This study addresses the retraction of sibilants in pre-consonantal position, a sound change that has taken place in the historical development of a small number of unrelated languages (e.g. German slagen > schlagen [ˈʃlaːɡən] ‘to beat’) but has attracted very little experimental attention. Whereas the origins of many sound changes can be readily identified based on known phonetic principles, the reasons why sibilants should come to be retracted in /sC/ clusters are not immediately obvious (cf. Bladon & Nolan 1977). Here we take English /st/ as a test case and consider acoustic, perceptual and physiological data in an effort to shed light on whether retraction in /sC/ can be considered a phonetically motivated sound change and, if so, to locate its origins in speech production or perception. By considering dynamic parameters (described below) we also aim to document how retraction unfolds over time, both acoustically and articulatorily. This study builds on a history of laboratory phonology investigations into other assimilation processes (e.g. Zsiga 1995) that have been fundamental to improving our understanding of the relationship between cognitive and physical aspects of speech.

The acoustic data involved native Australian English speakers’ productions of seam and steam in the carrier phrase Any ____ (20 Speakers x 10 repetitions x 2 words = 400 tokens). We calculated the first spectral moment (M1) over 500-12000 Hz on the power spectrum between the acoustic onset and offset of each sibilant token and also averaged M1 over the temporal middle half (25-75%) (henceforth ‘mean M1’). Figure 1 shows that M1 for the sibilant in /st/ (dashed line) was slightly lower, indicative of /s/-retraction, over the entire temporal middle half. A linear mixed model with mean M1 as the dependent variable, Context (/s, st/) as the fixed factor and speaker as a random factor confirmed the effect of the following stop on M1 ($\chi^2[1] = 9.5$, $p < 0.01$). Note that M1 stays relatively high towards the temporal end for /st/ i.e. acoustic retraction does not increase with proximity to the stop (as is instead the case for assimilation in /s#j/ clusters in English e.g. Zsiga 1995).

![Figure 1. M1 during the sibilant in seam and steam over linearly normalized time. Data were averaged for 20 speakers.](image)

We then tested in a perception experiment whether the acoustic retraction in /st/ seen in Figure 1 is audible to listeners and affects their categorization of the sibilant. To do so, we spliced the sibilant from steam and seam (i.e. the same tokens from the acoustic study just described) and prepended them to an /i:t/ produced by the same speaker. These /s/+/i:t/ tokens were presented to 22 native Australian English listeners in a forced choice perception test, with listeners choosing between SHEET/SEAT. Sibilants spliced from steam elicited more SHEET responses than sibilants spliced from seam (7% of responses vs. 3%, respectively). A logistic regression within a glmm with Response (SHEET/SEAT) as the dependent variable, Word (3 levels: seam, steam) as fixed factor and Speaker Voice (20 levels) and Listener Participant (22 levels) as random factors showed that listener responses were significantly affected by the word from which the sibilant was spliced ($\chi^2[1] = 13.8$, $p < 0.001$). In other words, acoustic retraction was audible to listeners, who occasionally categorized the sibilant in /st/ as /ʃ/.

However, the question remains as to the cause for acoustic retraction of sibilants in /st/ in the first place. On the one hand, there could be physiological (coarticulatory) conditioning for /s/-retraction in
/st/. Another possible explanation involves retraction in /str/, a common tendency for many English speakers which is motivated phonetically because of long-distance assimilation to the rhotic (e.g. Rutter 2011). That is, retraction in English /st/ might come about because an individual speaker generalizes their (phonetically-motivated) retracted sibilant pronunciation in /str/ to other similar clusters such as /st/ (cf. Janda and Joseph 2003). To investigate this issue we recorded both articular (EMA) and acoustic data from five native British English speakers who read ten times a randomized set of real English words containing word-initial /#sV, #stV, #strV/ in the carrier phrase another ____ app. We focused on British English so that we could examine the earliest stages of the /str/ assimilation process (i.e. before phonological reanalysis to /ʃtr/ which has arguably happened for some American English speakers). Acoustic analysis followed the same procedures described above; for the articulatory analysis we calculated the location of the tongue tip sensor on the horizontal plane (likely to carry the most relevant information) between the acoustic onset and offset of each sibilant.

Inspection of the physiological data showed that acoustic retraction in /st/ can be attributed to articulatory tongue tip retraction. Compare the data for two speakers in Figure 2: eng5 showed acoustic retraction in /st/ (the dashed line is lower than the solid line) whereas eng3 did not. These differences correspond well with the tongue tip data on the right: eng5 showed retraction on the horizontal plane for /stV/ (compared to /sV/) whereas speaker eng3 does not. Both speakers produce considerable acoustic retraction in /str/ (dotted lines) and that there is dynamic articulatory tongue-tip retraction over the course of the sibilant in this context (but not in /st/). Thus these data suggest that the acoustic retraction to which listeners were shown to be sensitive may have an articulatory basis. Nonetheless, we cannot offer an explanation as to why speakers retract the tongue tip for the sibilant in /stV/.

In conclusion, the acoustic and perceptual evidence of a bias towards retraction in /st/ described thus far may help to explain historical developments in a small number of languages in which /st/ > /ʃt/ has occurred: that is, these acoustic, perceptual and articulatory results show converging evidence that /s/-retraction in /st/ is a phonetically motivated sound change. The strong tendency towards retraction in English /str/ might help to promote phonologization of retraction in /st/ as individuals generalize over remembered tokens.