

# Phonological and prosodic conditioning of /s/-retraction in American English

*Jacob B. Phillips*, University of Chicago

**Introduction:** Individuals vary significantly in sibilant production, making inter-speaker and cross-linguistic comparison of sibilants particularly challenging. One notable source of variation in American English is /s/-retraction, the process by which /s/ is articulated approaching [ʃ] in the context of /r/ (Shapiro 1995; Lawrence 2000; Mielke et al. 2010). This is particularly robust in /str/ clusters, such as ‘street’, and less commonly observed in /spr/ (‘spree’) and /skr/ (‘scream’) clusters (Baker et al. 2011). There is remarkable variance not only between individuals in their articulation of /s/ in these clusters, but also within individuals. At question in the literature is whether /s/-retraction can be described simply as coarticulation from /r/ to /s/, since many individuals are reported to exhibit categorical distinctions between /ʃ/ and /str/, on one hand, and prevocalic /s/ and /s{p,k}r/ clusters, on the other (Baker et al. 2011). This study seeks to question the treatment of /s/-retraction as coarticulation to better understand the observed variation.

One possible mechanism for targeting coarticulation but not phonologized sound changes is the manipulation of prosodic position. Much work has examined the effects of prosodic strengthening and lengthening, whereby segments adjacent to prosodic boundaries are articulated with longer durations and other acoustic cues. Particularly, coarticulation has been found to diminish adjacent to prosodic boundaries, as a result of lengthening of those phones and perhaps due to a speaker’s effort to mark the position and aid in lexical access (Cho 2004: for English vowel-to-vowel coarticulation). This suggests that /sCr/ clusters adjacent to a boundary should exhibit less retraction in these positions if it is best described as coarticulation or exhibit little or no effect if the change has been phonologized.

**Methods:** Stimuli were designed to manipulate the prosodic position of the target word (IP-initial and IP-medial), while controlling for syllable count, syntactic position of the target, and the adjacent phonological material. The target segments were /s/, prevocalic and preceding /{p,t,k}r/, and prevocalic /ʃ/. For each of the 4 target and 11 filler onsets + /eɪ/, the lexical word of the form C(CC)VC was selected with a frequency nearest 5 instances per million words (mean 4.81) following SUBTLEX<sub>US</sub>. There were a total of 32 stimuli ((4 target /s/-onsets + 1 target /ʃ/-onset + 11 fillers) × 2 prosodic positions). Each stimuli was repeated four times.

The participants were 31 native speakers of American English between the ages of 18 and 22, recruited at an American university, with 21 males and 10 females. Participants were seated in an isolated double-walled soundbooth and were asked to read the stimuli at a normal speaking rate with no specific emphasis, stress, or focus. The stimuli were presented using PsychoPy in four blocks of 32, with each block containing each target in each position. The order of the stimuli within each block was randomized. The pace of the study was determined by the participants.

Recordings were manually checked against the elicitation list for errors or disfluencies and were force-aligned using FAVE. A custom Praat script extracted the duration values for all FAVE-aligned segments. Center of Gravity (CoG) measurements were calculated using a 40 ms Hamming window with preemphasis at 80 Hz and a frequency range from 500 to 12000 Hz, centered at eleven points (at 10% increments of the fricative’s duration from 0% to 100%) during the fricative. Measurement from the first and last 20% of the fricative were not included in the analysis.

**Results & Discussion:** Center of gravity of /s/ was modeled using the linear mixed-effects regression function `lmer()` in the `lme4` package of R. TRIALNUMBER (1-128, scaled), PROSODIC-POSITION, (IP-Initial vs. IP-Medial, contrast coded), FOLLOWINGSEGMENT (NULL,p,t,k, treatment coded with NULL as base), TIMEPOINT (3-9, scaled), CoG-/sh/ (scaled) and SEX (M vs. F, contrast coded) were entered as fixed effects with by-subject intercepts and by-subject random slopes for PROSODICPOSITION\*FOLLOWINGSEGMENT.

The center of gravity of /s/ is significantly lowered with each following consonant /p,t,k/. The effect of following segment is temporally sensitive as indicated by the significant interaction TIME-POINT\*FOLLOWINGSEGMENT (for /p/,  $t = -9.49, p < 0.001$ ; for /t/,  $t = -9.71, p < 0.001$ ; for /k/,  $t = -7.46, p < 0.001$ ), illustrating the coarticulatory dynamics of /s/-retraction (1: Top Left).

Consistent with the findings of Jang (2011: for Korean sibilants), CoG is found to be lowered in IP-initial position across the board ( $t = -2.78, p < 0.01$ ). This contrasts with the findings of Cho (2004: for English vowels) that cues of coarticulation are diminished in phrase-initial positions. Additionally, while there was not a group level effect for the interaction of prosodic position and the following segment, the inclusion of by-subject random slopes for that interaction, which significantly improves model likelihood, illustrates that individuals vary with respect to the effects of prosodic conditioning of /s/-retraction in different phonological environments (1: Bottom Left & Right).

Additionally, the center of gravity of /f/ showed a significant effect in modeling the center of gravity of /s/ ( $t = -5.49, p < 0.001$ ), with center of gravity of /s/ lowering when center of gravity for /f/ was lower. This suggests that /s/-retraction is influenced by the phonological system of the individual, maximizing contrast between the sibilants (1: Top Right).

These findings suggest that higher-order phonological structure does indeed play a role in understanding the variation of /s/ production generally, but also illustrate that there is additional inter-speaker variation in how these structure affect articulations of /s/ in different phonological environments. Taken together, these findings demonstrate that while /s/-retraction may accurately described as coarticulation, there are also effects of prosodic structure and the phonological system in determining the distribution and degree of retraction.

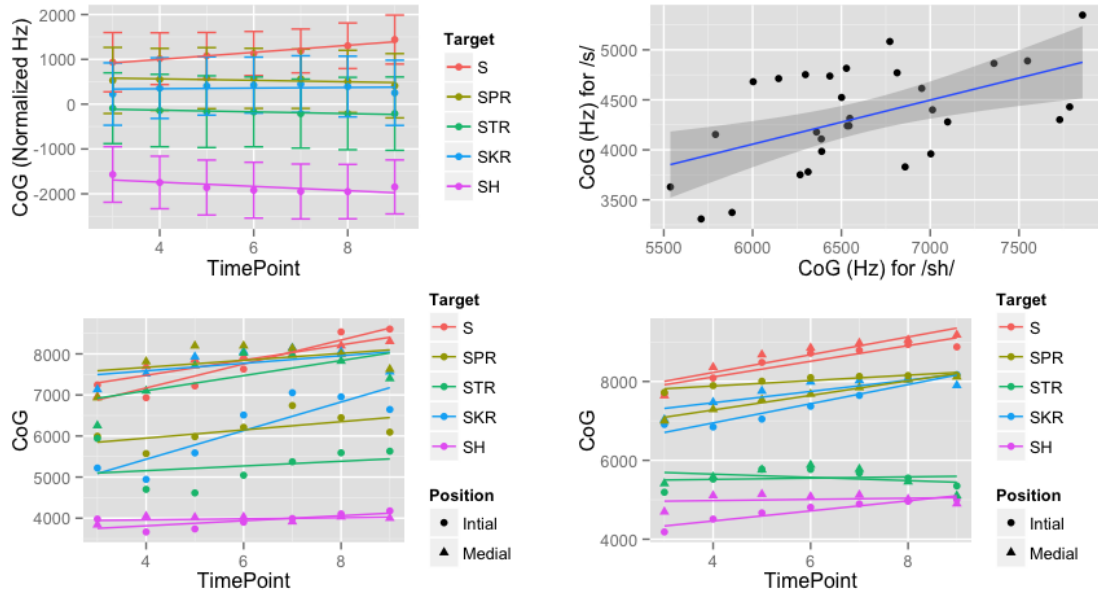


Figure 1: **Top Left:** Timecourse of retraction across individuals, showing the dynamic representations of /s/ in different phonological environments. **Top Right:** CoG for /sh/ and /s/, illustrating the effect of the phonological system on sibilant production. **Bottom Left:** Timecourse of retraction for an individual who shows an effect of PROSODIC-POSITION\*FOLLOWINGSEGMENT, with /sCr/ clusters showing very different trajectories in Initial vs. Medial positions. **Bottom Right:** Timecourse of retraction for an individual who does not show such an effect.

**Selected references** • Baker, A., D. Archangeli & J. Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *LVC* 23. 347–374. • Cho, T. 2004. Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *JPhon* 32(2). 141–176. • Jang, M. 2011. Prosodically driven phonetic properties in the production of Korean fricatives. *SPPM* 17(1). 65–86. • Lawrence, W. 2000. /str/ → /ftr/: Assimilation at a distance? *American Speech* 75(1). 82–87. • Mielke, J., A. Baker & D. Archangeli. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *LabPhon*. • Shapiro, M. 1995. A case of distant assimilation: /str/ → /ftr/. *AS* 70(1). 101–107.