Durations of Articulator Movements for /s/-Stop Clusters*

Gloria J. Borden† and Thomas Gay‡

ABSTRACT

Cine X-ray data derived by pellet tracking from three subjects indicate that there are certain basic strategies used for initial /sp/, /st/, and /sk/ clusters that reveal a remarkable economy of effort in tongue movement. In addition, the known acoustic shortening of /s/ before /p/ is seen to be a result of early lip closure relative to the tongue closure delayed by its involvement with the /s/.

In this paper we present some findings from an ongoing study of the fricative /s/ as it is produced by normal speakers either as a single consonant or in cluster with other consonants. In general, consonant clusters in speech are interesting to study because the individual events are so closely timed that they presuppose a complex motor program. The long-range purpose of the study is to investigate the relationships among electromyographic (EMG) recordings taken from tongue muscles, movement data recorded in the form of X-ray motion pictures, and acoustic measurements. Data have been collected thus far from three subjects; from two of these subjects we have acoustic and movement information and from one we have simultaneous EMG and cine X-ray data. The focus of this paper will be on a comparative analysis of the clusters /sp/, /st/, and /sk/ achieved by inspection of acoustic information and movement information as recorded on X-ray motion pictures.

*This is a combination of oral papers presented by the authors in 1975: "Production of /s/ in Clusters," presented at the 8th International Congress of Phonetic Sciences, Leeds, England, 17-23 August, and "The Relative Durations of Articulator Movements for /s/-Stop Clusters," presented at the 90th meeting of the Acoustical Society of America, San Francisco, 4-7 November.

† Also City College, City University of New York.

‡ Also University of Connecticut Health Center, Farmington.

Acknowledgment: This research is sponsored by the National Institute of Dental Research and the National Institute of Neurological Diseases and Stroke. We are grateful to them, to Dr. Seiji Niimi from the University of Tokyo, and to Dr. J. Daniel Subtelny in Rochester.

[HASKINS LABORATORIES: Status Report on Speech Research SR-44 (1975)]
Before noting the features common to all of the speakers, a few of the individual variations should be mentioned. Subject 1 produces /s/ with the tongue tip low behind the lower incisors and the blade high to form the tongue-palate constriction, whereas Subjects 2 and 3 elevate the tongue tip to a position posterior to the upper incisors. In all cases, it was the blade of the tongue, not the tip, forming the constriction, but for the tongue-tip-high subjects a more anterior portion of the tongue blade was involved. These two variant articulations of /s/ are common.

It was noted too that Subject 3 moves his tongue front and back more than the other subjects, but for all subjects vertical movement was more extensive than horizontal movement; that is, all subjects move their jaws, lips, and tongue primarily up and down with much less front-back movement.

The instrumentation used for the first subject has been reported by Borden and Gay (1975). The two subjects recorded since then were recorded and analyzed under a newer system. It is the more recent recording and analysis system that we shall describe in this paper.

Figure 1 shows the position of the subjects for the X-ray films. The subject's head was stabilized and lead pellets of 2.5-mm diameter were attached by means of a cyanoacrylate adhesive to the upper and lower lips and to three positions on the tongue—the tip, the blade, and the dorsum—approximately one inch apart, and a reference pellet was attached at the embrasure of the upper central incisors. The X-ray generator delivered 1-msec pulses at 100 kV to a 9-in. image intensifier tube. The film was recorded with a 16-mm camera at a speed of 60 frames per second.

---

Figure 1: Lateral X-ray motion pictures record movement of lead pellets attached to the tongue, with pellet on incisors used as reference.
The film was analyzed frame by frame. The method used will be described later, but there are certain strategies used for the production of /sp/, /st/, and /sk/ before the low vowel /a/ that should be noted here. The simplest way to view the strategies used by the subjects is to refer to a schematic representation of relative jaw, tongue, and lip movement. The acoustic signal and the movement data are lined up at the moment of /p/ closure.

The subjects had similar strategies for /st/ and /sk/, but one subject differed for /sp/. Therefore, two strategies will be shown for this cluster. The subjects agreed essentially on articulator movement irrespective of the high or low tongue tip difference.

Figure 2 represents the first strategy for producing the utterance /spop/: the simple synchronous opening of the jaw with the tongue and lower lip. The upper lip moves relatively little. Notice that the jaw that carries the tongue is free to lower for /a/ immediately upon lip closure for the preceding /p/.

Subjects 1 and 2, however, used a less straightforward strategy for /spop/ in which the tongue lowered before the jaw and lips. In Figure 3 the lineup point for the acoustic and movement data is lip closure for the first /p/. Again, the upper lip is steady as the lower lip is closing for the /p/. Note that as the lower lip is being elevated, the tongue is beginning to lower for /a/. The movements are not only asynchronous but in opposite directions, the lip coming up as the tongue is going down. The difference in strategy between these subjects and the first shown may be attributed to speaking rate, since the first simpler strategy was used with a relatively slow utterance of /spop/. Notice here too that the tongue lowers a bit on its own before being carried the rest of the way by the lowering of the jaw.

For /st/ all three subjects used the same strategy, which is abstracted in Figure 4. For /st/, instead of the tongue lowering as a unit, as was the case for /sp/, the tongue tip and the blade rise to form the blade-alveolar ridge occlusion for /t/, but the more posterior portion of the tongue is free to descend early for the low vowel. Note too that the blade is high for the /s/ as the tongue tip is rising to take a more active part in the stop closure. (Subject 1, who produced /s/ with the tongue tip low, uses this basic strategy for /st/, but the tongue tip is kept relatively low for the alveolar stop as well.) If we look at one example of this strategy for /st/, we can see the differential movement of the tongue pellets. Figure 5 illustrates the film analysis system used in this study. It was produced by first projecting the film image onto a writing surface via an overhead mirror system in order to mark pellet positions and frame numbers that are input to a small laboratory computer via a digitizing tablet. The computer measures the x and y coordinate positions relative to the reference pellet position, converts the measurements to millimeter distances, and stores the data on digital magnetic tape. A second program draws axes and plots the measured data on a display scope from which a hard copy is made.

The horizontal line represents the position of the reference pellet that was placed between the upper central incisors. The vertical line represents the lip closure for the /p/. The utterance here is /stpos/. The pellet on the tip of the tongue represented by a square can be seen to elevate slowly for the fricative /s/, rise even more for the alveolar stop /t/, and rapidly descend for the low vowel /a/. The blade pellet (the circles) also remains high for the /st/ and descends upon release of the stop. The more posteriorly placed tongue
/spap/

Strategy 1: Synchronous opening of jaw with tongue and lower lip.

```
Upper Lip steady
```

```
Jaw steady  Opening  Closing
```

```
Tongue high  Lowering  Rising
```

```
L. Lip steady  Closing  Opening  Closing
```

Figure 2: Schematic of relative movement of articulators for /spap/. Jaw lowers for /a/ immediately upon lip closure for /p/.
/spap/

Strategy 2: Tongue lowers before jaw and lips.

Upper Lip steady

Jaw steady Opening c 2cm Closing

Tongue high Lowering 2.5cm Low rising

L. Lip closing Opening c 2cm Closing

Figure 3: Schematic of asynchronous movements of articulators for /spap/. Tongue lowers as a unit.
/stəpə/

Strategy: Back of tongue descends early for /a/.

- **Upper Lip Steady**

- **Jaw steady** | **Lowering** | **Low rising**

- **L. Lip steady** | **Lowering** | **Rising**

- **Tongue Tip rising** | **Lowering** | **Low rising**

- **Blade high** | **Lowering** | **Low rising**

- **Dorsum up** | **Lowering** | **Low rising**

**Figure 4:** Schematic of differential tongue movement for /stəpə/. Posterior pellet lowers ahead of anterior pellets.
UP-DOWN MOVEMENT

/stapə/

- (TNG 1)
- (TNG 2)
- (TNG 3)

DISTANCE IN CENTIMETERS

Figure 5: Raw data of pellet positions for one subject during utterance /stapə/.
pellet, however, represented by the small triangle, can be seen to start its
descent for /a/ during the /st/ production, in this case three frames or 50 msec
before the more anterior portions of the tongue are free to move. The raw data
presented in the figures have not been smoothed. The occasional jumps in data
points are spurious and not physiologically.

Figure 6 shows that for /skap/ as well the tongue moves differentially and
not as a unit. In this case the more posterior part of the tongue is involved
with the stop and the anterior portion with the fricative, so that while the
tongue tip and blade are up for the /s/ constriction, the dorsum is rising for
/k/ and is not free to lower until the release of the occlusion. Note, however,
that first the tip lowers, then the blade. The tongue-tip-low subject followed
the same strategy as the other subjects with the exception that the tip remained
low, lowering further simultaneously with the blade as the more posterior dorsum
of the tongue concluded the /k/ gesture. The jaw movement is interesting in the
case of the /sk/ cluster. Recall that for /sp/ and /st/ the jaw remained high
during /s/, whereas here for /sk/ it starts to lower during the fricative. Not
only do the high lip position necessary for /p/ and the high tongue position
necessary for /t/ put constraints upon jaw lowering, but jaw lowering may facili-
tate the elevation of the back of the tongue necessary for the /k/. To take
one example of the differential tongue movement for /sk/, it can be seen in
Figure 7 that the more anteriorly placed pellets lower before the dorsal
pellet, in the case of the tip, by 50 msec and the blade, by 17 msec. Figure 8
represents horizontal tongue movement. This is the subject who used more front-
ing and backing of the tongue than the other subjects. Here it can be seen that
the tongue tip started back for /a/ 83 msec before the dorsum and the blade,
33 msec before the dorsal pellet. It seems then that when the tongue is not
involved in the stop for an /s/-stop cluster, it is free to move toward the
vowel immediately after its involvement with /s/ production in /sp/, whereas for
/st/ and /sk/, since the tongue is involved in both parts of the cluster, only
the uninvolved portions of the tongue are free to move toward the vowel position
during the consonant cluster production.

The strategies used by all three subjects for /st/ and /sk/ indicate a re-
markable economy of effort in that only those portions of the tongue primarily
involved in a high tongue gesture remain elevated. Other parts of the tongue
lower for the following vowel as soon as they are free to do so; for /st/, the
more anterior portion of the tongue is involved for both /s/ and /t/, but the
dorsum lowers early; whereas for /sk/, the reverse is true, with the anterior
portions of the tongue lowering for the vowel during the dorsal elevation for
the /k/. These observations are compatible with Öhman's (1966) ideas about con-
sonant and vowel articulation; that is, movements for the vowel can occur during
production of the preceding consonant.

It has been reported by Schwartz (1970) and by Klatt (1971), among others,
that the duration of the acoustic noise for /s/ is shortened when clustered with
a stop, and also that there is a tendency for the /s/ in /sp/ to be shorter than
the /s/ in /st/ or /sk/. To assess acoustic durations, spectrograms of all the
subjects' speech were measured to the nearest one-tenth of an inch and the mea-
surements were converted into milliseconds. The results are shown in Figure 9.
The subjects were consistent in the shortening of the /s/ before the labial stop
relative to the palatal stops. The differences were approximately 20 and 10 msec
for the first subject, 9 msec for the second, and 18 and 37 msec for the third
/skæpə/

Strategy: Tongue starts lowering (and backing) tip first, then anterior blade, then post blade.

U. Lip steady

Jaw raised  Lowering  Rising

L. Lip steady  Lowering  Rising

Tongue Tip up  Lowering  Rising
Blade up  Lowering  Rising
Dorsum rising ø up  Lowering  Rising

Figure 6: Schematic of differential tongue movement for /skæpə/. Anterior pellets lower before posterior pellet.
Figure 7: Raw data of pellet positions for one subject during utterance /skapə/.
Figure 8: Raw data of horizontal tongue pellet movement for one subject during utterance /skæpə/.
subject. Note that the third subject, GS, was remarkably consistent in both stop-closure duration and vowel durations.

<table>
<thead>
<tr>
<th></th>
<th>s</th>
<th>p</th>
<th>a</th>
<th>s</th>
<th>t</th>
<th>a</th>
<th>s</th>
<th>k</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB</td>
<td>130</td>
<td>112</td>
<td>168</td>
<td>150</td>
<td>70</td>
<td>200</td>
<td>140</td>
<td>103</td>
<td>200</td>
</tr>
<tr>
<td>FC</td>
<td>94</td>
<td>112</td>
<td>140</td>
<td>103</td>
<td>131</td>
<td>131</td>
<td>103</td>
<td>103</td>
<td>159</td>
</tr>
<tr>
<td>GS</td>
<td>94</td>
<td>75</td>
<td>150</td>
<td>112</td>
<td>75</td>
<td>150</td>
<td>131</td>
<td>75</td>
<td>150</td>
</tr>
</tbody>
</table>

Figure 9: Acoustic durations in msec.

Figure 10 relates the acoustic data to the movement data by showing the critical movements for the stop closure for /p/, /t/, and /k/. The subject in this figure is the one who conveniently produced stop closures of 75 msec and vowels of 150 msec across these clusters. Using the closure of the lips after /a/ as the lineup point for the movement comparisons, we can see that the back pellet (marked T. Dorsum) rises during /s/, continues rising to a maximum, during the /k/ closure, and then moves down for /a/. The pellet on the tongue tip seemed critical for the /t/, and the pellet on the lower lip, for the /p/. The difference in timing is evident here but it is easier to compare them if we track from left to right by changing the lineup point.

In Figure 11, we have lined up the tongue- and lip-movement curves at the onset of /s/, since that is the segment differing in duration, and it is obvious that the movement toward closure for the stop is temporally different for /p/, /t/, or /k/. The peak amplitude for the lower-lip movement for /p/ occurs closer in time to the onset of the /s/ noise than does the peak amplitude of movement of the tongue for either /t/ or /k/.

The /s/ seems to be shorter in the case of /sp/, not because the lips close with any more force or velocity, but because they start earlier and are not involved in any conflicting gesture. The lips are free to close. The tongue, however, is involved with the /s/ constriction and movement toward closure is delayed.

This effect is reflected in the EMG data recorded from the first subject. Unfortunately, we are unable to include the EMG data in this paper owing to some inconsistencies that must be rechecked. However, we can report that the orbicularis oris muscle is active in the case of /sp/ before the superior longitudinal or styloglossus muscles show activity for the /t/ and /k/ closures, respectively. The observation is contaminated by the fact that, for that subject at least, the orbicularis oris is active not only for the /p/ but for the /s/.
Figure 10: Movements of pellets critical for /p/, /t/, and /k/ closure plotted against acoustic signal.

There are many problems in this combined-techniques approach. One is obtaining clear spectrograms in the presence of the noise of EMG and cine X-ray recording. The measurements are difficult and possible only because of the automatic gain control of the voiceprint that suppresses the background noise in the presence of a strong speech signal.

Another problem in making durational measurements is that both movement and EMG signals are difficult to segment. A tongue blade can move forward and up for both the /s/ and /t/, and one cannot determine clearly the durations of events attributable to each phoneme. For EMG signals there is a problem similar to the one we alluded to in the case of the lip muscles for /s/ and /p/.

Finally, some of the extremely fast adjustments that we find interesting are difficult to capture with techniques that only sample every 20 msec. Despite the difficulties, however, we feel this is an effort worth pursuing in order to explore the temporal aspects of what may be firmly time-locked speech events. We plan another simultaneous EMG and X-ray experiment to further that effort.
Figure 11: Movements of pellets critical for /p/, /t/, and /k/ closure plotted against acoustic signal. Lip closure for /p/ precedes tongue closure for /t/ or /k/.

REFERENCES


