Language Variation and Change, **23** (2011), 347–374.
© Cambridge University Press, 2011 0954-3945/11 $16.00
doi:10.1017/S0954394511000135

Variability in American English s-retraction suggests a solution to the actuation problem

ADAM BAKER
Academy of Sciences of Afghanistan

DIANA ARCHANGELI
University of Arizona

JEFF MIELKE
University of Ottawa

ABSTRACT

Although formulated by Weinreich, Labov, and Herzog in 1968, the actuation problem has remained an unsolved problem in understanding sound change: if sound change is conceived as the accumulation of coarticulation, and coarticulation is widespread, how can some speech communities resist phonetic pressure to change? We present data from American English s-retraction that suggest a partial solution. S-retraction is the phenomenon in which /s/ is realized as an [ʃ]-like sound, especially when it occurs in an /stɹ/ cluster ('*street*' pronounced more like [ʃtɹit] than like [stɹit]). The speech of English speakers judged *not* to exhibit s-retraction shows a large coarticulatory bias in the direction of retraction. Further, there is also substantial interspeaker variation in the extent of this bias. We propose that this interspeaker variation, coupled with the coarticulatory bias, facilitates the initiation of sound change. In this account, sound change begins when a listener accidentally interprets an extreme case of a phonetic effect as an articulatory target and then adjusts her own speech in response. This adoption of a new target requires phonetic variation that predates the change. Thus, sound change is predicted to be biased toward phonetic effects that exhibit interspeaker variability, and if sound change requires an accident that is rare, then sound change itself is correctly predicted to be rare as well.

INTRODUCTION

We report on an investigation into the phonetic motivation of s-retraction in American English, with emphasis on the relevance of the data to the actuation

Supported by the College of Social and Behavioral Sciences at the University of Arizona, and James S. McDonnell Foundation grant #220020045 BBMB, PI Archangeli. We are grateful to Robin Dodsworth and four anonymous reviewers for comments on earlier drafts.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

problem in the theory of sound change. We show that the phonetic motivation for s-retraction can be demonstrated generally, and that the sound change can be explained as phonologization of coarticulation. This merely illustrates and accentuates the actuation problem, however, because it creates the opposite problem of predicting that s-retraction should occur in all speakers, whereas in fact it does not. We are able to propose a solution to the problem, however, by considering speaker-specific variation in the phonetic data. We follow Janda and Joseph (2003) in understanding the initiation of sound change as a phonetically abrupt event rather than a gradual one. We further propose that sound change can be initiated when a hearer interprets an extreme instance of a phonetic effect as a distinct production target, an exaggeration of the normal coarticulation. If the speaker adopts the new production target, the sound change has begun to propagate around the speech community. Because the distribution of extreme coarticulation in a way that favors its interpretation as a distinct target is predicted to be rare, sound change is correctly predicted to be a rare occurrence.

We take as uncontroversial that properties of a language user's phonological system are largely driven by facts about his/her social network and the social meaning associated with phonetic variables. In this article paper, we are concerned specifically with how a production target can *start* to shift. By the actuation of sound change, we refer specifically to the inception of a *change from below*.[1] We have nothing to say about language changes that can be understood in terms of language contact or pre-existing social meaning.

We examine the actuation problem particularly in relation to English s-retraction. This is a well documented sound pattern found in several varieties of spoken English, in which /s/ takes on a more [ʃ]-like pronunciation, generally in the vicinity of /ɹ/ (Armstrong, 2003; Durian, 2007; Janda and Joseph, 2003; Janda, Joseph, & Jacobs et al., 1994; Labov, 1984; Lawrence, 2000; Rutter, 2011; Shapiro, 1995). S-retraction is observed in varieties of English spoken in New Zealand, the United Kingdom, and the United States. It is possible that the sound change giving rise to s-retraction has occurred more than once, in parallel.[2] The most frequently encountered pattern involves the realization of /s/ as an [ʃ]-like sound in the /stɹ/ cluster: '*street*' sounds more like [ʃtɹit] than [stɹit]. In word-initial consonant clusters, s-retraction occurs less frequently in /skɹ/ clusters, and even less frequently in /spɹ/ clusters (Shapiro, 1995).[3] Retraction can also occur in word-medial /sɹ/ (in words such as '*grocery*'), and marginally in some other phonetic contexts (Janda and Joseph, 2003). In this paper, we limit our investigation to word-initial consonant clusters. Note that although s-retraction introduces an allophonic alternation, there is no phonemic consequence for the change.

We show that s-retraction, like many other phonological patterns, is paralleled by coarticulatory effects present in the speech of speakers of English varieties that have not undergone the sound change or participated in its spread (henceforth "non-retractors").[4] We also show that phonetic motivation for s-retraction has two components. The greater part of this phonetic motivation is

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

common to all speakers, but there is in addition a narrower coarticulatory motivation that varies by speaker. These facts bear significantly on the actuation problem, a challenge for theories of language change first posed by Weinreich, Labov, and Herzog in their landmark 1968 theoretical work. The actuation problem concerns the initiation of sound changes. It is difficult for a theory to include a mechanism of change that does not underapply (failing to predict cases where change occurs) or overapply (failing to predict cases where change does not occur). For well over a century, linguists have espoused the position that sound change is the gradual accommodation of language to the needs of speakers and hearers. Numerous such approaches—which we will refer to as *accumulation-of-error theories* are to be found in the literature. Whitney is typical in presenting such a view:

> The first and most important principle [of linguistic change] that we have to notice, the one which lies at the bottom of nearly all phonetic change in language, is the tendency . . . to make the work of utterance easier to the speakers, to put a more facile in the stead of a more difficult sound or combination of sounds. . . . . All articulate sounds are produced by effort, by expenditure of muscular energy, in the lungs, throat, and mouth. This effort, like every other which man makes, he has an instinctive disposition to seek relief from, to avoid: we may call it laziness, or we may call it economy . . . . (1896, 69–70)

The idea of a gradual accumulation of errors in speech is an intuitive one, and has been adopted implicitly or explicitly by many subsequent researchers. We use the term '*error*' here very generally, including minor changes in language that could be considered functional or teleological. Our critique of these approaches ("The Actuation Problem in s-retraction") does not require that these changes be intentional or unintentional. Ohala has probably been the most significant exponent of the accumulation-of-error view, explaining sound changes variously as accumulated errors in production (e.g., Ohala, 1993; Ohala and & Ohala, 1991a; Ohala, 1993) or perception (e.g., Ohala, 1981, 1983). Ohala and Ohala (1991b) describe this research focus as follows:

> [I]t is probably not possible to explain why a given sound change happens in one language but not another and it is generally not a fruitful question to pursue. Phonetics can at best explain a sound change by showing, first, that precursors to it are found widely in languages spoken today . . . and, second, that there is a phonetic basis for these precursors.

The work on the historical basis for synchronic sound patterns has focused in particular on the types of changes that are observed, and why those changes, instead of other changes, are more frequent. The related questions, such as why changes are rare, even though phonetic precursors are common, and why changes happen where and when they do, even though similar conditions exist in other places and at other times, have been met with more pessimism. This has

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

produced a situation where work on sound change either implicitly or explicitly treats accumulation of error as the mechanism by which a phonetic variable is selected to become a sound change,[5] or more frequently, sidesteps the question entirely.[6]

It is within the context of one accumulation-of-error theory—specifically that of the Neogrammarians—that Weinreich et al. (1968) describe the actuation problem.

> [I]f the pursuit of ease is the cause of sound change in idiolects, the fundamental questions arise: why do not speakers go about it more quickly, and why do Language Customs split in that some speakers set out on a particular ease-seeking path whereas other retain their less comfortable pattern?. . . . For even when the course of a language change has been fully described and its ability explained, the question always remains as to why the change was not actuated sooner, or why it was not simultaneously actuated wherever identical functional properties prevailed. The unsolved actuation riddle is the price paid by any facile and individualistic explanation of language change. It creates the opposite problem—of explaining why language fails to change. (Weinreich et al., 1968, 111–112).

As we shall discussed further below, the difficulty lies in creating a theory in which sound change can plausibly occur, but without making sound change inevitable. As an example, the development of Proto-Indo-European witnessed two developments of the labiovelar stop. In Cumbrian, *k^w changed to p, while whereas Latin *k^w remained k^w.[7] The challenge posed by the actuation problem is to determine a theory of sound change where the change *k^w > p is a possible outcome—because *k^w *can* change to p—but not an inevitable outcome, because a language can remain unaffected. The actuation problem has long plagued accounts of sound change.; Labov (2001, 466) calls called the problem "unsolved." The data from English s-retraction cast the actuation problem in sharp relief, because even non-retractors display a strong coarticulatory tendency towards retraction. Given these facts, theories that represent sound change as the accumulation of error have difficulty explaining how any speaker can fail to participate in s-retraction.

The phonologization of coarticulation occurs when a phonetically motivated effect is so strong that it is interpreted as a distinct phonetic target. It is already known that phonologization depends on imitating a distinct phonetic target (e.g., Janda, 1999, 2003; Janda and & Joseph, 2003), but that in most cases, there is not sufficient exaggeration to produce the novel targets that can spur a sound change. It is clear that once a phonetic effect has been so interpreted, it may then be imitated by other speakers, leading to the spread of sound change. However, if every speaker produces essentially the same amount of coarticulation, then there is nothing to imitate in the first place. We propose that inter-speaker differences in the *degree* of a phonetic effect facilitate identification of new targets. Specifically, we argue that coarticulation patterns that *already vary* from speaker to speaker for purely phonetic reasons can create the situation of extreme articulation that gets interpreted as a distinct phonetic target. That is, the

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

potential for sound change arises when inter-speaker variability is so great that some speakers actually produce a sound that can be perceived as a distinct target. Given the appropriate social conditions, this potential can be realized through another speaker adopting the novel target in his/her speech.

In the case of s-retraction, the source of phonetically motivated variation in /s/ is variation in the production strategy used to produce the /ɹ/ that conditions retraction. A listener who hears tokens of /s/ that have been produced with a relatively high degree of retraction (for idiosyncratic phonetic reasons that are unknown to him/her), may interpret these tokens as the target, and may potentially shift his/her production target to involve a degree of coarticulation that was not actually motivated by his/her own idiosyncratic phonetic motivations.

Although the magnitude of phonetically motivated coarticulation is dependent on internal factors, the likelihood of a speaker with extreme coarticulation being imitated by other speakers depends on social factors. Importantly, not all members of a speech community are equally influential, and we argue that the ability of phonetically motivated coarticulation to lead to sound change depends on the chance alignment of extreme coarticulation with extreme influence. Labov (2002) decribes described a typical leader of linguistic change as follows:

> More detailed studies of the social characteristics of the leaders of linguistic change show that they are female members of the highest status local group, upwardly mobile, with dense network connections within the local neighborhood, but an even wider variety of social contacts beyond the local area (Labov, 2002, Ch. 11).

> Moreover, these leaders of change show a history of nonconformity in other respects, and show a life-long history of resistance to repressive or unjust authority (Labov, 2002, Ch. 12).

Although there is little evidence that speakers use social meaning to guide their production of variables that have not already achieved a high degree of social salience, we are not in a position to rule out the possibility that imitation is unconsciously guided by a speaker's social estimation of other community members. For instance, Babel (2009) has recently reported that subjects in an imitation experiment exhibit more imitation of speakers with whom they identify. If future research shows evidence that social identity is relevant for the spread of change from below, then we would add that sound change should be more likely in cases where a phonetic variable happens to coincide with a social variable. Another possibility is that a listener would mistakenly interpret as socially meaningful a variable that has not achieved salience in the community at large, because of the way it happens to be distributed among speakers, and imitate it on that basis. Because change from below is not believed to involve social factors until after it has propagated, this small-scale perceived confluence of phonetic and social factors would probably not be directly related to any social value that the new variant might acquire if it does eventually spread through the

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

community. The point here is that there are many possible ways in which spurious correlations between phonetic and social factors could conceivably occur, but for our present purposes, we are only concerned with how the overall measure of a speaker's influence coincides with phonetically-motivated variability.

The idea that non-contact-induced change is dependent at least in part on pre-existing phonetically motivated inter-speaker variability has implications for the predicted frequency of sound change. This would be a step toward a theory that correctly predicts sound change to be sporadic and rare, rather than an inevitable consequence of the accumulation of error. Sound change is predicted to be rare and sporadic because change depends on how the necessary excessive variable phonetic variants happen to be distributed in the speech community, and most of the logically possible distributions do not seem likely to favor adoption of a new sound pattern. An additional prediction is that phonetic effects that lack noticeable inter-speaker variability would be poor candidates for sound change. That is, for change to occur, the range of the inter-speaker variability must be so great that at one extreme the result can be interpreted as a new target.

Coarticulation is largely language-specific (Beddor, Brash, & Narayan, 2007; Beddor, Harnesberger, & Lindemann, 2002a; Ladefoged, 1984), and this is seen in the fact that languages with similar coarticulation patterns exhibit them to different degrees. This suggests that conventionalized coarticulation (as opposed to the bare minimum of phonetically-determined coarticulation) is probably the norm rather than the exception. Our claim does not require the starting point to be strictly phonetically determined coarticulation, as the our results below will show that inter-speaker variability in coarticulation degree is present even among individuals with s-retraction. Even if word-internal coarticulation is completely encoded in the lexicon (Coleman, 2000), inter-speaker variability still provides the opportunity for excessive productions, if other conditions are right.

This proposal is closely related to a recent study by Beddor (2009) that the phonologization of coarticulation is facilitated by covariance between the cause and effect of coarticulation. Beddor showeds that in the case of coarticulatory vowel nasalization, the duration of the nasalized portion of the vowel is inversely correlated with the duration of the nasal consonant. This variation provides an opportunity for listeners to interpret nasalization as a property of the vowel instead of as a property of the consonant, and speakers vary in the timing of their velum lowering and listeners vary in their interpretation of other speakers' productions.

The proposal that inter-speaker phonetically motivated variation is crucial allows us to make clear predictions about the role of phonetic effects in initiating sound change. Specifically, we claim that under standard assumptions, phonetic effects that are shared by all the members of a community—however large—should not be expected to produce sound change because, by their very commonness, they are not good candidates. Further, phonetic effects that occur to the same degree across speakers should be easier to perceptually compensate

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

for. We argue instead that *speaker-specific phonetic motivations* are favored to be identified as novel targets, and can thus actuate sound change.

We begin by presenting the phonetic data pertinent to our claims about the actuation problem in the discussion section. The goal of this section is to demonstrate clearly: (i) that in the speech of people who do not exhibit the s-retraction change, there is nevertheless a bias toward retraction, (ii) that this bias parallels the phonetic environments in which s-retraction occurs, and—crucially for our subsequent argument—(iii) that the magnitude of the phonetic effect varies between speakers. We then review the actuation problem as it relates to the facts of English s-retraction ("The Actuation Problem in s-retraction"). In so doing, we set the phonetic data to one side, and argue from logic and mathematical probability. The assumptions behind accumulation-of-error theories yield themselves readily to basic mathematical analysis, and we use this to uncover their fundamental flaw: that they predict sound change to be too likely to occur.[8] We present our solution in the discussion section, noting that some standard assumptions in the field suggest a solution to the actuation problem, along the lines that we have already outlined above. Following consideration of some final issues, we conclude.

PHONETIC MOTIVATION FOR S-RETRACTION

In order to determine the extent of phonetic motivation for s-retraction, we collected data from both retractors (people who tend to say things like [ʃtɹit]) and non-retractors (people who tend to say things like [stɹit]). We treated present-day non-retractors as models of past "pre-retractors": the speakers in the past, before the sound change began. This allowed us to study the phonetic properties of s-retraction in isolation of other confounding factors, namely the confound of retractors having a different production target.

METHODS

Subjects

Subjects were undergraduates from the University of Arizona who received extra credit for participating in the experiment. 32 Thirty-two subjects were recorded, though 4 were discarded due to technical difficulties. The remaining subjects were 18–22 years of age (mean 18.6 years). Two experiments were run simultaneously, drawing from the same subject pool. Retractors were in relatively short supply, so retracting speakers were identified by having them read a short word list, and were selectively assigned to the present experiment. Additionally, subjects from regions where s-retraction was thought to be more prevalent were assigned to this experiment. Consequently, the proportion of retractors to non-retractors that we report should not be taken to be representative of any population of speakers. For each subject included in the experiment,

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

TABLE 1. *Self-reported information from subjects included in the analysis*

Judgment	Sex	Age (yrs)	Where subject grew up
Retractor	F	19	Philadelphia, Pennsylvania; New Jersey; Arizona
Retractor	F	18	Los Angeles, California; Dallas, Texas; South Dakota; Arizona
Retractor	F	19	Irvine, California; Wisconsin
Retractor	F	18	Tucson, Arizona
Retractor	M	20	Pasadena, California
Retractor	M	18	Tarzana, California; Calabasas, California
Retractor	M	22	Brooklyn, New York
Retractor	M	18	Spokane, Washington
Non-retractor	F	19	Tempe, Arizona
Non-retractor	F	18	Nogales, Arizona
Non-retractor	F	20	Phoenix, Arizona
Non-retractor	F	19	St. Paul, Minnesota; Arizona
Non-retractor	F	19	Tucson, Arizona
Non-retractor	M	18	Seattle, Washington; Boulder, Colorado; Iowa City, Iowa
Non-retractor	F	18	Lexington, Kentucky
Non-retractor	F	19	San Francisco, California
Non-retractor	M	18	San Jose, California; Thousand Oaks, California
Non-retractor	M	19	Arizona; Wyoming; South Dakota
Non-retractor	M	20	Royal Oak, Michigan; Houston, Texas; Hackettstown, New Jersey

Note: Self-reported information from subjects included in the analysis. "Judgment" indicates whether the subject was classified as a retractor or non-retractor; see "Protocol" section for more on this classification. The order of subjects from top to bottom in this table corresponds with the order of subjects from left to right in Figure 3.

Table 1 shows sex, age, and where the subject grew up. After data collection, speakers were reclassified as retractors or non-retractors according to a procedure described in "Protocol"; the preliminary classification was not carried forward into subsequent analysis.

Experimental materials

Items consisted of monosyllabic words with sibilants in word-initial position. Items were collected with initial /s/, /sC/, /sCɹ/, /ʃ/, and /SɹC/ (C = {p,t,k}), with following vowels /æ, i, ɪ, u/. Two words from each phonetic context were included, and four tokens of each item were collected. Items are listed in the appendix, along with filler words.

Protocol

After filling out a language background and speech pathology questionnaire, subjects were seated in a comfortable chair in front of a computer screen, with a microphone placed about 6 inches in front of the mouth. A simultaneous ultrasound recording of the tongue was also made, with an ultrasound transducer placed under the chin. The subject also wore a pair of sunglasses with head-tracking markers on them. The articulatory data provide an account for the interspeaker variation, but they are not analyzed in depth here. See Mielke,

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

Baker, and Archangeli (2010) for a more complete articulatory analysis. Tokens were presented on the computer screen in the frame "Please say [token] again." Subjects read the entire frame sentence, advancing through the series of prompts at their own pace by clicking a mouse button.[9] The order of the stimuli was random, with the items occurring four times apiece. Subjects were encouraged to read quickly, though the actual rate of speech varied noticeably between speakers.

Speakers were classified as either retractors or non-retractors after the experiment, based on a classification of their tokens as retracted or non-retracted. Two of the authors independently listened to each /stɹ/ token in the context of the frame sentence and judged the token to be either retracted or non-retracted. The first round of identification produced agreement between the experimenters for 83.8% of the tokens. To improve this figure, each experimenter judged the disputed tokens a second time. In the second round of identification, each experimenter knew that the token was disputed, but was not aware of his earlier choice. After the second round, there was agreement for 94.7% of the stimuli. The remaining stimuli, which had produced disagreement twice, were discarded from the study.

Previous studies of s-retraction have identified categorical intraspeaker variation in s-retraction (e.g., Durian, 2007); this was our finding as well. Although there were instances of speakers who always retracted or never retracted, there was variability in the speech of most subjects. A break occurred in the data between the 35.48% and 62.50% rates of retraction (number of retracted /stɹ/ tokens divided by the total number of /stɹ/ tokens), and we treated this as the threshold for classifying a speaker as a retractor or non-retractor. Of the 28 subjects, 8 were retractors and 20 were non-retractors. Our goals for the study were to identify the phonetic correlates of retraction in retractors' speech, and to look for coarticulatory antecedents to retraction in non-retractors' speech. As such, we were interested in tokens in which retractors were retracting. It is not helpful to examine a non-retracted token to discover the phonetic correlate of retraction. Therefore, we discarded the tokens in which retractors did not retract. There is also a possibility that speakers who retract for a minority of tokens would have an affected pronunciation in their non-retracted tokens as well. We therefore, we also discarded data from non-retractors who exhibited variable behavior, limiting the non-retractor data to the 11 subjects who were judged never to have a retracted token.[10]

Acoustic analysis

For acoustic analysis, the boundaries of each fricative were determined manually, using both the waveform and a spectrogram as a guide. Fricatives were characterized by their centroid frequencies, a measure commonly used to describe sibilants (e.g., Baum and & McNutt, 1990; Jesus and & Shadle, 2002; Jongman, Wayland, & Wong, 2000; Nittrouer, Studdert-Kennedy, & McGowan, 1989). The centroid frequency of /s/ is characteristically high, whereas the centroid frequency of /ʃ/ is lower. Because s-retraction involves the change of /s/

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

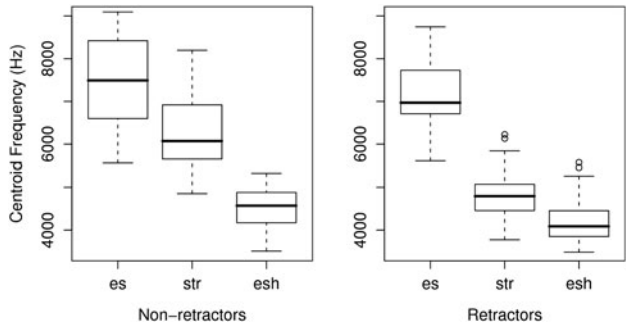


FIGURE 1. A comparison of the sibilants of non-retractors (left) and retractors (right). The [s]tr of retractors is noticeably lower, both in absolute frequency, and in its relative position between [s] and [ʃ].

into an [ʃ]-like sound, this measure was expected to be appropriate for quantifying the sound change, and the retracted allophone of /s/ is expected to have a lower centroid frequency than an unaffected /s/ production. The spectrum from the middle half of each fricative was calculated, and the center of gravity of each spectrum was computed as the average frequency, weighted by the power at each frequency. Tokens were averaged across item before statistical analysis.

RESULTS

Test of centroid frequency as a measure of s-retraction

To verify that the centroid frequency is an appropriate measure for s-retraction, the centroid frequencies of [s], [ʃ], and [s] from the /stɹ/ context (henceforth, [s]tr) were compared for both retractors and non-retractors. A plot is shown in Figure 1. Consistent with previous research, [s] and [ʃ] were well-differentiated by their centroid frequencies, with [s] having the higher centroid. As was expected, retractors' [s]tr had lower centroid than non-retractors' [s] in the same context. This confirms the auditory impression that retracted [s] is more [ʃ]-like than plain [s], and validates centroid frequency as a measure of s-retraction.

Phonetic effect of a following stop and [ɹ]

To assess the phonetic effect of a following consonant and a following [ɹ] on the centroid frequency of [s], the speech of non-retractors was investigated. A within-subjects analysis of variance was performed on tokens of [s] from the sC and sCɹ clusters. The dependent variable was the centroid frequency of the fricative, with independent variables ConsonantContext (p, t, k) and presence (whether [ɹ] was present or not present). There was a significant main effect of ConsonantContext ($F(2,36) = 22.3$, $p < .0001$) and of presence

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

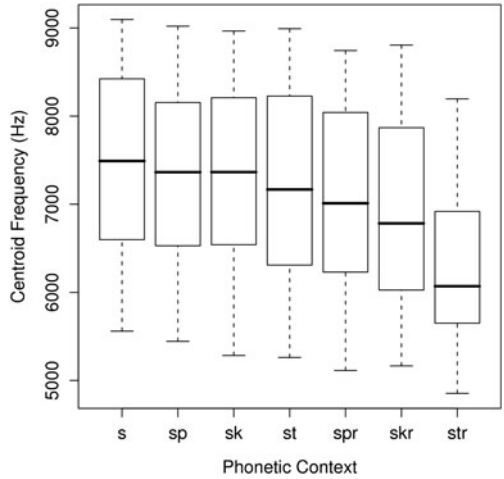


FIGURE 2. Realizations of [s] in initial clusters, in non-retractors. The presence of a consonant has a depressing effect on the centroid frequency, and the presence of [ɹ] in the cluster depresses the centroid frequency even more. [s]tr is most heavily affected.

($F(1,16) = 66.5$, $p < .0001$). The interaction of the two factors was also significant ($F(2,38) = 19.7$, $p < .0001$). The data were therefore broken across the factor ConsonantContext. The effect of presence was significant for each of /p/, /t/, and /k/ (respectively, $F(1,19) = 29.8$, $p < .0001$; $F(1,19) = 43.3$, $p < .0001$; $F(1,19) = 55.5$, $p < .0001$).

The data are presented graphically in Figure 2. The [s] from a simple onset had the highest centroid frequency. Next were consonants with a following stop with a following vowel (and not a following [ɹ]), in descending order: [sp], [sk], and [st]. Finally, the lowest centroid frequencies were those of [s] followed by both a consonant and [ɹ]. Of these, [s]tr clearly has the lowest centroid frequency; it is the most [ʃ]-like of all.

The presence of [ɹ] clearly has a depressing effect on /s/, then, which indicates a potential phonetic motivation at the *initiation* of the sound change. It is of interest to know whether [ɹ] would exert similar phonetic pressure throughout the course of the sound change, that is, after /s/ had taken on a more [ʃ]-like pronunciation. Although we cannot test this directly with the data we collected, we can compare the productions of a simple [ʃ] with those of an [ʃ] that precedes [ɹ]. From this basis. we can generalize reasonably safely: if a following [ɹ] depresses both /s/ and /ʃ/, then it is quite plausible that it would depress intermediate sibilants as well (and the current non-retractors may well already be at an intermediate stage). To see whether this is indeed the case, a post hoc one-tailed *t*-test was performed comparing tokens with initial ʃ and those with initial ʃɹ. The test indicated at $p < .004$ that the centroid frequency of /ʃ/ is lower when an [ɹ] follows. We therefore conclude that the presence of [ɹ] would motivate the

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

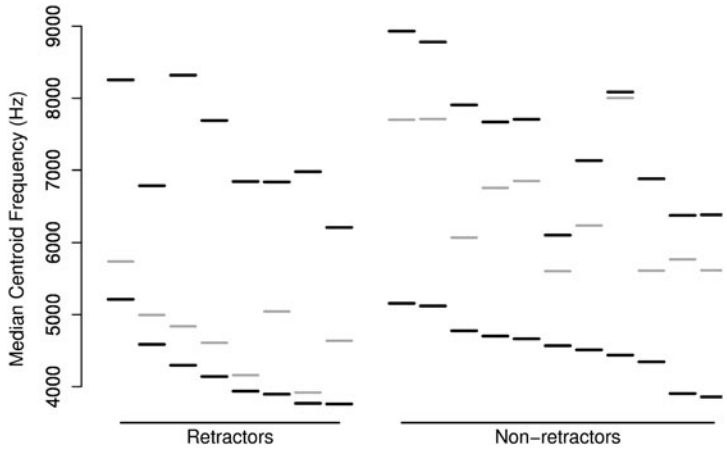


FIGURE 3. Median centroid frequencies from all subjects: each column represents one subject. In each column, the top bar represents the median centroid frequency of plain [s], the middle gray bar the median centroid frequency of [s]tr, and the lowest bar the median centroid frequency of plain [ʃ]. In the speech of non-retractors, there is wide variation in the centroid frequency of [s]tr, both absolutely, and in its relative position between [s] and [ʃ].

sound change not just at its initiation, but also at every subsequent stage of the change.

Speaker-specificity of the phonetic motivation

The final empirical point to be emphasized is the speaker-specificity of the phonetic motivation for s-retraction: non-retractors exhibit varying phonetic tendencies toward s-retraction. Figure 3 presents a condensed overview of the data. The median centroid frequencies for three allophones of plain [s], [s]tr (light graey), and the plain [ʃ], are shown for each speaker. It is evident that there is variation in the centroid frequencies of [s]tr, both in absolute and relative terms. Absolute variation is readily apprehended from the graph. Relative variation can also be observed. A typical non-retractor has an [s]tr that is closer to plain [s] than it is to plain [ʃ]; in contrast, a typical retractor's [s]tr is closer to plain [ʃ]. Clearly then, even speakers who appear not to have retracted [s] as a production target exhibit a phonetic tendency towards s-retraction. There is also a substantial range of variation among non-retractors, however, both in relative and absolute terms.[11]

Mielke et al. (2010) attributed inter-speaker variability in s-retraction to inter-speaker variability in /ɹ/ production strategy. In brief, American English /ɹ/ is known to have several articulatory allophones, which for the production of /ɹ/ itself have minimal acoustical and perceptual consequence (e.g., Delattre and & Freeman, 1968). Although many speakers produce distinctly different versions

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

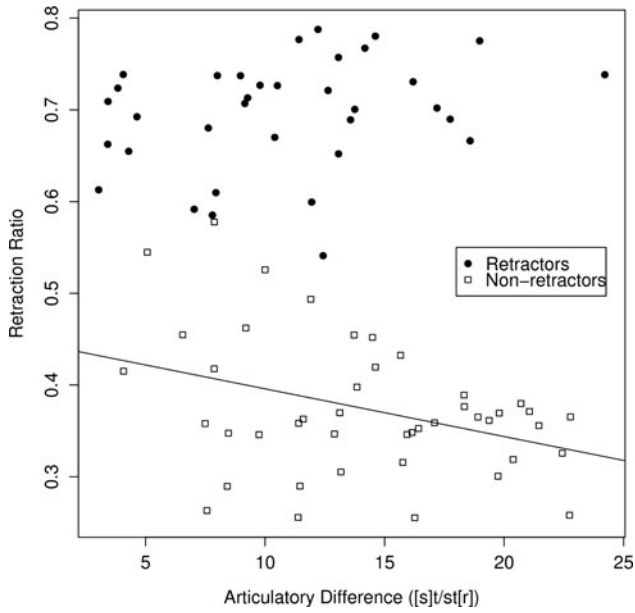


FIGURE 4. S-retraction ration plotted against [s]-[ɹ] articulatory difference (Mielke et al., 2010).

of /ɹ/ in different segmental contexts (Mielke et al., 2010), /stɹ/ strongly favors a bunched /ɹ/ over a retroflex /ɹ/ and all of the subjects in this study produced a bunched /ɹ/ variant (Delattre and Freeman's types 3–6) in /stɹ/ clusters. Nevertheless, the subjects did show variation within this broad category of bunched /ɹ/ types, and different articulatory allophones of /ɹ/ do in fact cause different degrees of coarticulatory retraction of /s/ when it occurs in the /stɹ/ cluster. Specifically, more similarity between a speaker's [s] and [ɹ] tongue shapes was found to be associated with more coarticulation between those segments when they co-occur in /stɹ/ clusters. Figure 4 shows the relationship between [s]-[ɹ] similarity and retraction ratio, which is defined as the position of [s]tr along the scale from [s] (ratio = 0) to [ʃ] (ratio = 1). This 0–1 range corresponds to the interval between each pair of black lines in Figure 3, and the location of the gray lines in that figure matches the retraction ratio for each speaker. Non-retractors whose canonical [s] and [ɹ] involve more similar tongue shapes exhibit more coarticulation and hence more coarticulatory s-retraction. Retractors do not show a significant effect of articulatory similarity on s-retraction, apparently because their retraction is due to a distinct production target and is not dependent on coarticulation. The higher retraction ratio of retractors indicates that their [s] centroid frequency is close to [ʃ] and not influenced by [s]-[ɹ] distance, whereas non-retractors have an [s] that is less similar to [ʃ] and whose centroid frequency is affected by [s]-[ɹ] distance.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

This speaker-specific variation in /ɹ/ tongue shape is the cause of the interspeaker variability in s-retraction.

SUMMARY

The preceding findings contain three important facts about the phonetic motivation for s-retraction in American English, which are pertinent to the subsequent discussion.

First, non-retractors as a whole have a substantial coarticulatory bias toward s-retraction, which is shared by all speakers. The speaker-specific portion of the variation is much smaller. From Figures 1 and Figure 3, it is evident that even in the speech of non-retractors, there is a substantial bias toward retraction. The bias toward s-retraction can be quantified either with absolute measurements, or with a relative measure. We quantify s-retraction by calculating the position of [s]tr between [s] and [ʃ]. The pertinent numbers are provided in Table 2. A retracted [s] is characterized by a centroid frequency that is 75.5% of the distance from [s] to [ʃ]. But in non-retractors, [s]tr is already 48.6% of the way to [ʃ]. Expressed as a ratio, based on coarticulation alone, non-retractors are about 65% of the way to being retractors (48.6%/75.5% = 64.6%). This effect of /ɹ/ on /s/ could be purely phonetic or it could be phonological:, i.e., the result of a production target that is distinct from /s/ in other contexts, more than is mandated by articulatory demands. The negative correlation between retraction ratio and articulatory distance suggests that at least some of the coarticulation is phonetic.

Secondly, the coarticulatory effect observed in non-retractors parallels the retracted /ɹ/ in retractors. A following consonant depresses the centroid frequency of [s], but the presence of /ɹ/ depresses the centroid frequency further. This is consistent with the s-retraction pattern, where /ɹ/ is a crucial element of the conditioning environment.[12] The case for the presence of phonetic motivation is made stronger when we consider the specific effects of following consonants. As noted previously, s-retraction is most common in /stɹ/ clusters, less common in /skɹ/ clusters, and less common still in /spɹ/ clusters (Shapiro, 1995; also the present study). This mirrors precisely the pattern of coarticulation in Figure 2. The coarticulatory effect is greatest in the /stɹ/ cluster, lesser in the /skɹ/ cluster, and least in the /spɹ/ cluster. These facts support the claim that s-retraction has a clear phonetic motivation.

The final important fact is that the phonetic motivation for s-retraction is variable across speakers (Figure 3). Non-retractors have individualized phonetic biases toward s-retraction, which is true if we consider the data in absolute or relative terms. Structured heterogeneity is indeed the norm for language data, but we call attention to the speaker-specific coarticulatory effects here because this heterogeneity is critical, in our view, to solving the actuation problem.

The phonetic effects observable in the speech of non-retractors admit an elegant accumulation-of-error account of the s-retraction sound change. Such an account would resemble the following: speakers begin with a characteristic [s] target in

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

TABLE 2. *The values in the first three columns represent the median centroid frequencies of [s], [ʃ], and [s]tr for non-retractors and retractors. The rightmost column shows and the distance of [s]tr from [s], as a percentage of the distance from [s] to [ʃ]. A greater percentage value means closer to [ʃ] and further from [s]*

				[s]tr expressed as a proportion of the distance from [s] to [ʃ]
	[s]	[s]tr	[ʃ]	
Non-retractors	7491 Hz	6070 Hz	4570 Hz	0.486 (48.6%)
Retractors	6969 Hz	4794 Hz	4087.5 Hz	0.755 (75.5%)

Note: The values in the first three columns represent the median centroid frequencies of [s], [ʃ], and [s]tr for non-retractors and retractors. The rightmost column shows the distance of [s]tr from [s], as a percentage of the distance from [s] to [ʃ]. A greater percentage value means closer to [ʃ] and further from [s].

the /stɹ/ cluster. Coarticulatory influence from the /ɹ/, however, lowers the centroid frequency of [s] in the actual production. Those listening to the speaker hear a more [ʃ]-like production of /s/ and may adjust their own production targets slightly in response. Over time, these adjustments accumulate and are eventually recognized as a sound change. S-retraction is expected to be most common in the /stɹ/ context because the phonetic motivation for retraction is strongest there. These same phonetic facts about s-retraction, however, also form a particularly pointed illustration of the actuation problem, which we discuss in the next section.

THE ACTUATION PROBLEM IN S-RETRACTION

In presenting the actuation problem specifically for the s-retraction sound change, we repeat the question asked by Weinreich et al. (1968): "why do Language Customs split in that some speakers set out on a particular ease-seeking path whereas other retain their less comfortable pattern?" In the present day, a population of speakers remains in which the s-retraction sound change has not taken hold. A split has occurred in English dialects. In spite of this, the speech of today's non-retractors exhibits a clear tendency toward s-retraction. It was remarked in the last section that, taken as a whole, non-retractors' [s]tr is already more than half of the way to the retractors' [s]tr. Why do these speakers not become retractors themselves, in the face of such a strong phonetic motivation to do so? If sound change consists in the gradual accumulation of production error, why does production error *not* accumulate in non-retractors? Accumulation-of-error theories offer no answers to these questions. Rather, they predict that sound change should occur for all speakers, because all speakers have similar phonetic biases. This is the actuation problem. The main argument of this section is that the actuation problem can be expressed in terms of the probability of change,

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

that is, the probability of a given change occurring in a given time period. A theory of change which that predicts change to be likely runs afoul of the actuation problem. A theory of change which that predicts change to be unlikely escapes the problem of actuation. We demonstrate that accumulation-of-error theories always predict change to be likely. To solve the actuation problem, a mechanism of change is required that is rare this is proposed in the next section.

In this section, we consider below the statistical properties of some models of sound change. We note that the probability of a sound change occurring, P_C, influences not only the likelihood of a sound change, but also the likelihood of splits in speech communities. We then consider the statistical properties of accumulation-of-error theories, to create a rigorous description of why they overpredict sound change. At this point in our argument, we leave off arguing from our specific empirical facts, and argue instead from logic and probability. There are two other ways the argument might have progressed, by providing evidence either through either additional empirical data, or through computer simulations. The methodological issues involved in being able to observe the initiation of a real sound change are clearly prohibitive. An alternative would be to create computer simulations of populations of speakers whose speech may or may not change, given certain simulation parameters; Baker (2008a) explores the actuation problem with this method. Since the assumptions involved in theories of sound change are at this level of analysis very basic, however, a computer simulation is something of an overkill. We can obtain the same results with a few basic statistical equations.

Calculating the probability of the change of a given sound change at a given time is a fraught problem. A particular sound change occurs under conditions that probably arise only once in history, given the possible variations in linguistic and social factors. How can a probability estimate be calculated? Instead of going down this road, we follow Weinreich et al. (1968) in examining splits in language communities, from which we can draw conclusions about how probable sound change must be, without having specific probability estimates in hand. We consider the simple scenario in which a single speech community divides into two, Group A and Group B, with each daughter community developing independently. Such a scenario is probably the minority in history—perhaps arising only through abrupt population splits due to migration—but it serves as a useful starting point for the present discussion assuming independent development permits simple mathematical comparisons to be made. In such a scenario, a split must be predicted to be the more likely outcome, in which one group undergoes the change and the other does not. For two groups (A and B) and one sound change, there are four possible outcomes: both Group A and Group B change, only Group A changes, only Group B changes, or neither group changes. If we suppose that the probability of sound change is high—say, $P_C = .95$—then the probabilities of the possible outcomes are those given in Table 3. The typical scenario is the one in which both groups undergo the sound change; this is predicted to occur 90.25% of the time. The next most likely scenario is that one group would change but not the other (9.5%), followed by

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

TABLE 3. *Probabilities of sound change scenarios if the probability of change* (P_C) *is high*

Scenario	Probability calculation		Probability
Both groups change	$P_C \cdot P_C$	=	.9025
Only one group changes	$P_C \cdot (1 - P_C) + (1 - P_C) \cdot P_C$	=	.095
Neither group changes	$(1 - P_C) \cdot (1 - P_C)$	=	.0025

TABLE 4. *Probabilities of sound change scenarios if the probability of change* (P_C) *is low*

Scenario	Probability calculation		Probability
Both groups change	$P_C \cdot P_C$	=	.0025
Only one group changes	$P_C \cdot (1 - P_C) + (1 - P_C) \cdot P_C$	=	.095
Neither group changes	$(1 - P_C) \cdot (1 - P_C)$	=	.9025

the scenario in which neither group changes (0.25%). If sound change is very likely to occur, then it should be (1) initiated promptly, (2) very likely to occur in both groups, and (3) very *unlikely* to occur in neither group. This is clearly the incorrect prediction because, among other reasons, it is simply not the case that every language undergoes almost every possible sound change, all of the time.

Conversely, if we suppose a low probability of sound change ($P_C = .05$), then the situation is reversed, as shown in Table 4. The most common outcome (90.25%) is that sound change does not occur; although sound change is reasonably probable to occur in only one group (9.5%), it is very unlikely to occur in both groups (0.25%). This is the familiar scenario languages usually do not undergo a particular change, splits do occur, but parallel development is very unlikely. To avoid predicting across-the-board sound change, then, the predicted probability of change (P_C) needs to be small. If P_C is high (as in Table 3, above), then (1) sound change is expected to be initiated frequently (i.e., immediately), and (2) parallel development between language groups that have no contact is predicted to be the rule rather than the exception.

Having established that an adequate theory of sound change must predict that the probability of change be low, we can give a more mathematically precise formulation to the problems that accumulation-of-error theories run into. The incorrect predictions of accumulation-of-error theories derive from the high probability of change that they produce. This fact can be illustrated with reasonably simple statistics. Suppose that the accumulation of k errors in a speech community is sufficient to complete a sound change. Then the probability of change (P_C) is equal to the probability of k errors accumulating. We specify further that the probability of an error occurring within a given time increment is P_e, and that each error occurs independently of other errors.[13] Figure 5 shows the probability that a particular number of errors (in this case,

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

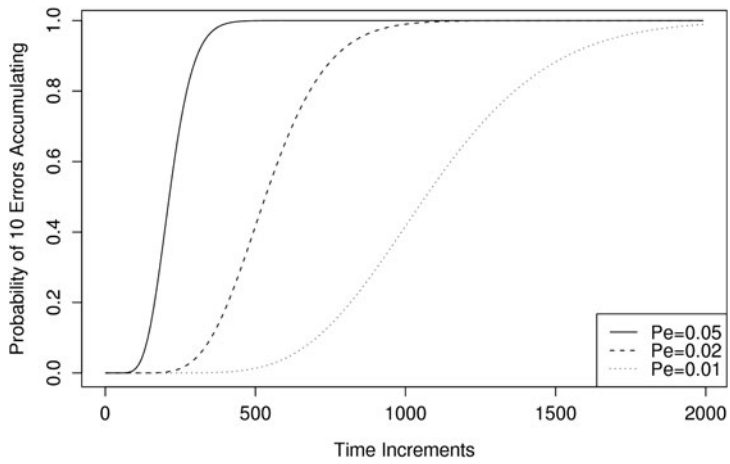


FIGURE 5. Probability over time of 10 errors accumulating, for various values of P_e.

ten10) will have accumulated, as a function of the number of time increments.[14] Curves are shown for different values of P_e: .01, .02, and .05. The more likely the error is to occur, the more quickly the sufficient number of errors are accumulated. Even if P_e is just .01, still after 2000 time increments, the probability of there being a sufficient number of errors to induce change is $P_C = .9895$. The errors *inevitably* accumulate, as long as there is *some* probability of an error occurring ($P_e > 0$). Plugging this value of P_C into the calculations shown in Tables 3 and 4, this means that the chances of change occurring in both groups after 2000 time increments is nearly 98%, the chance of a split is 2.18%, and the chance of the change not occurring in either group is less than 2 in ten thousand.

These calculations illustrate mathematically the empirical weaknesses of an accumulation-of-error approach. Over time, the probability of change (P_C) is high even if the probability of an error occurring (P_e) is low. Therefore, sound change is predicted to occur nearly across the board, and splits between languages are consequently predicted to be extremely rare. The overwhelming probability is that both groups will undergo the sound change. Consequently, it is appropriate to note that accumulation-of-error theories of sound change tend to be just that: theories that explain why languages change. The neglected question, of why languages *fail to change*, is just as significant in historical phonology. A theory that predicts sound change across the board is of no greater explanatory adequacy in this regard than a theory that permits no sound change at all. An adequate theory of sound change must allow for cases where a sound change does and does not occur. In the next subsection, we propose the alternative: a theory of the actuation of sound change in which change is expected to be very infrequent. We propose that sound change is not initiated by the gradual accumulation of error over time, but rather by a single, rare event.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

DISCUSSION

Toward a solution to the actuation problem

The challenges posed by the actuation problem can be addressed by considering the importance of inter-speaker variation that exists *prior* to the start of a sound change. We claim that sound changes are initiated when there is a spurious correlation between a phonetic variable and a social variable, and a speaker adopts the change into his/her own speech. If the exaggerated form continues to propagate through the speech community, a new change will have been initiated.

Because different degrees of coarticulation are expected only rarely to be distributed in a community in a way that facilitates change from below—and because of other conditions on the spread of a change, which are we will discussed below—this account of the mechanism of the initiation of sound change meets the crucial criterion of predicting sound change to be a rare event, and thus offers a solution to the actuation problem. Our claims about the initiation of sound change are very nearly corollaries to the first and third principles outlined by Weinreich et al.

> "1. Linguistic change is not to be identified with random drift proceeding from inherent variation in speech. Linguistic change begins when the generalization of a particular alternation in a given subgroup of the speech community assumes direction and takes on the character of orderly differentiation.
>
> ....
>
> 3. Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity." (Weinreich et al., 1968, 187, 188).

The contribution of the present paper is to suggest a mechanism by which phonetic variation "assumes direction." Following Weinreich et al. (1968), we do not identify sound change with an accumulation over time of production or perception errors (e.g., contrary to Paul, 1880; Whitney, 1896, etc.). We also follow them in emphasizing the heterogeneity of the speech community and assert that sound change is facilitated by particular random distributions of variants. Earlier we described the speaker-specific variation that can be observed in [s]tr. Depending on a speaker's idiosyncratic articulatory behavior, the coarticulatory effect of [ɹ] onto [s] can be larger or smaller. Members of a speech community are obviously also heterogeneous with respect to social position, and, crucially, this heterogeneity should not be expected to correlate with articulatorily-motivated speaker-specific variation in coarticulation.

The distribution of degrees of coarticulation in a way that favors exaggeration is not expected to arise very frequently. Therefore, this account of sound change meets the critical desideratum of the previous section: predicting that sound change is infrequent. (We consider factors that might affect the frequency of change in the discussion.) Consequently, this account is able to explain why sound change is relatively rare, and why splits occur between speech communities.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

In placing the emphasis on the outset of sound change, we are closely following Janda and Joseph (2003), who make made a clear distinction between the factors that cause a sound change to be initiated, and the factors that cause the change to be propagated. In broad strokes, Janda and Joseph claimed that sound change is initiated by phonetic factors, but it is spread primarily by social factors. They claim that the phonetic factors in sound change, as identified by the Neogrammarians, hold necessarily at the outset of sound change, but not necessarily during the spread of the sound change. We do not believe that isolated phonetic factors have any bearing on sound change. Indeed, every word uttered by a speaker is influenced by coarticulation, and every word perceived is subject to coarticulatory effects; these phonetic factors overwhelmingly *do not* give rise to sound change. Rather, it is when phonetic factors are arranged in a way that facilitates a wide range of speaker variability that sound change is more likely to be initiated. In fact, our proposal makes the testable prediction that not even every phonetic variable is a good candidate for sound change, but only those with substantial phonetically-motivated speaker-specific variation.[15]

How do these claims relate specifically to s-retraction? What is the relationship between the phonetic precursors to retraction that we see in non-retractors, and the sound change? Emphatically, we reject as empirically inadequate the suggestion that s-retraction came about as the gradual incorporation of production error. The phonetic facts we describe in the results section make this a possible way to explain the change, but it is not possible to explain how the change could ever have failed to have occurred. Instead, we consider the crucial phonetic condition to be the existence of *speaker-specific tendencies* toward s-retraction (also described in the results). It is only *speaker-specific* effects that have the potential to be correlated to social variables (see the discussion in the next section). Note, however, that even then there is not a guarantee of change, but merely the possibility. This description of the initiation of sound change arises from the phonetic facts, but relies on probabilistic argument. Experimental testing of the proposal must await future work, though undoubtedly observing the initiation of a sound change in the very first speaker will be very difficult.

We now consider some implications of the proposed solution, particularly with regard to the role of speaker-specific phonetic variation and some factors influencing the probability of sound change under the current proposal.

Sources of sound change: Speaker-specific phonetic variation

This theory of sound change has implications for the role of phonetic bias in producing sound change. Contrary to previous explanations of sound change, notably the Neogrammarian accumulation-of-error accounts, phonetic variables are expected to play a role only to the extent that they can become exaggerated. A corollary of this principle is that phonetic factors owing to coarticulation and perceptual effects are not predicted to be involved in sound change at all, except to the extent that these effects vary in magnitude across speakers.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

There are other reasons not to expect larger phonetic effects to lead more frequently to sound change. Perceptual compensation for coarticulation (e.g., Beddor, Harnsberfer, & Lindemann, 2002b; Ohala, 1993; Lotto and Kluender, 1998; Ohala, 1993, inter alia) is one way that a phonetic effect can remain stable without being exaggerated and phonologized. Ohala (1994) suggesteds that larger phonetic effects may be *less* likely to lead to sound change, because they may be easier for listeners to compensate for.[16] We suggest that interspeaker variability may *favor* the phonologization of a phonetic effect, no matter the magnitude. In this regard, recall the interspeaker variability of [s]tr among non-retractors, seen in Figure 3. In addition to creating the possibility of exaggeration, interspeaker variability may also render the task of perceptual compensation for coarticulation more difficult, thus making the phonetic effect more noticeable in the first place. This more complicated perceptual compensation task is closely related to the effect of variability in the magnitude of the coarticulatory cause and effect that is reported by Beddor (2009).

We have noted that all non-retracting speakers produce [s]tr with substantial coarticulation (almost two-thirds of the way to [ʃ]), relative to simple [s]. But since as this coarticulatory bias is common to all speakers, it is not likely to result in a sound change. If everyone produces [s]tr clusters in the same way, the coarticulation cannot be a target for sound change because no speaker appears to have a particularly different amount of coarticulation. Moreover, because the coarticulation is predictable, it is easy to perceptually compensate for. Therefore, the presence of a stable coarticulatory pattern, which creates a difficulty for accumulation-of-error theories, is expected. The speaker-specific component of coarticulation also provides a candidate for a sound change. Knowing that no two speakers, or indeed two productions, are exactly alike, we might further stipulate that only speaker-specific patterns that *vary noticeably* from one another are candidates for sound change. It is not the case that all measurable phonetic detail is perceptible, and consequently, we should not expect all speaker-specific coarticulatory patterns to be so. How common are such perceptibly different patterns? To our knowledge, studies of speaker-specific coarticulatory effects are rare. As discussed previously, Mielke et al. (2010) examined the causes of speaker-specific motivation for s-retraction, specifically with reference to speaker-specific articulatory allophones of /ɹ/. Sound change is of course not limited to contexts that involve such exceptional segments, but it remains to be seen how widespread speaker-specific phonetic effects actually are.

A commonly recognized property of sound change is that, on the whole, sound changes favor the development of phonetically natural productions. This was one of the early intuitions of the Neogrammarians, and has found empirical support in a variety of studies. Word-final devoicing, for instance, is a common sound change, and has straightforward articulatory and perceptual motivations (Blevins, 2004). We associate the initiation of sound change with an accidental distribution of phonetically-determined variants. We have no reason to expect a priori, however, that social organization is related to phonetic variables in a way that

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

encourages sound changes that favor phonetic ease. For example, we do not expect it to be true that speakers who are socially situated to be leaders in change are naturally inclined to more coarticulated productions. In isolation, this leads to the incorrect prediction that the direction of sound change should not be biased toward phonetically natural changes. The accidental distribution of speakers who produce a relatively high degree of coarticulation in a way that favors exaggeration in the direction of more coarticulation is just as likely as the accidental distribution of speakers who produce a relatively *low* degree of coarticulation in a way that would favor exaggeration in the opposite direction (less coarticulation). A variant with low coarticulation, however, would sound more similar to the norm for the same sound in contexts that do not induce coarticulation, or in more hyperarticulated speech. For this and probably other reasons, it seems reasonable that variants with more coarticulation are more likely to get exaggerated than variants with low coarticulation, which is consistent with the observed asymmetric tendency of sound change.

Difficulties in quantifying the likelihood of sound change

We have claimed that under the current proposal, sound change is predicted to be unlikely. But exactly how rare is sound change expected to be? Establishing a quantitative probability would be quite difficult, but we can note that three conditions must be met, which intuitively seem as if they would be unlikely to co-occur within a given time period.

First, a certain phonetic variable must be spuriously related to a socially relevant factor. Although one might imagine that this would be rare, a quantitative answer would be difficult to give. In order to do so, we would have to know how much speech data a speaker typically evaluates, and how he/she evaluates it.[17] Next, it would be necessary to determine how likely a speaker is to change his/her behavior, once a sufficient number of apparently innovative tokens have been encountered. This might have to do with the speaker's willingness or aversion to risking innovative behavior—however unconscious—in a social context. If these two conditions are met, the speaker will change his/her behavior. Thus, an apparent case of exaggeration will have led to a real case of exaggeration, and, in an important sense, a sound change will have occurred. Generally the sound changes that linguists study are evidenced in the speech of many or all speakers of a community, however. For that purpose, we can consider the probability of a sound change being propagated in a social network. We might simply evaluate this as the probability of a given person's being in an appropriate social position to initiate change. A more complex analysis might consider the decision of each subsequent speaker in the community—which might involve a calculation of the probability of that speaker detecting the shift in the production target, and perhaps a calculation of that speaker's relationship to his/her interlocutors.

If these probabilities are considered to be independent, then the probability of sound change would be the product of the three probabilities mentioned

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

earlier. Even if these probabilities were relatively large (say, one in twenty 20), the probability of sound change occurring would become quite small ($P_C = .05 \times .05 \times .05 = .000125$). Therefore, despite a clear need for further research into the probability of sound change, we are nevertheless confident in asserting that the proposed limitation on sound change predicts that sound change will be quite rare.

We also may wish to consider the possibility of other speaker-specific factors that are related to an individual's likelihood of participating in the actuation of a sound change. Babel (2009) reported that speakers are more likely to imitate (subconsciously) speakers with whom they identify, but little is known about speaker-specific differences in the tendency or ability to imitate the speech of others. Further, Yu (2010) found that females who exhibit the fewest autistic-like traits are less likely to perceptually compensate for coarticulation. This result indicates that cognitive factors may also play a role. If the actuation of sound change relies additionally on the confluence of cognitive factors with social and phonetic factors, it is predicted to be an even rarer event, but much more research is required in this area.

CONCLUSION

We have presented evidence from American English s-retraction, which suggests a solution to the actuation problem, a long-standing problem in historical linguistics. The s-retraction data cast the actuation problem into sharp relief, because there is a clear phonetic tendency toward retraction in the speech of speakers who have not undergone the sound change. The actuation riddle of s-retraction is why any group of speakers would not undergo the sound change, in spite of clear phonetic pressure to do so.

In addressing the problem, we emphasized the role of the speaker-specific phonetic variation that is present in our data. Following the work of other researchers, we also emphasize the difference in factors that drive the initiation and propagation of sound change (Janda and Joseph, 2003). In our view, sound change can be initiated when phonetically-determined inter-speaker variation happens to result in variants being distributed in a community in a way that favors the interpretation of extreme phonetically-motivated coarticulation as an exaggeration *beyond* the phonetic motivation. If the speaker adopts a new production target, and this propagates around the speech community, then a sound change is born. Phonetic variables that exhibit inter-speaker variation—in this case, a variable tendency toward s-retraction—have an advantage. Consequently, phonetic effects which that vary between speakers are more likely candidates for new sound changes.

This proposal avoids the problems created by accumulation-of-error theories of sound change. Such theories overpredict the occurrence of sound change, and do not predict the occurrence of splits in communities, because they predict sound change to be so frequent. The present proposal predicts sound change to be far

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

less common, because several events that are themselves low-probability are required to coincide for the sound change to be successfully initiated.

The proposal, that sound change depends on phonetically-motivated inter-speaker variation that precedes socially motivated inter-speaker variation, permits a solution to the actuation problem. Because sound change is expected to be a rare occurrence, it is predicted that possible sound changes typically will not occur, and that splits in speech communities can happen in the rare instances where a sound change does occur. In general terms, then, our empirically motivated model makes correct predictions about the frequency of occurrence of sound change, by looking at the conditions necessary for the exaggeration of a phonetic effect at its initiation.

NOTES

1. Labov (1994, 78) defines change from below and change from above as follows:

> *Changes from below* are systematic changes that appear first in the vernacular, and represent the operation of internal, linguistic factors. At the outset, and through most of their development, they are completely below the level of social awareness.
>
> *Changes from above* are introduced by the dominant social classes, often with full public awareness. Normally, they represent borrowings from other speech communities that have higher prestige in the view of the dominant class.

2. If any of these instances of s-retraction can be attributed to the influence of a Germanic substrate (Joseph, 2005), then they would not require an explanation in terms of the mechanism we propose here (though the initiation of variation in the Germanic substrate might).

3. In our data as well, we encountered one speaker with retraction in the /skɹ/ context, and none with retraction in the /spɹ/ context.

4. We do not intend to claim that the subjects in our study had themselves "converted" to become retractors at some point in the past, or indeed that s-retraction is a change in progress. We only refer to the subset of the population that produces a retracted /s/ in the /stɹ/ context (and other contexts) because they have a different production target from other speakers. In this usage all modern day German speakers could be referred to as "Umlauters," being speakers who evidence in their own speech changes such as Old High German *gasti* > *Gäste*: not because the speakers themselves have changed, but because their speech shows evidence of the change.

5. It may be noted that nearly every paper in Baker (2008b)'s review of computational models of language change crucially relies on the accumulation of speaker error.

6. A recent work that addresses these questions more directly is Blevins (2004, especially 41–44), who cites Watson (1999), Bybee (2001), and Pierrehumbert (2001) on the potential role of frequency in determining when sound changes do or do not occur.

7. Thanks for an anonymous reviewer for providing this helpful example.

8. Here and throughout the paper, when we refer to the likelihood or rarity of a sound change, we refer to the statistical probability that a particular change will occur, in a particular time frame, in a particular language.

9. The fact that the stimuli were presented orthographically leads to the potential of eliciting read speech, which might be expected to reduce subjects' tendencies to retract. This might also have been exacerbated by the laboratory context.

10. These subsets of the data were eliminated to remove any doubt as to the validity of the results of the study. In fact, since a relatively small number of tokens was discarded, none of the results reported here change qualitatively even if all the data are included in the analysis.

11. See Table 1 for demographics of the speakers reported in Figure 3.

12. Janda and Joseph (2003) cite occurences of s-retraction in /st/ and /sk/ clusters (in '*school*', '*still*'), but s-retraction in these contexts seems to be marginal, or at least unsystematic. Janda and Joseph treat these instances as subsequent generalizations from an original retraction pattern that only involved [s]tr.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

Given the current data, we might speculate that s-retraction in these contexts is encouraged by the effect of a following consonant on the centroid frequency of [s]. Setting that aside, however, the crucial observation is that s-retraction is far more common when an /ɹ/ is present. There is an attested pattern of retraction in /stɹ/ but not /st/ clusters (e.g., all eight retractors in this study), but no attestation in the literature of retraction in /st/ clusters and not /stɹ/ clusters.

13. For the present purposes the precise nature of the errors and the change to which they lead is unimportant. Errors have been considered to be quantal entities here for the sake of convenience, not necessity.

14. This figure is derived from reasonably simple statistics. If each error occurs independently, then each error is a Bernoulli trial, and the probability of errors occurring over time increments is expressed by:

$$1 - I_{(1-Pe)}(n-k, k+1) \tag{1}$$

Where

$$I_x(a, b) = \sum_{j=a}^{a+b-1} \frac{(a+b-1)!}{j!(a+b-1-j)!} x^j (1-x)^{a+b-1-j} \tag{2}$$

15. In saying that not even every phonetic variable is a candidate for sound change, we are referring specifically to change from below, and not to contact-induced change. Since we are concerned specifically with the path to the eventual acquisition of indexical meaning by a phonetic variant that previously had none, we make no claims about the spread of phonetic variants that already have social value.

16. Another reason for the magnitude of a phonetic effect not to be directly related with its tendency to lead to sound change is proposed by Moreton (2008), who argues that large phonetic effects may be "underphonologized" if they are disfavored by a learner's analytic bias, whereas smaller phonetic effects that seem to be favored by this bias are phonologized more often. See also Hansson (2008) for a discussion.

17. In a computer simulation, it would be easy to model speakers as mini-statisticians, tabulating data and making conclusions upon an individually-specified -value. For a real human, the process of evaluation is bound to be more complex, and perhaps rely more on informal intuition.

REFERENCES

Armstrong, Mike. (2003). */s/-retraction in the ViC corpus.* ms, The Ohio State University ms.

Babel, Molly. (2009). *Phonetic and Social Selectivity in Speech Accommodation*. Ph.D. thesis, University of California, Berkeley.

Baker, Adam. (2008a). Addressing the actuation problem with quantitative models of sound change. *Proceedings from PLC 31.* Special issue of vol. 14.1 of *U. Penn Working Papers in Linguistics* 14 (1):29–41.

_____. (2008b). Computational approaches to the study of language change. *Language and Linguistics Compass* 2(March):289–307.

Baum, Shari R., & McNutt, James C. (1990). An acoustic analysis of frontal misarticulation of /s/ in children. *Journal of Phonetics* 18:51–63.

Beddor, Patrice Speeter. (2009). A coarticulatory path to sound change. *Language* 85:785–821.

Beddor, Patrice Speeter, Brasher, Anthony, & Narayan, Chandan. (2007). "Applying perceptual methods to the study of phonetic variation and sound change." In M.-J. Solé, P. S. Beddor, and M. Ohala (eds.), *Experimental approaches to phonology*. Oxford: Oxford University Press. 125–143.

Beddor, Patrice Speeter, Harnsberger, James D., & Lindemann, Stephanie. (2002a). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30:591–627.

_____. (2002b). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30:591–627.

Blevins, Juliette. (2004). *Evolutionary phonology*. Cambridge, Mass: MIT Press.

Bybee, Joan. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

Coleman, John. (2000). "Where is coarticulation? " In M. Broe & J. Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon*. Cambridge: Cambridge University Press. 102–117.
Delattre, Pierre, & Freeman, Donald C. (1968). "A dialect study of American R's by X-ray motion picture." *Language* 44:29–68.
Durian, David. (2007). "Getting [s]tronger every day? : Urbanization and the socio-geographic diffusion of (str) in Columbus, OH." *University of Pennsylvania Working Papers in Linguistics*. vol. 13:65–79.
Hansson, Gunnar Ólafur. (2008). "Diachronic explanations of sound patterns." *Language and Linguistics Compass* 2(5):859–893.
Janda, Richard D. (1999). "Accounts of phonemic split have been exaggerated—but not enough." *Proceedings of ICPhS* 14.
_____. (2003). ""Phonologization" as the start of dephoneticization—or, on sound-change and its aftermath: Of extension, generalization, lexicalization, and morphologization." In B. D. Joseph & R. D. Janda (eds.), *The handbook of historical linguistics*. Oxford, UK: Blackwell. 401–422.
Janda, Richard D., & Joseph, Brian D. (2003). "Reconsidering the canons of sound-change: Towards a "big bang" theory." In B. Blake & K. Burridge (eds.), *Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*. Amsterdam: John Benjamins Publishing Co. 205–219.
Janda, Richard D., Joseph, Brian D., & Jacobs, Neil. (1994). "Systematic hyperforeignisms as maximally external evidence for linguistic rules." In R. Corrigan Lima & G. Iverson (eds.), *The reality of linguistic rules*. Amsterdam: John Benjamins Publishing. 67–92.
Jesus, Luis M. T., & Shadle, Christine H. (2002). "A parametric study of the spectral characteristics of European Portuguese fricatives." *Journal of Phonetics* 30:437–464.
Jongman, Allard, Wayland, Ratree, & Wong, Serena. (2000). "Acoustic characteristics of English fricatives." *Journal of the Acoustical Society of America* 108(3):1252–1263.
Joseph, Brian D. (2005). "On projecting variation back into a proto-language, with particular attention to Germanic evidence and some thoughts on "drift"." In Thomas D. Cravens (ed.), *Variation and reconstruction*. John Benjamins Publishing Company. 103–118.
Labov, William. (1984). "Field methods of the project on linguistic change and variation." In J. Baugh & J. Sherzer (eds.), *Language in use: Readings in sociolinguistics*. Englewood Cliffs, NJ: Prentice-Hall. 28–53.
_____. (1994). *Principles of linguistic change, Volume 1: Internal factors*. Chichester, UK: Blackwell.
_____. (2001). *Principles of Linguistic Change, Volume 2: Social Factors*. Chichester, UK: Blackwell.
_____. (2002). "Driving forces in linguistic change." *2002 International Conference on Korean Linguistics*.
Ladefoged, Peter. (1984). "'Out of chaos comes order'": Physical, biological and structural patterns in phonetics." In M. P. R. Van den Broecke & A. Cohen (eds.), *Proceedings of ICPhS 10*. Dordrecht: Foris Publications. 83–95.
Lawrence, Wayne. (2000). "/str/→ /ʃtr/: Assimilation at a distance? " *American Speech* 75(1):82–87.
Lotto, Andrew J., & Kluender, Keith R. (1998). "General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification." *Perception & Psychophysics* 60(4):602–619.
Mielke, Jeff, Baker, Adam, & Archangeli, Diana. (2010). "Variability and homogeneity in American English /ɹ/ allophony and /s/ retraction." In B. Kühnert (ed.), *Variation, detail, and representation. LabPhon 10*. Berlin: Mouton de Gruyter. 699–729.
Moreton, Elliott. (2008). "Analytic bias and phonological typology." *Phonology* 25(1):83–127.
Nittrouer, Susan, Studdert-Kennedy, Michael, & McGowan, Richard S. (1989) . "The emergence of phonetic segments evidence from the spectral structure of fricative-vowel syllables spoken by children and adults." *Journal of Speech, Language and Hearing Research* 32(1):120–132.
Ohala, John J. (1981). "The listener as a source of sound change." In C. S. Masek, R. A. Hendrick, & M. F. Miller (eds.), *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistics Society. 178–203.
_____. (1983). "The origin of sound patterns in vocal tract constraints." In P. F. MacNeilage (ed.), *The production of speech*. New York: Springer-Verlag. 189–216.
_____. (1993). "Coarticulation and phonology." *Language and Speech* 36(2–3):155–170.
_____. (1994). "Hierarchies of environments for sound variation; plus implications for "'neutral"' vowels in vowel harmony." *Acta Linguistica Hafniensia* 27:371–382.
Ohala, John J., & Ohala, Manjari. (1991a). "Nasal epenthesis in Hindi." *Phonetica* 48:207–220.
_____. (1991b). "Reply to commentators." *Phonetica* 48:271–274.
Paul, Hermann. (1880). *Prinzipien de Sprachgeschichte*. Halle: Niemeyer.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

Pierrehumbert, Janet B. (2001). "Exemplar dynamics: Word frequency, lenition and contrast." In J. Bybee & P. Hopper (eds). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 323–418.
Rutter, Ben. (2011). "Acoustic analysis of a sound change in progress: The consonant cluster /stɹ/ in English." *Journal of the International Phonetic Association* 41(1):27–40.
Shapiro, Michael. (1995). "A case of distant assimilation: /str/ → /ʃtr/." *American Speech* 70(1):101–107.
Watson, Ian. (1999). "Phonetics, phonologization, and French nasal vowels." *Oxford University Working Papers in Linguistics, Philology & Phonetics* 4:157–173.
Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). "Empirical foundations for a theory of language change." In W. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press.
Whitney, William Dwight. (1867/1896). *Language and the study of language*. New York: Charles Scribner's Sons.
Yu, Alan C. L. (2010). "Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change." *PLoS ONE* 5(8): E11950. doi:10.1371/journal.pone.0011950.

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press

APPENDIX

Stimuli

/s(C)/

sap	sat	seep	seat	soup	suit	sit	sip
Span	spat	speed	speak	spoon	spoof	spin	spit
Stab	stack	steep	steel	stoop	stool	stick	stiff
Scab	scam	ski	scheme	school	scoop	skip	skit

/sʃ(ɹ)/.

shall	shack	she	sheet	shoe	shoot	Ship	shift
Shrank	shrapnel	shriek	shree	shrew	shrewd	Shrill	shrimp

/sCɹ/.

sprat	sprang	spree	Spreel	spruce	sprew	sprint	sprig
Strap	strand	street	stream	strew	strewn	strip	strict
Scram	scratch	screen	screech	screw	scrooge	script	scrimp

/(C)ɹ/.

rat	Ram	reed	ream	rude	rune	rinse	rig
Bran	brat	breed	breeze	brew	bruise	brick	brisk
Drab	draft	dream	Dreese	drew	droop	drip	drift
Gram	grass	green	greet	grew	group	grip	grin

/ʃ C(ɹ)/ & /sɹ/

spiel	schtick	spritz	Sri Lanka

Filler

bag	bog	cash	cash	cat	chief	dog	fast
fish	goat	Jeep	lack	mouse	pack	page	pat
pop	shake	sheep	shop	sock	take	tape	took

https://doi.org/10.1017/S0954394511000135 Published online by Cambridge University Press