

Sonorant devoicing and the phonetic realization of [spread glottis] in English*

Ayako Tsuchida, Abigail C. Cohn and Masanobu Kumada¹

While sonorant devoicing in English is often described as an allophonic rule, Browman and Goldstein (1986) and Iverson and Salmons (1995) show that such an approach misses many aspects of the observed patterns. Both argue for an analysis which focuses instead on spread glottal gestures. We test their specific claims about the realization of laryngeal gestures phonetically, using fiberoptic and acoustic data. Results show that while Iverson and Salmons (1995) and Browman and Goldstein (1986)'s accounts capture the broad strokes of the patterns, there are additional complexities including lack of a fixed duration for the glottal opening gesture and variation in the coordination of peak glottal opening with oral articulations. We offer an analysis which assumes both phonological and phonetic components of the grammar.

1 Introduction

It is widely known that voiceless stops in English are aspirated in certain positions, notably in the onset of a stressed syllable and word-initially. Another well-known fact about the sound pattern of English is that sonorants devoice after voiceless obstruents (both stops and fricatives). Neither aspiration nor sonorant devoicing occurs in clusters with an initial [s]. Examples are given below (transcriptions following Iverson and Salmons 1995):

(1) pea [ph]	plea [p]	speak [sp]
fee [f]	flee [fl] ²	spleen [spl]

It is generally observed that the devoicing of a sonorant may be partial and that the extent of devoicing differs depending on the manner of articulation of the preceding consonant. Namely, there is less devoicing following fricatives than following stops (Klatt 1975). It should be noted that this difference is not captured in the transcriptions in (1).

* This work was supported in part by NSF Research Planning Grant #BNS-9111155 to A. Cohn. We would also like to thank Anders Löfqvist for his generosity in sponsoring the acquisition and analysis of these articulatory data at Haskins Laboratories and acknowledge the support of GRANT DC-00865 to A. Löfqvist. A version of this paper was presented at the 73rd meeting of the Linguistic Society of America in Los Angeles and we would like to thank the audience as well as members of the Cornell Phonetics Lab for their helpful questions and comments.

¹ Haskins Laboratory, New Haven CT, and Department of Physiology I, National Defense Medical College, Saitama, Japan.

² Our own data as well as close reading of previous literature suggest that a more appropriate broad transcription is [fl], as the amount of devoicing, when observed, is quite small.

Traditional analyses treat both stop aspiration and sonorant devoicing as allophonic variation: aspiration as the addition of a feature, and sonorant devoicing as the spread of [-voice]. Yet, as discussed by Browman and Goldstein (1986) and Iverson and Salmons (1995) among others, such an approach misses a number of aspects of the observed patterns. The parallel between the distribution of aspiration and devoicing is not captured. There are additional problems including choice of feature, use of [-voice] argued by some to be privative, and so forth.

Browman and Goldstein (1986) and Iverson and Salmons (1995) argue that if we look at these patterns in terms of glottal gestures, a more insightful analysis emerges, accounting for the basic pattern of sonorant devoicing in both stops and fricatives. While there are some differences between the two analyses (based in large part on the assumptions about the nature of phonology), they offer similar accounts and make similar claims about the observed events. Under both of these approaches, the role of phonology is minimal or non-existent and the phonetic implementation is predicted to be transparent. Notably it is assumed that the glottal opening gesture has constant duration (Kim 1970) and is aligned differently for stops (at releases) and fricatives (midpoint or onset). Overall, close correlations between duration of glottal opening and aspiration/devoicing are posited.

While neither work offers actual derivations for a full range of cases, the following schematic illustration, showing glottal opening for each case, is consistent with these proposals:

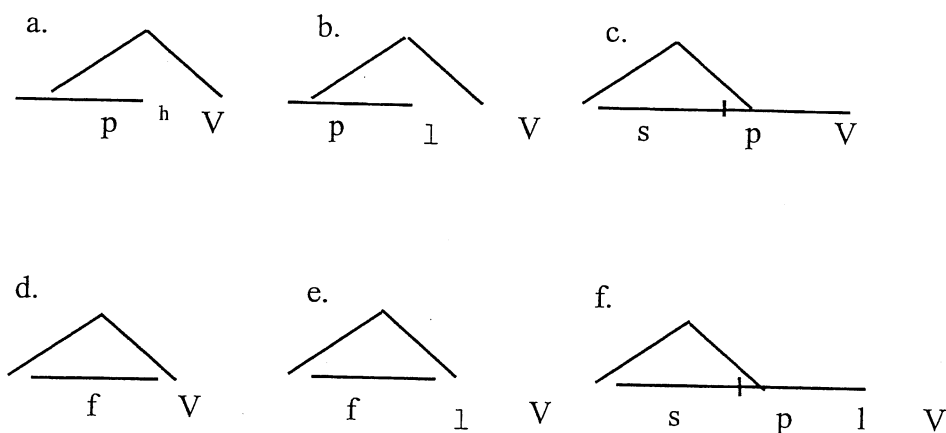


Figure 1. Schematic illustration

Crucially, it is assumed that there is only one glottal opening per onset and the opening is constant for all cases. Single voiceless stops, represented in (a), are aspirated since the glottal opening peak occurs at the release. As illustrated in (b), when a liquid follows a

stop, it is devoiced since the glottis is still open during its production. When there is a fricative, the glottal opening is aligned at its midpoint—this is exemplified in (c) through (f). No aspiration occurs after a fricative, since the glottis is closed by the time of the stop onset (c) or the vowel onset (d). Similarly, little or no sonorant devoicing occurs in fricative-liquid clusters or in triconsonantal clusters, since the glottis is mostly closed during the production of sonorants as can be seen in (e) and (f).

It should be noted that the illustrations presented in Figure 1 are interpretations of assumptions and that neither Browman and Goldstein (1986) nor Iverson and Salmons (1995) present actual glottal gestures in running speech. In the present paper, we test these claims phonetically. Since it is very hard to fully investigate glottal gestures from acoustic data alone (see Tsuchida 1997), we present fiberoptic and acoustic data from a speaker of American English. It will be shown that the actual glottal patterns are more complicated than the assumptions made by either Browman and Goldstein (1986) or Iverson and Salmons (1995). Our goal is to provide an adequate account of both the phonological patterns and phonetic realization of these facts.

2 Method

The experimental method we used was transillumination, where a flexible fiberoptic scope was inserted through the subject's nasal cavity and the amount of light from the scope passing through the glottis was collected as the indication of the glottal opening size (Löfqvist and Yoshioka 1980). A female speaker of American English produced five repetitions of stimuli, framed in a sentence *Say ... again*. The audio signal was recorded at the same time, thus we obtained both acoustic and articulatory data.

Earlier studies of spread glottal gestures investigated limited sets of voiceless consonants and obstruent clusters (Yoshioka et al. 1981; Löfqvist and McGowan 1992). We set out to look at a full set of single voiceless consonants and full combination of voiceless onset clusters, including obstruent-liquid clusters. The structure of the wordlist is given in (2):

- | | |
|-------|-----------------------------------|
| (2) T | T=voiceless stop ([p, t, k]) |
| S | S=voiceless fricative ([f, s, ʃ]) |
| TL | L=liquid ([l]) |
| SL | |
| ST | |
| STL | |

These clusters occurred both in initial and medial positions, in stressed and unstressed syllables. In this paper, we focus on the initial stressed position. For clusters containing liquids, we looked at both [l] and [r], but we will focus here on [l], as stop-[r] clusters showed some further complications due to durational properties which are beyond the scope of the present discussion.

The recordings were digitized on a VAX computer and stored as files to be processed by the HADES software developed at Haskins Laboratories. Wideband spectrograms and waveforms were made for each stimulus. Segmental boundaries in each token were labeled. The onset of the word-initial obstruent was determined as the point immediately following the preceding vowel's periodicity. In single obstruents, their offset was defined as the point immediately prior to the initiation of the following vowel's first periodicity. In clusters SL, ST and STL, the fricative offset (which at the same time served as the onset of the following segment) was the point where frication in the waveform and spectrogram ended. In clusters TL and STL, the stop offset was the point where aspiration in the waveform and spectrogram ended. For liquids, when they were not voiced at their onset, the point where voicing started was also labeled. The [l] offset was clearly observed as a discontinuity in the spectrogram. Finally, stop releases were also labeled.

An example of our data is provided in Figure 2. The top two panels show the waveform and spectrogram for the word *pea* uttered in the frame sentence. The vowel onset and offset, as well as the stop release are labeled. The third panel shows the transillumination data. The vertical axis shows the amount of light. The graph reveals that there was much light going through the glottis during the production of [p], indicating that the glottis was open. The fourth panel shows the derivative of the transillumination data — this could be considered as the indication of glottal opening and closing velocity. We determined the onset (ON) and offset (OFF) of the glottal opening as the points where the velocity was 10% of the peak velocity value (P1) and minimum velocity value (V1), respectively.³ All of these points are marked in the fourth panel. In what follows, we use the term *opening phase* for the time it took for the glottis to reach peak opening, that is, the duration between ON and Z2, and *closing phase* for the amount of time it took for the glottis to close, namely, the duration between Z2 and OFF.

³ Velocity was used since it provides a systematic way of locating the point where the transillumination data shows a peak and a valley (namely, where the velocity is 0).

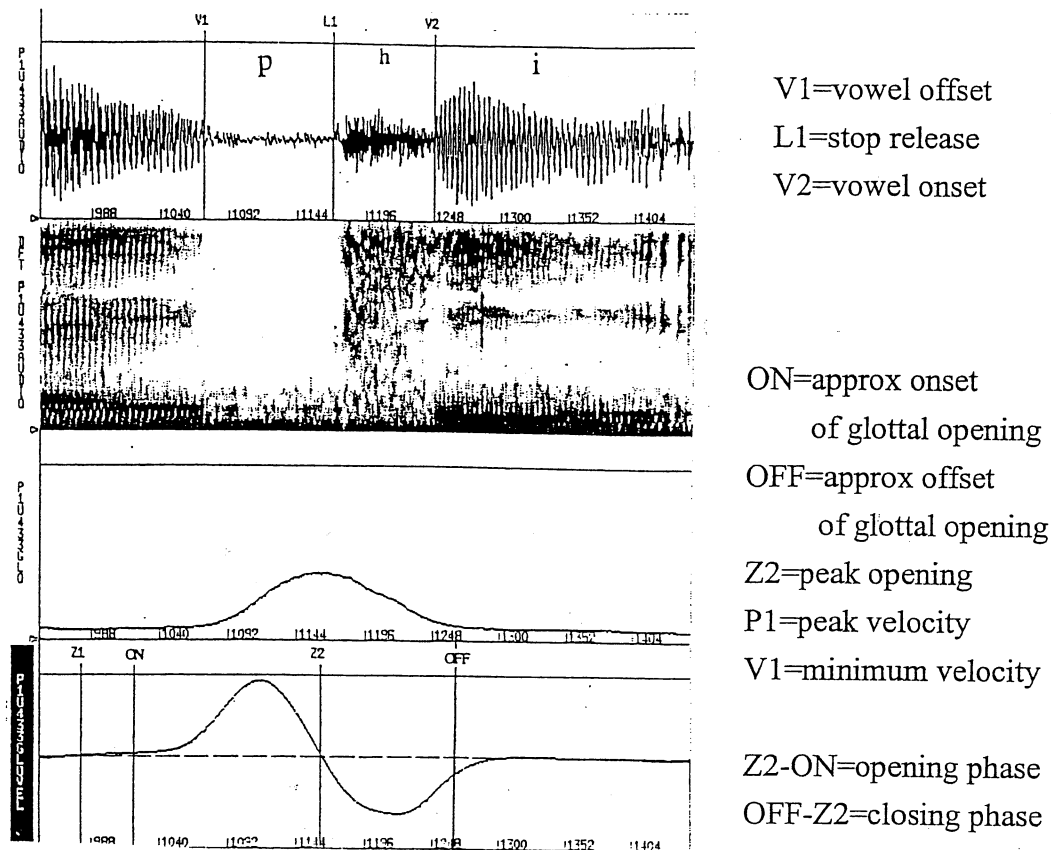


Figure 2. Example of acoustic and transillumination data

3 Results

3.1 Duration of glottal opening gestures: single stops and fricatives

First, let us see if the duration of glottal opening gestures is constant. The measurements from single stops and single fricatives (T vs. S) are presented in Figure 3. Contrary to what Browman and Goldstein (1986) and Iverson and Salmons (1995) assume, the glottal opening duration is quite different between stops and fricatives in that fricative gestures are much longer than stop ones (a similar finding has been reported by Löfqvist and Yoshioka 1981). When we examine the opening phase and closing phase, we see that glottal gestures are not symmetrical and that most of the difference is in the closing phase, while the opening phase is almost constant.

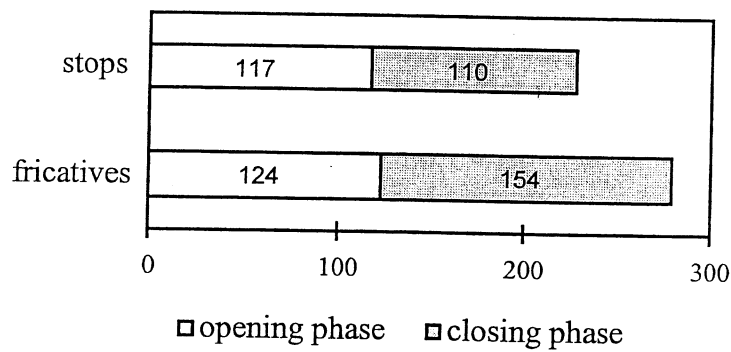


Figure 3. Duration of glottal opening gestures: single stops vs. fricatives (in ms.)

From these results, it is clear that glottal gestures are not simple ballistic events. The difference between stops and fricatives rather indicates that much of glottal gesture is under speaker control (see also Yoshioka et al. 1981).

Of interest is also the question of magnitude of the glottal opening, especially since it has been suggested that there is a direct link between magnitude and duration. We cannot directly investigate magnitude with the current data since amount of light cannot be correlated directly with degree of glottal opening. However, if velocity is not predictable (and we know it is not since the glottal gestures are not symmetrical), then there cannot be a complete linkage between magnitude and duration — the relationship must be more complicated.

3.2 Alignment of glottal gestures to oral events: single stops and fricatives

Let us now turn to the alignment of glottal gestures to oral events. Figure 4 provides oral event durations and alignment of glottal gestures in single stops:

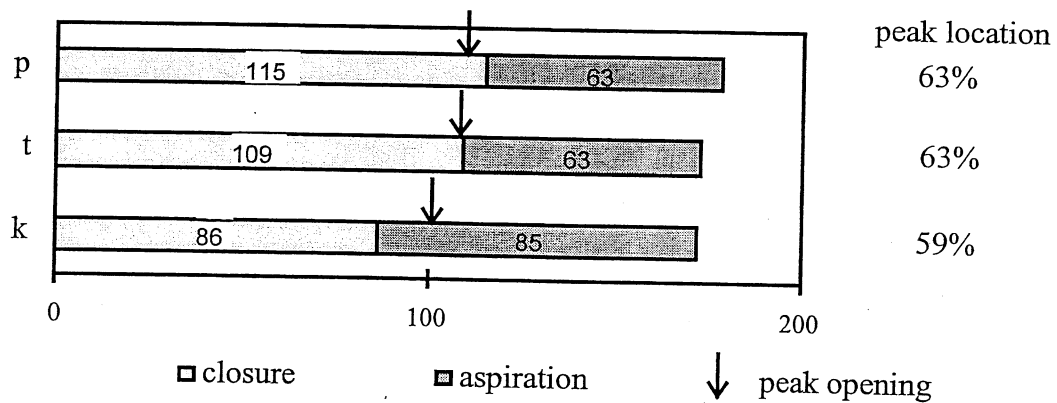


Figure 4. Alignment of glottal gestures to oral events: single stops

Of interest here is whether the peak opening is coordinated with the stop release, as has been assumed by Browman and Goldstein (1986). The results show that peak glottal opening is only loosely aligned with stop release. Variation occurs based on place of articulation: [p]'s peak opening often occurs before the burst, whereas [t] and [k]'s peak opening occurs either at or after the burst. It should be noted that there is an inverse relationship between closure duration and aspiration from front to back of the oral cavity (Hutter 1985), resulting in relatively stable combined duration. Interestingly, when we calculate the location of the peak with respect to total duration (i.e. closure and aspiration combined), the results are almost constant — 63% for [p] and [t], 59% for [k]. This stability of location suggests coordination with the onset of stop closure rather than the release (consistent with Kingston 1990).

Let us now consider the alignment in fricatives. Figure 5 provides the durational measurement for [f], [s] and [ʃ] with the location of the glottal peak:

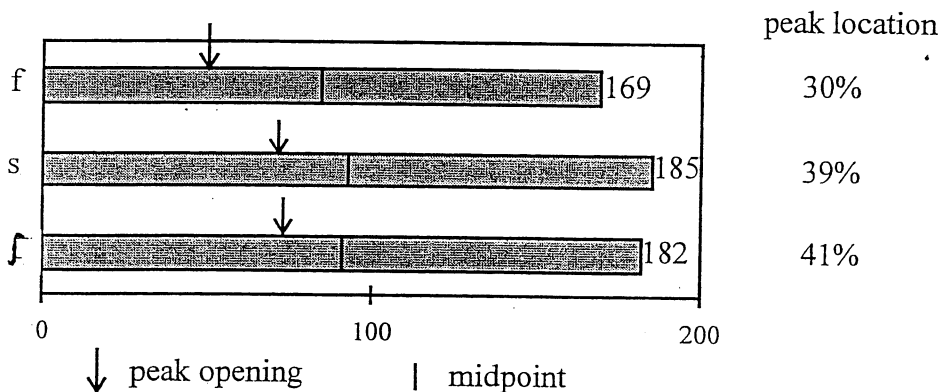


Figure 5. Alignment of glottal gesture to oral events: single fricatives

The question to be addressed here is whether the peak opening is aligned with the midpoint of fricative duration. The results clearly show that it is not the case. The peak occurs more towards the beginning. When we calculate the location of the peak with respect to total fricative duration, the opening is also not consistently aligned with the onset of the fricative, unlike stops. There is greater variation in alignment of glottal gestures in fricatives than in stops, despite relatively constant glottal gesture duration. However, we also observe some similarity between stops and fricatives. Peak glottal opening occurs later going from front to back of the oral cavity.

To summarize, these results show that the assumptions Browman and Goldstein (1986) and Iverson and Salmons (1995) make for their analyses are overly simple. The

durations of glottal gestures are not constant between stops and fricatives. Peak glottal opening for stops seems to be coordinated with the onset of the stop, rather than the release. Peak opening occurs towards the beginning in fricatives.

It is interesting to note that the average durations of stops and fricatives are almost identical (174 ms and 179 ms, respectively), even though the glottal opening gestures show 51 ms difference. This indicates that segmental duration and glottal opening duration are not correlated in a straightforward fashion.

3.3 Duration of oral events

We now examine the degree of similarity between glottal opening gestures and their alignment in simple CV and (C)CCV cases.

We have already seen that glottal gestures for single stops and single fricatives are different. What about stop-liquid and fricative-stop and fricative-stop-liquid cases? First, we consider durations of oral events:

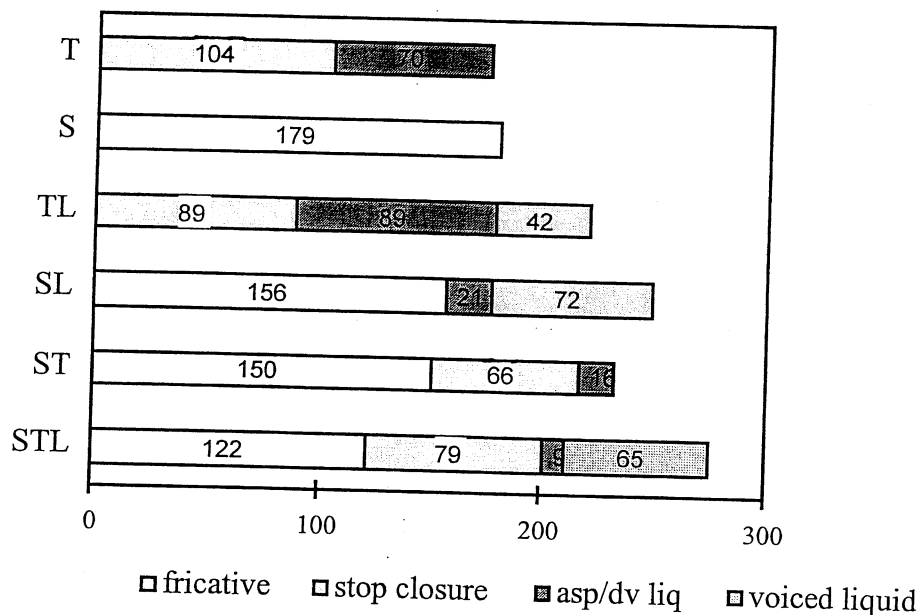


Figure 6. Oral durations in single consonants and clusters

In general, shortening is observed in clusters, most consistently so for fricatives. On average, single fricatives were 179 ms, while those in biconsonantal clusters were 150/156 ms and those in triconsonantal clusters were 122 ms. Haggard's (1973) observation of roughly 15% shortening seems to hold for the present data. Stops also

showed similar shortening, although our results were affected by unbalanced number of tokens per place of articulation.⁴

Combining the duration of the voiceless and voiced portions of [l], we find that [l] was notably more devoiced after stops than fricatives and that it was shorter in fricative-[l] than stop-[l] cases and shortest in the triconsonantal STL case. Hypothesizing that the [l] shows the expected shortening in the fricative-[l] case, the long duration of the [l] in the stop-[l] case is unexpected.

3.4 Glottal gestures and their alignment: all cases^{footnote}

Let us now turn to the glottal gestures and peak location in clusters:

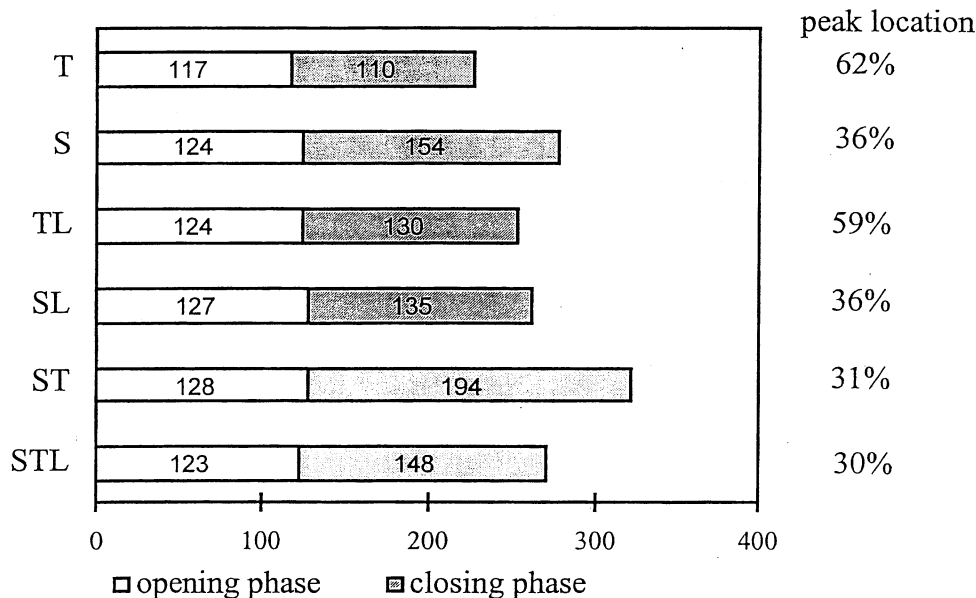


Figure 7. Duration of glottal gestures and peak location, all cases

We note variation in total glottal gestures, but the opening phase is strikingly constant across all cases. Observed differences occur during the closing phase.

In fricatives, decrease in duration of closing phase closely parallels decrease in fricative duration. For example, in the fricative-[l] case, the fricative was 23 ms shorter as we have seen in Figure 6 and the closing phase is reduced by 19ms. The closing phase in stop-[l], however, is unexpectedly long. Based on the patterns observed in fricatives,

⁴ It should be noted that in order to fully understand these data we need a model of oral duration, which is beyond the scope of the present project — specific models of consonant coordination in clusters are provided by Browman and Goldstein (1988), Byrd (1995), and Byrd and Tan (1996) among others.

we expect the closing phase to be shorter compared to single stop cases, but it is actually longer. Similar lengthening of the voiceless portion has been observed by Klatt (1975).

Another thing we note is marked lengthening of closing phase in the fricative-stop case. It is in fact longer than the triconsonantal clusters which contains the identical fricative-stop sequence. It should be recalled that both Browman and Goldstein (1986) and Iverson and Salmons (1995) have argued that [l] in triconsonantal clusters is not devoiced since the glottis opening gesture, which they assume to have constant duration, cannot extend to the third segment. The long closing phase in fricative-stop clusters, however, shows that things are more complex. It should be possible to sustain the opening gesture for a longer period in triconsonantal clusters as well (thus making at least part of [l] devoiced). The glottis must be *controlled* to close by the onset of [l] in triconsonantal clusters.

As far as peak location, calculated as a percentage of the entire voiceless portion, that for obstruent-[l] clusters is similar to that observed for single consonants of the same manner (roughly 1/3 for fricatives and 2/3 for stops). In fricative-stop and triconsonantal cases, the peak falls even further towards the beginning of the voiceless portion. When their peak location is recalculated with respect to the fricative duration only, it falls very close to the middle of the fricative (48% for ST and 52% for STL). This timing is much later compared to the single fricative case, suggesting a temporal compromise between the demands of the fricative and the stop.

In summary, there are numerous aspects of phonetic patterns that are more complex than expected based on Browman and Goldstein (1986) and Iverson and Salmons (1995)'s accounts. There are differences in glottal gestures between stops and fricatives, which carry over into clusters. The relationship between aspiration and sonorant devoicing is not as straightforward as predicted.

4. Proposed analysis

In this section, we sketch out an account of these facts. Crucial to our analysis is the assumption that phonology and phonetics are distinct (Cohn 1990, 1998; Tsuchida 1997, 1998; Zsiga 1999). We will first consider the phonological specifications of English consonants, then discuss their phonetic realizations.

For laryngeal specifications in the phonology, we assume that there are three privative features: [Voice], [Spread Glottis (SG)] and [Constricted Glottis (CG)] (Cho 1990; Lombardi 1991). While English obstruents have often been categorized as *voiced* vs. *voiceless*, Iverson and Salmons (1995) propose that the laryngeal contrast should be

captured in terms of the presence vs. lack of aspiration ([SG]). Specifically, they argue that voiceless obstruents in English are all marked as [SG], while their voiced counterparts lack any laryngeal feature. Under their analysis, the feature [Voice] does not play a contrastive role in English phonology.⁵ Their assumptions are summarized below:

(3) Laryngeal specifications in English (Iverson and Salmons 1995)

voiceless stops	[SG]
voiced stops	∅
voiceless fricatives	[SG]
voiced fricatives	∅

We agree with Iverson and Salmons (1995) that voiceless stops, at least in initial position, are contrastively specified as [SG] in English (we leave aside additional complications about specifications of stops in medial position). It is well known that the surface contrast in word-initial stops in English is in terms of aspiration, rather than voicing. However, assuming both voiceless stops and voiceless fricatives to have [SG] is problematic, given the different behavior of stops and fricatives in the present data. We argue instead that voiceless fricatives are unspecified for [SG]. The glottal gesture for fricatives arises in the phonetics and is *not* due to a (redundant) phonological specification. We argue that voiced fricatives are contrastively specified as [Voice] and that English is in effect a hybrid system using both [SG] and [Voice]. Further, we assume that the sonorants are unspecified. The system we are proposing is summarized below:

(4) Proposed specifications of English consonant (word-initial position)

	[SG]	[Voice]
voiceless stops	[SG]	∅
voiced stops	∅	∅
voiceless fricatives	∅	∅
voiced fricatives	∅	[Voice]
sonorants	∅	∅

⁵ It is not clear whether Iverson and Salmons (1995) are assuming redundant [Voice] specifications for sonorants.

not allowed in English. In the ST case (c), the glottal opening gesture arises from the [SG] specification as well as the aerodynamic requirement for the fricative. The combined effects result in the longest opening duration. Similar effects occur in the STL case, yet the opening is shorter than ST, since there are more segments in the onset and each duration is shortened.

We account for the quite distinct patterns of timing and glottal configuration between stop-[l] and fricative-[l] by assuming that fricative lacks a [SG] specification and thus fricative-[l] has no [SG] (see (e) above) — the more limited devoicing in this case is a phonetic effect.

It is crucial to make the distinction between phonology and phonetics to account for the present data, since otherwise we cannot explain why sonorants following aspirated stops and voiceless fricatives show distinct patterns. The liquid following an aspirated stop clearly shows robust devoicing, which cannot be explained simply in terms of glottal timing.

5. Conclusions

In the present study, it was shown that numerous factors affect the realization of spread glottal gestures and their coordination with oral events in English. While the approaches of Browman and Goldstein (1986) and Iverson and Salmons (1995) capture the broad strokes of the patterns, there are additional complexities which need to be accounted for. We have shown that the observed patterns can be adequately accounted for by assuming both phonological and phonetic components to the grammar.

Under our analysis, the glottal opening gestures in voiceless stops and voiceless fricatives arise differently in English. Specifically, we have argued that glottal opening in stops is a result of phonetic implementation of [SG], while that in fricatives is of aerodynamic requirements. Interestingly, Tsuchida (1997) has argued that in Japanese, a language which lacks aspirated stops, voiceless fricatives are phonologically specified for [SG]. Japanese voiceless fricatives are produced with a wide glottal opening, just as English ones. We assume that languages differ as to whether they promote the glottal opening for voiceless fricatives present in fricatives due to aerodynamic requirements to a phonological feature [SG] specification. Cross-linguistic investigation for understanding the typological patterns is warranted.

References

- Browman, C. and L. Goldstein (1986) Toward an articulatory phonology. *Phonology Yearbook* 3: 219-252.
- Browman, C. and L. Goldstein (1988) Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45: 140-155.
- Byrd, D. (1995) C-centers revisited. *Phonetica* 52: 285-306.
- Byrd, D. and C. Tan (1996) Saying consonant clusters quickly. *Journal of Phonetics* 24: 263-282.
- Cohn, A. (1990) Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics* 76.
- Cohn, A. (1998) The phonetics-phonology interface revisited: Where's phonetics? *Texas Linguistics Society 1998 Conference Proceedings*: pp. 25-40.
- Haggard, M. (1973) Abbreviation of consonants in English pre-and post-vocalic clusters. *Journal of Phonetics* 1: 9-24.
- Hutter, B. (1985) Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* 42: 1-24.
- Iverson, G. and J. Salmons (1995) Aspiration and laryngeal representation in Germanic. *Phonology* 12: 369-396.
- Kim, C.-W. (1970) A theory of aspiration. *Phonetica* 21: 107-116.
- Kingston, J. (1990) Articulatory binding. In J. Kingston and M. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press, pp. 406-434.
- Klatt, D. (1975) Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18: 686-706.
- Löfqvist, A., and R. S. McGowan (1992) Influence of consonantal environment on voice source aerodynamics. *Journal of Phonetics* 20: 93-110.
- Löfqvist A. and H. Yoshioka (1980) Laryngeal activity in Swedish obstruent clusters. *JASA* 68: 792-801.
- Löfqvist, A. and H. Yoshioka (1981) Interarticulator programming in obstruent production. *Phonetica* 38: 21-34.
- Tsuchida, A. (1997) *Phonetics and Phonology of Japanese Vowel Devoicing*. PhD dissertation, Cornell University.
- Tsuchida, A. (1998) Phonetic and phonological vowel devoicing in Japanese. *Texas Linguistics Society 1998 Conference Proceedings*: pp. 173-188.

Yoshioka, H., A. Löfqvist and H. Hirose (1981). Laryngeal adjustments in the production of consonant clusters and geminates in American English. *JASA* 70: 1615-1623.

Zsiga, E. (1999) *Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian*. Ms, Georgetown University.