed portion of a glottal cycle in normal voice production (see the voicing information of the adjacent vowels in the unsmoothed transillumination signal). Given what was said about the distribution of the glottal stop, the occurrence of a glottal constriction pattern beyond the vowel baseline is a likely indication of a word boundary (at least in terms of the notion "phonological word" mentioned in Notes 5 and 7).

So far glottal stop was discussed with respect to word-initial position. Kohler (1994) shows that a glottal stop or related forms of glottal constriction can also occur word-finally, in which case they accompany or even replace stop consonants. Kohler, however, also mentions that this glottal constriction pattern can extend to positions within a word, especially in combination with nasals, like in *Leutnant* 'lieutenant', *hinten* 'behind', *Punkten* 'point DAT PL' (Kohler 1994: 45; boldface indicating the stop that undergoes glottal replacement). In some cases the mentioned notion of "phonological word" might be of help, but not in all, which creates a false alarm problem (i.e. glottal stop, but no word boundary).

7. Conclusion

In this contribution we have been concerned with the question of whether information about the presence or absence of word boundaries is encoded in the way glottal opening gestures are organized in the production of voiceless obstruent sequences. An appealing argument in favor of the "Word Boundary Marking Hypothesis" was offered by the fact that the same sequence of voiceless obstruents of the type /st/ is produced with a single (monomodal) glottal opening word-internally, but with a double (bimodal) glottal opening if the two sounds are separated by a word boundary. From this finding it seemed to follow that the difference between a bimodal and a monomodal glottal opening pattern can be directly attributed to the presence versus the absence of a word boundary, respectively. The flaw in this argument lies in the fact that these two obstruent sequences are not really identical on the phonetic level: if a voiceless stop occurs alone in word-initial position, it is aspirated, but if it clusters with a preceding fricative it is unaspirated (at least this is the regular case in the Germanic languages). With this knowledge in mind it became possible to attribute the different glottal opening patterns to the particular phonetic properties of the segments involved, rather than to an autonomous influence of juncture. In this example, the presence of aspiration could be made responsible for one of the two peaks in the glottal opening pattern. This is the essence of what was referred to as "Löfqvist's rule" in § 2. Löfqvist's rule offered an alternative to the Word Boundary Marking Hypothesis, that did not only offer a different perspective to the interpretation of the /st/ case, but that could also account for cases that constitute a direct falsification of the Word Boundary Marking Hypothesis. Some of these cases were already known in the literature, while others were added by the new evidence on German that was presented here.

If we focus on the role of aspiration in the light of the information discussed and presented here, the most straightforward reasoning seems to be that it is aspiration in itself that carries information about word boundary marking, and that aspiration in turn requires a certain arrangement of glottal opening gestures. In other words, word boundaries are not directly reflected in the organization of glottal gestures, but only indirectly – mediated by aspiration as the most direct "demarcative feature" in this case (to use the terminology of Trubetzkoy 1939 or Jakobson & Waugh 1987). That the presence or degree of aspiration is an important boundary marker in German is for example claimed by the pronouncing dictionary *Duden* (Mangold 1990: 49). Duden claims /p,t,k/ to be strongly or even very strongly aspirated in word-initial position, even if the first syllable in the word is not stressed (as in the initial [t^h] of *Talént* 'talent'). Aspiration has been interpreted as juncture-dependent already by Moulton (1947).[8] Aspiration can also be an explanation for glottal opening behavior in other functions than word boundary marking. In § 5 we saw that the clarity with which a bimodal glottal opening pattern occurs is in part determined by the place of articulation of the following stop (being in the order k > t > p). Yet, still a large amount of free variation was observed. To the extent that glottal organization is interpreted as a "redundant feature" in the expression of place of articulation (cf. Jakobson & Waugh 1987 for this term), it must be acknowledged that it is again aspiration that does the better job as a redundant feature of place of articulation. In the acoustic signal of the transillumination study of Jessen (1995) the order k > t > p in aspiration duration turned out to be a reliable and statistically significant feature of place of articulation in word-initial position, which is a more stable effect than the degree of gestural aggregation, addressed in § 5.

Thus, one way we can conclude this paper is by hypothesizing that in the expression of word boundaries (as well as in some other functions) the *acoustics* (here: aspiration) has primacy over the *articulation* (here: organization of glottal opening gestures). Such a conclusion would emphasize the goal-oriented or listener-oriented nature of articulatory (here: gestural) organization in general (cf., among others, Jakobson & Waugh 1987, Lindblom 1990, Kohler 1994, Perkell et al. 1995).

8. References

Browman, C.P. & Goldstein, L. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3: 219-252.

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

Cooper, A.M. 1991. *An articulatory account of aspiration in English*. Ph.D. Dissertation, Yale University.

Davis, K. 1994. Stop voicing in Hindi. *Journal of Phonetics* 22: 177-193.

Fischer-Jørgensen, E. & Hutters, B. 1981. Aspirated stop consonants before low vowels, a problem of delimitation - its causes and consequences. *Annual Report of the Institute of Phonetics of the University of Copenhagen* 15: 77-102.

Frøkjaer-Jensen, B., Ludvigsen, C. & Rischel, J. 1971. A glottographic study of some Danish consonants. In Hammerich, L.L., Jakobson, R. & Zwirner, E. (eds.) *Form and substance. Phonetic and linguistic papers presented to Eli Fischer-Jørgensen*. Odense: Akademisk Forlag. 123-140.

Fukui, N. & Hirose, H. 1983. Laryngeal adjustments in Danish voiceless obstruent production. *Annual Bulletin. Research Institute of Logopedics and Phoniatrics. University of Tokyo* 17: 61-71.

Giegerich, H.J. 1989. *Syllable structure and lexical derivation in German*. Bloomington: Indiana University Linguistics Club.

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

Halle, M. & Stevens, K.N. 1971. A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Status Report* 101: 198-213.

Hartmann, E. 1963. Positionsbedingte phonetische Varianten des deutschen Hauchlautes. *Zeitschrift für Phonetik, Sprachwissenschaft und Komunikationsforschung* 16: 49-55.

Hayes, B. 1986. Inalterability in CV Phonology. *Language* 62: 321-351.

Hoole, P., to appear. Laryngeal coarticulation: coarticulatory investigations of the devoicing gesture. In Hardcastle, W.H. & Hewlett, N. (eds.) *Instrumental studies of coarticulation*. Cambridge etc.: Cambridge University Press.

Hoole, P., Pompino-Marschall, B. & Dames, M. 1984. Glottal timing in German voiceless obstruents. *Proceedings of the International Congress of Phonetic Sciences* 10, 2b: 399-403.

Iverson, G.K. & Salmons, J.C. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12: 369-396.

Jakobson, R. & Waugh, L.R. 1987 (2). *The sound shape of language*. Berlin etc.: Mouton de Gruyter.

Jessen, M. 1995. Glottal opening in German obstruents. *Proceedings of the International Congress of Phonetic Sciences* 13, 3: 428-431.

Jessen, M., to appear. *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Kenstowicz, M. 1994. *Phonology in generative grammar*. Cambridge, USA & Oxford, UK: Blackwell.

Kim, C.-W. 1970. A theory of aspiration. *Phonetica* 21: 107-116.

Klatt, D.H. 1976. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59: 1208-1221.

Knetschke, E. & Sperlbaum, M. 1987. *Zur Orthoepie der Plosiva in der deutschen Hochsprache. Eine auditiv-komparative Untersuchung*. Tübingen: Niemeyer.

Kohler, K.J. 1994. Glottal stops and glottalization in German. *Phonetica* 51: 38-51.

Kohler, K.J. 1995 (2). *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag.

Krech, E.-M. 1968. *Sprechwissenschaftlich-phonetische Untersuchungen zum Gebrauch des Glottisschlageinsatzes in der allgemeinen deutschen Hochlautung*. Basel & New York: S. Karger.

Krech, E.-M. et al. 1982. *Großes Wörterbuch der deutschen Aussprache*. Leipzig: VEB Bibliographisches Institut.

Ladefoged, P., Williamson, K., Elugbe, B. & Sister A.A. Uwalaka. 1976. The stops of Owerri Igbo. *Studies in African Linguistics* Supplement 6: 147-163.

Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W.J. & Marchal, A. (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer. 403-439.

Lindqvist, J. 1972. Laryngeal articulation studied on Swedish subjects. *Speech Transmission Laboratory Charterly Progress and Status Report. Royal Institute of Technology, Stockholm* 2-2: 10-27.

Lisker, L. & Baer, T. 1984. Laryngeal management at utterance-internal word boundary in American English. *Language and Speech* 27: 163-171.

Löfqvist, A. 1978. Laryngeal articulation and junctures in the production of Swedish obstruent sequences. In Gårding, E., Bruce, G. & Bannert, R. (eds.) *Nordic Prosody*. Lund: Lund University Press. 73-83.

Löfqvist, A. 1992. Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. *Language and Speech* 35: 15-28.

Löfqvist, A. & Yoshioka, H. 1980. Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America* 68: 792-801.

Löfqvist, A. & Yoshioka, H. 1981. Laryngeal activity in Icelandic obstruent production. *Nordic Journal of Linguistics* 4: 1-18.

Löfqvist, A. & McGowan, R.S. 1992. Influence of consonantal environment on voice source aerodynamics. *Journal of Phonetics* 20: 93-110.

Lotzmann, G. 1975. *Zur Aspiration der Explosivae im Deutschen. Ein sprechwissenschaftlich-phonetischer Beitrag zur deutschen Hochlautung*. Göppingen: Kümmerle.

Mangold, M. 1990 (3). *DUDEN Ausspracheworterbuch*. Mannhein etc.: Dudenverlag.

McCarthy, J.J. 1988. Feature geometry and dependency: a review. *Phonetica* 43: 84-108.

Moulton, W.G. 1947. Juncture in Modern Standard German. *Language* 23: 212-226.

Munhall, K. & Löfqvist, A. 1992. Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics* 20: 111-126.

Nespor, M. & Vogel, I. 1986. *Prosodic phonology*. Dordrecht: Foris.

Perkell, J.S., Matthies, M.L., Svirsky, M.A. & Jordan, M.I. 1995. Goal-based speech motor control: a theoretical framework and some preliminary data. *Journal of Phonetics* 23: 23-35.

Pétursson, M. 1977. Timing of glottal events in the production of aspiration after [s]. *Journal of Phonetics* 5: 205-212.

Sawashima, M. & Hirose, H. 1983. Laryngeal gestures in speech production. In MacNeilage, P.F. (ed.) *The production of speech*. New York etc.: Springer. 11-38.

Stock, D. 1971. *Untersuchungen zur Stimmhaftigkeit hochdeutscher Phonemrealisationen*. Hamburg: Buske.

Trubetzkoy, N.S. 1939. *Grundzüge der Phonologie*. 1958 by Göttingen: Vandenhoeck & Ruprecht.

Wiese, R. 1996. *The phonology of German*. Oxford: Oxford University Press.

Yoshioka, H., Löfqvist, A. & Hirose, H. 1981. Laryngeal adjustments in the production of consonant clusters and geminates in American English. *Journal of the Acoustical Society of America* 70: 1615-1623.

Yoshioka, H., Löfqvist, A. & Hirose, H. 1982. Laryngeal adjustments in Japanese voiceless sound production. *Journal of Phonetics* 10: 1-10.

Yoshioka, H., Löfqvist, A. & Collier, R. 1982. Laryngeal adjustments in Dutch voiceless obstruent production. *Annual Bulletin. Research Institute of Logopedics and Phoniatrics. University of Tokyo* 16: 27-35.

Yu, S.-T. 1992. *Unterspezifikation in der Phonologie des Deutschen*. Tübingen: Niemeyer.

---

* The entire transillumination experiment, of which only certain aspects are presented and discussed here, was supported by Grant DC-00865 from the National Institute on Deafness and Other Communication Disorders to Haskins Laboratories (cf. Jessen 1995, Jessen, to appear).

[1]Sequences of identical adjacent obstruents are refered to as "geminates" in the literature discussed here, even if a word boundary intervenes (see in particular Yoshioka, Löfqvist & Hirose 1981). In most of the phonological literature the term "(true) geminate" is restricted to the case that no morpheme- or word boundary intervenes between the identical obstruents (see Hayes 1986). In this sense the word-internal identical sequences of Japanese, investigated with transillumination and other methods by Yoshioka, Löfqvist & Hirose (1982), constitute true geminates in contrast to the "fake geminates" of English, Swedish etc. However, with respect to the presence of a single glottal opening gesture the Japanese true geminates in most cases behave like the fake geminates reported for the Germanic languages.

[2]The author favors a definition of aspiration duration in which the end of aspiration is not measured as the beginning of voicing in the following vowel (i.e. positive Voice Onset Time), but as the end of aspiration turbulence into the following vowel. One practical approximation of this latter event is the beginning of the second formant in the following vowel (cf. Fischer-Jørgensen & Hutters 1981, Davis 1994, Jessen to appear). Among the disadvantages of the VOT concept is the difficulty of measuring aspiration in languages with a voiced aspirated stop category (Ladefoged et al. 1976, Davis 1994 about this point).

[3]Another timing parameter – P-R (interval between stop release and peak glottal opening) – was statistically more robust than OG-OC and P-OC in the expression of place of articulation. For the word-initial context discussed here P-R was in the order velar > alveolar > labial and turned out to be well correlated with aspiration duration (Jessen 1995, to appear; similarly Cooper 1991 for English). P-R was not listed in Table 1 because it is not applicable to fricatives.

[4]An early coordination of glottal opening in /p/ and the subsequent merger with the glottal opening of a preceding /s/ across a word boundary is also reported by Lisker & Baer (1984) for English.

[5]To say that /h/ is common word-initially is insufficient insofar as /h/ is also common at the beginning of certain prefixes and suffixes (e.g. *hin-*, *her-*; *-haft*, *-heit*), as well as at the beginning of lexical stems, whether occuring in isolation or in combination with other morphemes. One might argue on the basis of these facts that /h/ is more of an index for morpheme boundaries than for word boundaries. However, there is evidence that the appropriate domain is larger than the morpheme and constitutes what is referred to as the "phonological word" by several phonologists.

The occurrence of /h/ is also influenced by the stress level of the syllable it initiates (although there are exceptions to this generalization). For this reason it has been suggested that the "foot" might be more adequate to characterize the domain of the occurrence of /h/ than the phonological word (see Wiese 1996: 60). Althogether, word boundary sensitivity and stress sensitivity are often related in intricate ways in languages such as English or German (Wiese 1996: 72; cf. also Cooper 1991 and Note 7). Notice also that in prosodic phonology the "foot" is not independent of the "phonological word", but strictly dominated by it. This implies that the beginning of the phonological word is also the beginning of a foot (under the application of the "strict layer hypothesis"; see Nespor & Vogel 1986). Thus, both word boundary sensitivity and stress sensitivity are encoded in the concept of the foot in prosodic phonology.

[6]In the example shown in Figure 7 word-initial /h/ occurs between two vowels. Firstly, we need to keep in mind that in words like *Uhu*, *Ahorn*, mentioned above, /h/ also occurs intervocalically and is likely to be produced with substantial voicing. Thus, these words would constitute false alarms to word boundary marking in the sense that in those cases a voiced glottal opening gesture does not indicate a word boundary. Secondly, word-initial /h/ is not voiced to the same extent or not at all when preceded by a voiceless obstruent such as /ʃ/. This is not shown here, but found in productions of the utterance *rasch hier* 'quickly here' (but cf. Stock 1971: 100ff. for several cases of voiced /h/ even in this context).

[7]Again, we need to keep in mind that this is a simplification insofar as the glottal stop can also occur at the beginning of prefixes and stems (though not suffixes) that occur word-internally. Some phonologists generalize the appropriate position as the "phonological word" (see the discussion in Wiese 1996: 72), while others prefer the "foot" as the most appropriate domain for glottal stop insertion, due to its stress sensitivity (cf. Note 5). Proponents of this solution include Giegerich (1989: 62ff.), Hall (1992: 58f.), Yu (1992: 84ff.), and Wiese (1996: 58ff.).

[8]However, the value of aspiration as a word boundary marker in German has to be seen as a gradient, not a categorical phenomenon. That is, aspiration is usually longer word-initially than word-medially, but it is not the case that it is present in the former and absent in the latter context. There is sufficient evidence that a non-negligible amount of aspiration duration exists in word-medial position before schwa (see Jessen, to appear for data and discussion of the literature). It should also be mentioned that along with glottal stop insertion, aspiration has been analyzed as foot-dependent by several phonologists (see the references in Note 7).