Language Variation and Change (2023), **35**, 79–105
doi:10.1017/S0954394523000078

"I can't see myself ever living any[w]ere else": Variation in (HW) in Edinburgh English

Nina Markl

The University of Edinburgh, UK
Email: nina.markl@ed.ac.uk

Abstract

Sociolinguistic research across Scotland in recent decades has documented an erosion of the phonemic contrast between /ʍ/ (as in *which*) and /w/ (as in *witch*). Based on acoustic phonetic analysis of 1,400 <wh> realizations produced by eighteen Edinburgh women born between 1938 and 1993, I argue that in the context of Edinburgh this is best understood as a complex sociolinguistic variable (HW) encompassing (at least) six fricated and fricationless variants. Realizations vary in type and relative duration of frication, voicing, and glide quality. Bayesian statistical analysis suggests that choice and realization of variants is conditioned by speaker's social class, style, and phonetic context. Unlike some prior work, I do not find evidence of ongoing (apparent-time) change or an effect of contact with Southern British English. Fricated variants are most prevalent in formal speech styles and in the speech of middle-class women, while working-class speakers favor fricationless variants.

Historically, English pronunciations of the digraph <wh> differed from those of <w>, yielding minimal pairs such as *which* ~ *witch* and *whine* ~ *wine*. While this contrast has been lost in many varieties of English (e.g., in North America; Labov, Ash & Boberg, 2008:45), Scottish Englishes have traditionally been described as retaining it (Giegerich, 1992:36; Jones, 2002; Wells, 1982:409). However, sociolinguistic research across Scotland suggests this is changing (Brato, 2014; Chirrey, 1999; Lawson & Stuart-Smith, 1999; Macafee, 1983; Reiersen, 2013; Robinson, 2005; Schützler, 2010; Stuart-Smith, Timmins, & Tweedie, 2007). Based on an analysis of 1,400 tokens of <wh> produced by eighteen female speakers of Edinburgh English, I propose that, in Edinburgh at least, this is best conceptualized as variation within the sociolinguistic variable (HW).

(HW) denotes a sociolinguistic variable that encompasses all pronunciations of <wh>. Variants of (HW) are often described as labial-velar fricative and labial-velar approximant. In contrast, I use the terms "fricated" to refer to tokens characterized by a period of frication preceding a glide, and "fricationless" to refer to tokens that only consist of a glide. This distinction is useful because, as I show in this study, variants of (HW) differ with respect to type and duration of frication and quality of glide, voicing, and phonation. I find that middle-class speakers (and those who orient toward

Standard Scottish English) produce higher rates of fricated tokens, while working class speakers (and those who orient toward Scots) favor fricationless tokens. As with other sociolinguistic variables, the phonetic context is also a meaningful predictor of variant choice and variant realization. Notably, I do not find evidence of a change in progress. Given the broader context this does not support an effect of contact with Southern British English(es) as suggested in prior work and points instead to variation based on social class and speech style.

Background

Setting the scene: Edinburgh, Leith, and linguistic diversity

Home to the Scottish parliament, several universities and a finance sector, Edinburgh is a cosmopolitan city associated with power and wealth. The most prominent spoken varieties are Gaelic, Scots, Scottish Standard English (SSE), Southern British English, Polish, and Urdu.[1] Gaelic, which has been repressed like other Celtic languages in the British Isles, has been the focus of revitalization efforts in recent years and is visible on official signage (Lawson, 2014). Scots differs from English on the levels of syntax, phonology, and lexis (Jones, 2002; Lawson, 2014). SSE features some Scots lexis and syntax and differs from Standard Southern British English (SSBE) in terms of phonology (Giegerich, 1992; Schützler 2015). Like other standard varieties, SSE is strongly associated with formal contexts and middle- and upper-class speakers rather than a particular place. Scots, on the other hand, is generally associated with working-class speakers, and varieties of Scots are spoken in urban and rural areas of Scotland. Many speakers shift between Scots and SSE depending on the social context (Stuart-Smith, 2004). In Edinburgh, most people likely encounter Scots, SSE, and SSBE every day.

Most of the women interviewed for this study ($n = 16$) have grown up in Leith. Historically an independent port town, Leith retains a distinct identity from Edinburgh (Doucet, 2009; Marshall, 1986). In recent years, deindustrialization and gentrification have changed it dramatically. Today, the area around Leith Walk, a thoroughfare connecting Leith and Edinburgh city center, and Easter Road, home to Leith's football team, is one of the most densely populated in Scotland. It features small international supermarkets and tailors between pubs, bars, and restaurants. The other two participants lived in Morningside, a neighborhood of Edinburgh long perceived as middle class whose high street is dominated by upmarket boutiques and supermarkets, cafes, and pubs.

Historical perspective: an unstable contrast

Old English featured several <h>-initial clusters including <hw>, whose patterns of alliteration and rhyme suggest was pronounced as the voiceless labial-velar fricative [ʍ] (Minkova, 2004:16). This contrasted with <w>, produced as the labial-velar approximant [w]. This distinction was preserved in modern spelling as <wh> and <w>. While at first glance it appears that some varieties of English retained a phonemic contrast until recently, Minkova (2004) suggested that this contrast has been unstable for a long time. Considering Old English texts, Minkova (2004:17) argued

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

that <h>-insertions in etymologically <h>-less words indicate variation or confusion on the part of authors, while <h>-less spellings of etymological <hw>-words suggest reduction (e.g., <wistle> instead of <hwistle> and <bilhwit> instead of <bilewit>). The contrast later reappeared in the speech of upper-class Southern English speakers but has since been lost in all Anglo-English varieties, including SSBE (Wells, 1982)[2] and North American varieties (Bridwell, 2019; Labov et al., 2008:49; Thomas, 2019). Minkova (2004) traced the fricated variant in Scottish English to an allophone of Old English /hw/, [xʍ], which developed to [hw̥].

Variation in type and duration of frication in fricated (HW) tokens found in several varieties of English today could indicate that the distinction between different allophones was never as clean as Minkova (2004) described. Similarly, the apparent "reappearance" of the contrast in Early Modern English could be evidence for a "reconstruction" of the contrast based on spelling (Minkova, 2004) or variable retention (Milroy, 2004). In any case, the diachronic perspective highlights that any variability in choice and realization of variants found today is not necessarily new.

A contact-induced merger in progress?

Over the last forty years, variable use of fricated and fricationless variants of (HW) in Scotland has been found to be conditioned by age, gender, socioeconomic class, educational background, contact with SSBE, and linguistic factors such as phrasal position and phonetic context. (HW) has been described in Glasgow (e.g., Lawson & Stuart-Smith, 1999; Macafee, 1983; Stuart-Smith et al., 2007), Livingston (e.g., Robinson, 2005), Edinburgh (e.g., Chirrey, 1999; Fruehwald, Hall-Lew, Eiswirth, Boyd, & Elliot, 2019; Reiersen, 2013; Schützler, 2010), and Aberdeen (e.g., Brato, 2014).

Beginning with social factors, Schützler (2010) interpreted differences by age and gender among twenty-seven middle-class speakers in Edinburgh as a change in progress and argued, based on effects of contact with SSBE and level of education, that the loss of fricated (HW) was a contact effect. In forty-four Aberdeen speakers, Brato (2014) found that middle-class teenagers (in particular girls) and older speakers shifted from the traditional, local, Scots variant [f] of (HW) to the fricated supraregional Scottish English [ʍ], while younger working-class speakers shifted toward the fricationless variant [w]. Like Schützler (2010), Brato (2014) argued that contact with non-Scottish varieties played a role in this shift. Drawing on formal speech from 138 speakers in the ICE-Scotland corpus (e.g., parliamentary debates and television broadcasts), Li and Gut (2022) showed that even in formal SSE none of the speakers fully retained the original (HW) phonemic contrast, while 12% exclusively produced fricationless variants. Similar to Schützler (2010) and Brato (2014), they noted differences by age and gender, with women and younger speakers more likely to produce fricationless tokens (Li & Gut, 2022). In Glasgow, Stuart-Smith et al. (2007) and Lawson and Stuart-Smith (1999) described effects of age and social class on (HW). Working-class speakers, in particular young working-class speakers, preferred fricationless variants, while middle-class speakers favored fricated ones. The putative role of contact with Anglo-English varieties is particularly interesting among these Glaswegian speakers, as it highlights the complex relationship between social class,

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

local and nonlocal linguistic standards, and contact. Middle-class speakers retained the fricated variant despite contact with Southern British varieties (Stuart-Smith et al., 2007). Young working-class speakers adopted the "nonlocal" fricationless variant not due to direct face-to-face contact with Southern British speakers or a positive orientation toward Southern British varieties or a non-Scottish identity, but rather, Stuart-Smith et al. (2007) argued, to distinguish themselves from middle-class Glaswegians. As to linguistic factors, while fricated realizations seem to be straightforwardly favored after pauses and in lexically stressed positions, the effects of lexical frequency and lexical category, phonetic context, and word-specific effects are difficult to disentangle because the incidence of lexical items is so heavily skewed toward *what*, *when*, *why*, *where*, *which* (Schützler, 2010). Fricated tokens have been reported least likely to occur in word-internal contexts (e.g., *somewhere*), and fricationless variants to be favored word-initially, after vowels, and (less strongly) after consonants (Brato, 2014).

A range of variants

Many discussions of (HW) treat it as a merger (*which* ~ *witch*) (e.g., Fruehwald et al., 2019; Labov et al., 2008; Macafee, 1983; Reiersen, 2013; Schützler, 2010) between the voiceless labial-velar fricative [ʍ] that is characterized by a period of frication and the fully voiced labial-velar approximant [w]. However, in their study of Glaswegian children's speech, Lawson and Stuart-Smith (1999:2542) described an additional intermediate variant perceived as voiceless but lacking the characteristic period of frication and a category of tokens which "[are] neither like [hw] nor like [w]" but a "breathy [w̤]." Such tokens are found in a 1997 corpus of adult speakers from Glasgow, too (e.g., Stuart-Smith et al., 2007). Among children in Livingston, a town between Edinburgh and Glasgow, Robinson (2005:186) also found a "continuum of phonetically intermediate forms," the most "traditional" of which was a "voiceless lip-rounded consonant with audible friction at both velar and bilabial articulations," while an intermediate variant included voiced fricated tokens. Working on Southern White American English in South Carolina, Bridwell (2019:104) also described "voiced [hw] tokens" featuring both frication before the glide and voicing throughout the entire segment.

Acoustic phonetic variation and (HW)

Most prior studies of (HW) relied on auditory coding of variants (e.g., Brato, 2014; Schützler, 2010; Stuart-Smith et al., 2007) or minimal pair tests (e.g., Labov et al., 2008). However, acoustic phonetic analyses can reveal patterns that are not necessarily auditorily perceptible. Variants of (HW) consist of a glide, which can be characterized by formants, and an optional period of frication preceding the glide. Li and Gut (2022) examined this frication by measuring harmonicity, or the ratio of harmonics to noise. They found that while [ʍ] and [w] generally differed in harmonicity, there was considerable overlap of harmonicity values and that [w] tokens in <wh>-words were different from those in other words (Li & Gut, 2022). Complementing these findings, I analyze the center of gravity of the periods of frication preceding the glide.

Center of gravity (CoG), or spectral mean, is a measure commonly used in phonetic research to describe fricatives (e.g., Gordon, Barthmaier, & Sands, 2002; Jongman, Wayland, & Wong, 2000; Zimman, 2017). It represents a weighted average of the frequency components of a spectrum and shows the locus of high energy in the spectrum. In the only other study of (HW) considering CoG, Bridwell (2019:120) reported two categories of frication, one very similar to [h] (as in [hw]) and one associated with "true labiovelars" [ʍ]. For the former, the aperiodic noise was spread across frequencies and thus associated with a high CoG. For the latter, the noise was clustered at much lower frequencies and thus associated with a lower CoG.

Formants represent local peaks of acoustic energy that are estimated from the spectral envelope. In phonetics, the first three formants are commonly used to distinguish voiced sounds. Lawson and Stuart-Smith (1999) described distinctive formant patterns for the voiced glides of their variants: [w] tokens were characterized by low F1 and low and weak F2 contours before rising toward the expected formant loci of the following vowel. [ʍ] was characterized by an abrupt start of both formants at the onset of voicing without this period of lower formants. Their "intermediate" breathy tokens appeared to fall somewhere in between those extremes, with shorter periods of low F1 than the voiced variant. Notably, in their Glasgow-based study, Lawson and Stuart-Smith (1999) also found some apparently socially conditioned variation, with middle-class children producing a slightly longer period of low F1 than working-class children.

Data and Methods

Participants

The speech of eighteen women born between 1938 and 1994 was analyzed for this study. All had spent most of their life in Edinburgh, sixteen of them in Leith, a traditionally working-class neighborhood in North Edinburgh, and the remaining two in Morningside, a traditionally middle-class neighborhood in South Edinburgh. The group from Leith includes working- and middle-class speakers, and both Morningside speakers are middle-class.

Recordings

The data is comprised of eighteen semistructured sociolinguistic interviews. I conducted fifteen of those interviews between December 2018 and February 2019 in Leith as part of a research project on sociophonetic variation. These one-on-one conversations focused on the participants' experiences of growing up and/or living in Leith and other topics such as their work and hobbies. There was also a reading task (adapted from Schützler [2015]). The remaining three interviews were collected in 2014 by Jonathan Berk as part of a master's thesis exploring differences between Leith and Morningside and similarly focused on speakers' life in Morningside (Berk, 2014). All interviews were recorded using a portable digital recorder and a lavalier microphone in quiet, public spaces and digitized at 44kHz. Neither of the interviewers is from Scotland (or the UK), although we were both residents at the time of the interviews.

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Manual annotation
I orthographically transcribed all recordings, force-aligned them using the Montreal Forced Aligner (McAuliffe, Socolof, Mihuc, Wagner, & Sonderegger, 2017) and annotated tokens in *Praat* (Boersma, 2001). Using auditory perceptual information, spectrogram, and waveform, I first coded each token as fricationless or fricated. During this process, I discovered a range of variants. In addition to the "fully voiced fricationless" ($n = 808$) and "fully voiceless fricated" ($n = 464$), like Bridwell (2019) I also identified "voiced fricated" ($n = 29$) tokens where both frication and glide are characterized by voicing (see Table 1). While most of these occur after voiced segments, it is not clear that this is merely a coarticulation effect as six of the tokens occurred after voiceless obstruents or pauses. Conversely, there are also thirty-seven "voiceless fricationless" tokens featuring a voiceless glide (similar to Lawson & Stuart-Smith, 1999:2543), which appear distinct (both visually and auditorily) from "typical" fricationless tokens. Glides in fricated and fricationless tokens can be breathy ($n = 42$). In addition to these six variants, I also identified a small number ($n = 7$) of tokens that are more similar to [f], [v], and [h]. I annotated the duration of frication manually using changes in spectrogram and waveform from glide to the following vowel as cues (see Figure 2).

Acoustic phonetic measures

Formants were extracted from voiced parts of each <wh>-token (glides) and all "<w>-glides" (i.e., glides in words like *water*) using a semiautomated *Praat* script. Each <wh>-token was visually checked and the glide manually selected (for <w>-glides this process was fully automated). The script records the word and phonological environment of each token and segment duration. The first three to five formants (depending on trackability) were extracted in 5ms intervals, with the

Table 1. Counts of different (HW) variants ($n = 1400$); there were seven tokens annotated as [f] ($n = 5$), [v] ($n = 1$), and [h] ($n = 1$)

Variants	*N*
Frication	
voiceless frication	464
voiced frication	29
breathy	13
Frication total	506
No frication	
voiced glide	808
voiceless glide	37
breathy	42
No frication total	887
Total *N*	1393

maximum formant frequency set to 5500Hz and the window length set to 25ms. F1 and F2 measures taken between 45% and 55% of the voiced duration of the glide were retained in subsequent analysis in *R* (R Core Team, 2017). This narrow window effectively reflects the midpoint of the voiced glides and was preferable to a longer duration to avoid distortion from formant transitions. Furthermore, formants could not be reliably extracted for all tokens beyond this point. Measures were transformed from Hertz to Bark (Traunmüller, 1990).

CoG of the fricated portion of each fricated token was measured using a *Praat* script adapted from DiCanio (2017). This script creates a set number of spectral slices across the middle 80% of the segment (to minimize context effects). CoG measures for each slice are averaged across the segment. To avoid overlap between windows, which would bias the averaged CoG measure toward the middle of the segment, the original script was adapted to automatically adjust the window length based on the duration of the segment. To ensure that each window contained enough data to make inferences, a minimum window length of 5ms was implemented.

Data "tidying" and dataset construction

Three datasets were used for statistical analysis: the full manually annotated dataset of 1,400 tokens of (HW), and two subsets of that dataset. The first subset contains only fricated tokens and is used to analyze variation in CoG. There were no obvious outliers resulting from measurement errors, so no further tokens were excluded from analysis. The second subset contains measurements of F1 and F2 at the midpoint of the voiced portions of both fricated and fricationless glides as well as labial-velar approximants in <w>-tokens (e.g., in *water*). Some tokens were completely voiceless, in others the voiced portion of the segment was too short to reliably extract formants, and, in some, *Praat*'s formant tracking was inadequate. While formant trajectories would be interesting, not enough formant measurements could be extracted for most tokens to reliably explore these. Instead, I opted to only look at the measurement closest to the midpoint within 45% to 55% of the glide duration. This second subset contains 262 fricated tokens, 388 fricationless tokens, and 2,915 <w>-tokens.

Statistical analysis

Statistical analysis was conducted with the *R* package *brms* (Bürkner, 2017), which estimates generalized (non-)linear multivariate multilevel/mixed effects models using the probabilistic programming language *Stan* in R (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li, & Riddell, 2017). The key difference between popular frequentist regression models (as fitted with *lme4* [Bates, Mächler, Bolker, & Walker, 2015]) and their Bayesian equivalents is the underlying philosophical approach to statistics. Bayesian models combine prior information with observed data to estimate (posterior distributions of) model parameters.[3]

In this paper, four models are fitted to four dependent variables: proportion of frication (a rate, beta distribution), CoG (a numeric outcome variable [Hertz], lognormal distribution), F1 (a numeric outcome variable [Bark], lognormal distribution), and F2 (a numeric outcome variable [Bark], lognormal distribution). In this analysis, I use "weakly informative" priors (see Gelman, Jakulin, Grazia Pittau, & Su, 2008) to constrain the parameter space to appropriate estimates, for coefficients, intercept(s), and standard deviations. For example, we know that formant values are likely to fall within a specific range (see supplementary materials for full model specifications). Because Bayesian approaches estimate distributions of parameters, where some parameter values are more probable than others, we avoid asking whether or not there is an effect of a factor (Null Hypothesis Testing[4]) and instead ask what the most probable direction and magnitude of an effect is. In the context of this study, these questions are more relevant. Accordingly, I use the metrics "probability of direction" and "region of practical equivalence" to interpret results (see also Makowski, Ben-Shachar, Chen, & Lüdecke, 2019).[5] For completeness, I also include the median parameter estimate and lower and upper bounds of the 89% Highest Density Interval (HDI) that captures the most probable parameter values.

Interpretation: Probability of Direction

The Probability of Direction (PD) captures the certainty that an effect has the same direction as the median estimate of the posterior distribution (i.e., is positive or negative) (Makowski et al., 2019). The simplest method of computing PD is by counting all samples in the posterior distribution that share a sign with the median estimate and dividing by the number of total samples (i.e., the PD is equivalent to the percentage of positive/negative samples). It answers the question "What is the probability of the direction of the effect of independent variable A?" In the results below, I express PD as a positive/negative direction and the percentage of samples sharing that sign (e.g., -(100%) means that 100% of samples are negative). I also provide a description of this probability of direction along the following scale: "unclear" (<60.0%), "possibly positive/negative" (60.0-69.9%), "likely positive/negative" (70.0-89.9%), "very likely positive/negative" (90.0-100%).

Interpretation: Region of Practical Equivalence

The Region of Practical Equivalence (ROPE) describes an interval that is practically equivalent to zero, based on subject knowledge of what represents a meaningful difference. Effect size can be gauged by considering what percentage of a given posterior distribution falls within that interval. The intuition here is that while the coefficient might not be exactly zero, it may well be too small to be of any practical significance. ROPE directly answers the question "What is the probability that this effect is *not* of practical significance?" In the results below I express ROPE as the percentage of samples falling within ROPE. I also provide a description of the likely practical meaning of the effect along the following scale: "very unlikely meaningful" (>89.9% in ROPE), "unlikely meaningful" (50.0-89.9% in ROPE), "possibly meaningful" (20.0-49.9% in ROPE), "likely meaningful" (5.0-19.9%), "very likely meaningful" (<5.0%).

Variables and hypotheses

Social and linguistic independent variables are social class, year of birth, style, phonetic context, speech rate, and (for formants) type of glide. Speaker and word were included as random effects (intercepts and slopes where appropriate). Social class

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

is defined as “working class” (WC) or “middle class” (MC). The fifteen participants recorded in 2019 chose a social class label for themselves during the interview, while the three participants recorded in 2014 were assigned a social class label based on their occupation. Style was defined as either “conversation” or “reading,” as speakers in the 2019 sample also completed a reading passage. The definition of phonetic context depends on the statistical model and outcome variable. For the model looking at presence and duration of frication, the relevant phonetic context is the preceding context (pause or nonpause, as also used by Brato [2014] and Schützler [2011]). For the CoG model, the relevant phonetic context is the preceding manner of articulation (fricative, plosive, approximant, nasal, vowel, and pause). For F1, what matters is the following vowel height (high, mid, low) and, for F2, following vowel anteriority (front, central, back). Speech rate was operationalized as number of syllabic consonants or vowels per second (measured within chunks not interrupted by pauses of more than 3s). The glide type is only relevant for the formant models and comprises fricated, fricationless, and <w>-glides (e.g., the initial glide in *water*). Categorical variables (social class, phonetic context, style, glide type) were deviance-coded,[6] and numeric variables scaled and centered.

The hypotheses for this study are summarized in Table 2. I expect fricated tokens to be more likely among middle-class speakers and older speakers (see Brato, 2014; Robinson, 2005; Schützler, 2010; Stuart-Smith et al., 2007), and potentially formal styles. Based on work that associates lower CoG with “true labiovelars” (Bridwell, 2019:120; Robinson, 2005:186), I expect CoG to be lower in those formal contexts too. Otherwise, CoG is likely affected by the preceding manner of articulation (Bridwell, 2019). Fricated tokens are expected to be more likely after pauses (Brato, 2014; Schützler, 2010), which might translate to longer periods of frication. Formants are expected to be influenced by phonetic context, social class, and glide type (Lawson & Stuart-Smith, 1999).

Probability and proportion of frication: Zero-inflated beta regression

The probability (“is a given token fricated?”) and relative duration of frication (“how long is the period of frication in fricated tokens?”) can be modeled using a zero-inflated beta regression. Beta regressions are commonly used for proportions as they can model outcomes bounded by the open interval (0,1) (Douma & Weedon, 2019; Ferrari & Cribari-Neto, 2004; Stewart, 2013). Zero-inflated models can handle datasets containing many zeroes (i.e., in our case, many fricationless tokens), and are particularly suitable when there is theoretical reason to believe that the process generating a 0 or non-0 outcome (i.e., a fricationless or fricated token) is distinct from the subsequent process generating a positive outcome (i.e., a token with a particular rate of frication). A posterior predictive check, which simulates data based on priors and observed data, confirms that a zero-inflated beta regression is appropriate for the data.

Center of gravity: Bayesian linear mixed effects (log-normal)

To quantitatively analyze variation in CoG, I use a Bayesian linear effects model. These models are specified similarly to frequentist mixed effects models, including fixed and random intercepts as well as random slopes. CoG (in Hertz) is commonly log-normally distributed (validated via a posterior predictive check).

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Table 2. Summary of all hypotheses explored in this paper

Factor	Frication (rate)	Frication (relative)	CoG	F1	F2
Social class	MC: higher	MC: longer	MC: lower	MC: lower	??
Style	Reading: higher	Reading: longer	Reading: lower	Not tested	??
Year of Birth	Younger: lower	Younger: shorter	Younger: higher	No effect	??
Phonetic context	Post-pausal: higher	Post-pausal: longer	Preceding manner	Following vowel height	Following vowel anteriority
Speech Rate	Slow: higher	Slow: longer	No effect	No effect	No effect
Glide Type	NA	NA	NA	Fricated: higher	Fricated: higher
Statistical model	Zero-inflated beta regression		Linear mixed effects	Linear mixed effects	Linear mixed effects

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

F1 & F2: Bayesian linear mixed effects (log-normal)

F1 and F2 (in Bark) were modeled in two separate Bayesian linear mixed effects models.[7]

Results

Overall, fricationless tokens are more common than fricated tokens, though the exact distribution is conditioned by linguistic context, style, and social class. Interestingly, there is no clear effect of speaker year of birth, as would be expected in a straightforward change in progress. Proportion and type of frication varies depending on linguistic and social context. The glide differs depending on whether or not the token is fricated. In the discussion of the statistical results, I focus on the direction and magnitude of effects rather than exact coefficients. Recall that PD describes the probability of a particular direction of the effect, positive (+) or negative (−), on the dependent variable and that ROPE describes the probability that the effect is *not* of practical significance. For coefficient tables, visualizations, and details about the models, see supplementary materials.

Fricationless tokens are more common

Fricationless tokens are less likely among middle-class speakers, in a reading style, and following a pause; year of birth is likely not a meaningful predictor (Table 3). The zero-inflated component of the model is a logistic regression estimating the probability that a token is fricationless.

The proportion of frication varies

While the probability of direction indicates that the proportion of frication is conditioned by the same factors (and in the same way) as the presence or absence of frication, ROPE ([-0.18, +0.18]) suggests that the effects are not practically meaningful (see Table 4).

The type of frication varies

Frication noise is either diffused across the spectrum (as in Figure 1b) or clustered at low frequencies (Figure 1a). Most speakers produce both types of frication. For all fricated tokens, F1 and F2 start abruptly high at the onset of voicing and do not rise. The statistical analysis suggests that (1) context is a meaningful predictor (higher CoG after fricatives and lower CoG after vowels, approximants, and nasals); (2) style is a possible predictor with lower CoG in reading; and that (3) effects of year of birth (possibly positive) and social class (likely negative) are less likely to be meaningful (see Table 5).

F1 and F2 of glides vary

The qualitative analysis of spectrograms shows the formant patterns also reported by Lawson and Stuart-Smith (1999): fricationless tokens feature a period of low F1 and F2 before the appearance of higher formants and a movement toward the following vowel formants, while fricated tokens show an abrupt high start of F1 and F2. To confirm these observations quantitatively, F1 and F2 were measured at the midpoints of glides.

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Table 3. Probability of a fricationless (HW) token: summary of the zero-inflated component of the zero-inflated beta regression (*n* = 1400, overall rate of fricationless tokens = 64%)

Factors	Hypothesis	PD	ROPE	Interpretation	HDI (89%)
Intercept					Median = 0.421 [−0.461, 1.134]
Preceding (Non-pause, *n* = 966, 70%)					
Pause (*n* = 434, 50%)	less likely (negative)	−(100%)	0%	very likely negative, very likely meaningful	Median = −1.137 [−1.401, −0.872]
Style (Conversation, *n* = 1314, 65%)					
Reading (*n* = 86, 42%)	less likely (negative)	−(99.8%)	1.4%	very likely negative, very likely meaningful	Median = −0.847 [−1.319, −0.358]
Social Class (Working, *n* = 685, 75%)					
Middle Class (*n* = 715, 52%)	less likely (negative)	−(94.6%)	3.4%	very likely negative, very likely meaningful	Median= −1.731 [−3.595, −0.035]
Year of Birth *M* = 1963					
+1SD(YOB) = 16.9	more likely (positive)	−(59.0%)	24.0%	unclear, possibly meaningful	Median = −0.132 [−1.054, 0.872]
Speech rate *M* = 3.36 syllables/s					
+1SD(Speech rate) = 0.73	more likely (positive)	+(55.7%)	93.2%	unclear, very unlikely meaningful	Median = 0.014 [−0.142, 0.172]
Interactions					
Social Class*YOB	more likely (positive)	+(63.8%)	11.6%	possibly positive, likely meaningful	Median = 0.404 [−1.574, 2.312]

ROPE is defined as [-0.18, +0.18]. Numeric variables are scaled and centered, categorical variables are deviance-coded. Percentages in first column indicate rate of fricationless.

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Table 4. Proportion of frication in (HW) fricated tokens: summary of conditional component of the zero-inflated beta regression ($n = 1400$, mean overall proportion of frication = 28%)

Factors	Hypothesis	PD	ROPE	Interpretation	HDI (89%)
Intercept					Median = 1.077 [0.891, 1.261]
Social Class (Working Class, *n* = 685, 18%)					
Middle Class (*n* = 715, 37%)	longer (positive)	+94.9%	19.9%	very likely positive, likely meaningful	Median = 0.347 [0.008, 0.702]
Year of Birth *M* = 1963					
+1SD(YOB) = 16.9	shorter (negative)	+81.8%	76.7%	likely positive, unlikely meaningful	Median = 0.099 [−0.083, 0.279]
Style (Conversation, *n* = 1314, 27%)					
Reading (*n* = 86, 45%)	longer (positive)	+67.4%	93.1%	possibly positive, very unlikely meaningful	Median = 0.040 [−0.105, 0.183]
Preceding (Non-pause, *n* = 966, 23%)					
Pause (*n* = 434, 38%)	longer (positive)	−83.7%	99.1%	likely negative, very unlikely meaningful	Median = −0.054 [−0.140, 0.034]
Speech rate *M* = 3.36 syllables/s					
+1SD(Speech rate) = 0.73	shorter (negative)	−88.4%	100%	likely negative, very unlikely meaningful	Median = −0.043 [−0.102, 0.015]
Interactions					
Social Class*YOB	shorter (negative)	+81.8%	42.9%	likely positive, possibly meaningful	Median = 0.195 [−0.149, 0.567]

ROPE is defined as [-0.18, +0.18]. Percentages in first column indicate mean proportion of frication.

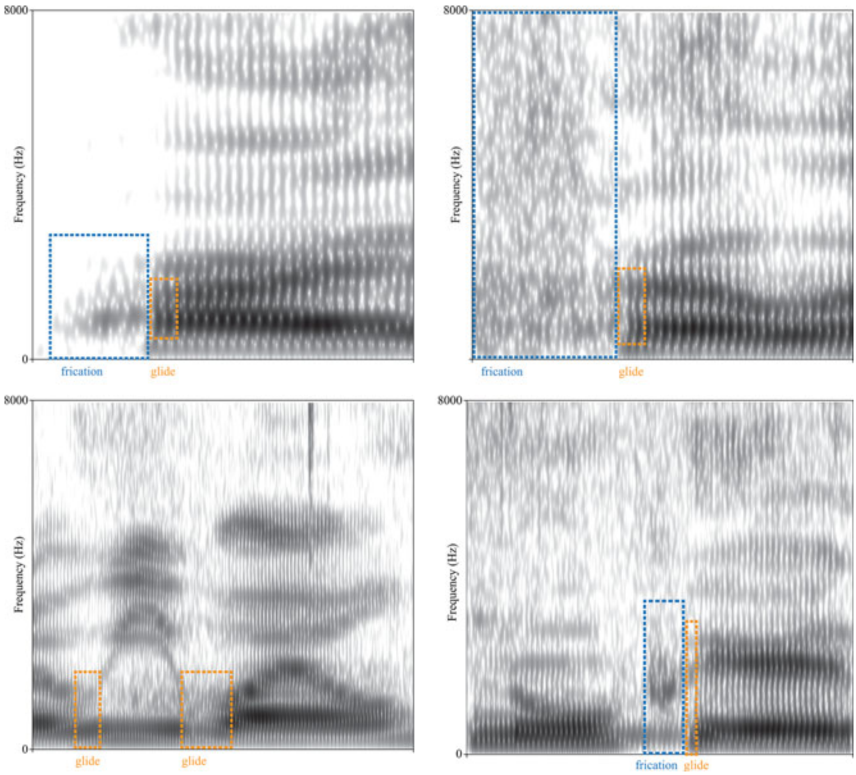

Figure 1. Frication is either spread across frequencies or clustered low. For fricated tokens, F1 and F2 start abruptly at the onset of voicing and do not rise. For fricationless tokens, F1 starts low and rises, and F2 is weak. Figure 1a (top left). Younger working-class woman (Lily) reading "why." Frication is clustered at low frequencies. Figure 1b (top right). Lily saying "while." Here the aperiodic noise is spread across frequencies. Figure 1c (bottom left). Young working-class woman (Fiona) saying "a wee while." Both glides are voiced and the (HW) realization is fricationless. Glides are characterized by a period of low F1 and low and weak F2. Figure 1d (bottom right). Lily saying "nowhere." The period of frication is clearly visible and audible as is the voice bar throughout.

In linear models, ROPE is often conventionally defined as $[-0.1\sigma, +0.1\sigma]$ where σ denotes the standard deviation of the dependent variable. For F1, this is equivalent to an absolute difference of 0.11 Bark from the intercept. For F2, this is equivalent to an absolute difference of 0.16 Bark from the intercept. For F1, I find that (1) glides in fricated (<wh>) tokens have likely higher F1, (2) glides with a shorter duration have lower F1, (3) middle-class speakers have possibly lower F1, and that (4) most effects of linguistic contexts are very small (see Table 6). For F2, glides in fricated tokens have lower F2 (mediated by following vowel anteriority), glides with a shorter duration have lower F2, and middle-class speakers have possibly lower F2 (see Table 7).

Discussion

Beyond the binary distinction of *which* and *witch* there are a range of other variants, whose use appears conditioned by phonetic context, social class, and style.

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

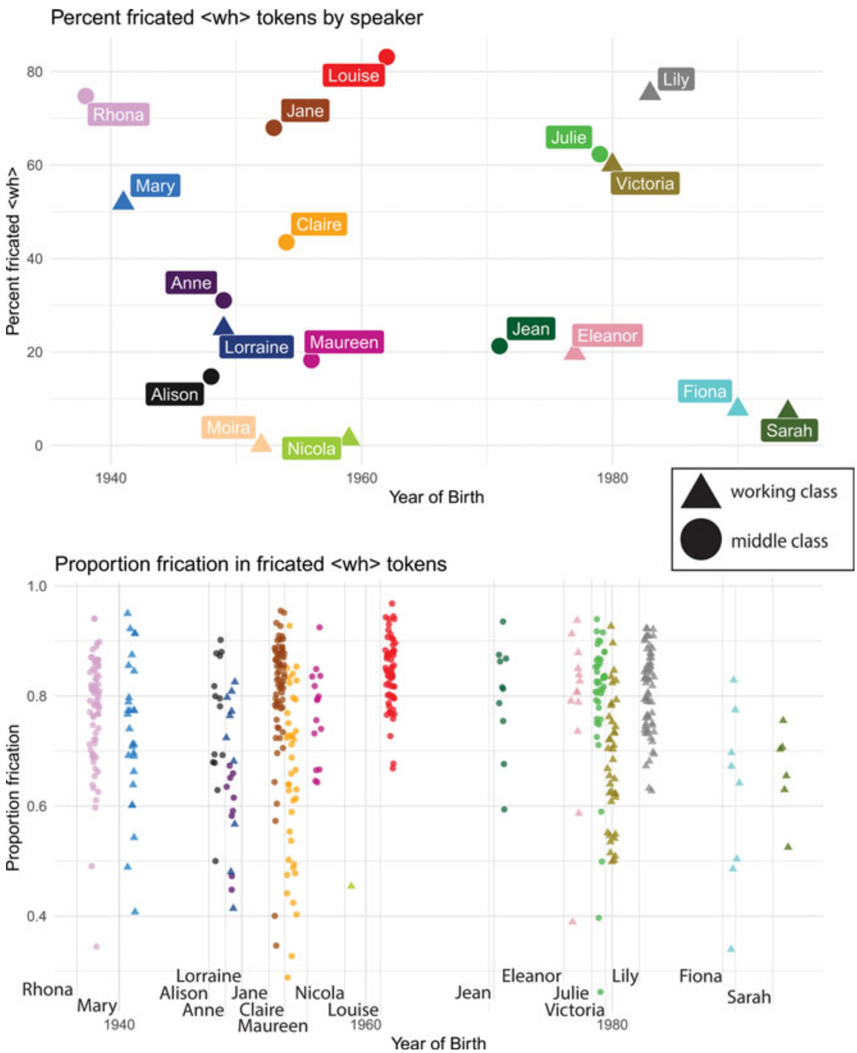

Figure 2. Rate and relative duration of frication varies greatly by speaker, social class, and year of birth. Figure 2a (top) shows the percentage of fricated tokens produced by speaker, year of birth, and social class. Lily and Victoria stand out with their exceptionally high rates given their age and social class. Figure 2b (bottom) shows the relative duration of frication in every fricated token by speaker, year of birth, and social class. The duration of frication is highly variable. Each shape represents a fricated <wh>-token and the y-axis indicates which proportion of the total segment duration was annotated as frication. (Note that Moira does not produce *any* fricated tokens.)

(HW) as a sociolinguistic variable: the variants

There are six main variants that differ in terms of frication (fricated/fricationless), glide quality, voicing (voiced/voiceless frication, voiced/voiceless glide), and phonation (breathy/modal).

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Table 5. Summary of log-normal mixed effects linear regression model for CoG ($n = 508$, grand mean CoG = 1243 Hz)

Factors	Hypothesis	PD	ROPE	Interpretation	HDI (89%)
Intercept					Median = 6.894 [6.779, 7.016]
Preceding (Pause, $n = 241$, 1224Hz)					
Approximant ($n = 34$, 970Hz)	lower (negative)	−100%	1.1%	very likely negative, very likely meaningful	Median = −0.241 [−0.335, −0.141]
Vowel ($n = 88$, 1119Hz)	lower (negative)	−100%	1.2%	very likely negative, very likely meaningful	Median = −0.197 [−0.266, −0.129]
Fricative ($n = 34$, 1953Hz)	higher (positive)	+100%	1.2%	very likely positive, very likely meaningful	Median = 0.239 [0.142, 0.338]
Nasal ($n = 34$, 1168Hz)	lower (negative)	−100%	1.8%	very likely negative, very likely meaningful	Median = −0.227 [−0.322, −0.130]
Plosive ($n = 77$, 1301Hz)	?	−53.9%	97.8%	unclear, very unlikely meaningful	Median = 0.004 [−0.065, 0.072]
Style (Conversation, $n = 458$, 1279Hz)					
Reading ($n = 50$, 915Hz)	lower (negative)	−98.6%	19.5%	very likely negative, likely meaningful	Median = −0.167 [−0.308, −0.037]
Year of Birth $M = 1961$					
YOB+1 SD(YOB) = 16.6	higher (positive)	−93.9%	40.7%	very likely negative, possibly meaningful	Median = −0.116 [−0.232, 0.001]
Social Class (Working Class, $n = 169$, 1012Hz)					
Middle Class ($n = 339$, 1358Hz)	?	+70.1%	45.5%	likely positive, possibly meaningful	Median = 0.077 [−0.158, 0.329]
Speech rate $M = 2.8$ syllables/s					
+1 SD(Speech rate) = 0.5	?	−57.1%	100%	unclear, very unlikely meaningful	Median = 0.003 [−0.026, 0.032]

Outcome variable: CoG of the period of frication in Hertz. ROPE is defined as an absolute difference of 100 Hz. Hz in first column indicates mean CoG.

Table 6. Linear mixed effects regression for F1 ($n = 3565$, grand mean F1 = 443Hz)

Factors	Hypothesis	PD	ROPE	Interpretation	HDI (89%)
Intercept					Median = 0.996 [0.924, 1.071]
Duration (log) *M* = 0.05s (–3.1)					
+1 SD(duration (log)) = 0.16	?	-100%	<0.1%	very likely negative, very likely meaningful	Median = –0.072 [–0.084, –0.059]
Type (<w>-glide, *n* = 2915, 437Hz)					
Fricated (*n* = 262, 497Hz)	higher F1 (positive)	+95.1%	29.3%	very likely positive, possibly meaningful	Median = 0.058 [0.002, 0.113]
Fricationless (*n* = 388, 458Hz)	lower F1 (negative)	+90.5%	63.4%	very likely positive, unlikely meaningful	Median = 0.031 [–0.007, 0.071]
Social Class (Working Class, *n* = 1846, 469Hz)					
Middle Class (*n* = 1719, 417Hz)	?	-86.3%	32.6%	likely negative, possibly meaningful	Median = –0.058 [–0.143, 0.03]
Year of Birth *M* = 1964					
+1 SD(Year of Birth) = 4.9	?	+91.5%	46.6%	very likely positive, possibly meaningful	Median = 0.042 [–0.009, 0.090]
Following vowel (Mid vowel, *n* = 1820, 445Hz)					
High vowel (*n* = 1213, 436Hz)	lower F1 (negative)	–92.8%	84.4%	very likely negative, unlikely meaningful	Median = –0.024 [–0.049, 0.003]
Low vowel (*n* = 532, 456Hz)	higher F1 (positive)	–54.4%	99.6%	unclear, very unlikely meaningful	Median = –0.002 [–0.024, 0.020]
F2 *M* = 10.03 Bark					
+1 SD(F2) = 0.36	higher F1	+100%	100%	very likely positive, very unlikely meaningful	Median = 0.025 [0.022, 0.029]

ROPE is defined as a difference of 0.11 Bark. Hz in first column indicates mean F1.

Table 7. Linear mixed effects regression for F2 ($n = 3565$, grand mean F2 = 1310Hz)

Factors	Hypothesis	PD	ROPE	Interpretation	HDI (89%)
Intercept					Median = 1.932, [1.886, 1.975]
F1 *M* = 4.36 Bark					
+1 SD(F1) = 0.14	higher F2	+100%	0%	very likely positive, very likely meaningful	Median = 0.034 [0.029,0.038]
Duration (log) *M* = 0.05s (-3.1)					
+1 SD(duration (log)) = 0.16	?	-100%	0.02%	very likely negative, very likely meaningful	Median = −0.041 [−0.051, −0.032]
Type (<w>-glide, *n* = 2915, 1313Hz)					
Fricated (*n* = 262, 1273Hz)	higher (positive)	-100%	0.14%	very likely negative, very likely meaningful	Median = −0.100 [−0.141, −0.058]
Fricationless (*n* = 388, 1307Hz)	lower (negative)	-91.3%	27.6%	very likely negative, possibly meaningful	Median = −0.034, [−0.072, 0.006]
Following vowel (central vowel, *n* = 994, 1341Hz)					
Front vowel (*n* = 1636, 1334Hz)	higher (positive)	+100%	0.3%	very likely positive, very likely meaningful	Median *M* = 0.105 [0.056, 0.151]
Back vowel (*n* = 935, 1234Hz)	lower (negative)	-96.7%	9.8%	very likely negative, likely meaningful	Median = −0.062 [−0.116, −0.008]
Interactions					
Fricated*front vowel (*n* = 163, 1436Hz)	?	+97.1%	5.5%	very likely positive, likely meaningful	Median = 0.097 [0.016, 0.177]

Fricationless*front vowel (*n* = 268, 1390Hz)	?	+88.6%	15.5%	likely positive, likely meaningful	Median = 0.061, [−0.019, 0.141]
Fricationless*back vowel (*n* = 95, 1126Hz)	?	+80.2%	20.0%	likely positive, possibly meaningful	Median = 0.047 [−0.041, 0.136]
Fricated*back vowel (*n* = 65, 969Hz)	?	−70.3%	24.6%	likely negative, possibly meaningful	Median = −0.03 [−0.116, 0.060]
Social Class (Working Class, *n* = 1846, 1325Hz)					
Middle Class (*n* = 1719, 1294Hz)	?	−90.6%	32.8%	likely negative, possibly meaningful	Median = −0.028, [−0.062, 0.007]
Year of Birth *M* = 1964					
+1 SD(Year of Birth) = 4.9	?	−99.4%	77.0%	likely negative, unlikely meaningful	Median = −0.031 [−0.049, −0.012]

ROPE is defined as a difference of 0.16 Bark. Hz in first column indicates mean F2.

Variation in frication

For tokens perceived as "fricated," frication accounts for at least 26% to 96% of token duration (see Figure 2b). Meaningful predictors conditioning this proportion of frication are speaker's social class and preceding phonetic environment. Notably, year of birth is not a meaningful predictor here, suggesting that there is no ongoing gradual loss of frication.

Variation in CoG is conditioned by style, preceding phonetic environment, and speaker's social class. If the preceding segment is fricative, CoG is significantly higher than after a pause. Conversely, CoG is significantly lower following an approximant, nasal, or vowel sound. While these coarticulation effects are not particularly surprising, the effects of social class and preceding pause on the realization of fricated tokens is interesting.

Variation in glides

Fricated and fricationless variants differ not just in frication but also in glide quality. The most meaningful predictors of F1 are frication and token duration. Glides in fricated tokens have a higher F1 at the midpoint than those in fricationless tokens. This confirms the observation (also made by Lawson & Stuart-Smith [1999]) that glides in fricationless tokens are characterized by a period of low F1, while fricated tokens show a very abrupt start of raised F1 and F2. I also identify a small effect of social class, whereby middle-class speakers produce tokens with lower F1 than working-class speakers. This echoes Lawson and Stuart-Smith's (1999) finding that middle-class children produce a longer period of low F1 than working-class children, corresponding, presumably, to a lower midpoint.

Fricated tokens have lower F2. Effects of phonetic environment are also likely meaningful and follow expectations: tokens preceding front vowels show higher F2 while those preceding back vowels show lower F2. There is also an interaction effect of frication and phonetic context where fricated glides appear more strongly influenced by their phonetic context than fricationless tokens. This could be due to the, on average, shorter duration of fricated glides. Context effects could be more pronounced as their midpoint is closer to the next segment than in a fricationless glide. A limitation of this analysis is that midpoints are not an ideal proxy for the formant trajectories considered in the qualitative analysis. While I have been assuming that tokens characterized by a rising F1 (fricationless) have a lower midpoint than those where F1 starts abruptly and high and remains stable (fricated), there could be a lot of variation regarding the formant trajectories.

Variation in voicing

I also observe tokens that are either fully voiced (including frication) or fully voiceless. Notably, these do not exclusively occur in environments that would give rise to coarticulation effects. Bridwell (2019) accounts for voiced [hw] by positing that the underlying representation of the voiceless labial-velar fricative is /hw/, which undergoes voicing in appropriate environments and surfaces as [w]. Since no participant in my study produces only voiced fricated tokens after voiced segments, this explanation

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

does not apply here. The devoiced [w] tokens are perceptually voiceless while lacking frication.

Variation in phonation type

A sizable subset of tokens are breathy. These variants are perceptually and acoustically hardest to pin down, as they vary in degree and duration of breathiness. While there are not enough of these tokens for a quantitative analysis, single spectrograms suggest that they are highly variable and either pattern more with "prototypical" fricated or fricationless tokens, depending on whether they show frication preceding the glide. Generally, they appear more frequently among younger speakers. Sarah (born 1993, working class) and Fiona (born 1990, working class) produced 10% ($n = 8$) and 5% ($n = 5$) of all their tokens as breathy, respectively. Notably, these two women have some of the lowest rates of frication (both under 10%). Similarly, Lily (born 1983, working class) produced 8% of breathy tokens ($n = 7$). However, unlike Sarah and Fiona, Lily shows a very high rate of frication (over 70%). It is therefore not clear whether the rate of frication is related to the rate of breathiness. Social class could be an explanatory factor here. The prevalence of breathy (HW) realizations among young working-class women mirrors Lawson and Stuart-Smith's (1999) observation that breathy fricated variants are more common among working-class children (who would have been born around the same time as Sarah and Fiona). However, the highest rate of breathy tokens is found in Jean's speech (17%, born 1971, middle class, low rate of frication) and the absence of any middle-class women born after 1980 in this sample means this hypothesis remains untested here.

No apparent time change in this sample?

Following the Apparent Time approach, speakers are expected to reflect the linguistic norms of their speech community when they acquired the variety (Sankoff, 2006:115). If there was a change in progress, we would expect younger participants to produce higher rates of fricationless tokens than older participants. The probability of direction of the effect of year of birth in the zero-inflated component of the zero-inflated beta regression in Table 3 would then be positive (later year of birth ~ higher probability of fricationless token). Sixty percent of speakers predominantly use fricationless tokens (Figure 2a), which suggests that some change probably has taken place since older descriptions of Edinburgh English have noted that (only) "some younger speakers" use fricationless variants variably (Chirrey, 1999:36). However, there is little evidence of ongoing change.

Style and lexical variation

As hypothesized, tokens produced after a pause (about a third of the dataset) are much more likely to be fricated than those following a nonpausal segment. Unfortunately, other factors such as lexical frequency and lexical category are exceptionally hard to disentangle from phonetic context. Highly frequent words featuring (HW) tend to be closed class items (e.g., *what*, *which*, *where*, *when*, *why*), while open

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

class items are much rarer (e.g., *whisky*, *whistle*, *whale*). Most instances (93%) of (HW) are furthermore word-initial (exceptions include *elsewhere*, *anywhere*, *nowhere*). Schützler (2010:5) argues that the preference for fricated tokens after a pause is an articulatory effect. At a lower speech rate or after a pause there is "more time" to articulate the "slightly more effortful" fricated variant after a pause. However, since I do not find such a speech rate effect, I would argue that the post-pausal context favors fricated tokens for the same reason that the reading style does: because fricated tokens are part of a more careful or formal speech style.

Style (conversation or reading) conditions whether tokens are fricated, relative duration, and CoG of frication. Tokens are more likely to be fricated in a reading style. This effect is one of the strongest predictors of frication. Tokens in a reading style are somewhat more likely to have a longer period of frication, but this effect is likely very small. Read tokens do, however, have a meaningfully lower CoG. They appear to be most similar to Robinson's "voiceless lip-rounded consonant with audible friction at both velar and bilabial articulations" (2005:184), which she posits to be the "traditional form," and to Bridwell's "true voiceless labiovelar glides" (2019:120). These can be contrasted with tokens in which frication is more diffuse across frequencies more similar to a glottal fricative. One interpretation of the effect of speech style on CoG is that tokens with lower CoG are produced by speakers when they pay more attention to their speech because they are more prestigious. This prestige may be the result of their association with SSE, as some of the speakers with the highest rates of fricated tokens clearly orient toward the standard and/or describe negative attitudes toward Scots. Conversely, speakers who use more Scots lexis favor fricationless tokens. The most common <wh>-noun featuring in this corpus is *whisky*. Family members of two informants, Louise and Jane, used to work in (now defunct) whisky companies in Leith. Of all the occurrences of the word *whisky* ($n = 8$), five are fricated. Jane produces two of three tokens with the fricated variant (a slightly higher rate than her average), Louise produces both tokens with frication (she also has the highest average rate of frication of all speakers at 83%), Julie produces one of two tokens with frication, and Mary produces one token without frication. Another locally salient (HW) word is *whaling*. Leith used to be an active whaling port, and the oldest informant, Rhona (born 1938), recalls whaling boats in the Leith docks, and Moira (born 1951) notes that her father used to work as a whaler. Rhona, a retired teacher (middle class) with one of the highest rates of frication, produces two tokens in this context with frication, while Moira, a retired laboratory technician (working class) with the second lowest rate of frication (7%), produces two without.

Social class, identity, and changing neighborhoods

Speaker social class is a predictor of both amount and type of frication. Middle-class speakers produce (HW) tokens that are both *more likely* to be fricated and, if fricated, more likely to be *more* fricated. Middle-class speakers furthermore tend to produce fricated tokens with a higher CoG than working-class speakers (though this effect is likely smaller). These findings echo Stuart-Smith et al. (2007), who also reported

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

that Glasgow middle-class speakers favored fricated variants, while working-class speakers favored fricationless tokens.

An exception to this pattern are Lily (born 1983) and Victoria (born 1980), both in Leith and identifying themselves as "working class." They produce higher rates of fricated tokens than all other working-class speakers (Victoria: 60%, Lily: 75%), and much higher rates than other working-class women their age. Notably, their educational and professional backgrounds suggest that both style-shift along the Scots-SSE continuum (Stuart-Smith, 2004), and that they could perhaps be described as upper working class or "new middle class" (Dickson & Hall-Lew, 2017). Lily works in finance administration at a university, having previously worked in insurance but not having attended university herself. In her interview, she talks about shifting from "speak[ing] Leith," a variety she describes as having "its own words and phrases," to "an Edinburgh accent," especially when interacting with colleagues from outside of Scotland. Victoria is a community officer in Leith who studied at Edinburgh University. Victoria notes that, in her perception, the way people speak in Leith has changed between generations (though she was not asked about [HW] specifically). She finds that "the older generation definitely have a different dialect from [her]self and [her] brothers" and that people her age in Leith speak very similarly to people elsewhere in Edinburgh (likely referring to SSE). These perceptions are potentially colored by her broader negative attitudes toward Scots: she explains that she believes that children should not be taught Scots in schools and that she does not want her grandmother to speak Scots to her children. This metalinguistic commentary reveals that Victoria is very concerned with "speaking properly." Fricated (HW) appears to be part of this targeted style. Victoria's comments about language (however inflected by her attitudes) and both women's relationship to social class also speak to a real ongoing change in Leith. Crucially, the apparent time construct assumes that adult speakers remain relatively stable over their lifetime and that the speech community they were raised in is fundamentally the same today as it was then. While real time and panel studies that consider data collected at different points in time provide strong support for a model of intergenerational language change and intragenerational stability (e.g., Denis, Gardner, Brook, & Tagliamonte, 2019; Fruehwald, 2017), changes within speakers across the lifespan and broader external changes affecting the speech community are likely also factors. As shown above, speakers like Lily style-shift frequently and have (somewhat consciously) accommodated to a variety or standard other than the one they spoke growing up. Furthermore, over the course of the twentieth century, Leith has undergone drastic changes as the result of deindustrialization. After a period of economic decline (somewhat infamously portrayed in Irvine Welsh's (1994) novel *Trainspotting*), Leith has become one of the most densely populated and diverse areas of Scotland, and, in recent years, has been rapidly gentrifying (Doucet, 2009). The speech community in which the oldest participant Rhona (born 1938) grew up is therefore very different from the one the youngest informant Sarah was born into in 1993.

Social class and local identity may also offer a better alternative account for variation than contact with Southern British English invoked in prior work. Looking only at middle-class speakers from Edinburgh, Schützler (2010) argued that the adoption of fricationless [w] is an effect of higher education and contact with Southern British varieties of English (which in Edinburgh are closely intertwined). Most SSBE speakers, he argued, did not produce fricated variants at all, and the observed "merger" was thus an effect of language contact. Contrary to Schützler's argument, though, the informants in this sample with the most contact with SSBE and higher education tend to also be the speakers with the highest rates of frication (i.e., the least likely to "merge"). This is especially apparent in young upwardly mobile working-class women like Victoria and Lily, who appear to orient toward SSE, or, as Lily puts it, "Edinburgh English." The idea that fricated variants are associated with SSE can also account for the stylistic effect, as speakers are generally more likely to use a "more standard" form while reading. The lowest rates of frication are found among working class women who have not had much contact with higher education, both older (Moira, Nicola, Lorraine) and younger (Fiona, Sarah). In the context of Glasgow, Stuart-Smith et al. (2007) argued that young working-class speakers used the fricationless variant, which may well have entered originally from Southern British varieties (though perhaps not face-to-face contact[8]) to index distance from (or opposition to) the middle-class norm. Like in Glasgow (Stuart-Smith et al., 2007), linguistic differences between social class groups observed in Edinburgh could be the result of changing neighborhoods and social networks, changes in Scots among working-class speakers and distinct linguistic norms for different social class groups.

Conclusion

(HW) as a complex sociolinguistic variable

In this study, I have described the so-called *which* ~ *witch* merger as a sociolinguistic variable with six internally heterogeneous variants. Realizations of (HW) differ most notably in presence or absence of frication, relative duration of frication, type of frication, glide quality, phonation, and voicing. Contrary to other studies on (HW) in Scotland (and Edinburgh), I do not find evidence for *ongoing* change or effects of contact with Southern British English. Rather, variant selection and realization are conditioned by social class, style, and phonetic environment. Fricated variants are particularly prevalent in the speech of middle-class and "new middle-class" or upwardly mobile working-class women, as well as in formal speech styles. A more specifically designed project using laboratory recordings could shed some light on effects of lexical frequency, phonetic environment, and semantic content. Another striking finding to explore further is that the variants identified here are very similar to those found in other varieties of English (in Scotland, but also in the United States).

Acknowledgments. This work was supported in part by the UKRI Centre for Doctoral Training in Natural Language Processing, funded by the UKRI (grant EP/S022481/1) and the University of Edinburgh, School of Informatics and School of Philosophy, Psychology & Language Sciences. I would like to thank Lauren Hall-Lew, Catherine Lai, Josef Fruehwald, Adam Lopez, two anonymous reviewers, the LVC editorial team, the audience at UKLVC 13, the University of Edinburgh LVC Research Group, and Lukas Eigentler for comments and advice. Most importantly, I would like to thank the research participants who generously shared their experiences and time with me.

Competing Interest. The author declares no competing interests.

Notes

1. For a detailed introduction to language in Scotland, see Lawson (2014).
2. As noted by Wells (1982), in very formal registers of RP some speakers do retain a contrast.
3. For a detailed introduction to Bayesian statistics, I recommend McElreath (2016) and Kruschke (2011). For an introduction to *brms* in the context of linguistics, see Nalborczyk, Batailler, Lœvenbruck, Vilain, and Bürkner (2019), and Vasishth, Nicenboim, Beckman, Li, and Kong (2018).
4. In a frequentist linear effects model, $p < 0.05$ attached to independent variable A means: "in any 100 samples, we would only expect fewer than five samples to contain data as extreme or more extreme if the null hypothesis ('there is no effect of A') is true."
5. I direct any interested reader to the supplementary materials that contain the data, full model specifications, full model results, and fitted models, as well as all other diagnostics mentioned in the text: https://github.com/ninamarkl/hw_edinburgh.
6. Deviance (or contrast or sum) coding is an alternative to the more commonly used dummy coding. The intercept of a deviance-coded model represents a weighted mean of different groups, and all effects are "main effects" (Schad, Vasishth, Hohenstein, & Kliegl, 2020).
7. The log-normal distribution was chosen as the distributional family. A posterior predictive check confirmed that this distribution approximates the data better than a Gaussian would.
8. See also Stuart-Smith, Pryce, Timmins, & Gunter (2013).

References

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Berk, Jonathan. (2014). *The Scottish vowel length rule in Morningside and Leith*. Master's thesis, The University of Edinburgh.

Boersma, Paul. (2001). Praat, a system for doing phonetics by computer. *Glot International* 5 (9/10):341–45.

Brato, Thorsten. (2014). Accent variation and change in North-East Scotland: The case of (HW) in Aberdeen. In R. Lawson (ed.), *Sociolinguistics in Scotland*. London: Palgrave Macmillan. 32–52.

Bridwell, Keiko. (2019). *The distribution of [HW]: An acoustic analysis of sociophonetic factors governing the wine-whine merger in Southern American English*. Master's Dissertation, University of South Carolina.

Bürkner, Paul-Christian. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80:1–28.

Carpenter, Bob, Gelman, Andrew, Hoffman, Matthew D., Lee, Daniel, Goodrich, Ben, Betancourt, Michael, Brubaker, Marcus, Guo, Jiqiang, Li, Peter, & Riddell, Allen. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software* 76(1):1–32.

Chirrey, Deborah. (1999). Edinburgh: Descriptive material. In G. Docherty & P. Foulkes (eds.), *Urban voices: Accent studies in the British Isles*. London: Arnold. 428–43.

Denis, Derek, Gardner, Matt Hunt, Brook, Marisa, & Tagliamonte, Sali A. (2019). Peaks and arrowheads of vernacular reorganization. *Language Variation and Change* 31(1):43–67.

DiCanio, Christian. (2017). Time averaging for fricatives. https://www.acsu.buffalo.edu/~cdicanio/scripts/Time_averaging_for_fricatives_2.0.praat.

Dickson, Victoria, & Hall-Lew, Lauren. (2017). Class, gender, and rhoticity: The social stratification of non-prevocalic /r/ in Edinburgh speech. *Journal of English Linguistics* 45(3):29–59.

Doucet, Brian. (2009). Living through gentrification: Subjective experiences of local, non-gentrifying residents in Leith, Edinburgh. *Journal of Housing and the Built Environment* 24(3):299–315.

Douma, Jacob C., & Weedon, James T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution* 10(9):1412–30.

Ferrari, Silvia, & Cribari-Neto, Francisco. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics* 31(7):799–815.

Fruehwald, Josef. (2017). Generations, lifespans, and the zeitgeist. *Language Variation and Change* 29(1):1–27.

Fruehwald, Josef, Hall-Lew, Lauren, Eiswirth, Mirjam, Boyd, Zac, & Elliott, Zuzana. (2019, September 3-5). *A phonetic analysis of the which-witch merger in Edinburgh, Scotland* [poster presentation]. 12th UK Language Variation and Change (UKLVC 12) Conference, London, UK.

https://doi.org/10.1017/S0954394523000078 Published online by Cambridge University Press

Gelman, Andrew, Jakulin, Aleks, Grazia Pittau, Maria, & Su, Yu-Sung. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* 2(4):1360–83.
Giegerich, Heinz J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.
Gordon, Matthew, Barthmaier, Paul, & Sands, Kathy (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32(2):141–74.
Jones, Charles. (2002). *The English language in Scotland: An introduction to Scots*. East Linton, UK: Tuckwell.
Jongman, Allard, Wayland, Ratree, & Wong, Serena (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 108(3):1252.
Kruschke, John K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. Burlington MA: Academic Press.
Labov, William, Ash, Sharon, & Boberg, Charles. (2008). *The Atlas of North American English: Phonetics, phonology and sound change*. Berlin: De Gruyter Mouton.
Lawson, Robert. 2014. *Sociolinguistics in Scotland*. London: Palgrave Macmillan.
Lawson, Eleanor, & Stuart-Smith, Jane. (1999). A sociophonetic investigation of the Scottish consonants /x/ and /Hw/ in the speech of Glaswegian children. 14*th International Congress of Phonetic Sciences*, 2541–44.
Li, Zeyu, & Gut, Ulrike. (2022). The Distribution of /w/ and /ʍ/ in Scottish Standard English. *Corpus Linguistics and Linguistic Theory*. https://doi.org/10.1515/cllt-2021-0052
Macafee, Caroline. (1983). *Glasgow*. Amsterdam: Benjamins.
Makowski, Dominique, Ben-Shachar, Mattan S., Chen, S. H. Annabel, & Lüdecke, Daniel. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology* 10:2767.
Marshall, James Scott. (1986). *The life and times of Leith*. Edinburgh: Donald.
McAuliffe, Michael, Socolof, Michaela, Mihuc, Sarah, Wagner, Michael, & Sonderegger, Morgan. (2017). Montreal forced aligner: Trainable text-speech alignment using Kaldi. In *Proceedings of Interspeech* 17:498–502.
McElreath, Richard. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: CRC Press/Taylor & Francis Group.
Milroy, Lesley. (2004). An essay in historical linguistics?: On Donka Minkova's "Philology, linguistics, and the history of [hw]~[w]." In A. Curzan & K. Emmons (eds.), *Studies in the history of the English language II: Unfolding conversations*. Bern: De Gruyter Mouton. 46–53.
Minkova, Donka. (2004). Philology, linguistics, and the history of [hw]~[w]. In A. Curzan & K. Emmons (eds.), *Studies in the history of the English language II: Unfolding conversations*. Bern: De Gruyter Mouton. 7–45.
Nalborczyk, Ladislas, Batailler, Cédric, Loevenbruck, Hélène, Anne, Vilain, & Bürkner, Paul-Christian. (2019). An introduction to Bayesian multilevel models using Brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research* 62(5):1225–42.
R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/
Reiersen, Øystein. (2013). *Edinburgh Jockney? A socio-phonological study of accent variation and change in Edinburgh English*. Master's thesis, The University of Bergen.
Robinson, Christine. (2005). Changes in the dialect of Livingston. *Language and Literature* 14(2):181–93.
Sankoff, Gillian. (2006). Age: Apparent time and real time. *Elsevier Encyclopedia of Language & Linguistics*. 110–116.
Schad, Daniel J., Vasishth, Shravan, Hohenstein, Sven, & Kliegl, Reinhold. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language* 110:104038.
Schützler, Ole. (2010). Variable Scottish English consonants: The cases of /HW/ and non-prevocalic /r/. *Research in Language* 8:1–17.
Schützler, Ole. (2011). Charting vowel spaces in Edinburgh middle-class speech. *English World-Wide* 32 (1):24–45.
Schützler, Ole. (2015). *A sociophonetic approach to Scottish standard English*. Amsterdam: John Benjamins.
Stewart, Connie. (2013). Zero-inflated Beta distribution for modeling the proportions in quantitative fatty acid signature analysis. *Journal of Applied Statistics* 40(5):985–92.
Stuart-Smith, Jane. (2004). Scottish English. In B. Kortmann, K. Burridge, R. Mesthrie, E. W. Schneider, & C. Upton (eds.), *A handbook of varieties of English*. Berlin: Mouton de Gruyter. 47–67.

Stuart-Smith, Jane, Pryce, Gwilym, Timmins, Claire, & Gunter, Barrie. (2013). Television can also be a factor in language change: Evidence from an urban dialect. *Language* 89(3):501–36.

Stuart-Smith, Jane, Timmins, Claire, & Tweedie, Fiona. (2007). 'Talkin' Jockney'? Variation and change in Glaswegian accent. *Journal of Sociolinguistics* 11(2):221–60.

Thomas, Erik R. (2019). Trends from outside. In E. R. Thomas, (ed.), *Mexican American English: Substrate influence and the birth of an ethnolect*. Cambridge: Cambridge University Press. 125–48.

Traunmüller, Hartmut. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* 88(1):97–100.

Vasishth, Shravan, Nicenboim, Bruno, Beckman, Mary E., Li, Fangfang, & Kong, Eun Jong. (2018). Bayesian data analysis in the phonetic sciences: A Ttutorial introduction. *Journal of Phonetics* 71:147–61.

Wells, John Christopher. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Welsh, Irvine. (1994). *Trainspotting*. London: Minerva.

Zimman, Lal. (2017). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society* 46(3):339–70.

Cite this article: Markl N (2023). "I can't see myself ever living any[w]ere else": Variation in (HW) in Edinburgh English. *Language Variation and Change* **35**, 79–105. https://doi.org/10.1017/S0954394523000078