

Testing for Frequency and Structural Effects in an English Stress Shift

MORGAN SONDEREGGER

University of Chicago

Introduction

In many cases of language change, there is exceptionality: only some forms have changed, or different forms are changing at different rates. For any given case of change where exceptions exist, there are often several types of plausible conditioning factors for the behavior of different forms, such as phonetic environment, social factors, or token frequency. Of interest is the relative contribution of proposed factors on a form's likelihood of changing (or rate of change).

This paper considers one such case, the English diatonic stress shift (DSS). We examine the role of frequency and phonological structure as conditioning factors for which of a set of noun/verb pairs have undergone the DSS between 1700 and the present. Previous work by Phillips (1984) has shown a role of frequency: on average, words which have undergone the DSS have lower frequency than those which have not. Using a new dataset, we show via multiple logistic regression that there is a significant effect of frequency in the direction shown by Phillips, as well as effects of phonological structure; for example, a closed initial syllable makes change more likely. There is also a strong *interaction* between the effects of frequency and structure; in particular, structure modulates the strength and direction of the frequency effect. Our use of multiple regression follows its widespread use in sociolinguistics (e.g., Labov 1994) for quantifying the relative effects of different conditioning factors in cases of language change.

The paper is structured as follows. Section 1 gives background on the DSS. Section 2 describes possible conditioning factors related to frequency and phonological

**I thank Max Bane, James Kirby, and Alan Yu for helpful comments on earlier drafts, and Andrew Garrett, Matt Goldrick, Peter Graff, and Partha Niyogi for useful advice. Betty Phillips kindly provided work in progress. Errors or misunderstandings are mine alone. This work was supported in part by a Department of Education GAANN Fellowship.

structure, and Section 3 considers each factor’s effect in isolation. Sections 4–5 discuss regression models for the joint effect of frequency and structure, and Section 6 sums up.

1 Background

1.1 N/V pairs

Of interest here are English disyllabic, homographic noun/verb pairs, referred to as *N/V pairs* throughout. We use the notation $\{N_{\text{stress}}, V_{\text{stress}}\}$ to denote an N/V pair’s stress, with $1=\acute{\sigma}\sigma$, $2=\sigma\acute{\sigma}$.¹ Of the four logically possible stress patterns, all current N/V pairs for which N and V have categorical stress follow one of 3 patterns: $\{1,1\}$ (*fracture*: N $\acute{\sigma}\sigma$, V $\acute{\sigma}\sigma$), $\{1,2\}$ (*convict*: N $\acute{\sigma}\sigma$, V $\sigma\acute{\sigma}$), $\{2,2\}$ (*cement*: N $\sigma\acute{\sigma}$, V $\sigma\acute{\sigma}$). The fourth possible form, $\{2,1\}$, is unattested both synchronically and diachronically (Sherman 1975; Sonderegger and Niyogi 2010). There is also variation in the pronunciation of a few N/V pairs (e.g., *research*, *perfume*).

Change in N/V pair stress was first studied systematically by Sherman (1975), and subsequently by Phillips (1984) and Sonderegger and Niyogi (2010). In dictionary data from 1570 to the present, four types of clear-cut change can be observed between the three stress patterns, at the level of individual pairs: $\{2,2\}\rightarrow\{1,2\}$, $\{1,1\}\rightarrow\{1,2\}$, $\{1,2\}\rightarrow\{1,1\}$, and $\{1,2\}\rightarrow\{2,2\}$.

1.2 The diatonic stress shift

By far the most common type of change observed is $\{2,2\}\rightarrow\{1,2\}$, from the *oxytone* pattern to the *diatone* pattern. We restrict ourselves to this change, the *diatonic stress shift* (DSS).² Only two diatones are recorded before 1570 (Minkova 1997:160), suggesting the DSS began in the 16th century. The DSS has been categorized as lexical diffusion (Sherman 1975; Phillips 1984, 2006) or analogical change (Kiparsky 1995), and indeed the relationship between the two is controversial (see Phillips 2006, vs. Kiparsky 1995; Janda and Joseph 2003). All that is important for our purposes is the basic fact that the DSS is lexically gradual.

Most N/V pairs currently stressed as $\{1,2\}$, and hence most pairs which have undergone the DSS, begin with a prefix, such as *pre* or *re* (Fudge 1984:32). Setting aside why this is the case, we can infer that the set of prefixed words contains most words which *could* shift stress, an important point for the dataset considered below (Sec. 3.1).

Two types of motivation for the DSS (i.e., why it occurs at all) have been pro-

¹ By “stress,” we always mean primary stress.

² The other changes are too uncommon to study potential conditioning factors: the next most common change, $\{1,1\}\rightarrow\{1,2\}$, only affects 3 words in the sources used here.

posed; as this paper is concerned instead with which words it has affected, we only describe these briefly.³ There is a strong tendency in English for nouns to have earlier stress than verbs (e.g., Ross 1973); one explanation for the DSS is that speakers regularize {1,1} N/V pairs to follow the dominant pattern of stress in the lexicon (Phillips 2006:37–9). A different account is that the DSS is driven by the interaction between stress perception and prosodic context (Kelly 1989; references therein).

2 Frequency and structural conditioning

We now describe two types of possible conditioning factors for how likely an individual N/V pair is to undergo the DSS: word frequency and phonological structure.

2.1 Frequency conditioning

The most detailed proposal for a conditioning effect comes from Phillips (1984), who argues that lower frequency words are more likely to undergo the DSS. Phillips uses two lists of words from Sherman (1975):

1. Changed: Words currently stressed as {1,2}, and known to have undergone the DSS. (44 N/V pairs)
2. Unchanged: Words currently stressed as {2,2}, designated by Sherman as “likely to change” to {1,2} in the future. (124 N/V pairs)

All words begin with one of 9 prefixes: *a, con/m, de, dis, es, ex, pre, re, sur*. Phillips considers the (combined N+V) frequency for each pair, from the *American Heritage Word Frequency Book* (Carroll et al. 1971), and finds that within each prefix class (words sharing the same prefix), the average frequency of changed words is lower than that of unchanged words; the data is shown in Fig. 1.

We note that the observed differences in means are not statistically significant; that is, they could simply be due to chance.⁴ However, the frequency differences are consistently in the predicted direction, suggesting that there may be other factors obscuring an underlying significant effect. One possible issue is the small size of some prefix classes (e.g., 4 words for *pre*), which can make a significant difference in means impossible. For the dataset considered below, we show a statistical technique to deal with small prefix classes, and consider the role of phonological structure as well as frequency; one finding is a significant frequency effect in the direction discussed by Phillips.

³ Of course, the DSS’s motivation and which words it has affected could be related, a possibility not considered here.

⁴ Using Mann-Whitney tests (two-sided) of whether the difference in frequencies of changed vs. unchanged pairs is different from 0 in individual prefix classes ($p \geq 0.19$ for each prefix), or across all words ($p = 0.5$). Because the Mann-Whitney test is sensitive only to the relative order of values, the p values are the same if frequencies are first log-transformed.

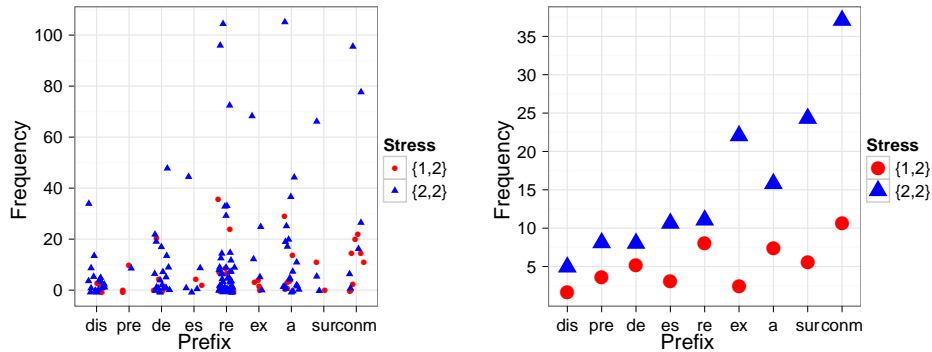


Figure 1: Frequency data from Phillips (1984). Left: frequencies for individual words. Right: means across all words of each type ($\{1,2\}$, $\{2,2\}$), within each prefix class.

2.2 Structural conditioning

We now discuss two types of possible structural conditioning factors for the DSS: syllable weight and final segment identity. σ_1 and σ_2 are used to denote the first and second syllables of an N/V pair.

Besides the general cross-linguistic correspondence between heavier syllables and stress, there is theoretical and experimental evidence for the importance of syllable weight to primary stress assignment in English. To a first approximation, English syllables are heavy if they contain a long vowel (diphthong or tense monophthong) or a coda, and light otherwise; only heavy syllables can bear stress (e.g., Giegerich 1992). For disyllabic nouns (the class of words which change in the DSS), primary stress is predicted to be final if σ_2 contains a long vowel, and initial otherwise (e.g., Chomsky and Halle 1968; Hayes 1982). In disyllabic nonsense words, Guion et al. (2003) find that syllables with long vowels are more likely to be assigned stress, but the presence of a σ_2 coda does not consistently attract stress; however, Baker and Smith (1976:21) show that a σ_1 coda increases the likelihood that stress is assigned to σ_1 (also in nonsense words). A more direct motivation for considering the role of syllable weight in the DSS comes from an observation by Poldauf (1984), that present-day N/V pairs pronounced $\{1,2\}$ often have a σ_1 coda. In general, we might hypothesize that *both heavier σ_1 and lighter σ_2 make change to $\{1,2\}$ more likely.*

Although both codas and vowel length contribute to syllable weight, below we will consider only codas as predictors of whether a word has undergone the DSS.⁵

⁵ Because σ_1 vowel length can change if stress shifts, it is unclear how it should be coded. We tried using σ_2 vowel length as a predictor, based on the verb pronunciation. There was no significant difference in this predictor for changed vs. unchanged words, and it did not prove useful in the

A different type of structural conditioning is suggested by an observation by Minkova (1997:161): many N/V pairs which at some point have followed the {1,2} pattern end in /t/ or /d/. Minkova specifically means Romance loans, which includes many N/V pairs affected by the DSS; the observation fits into a larger account of the role of Romance loans in changing the Middle English stress system.⁶ Based on Minkova’s finding, we predict that *words ending in final /t/ or /d/ will be more likely to shift stress to {1,2}*.

We have described two types of factors which may influence whether a word has undergone the DSS: word frequency and phonological structure. We now turn to the dataset used to test whether these factors do have any effect.

3 Data

As data, we wish to consider a set \mathcal{W} of words stressed {2,2} at some time t_1 , and still in use at a later time t_2 , when each word is stressed {1,2} or {2,2}. \mathcal{W} should be a representative sample of all words which could undergo the DSS between t_1 and t_2 .

3.1 Dataset

In our dataset, $t_1=1700$ and t_2 is the present day. These dates were chosen to maximize the timespan considered, and hence the number of words which have undergone the DSS, given that (Boyer 1700) is the first dictionary listing stress for a significant fraction of all N/V pairs (according to the data given by Sherman 1975). To find all words which were pronounced {2,2} in 1700, we would have to manually check an entire dictionary. To avoid doing so, we make use of the fact that the DSS seems to be largely restricted to prefixed words (Sec. 1.2); the set of prefixed words thus contains most words which *could* have changed, making it a good choice for \mathcal{W} .

We considered words beginning with one of 17 prefixes.⁷ Following previous work (Phillips 1984, 2006), “prefix” is used loosely here, to mean something between initial syllable and morphological prefix. Some prefixes are theoretically questionable; for example, *re* has at least two different senses. The important point is that most words which have undergone the DSS fall into groups of words with

regression models discussed below.

⁶ Briefly, an N/V pair ending in /t/ or /d/ will be trisyllabic $\sigma\sigma\sigma$ in the past tense form; this influences the speaker to assume the N/V pair belongs to a lexical stratum with a highly-ranked constraint (in the Optimality Theory sense) against final stress, which in turn makes the optimal stress of the N form $\sigma\sigma$.

⁷ The prefixes, and number of words for each, are: *re* (42), *a* (20), *dis* (17), *de* (16), *com/con* (9), *e* (8), *ex* (5), *sur* (4), *in/im* (3), *mis* (2), *sup* (2), *pro* (2), *per* (1), *pre* (1). *e* refers to words where the initial vowel is orthographically followed by a consonant other than ‘x’ (e.g., *escape*, *elect*).

similar initial segments, and the behavior of words in the same group may not be independent.

We found all N/V pairs (a) beginning with one of the 17 prefixes, (b) listed as {2,2} in Boyer (1700), a French–English dictionary, and (c) pronounced as {1,2} or {2,2} today, using CELEX (Baayen et al. 1996) and the OED as sources. There are 132 such words in this set, denoted L , of which 23 have changed: the 1700 pronunciation is {2,2} and the present-day pronunciation is {1,2}.⁸

3.2 Inputs

We are interested in the effect of 4 *input variables* (or *inputs*) on whether a word changed, corresponding to the conditioning factors considered above:

- **LogFreq**: The log of the total (N+V) lemma frequencies, from the British National Corpus (Leech et al. 2001).⁹
- **Coda1, Coda2**: The number of coda consonants of σ_1 and σ_2 , according to the CELEX syllabification.¹⁰ (Range: 0–1 for σ_1 , 0–3 for σ_2 .)
- **Coda2.td**: 1 if Coda2 is /t/ or /d/, 0 otherwise.

As exploratory data analysis, we first consider the effect of each input in isolation.

Frequency Considering the frequency of all words in L (across prefix classes), the result is the same as for Phillips’ data: the mean of **LogFreq** is lower for changed words than for unchanged words, but the difference is not significant (Mann-Whitney, $p = 0.11$). We would also like to test whether there is a frequency effect within prefix classes. However, many prefix classes have few (<5) words (Footnote 7), and the notion of a difference between changed and unchanged words does not make sense for such small sets. We thus cannot test within individual prefix classes without throwing out much of the data (from small prefix classes).

A solution is offered by mixed-effects logistic regression, via the concept of “partial pooling,” both described below (Sec. 4.3). For the moment, we note that in a simple mixed-effects model of the effect of **LogFreq** on whether a word changed,

⁸ For the N and V forms of each pair, if CELEX listed a pronunciation it was used; otherwise the OED pronunciation was used. To maximize the number of “changed” words in L , 2 words (*perfume*, *surname*) with variation in the V form (between $\acute{\sigma}\sigma$ and $\sigma\acute{\sigma}$) were counted as {1,2}. Space constraints prevent us from describing L in detail; the dataset can be found on the author’s web page (currently, people.cs.uchicago.edu/~morgan).

⁹ Six words had frequency 0, either because the N and V lemmas were not present in the BNC or because frequencies are in integer counts per million, rounded down; many low-frequency words are thus rounded to 0. To avoid taking log of 0, 0.25 was added to all N+V frequencies.

¹⁰ Syllabic liquids were not counted as codas.

Testing for frequency and structure

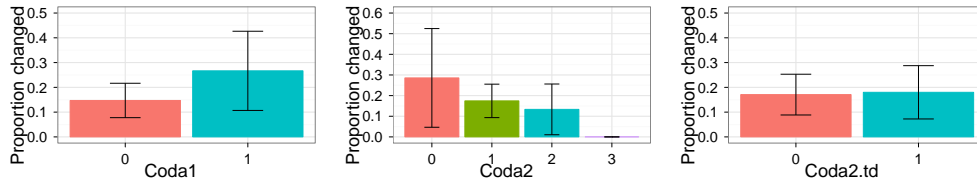


Figure 2: Empirical proportion of words which have changed, as a function of Coda1, Coda2, and Coda2.td.

controlling for prefix class, there is a significant effect of frequency in the predicted direction ($\beta = -0.28$, $z = -2.0$, $p < 0.05$): that is, lower frequency promotes change.¹¹

Structure Fig. 2 shows the proportion of words which have changed, plotted as a function of the 3 structural variables. In each case, the empirical means trend in the predicted direction—words with higher Coda1, lower Coda2, or higher Coda.td are more likely to change—but none of these trends are significant ($p > 0.2$).¹² For convenience, we refer to these three directions of the structural inputs as *hypothesized change-promoting structure*, with the caveat that for Coda2 it is actually the absence of structure that predicts change.

4 Regression models

Considered in isolation, the effects of individual input variables on whether a word changes are largely insignificant. However, they all trend in the expected directions, suggesting that more robust effects may be seen if all inputs are included in a single model. Including frequency and structural variables in the same model also allows us to check for interactions between the two.

4.1 Logistic regression

We model the effect of frequency and structure on whether a word changes using logistic regression (LR); which we briefly review to establish terminology. LR predicts the probability of a binary outcome (the *response*) given values of the *predictor variables*.¹³ Here, the response is changed: 0 for unchanged words, and 1

¹¹The model includes a fixed effect of LogFreq, with a random effect of Prefix on the intercept.

¹²Coda1: $\chi^2(1) = 1.55$, $p = 0.21$. Coda2: Spearman rank correlation $\rho = -0.11$, $p = 0.21$. Coda.td: $\chi^2(1) = 0.01$, $p = 0.92$.

¹³We follow Gelman and Hill (2007:37) in referring to the variables observed for each point—here, frequency and structure variables—as “inputs,” and the variables used in modeling as “predictors.” The distinction is that both interactions and main effects of the inputs are predictors.

for changed words. The data are (\vec{x}_i, y_i) pairs, where each \vec{x}_i is a vector of predictors and $y_i \in \{0, 1\}$ is the response, for an individual data point. The log-odds of $y = 1$ are modeled as a linear function of \vec{x} :

$$\log \frac{P(y = 1|\vec{x})}{P(y = 0|\vec{x})} = a + \vec{x} \cdot \vec{\beta} \quad (1)$$

where a is the intercept and $\vec{\beta}$ is a vector of coefficients.

We are interested in modeling the effects of 4 input variables on `changed`, as well as interactions between frequency and structural inputs, and controlling for prefix class, resulting in at least 8 degrees of freedom. Because we only have 132 data points, we cannot build a single model which considers all these factors at once without serious danger of overfitting.¹⁴ We instead consider two models, each using a different strategy to reduce the number of degrees of freedom; together, they give a good picture of what is conditioning the DSS.

4.2 Model 1: Across prefix classes

We first model the effect of all 4 inputs on `changed`, including possible interactions between frequency and structure, but disregarding prefix class.

The model was fitted using R, with all inputs centered to reduce collinearity.¹⁵ We began with a model including all interactions between `LogFreq` and the structural variables (`Coda1`, `Coda2`, `Coda2.td`). The interaction term between `Coda2` and `LogFreq` was removed because its contribution to model likelihood was not significant ($\chi^2(1) = 0.94$, $p = 0.33$).

The strategy used to address overfitting in Model 1 was *penalization* of model coefficients, following Baayen (2008:224–7). Penalization shrinks coefficients towards zero, and decreases coefficient significances. The resulting penalized model is summarized in Table 1.¹⁶

The main effects for structural inputs are all in the predicted directions, and either significant ($p < 0.05$: `Coda1`, `Coda2`) or approaching significance ($p < 0.1$: `Coda2.td`): on average, words with each type of hypothesized change-promoting structure were indeed more likely to change. The main effect of frequency is not significant: on average, there is no frequency effect when structure is controlled for.

¹⁴As a rule of thumb, overfitting is a concern for binomial data when the number of points with the minority outcome is smaller than 10–15 times the number of predictors (Harrell 2001:61); this is clearly the case for our data.

¹⁵`LogFreq` and `Coda2` were centered; `Coda1` and `Coda2.td` were sum coded. Collinearity is of particular concern because the structural inputs (e.g., `Coda2` and `Coda2.td`) are not independent. After centering, there was minimal collinearity (condition index=1.89).

¹⁶The penalty was 1.25. The penalized model validates fairly well, giving confidence that we have not overfit the data. Repeatedly taking bootstrap samples and refitting the model, all coefficients (except the main effect of `LogFreq`) are different from 0 with significances $p < 0.06$ (see Baayen 2008: 307–8).

Testing for frequency and structure

Predictor	Coef.	S.E.	z	p
Intercept	-0.73	0.55	-1.3	< 0.001
LogFreq	-0.16	0.13	-1.2	0.22
Coda1	1.06	0.53	2.0	0.047
Coda2	-0.98	0.48	-2.0	0.041
Coda2.td	0.96	0.55	1.8	0.078
LogFreq:Coda1	1.01	0.35	2.9	0.0041
LogFreq:Coda2.td	0.56	0.29	1.9	0.056

Table 1: Coefficients, standard errors, Wald z -scores, and significances for Model 1. Nagelkerke pseudo- $R^2 = 0.22$.

However, there is also an interesting interaction between frequency and structure. The presence of either a σ_l coda or a final t/d significantly changes the effect of frequency. When neither is present, the slope of LogFreq is negative (increasing frequency makes change less likely); when either is present, the slope of LogFreq increases, possibly becoming positive. We discuss this effect further below (Sec. 4.4).

4.3 Model 2: Controlling for prefix class

No main effect of frequency was found in Model 1. However, we found in exploratory data analysis (Sec. 3.2) that a significant frequency effect only emerged when prefix class was controlled for. In Model 2, we again consider the effects of both frequency and structure, but controlling for prefix class.

Mixed models and partial pooling The simplest way to control for prefix class would be to add a `Prefix` predictor to Model 1, with one level per prefix. However, several prefix classes have few (< 5) observations, and it is problematic to have levels of a categorical predictor with so few observations.

An elegant solution to this sort of situation, where the data come in groups of very different sizes, is offered by mixed-effects logistic regression, via what Gelman and Hill (2007:252–9) call “partial pooling.”¹⁷ In Model 2, in addition to *fixed effects* of frequency and structure, we will include a *random effect* of `Prefix` on the intercept. This means we assume that the intercept (a in Eqn. 1)—the intrinsic likelihood of change, in log-odds—differs by prefix class, and the intercepts for different prefix classes are normally distributed. Roughly speaking, the intercepts for small classes are inferred based on the proportion of changed words in large classes.¹⁸

¹⁷Baayen (2008) discusses mixed-effects logistic regression for linguistic data.

¹⁸The intercepts are fitted by starting with the empirical proportions (% changed words) for each

Morgan Sonderegger

Predictor	Coef.	S.E.	<i>z</i>	<i>p</i>
Intercept	-1.83	0.65	-2.83	0.005
LogFreq	-0.35	0.16	2.53	0.012
Structure	1.77	0.70	-2.11	0.035
LogFreq:Structure	2.4	0.74	3.18	0.0015

Table 2: Fixed effect coefficients, standard errors, Wald *z*-scores, and significances for Model 2. Nagelkerke $R^2 = 0.26$, calculated relative to a model with only an intercept term (and no random effects).

Model 2 Ideally, we would fit a mixed-effects logistic regression using the terms of Model 1 as fixed effects, and simply adding a random effect of `Prefix`; however, this model has too many degrees of freedom relative to the number of data points.¹⁹

We address overfitting in Model 2 by “data reduction” (Harrell 2001), and reduce the degrees of freedom by defining a single composite predictor for structure:

$$\text{Structure} = \text{Coda1} - \frac{1}{3}\text{Coda2} + \text{Coda2.td}$$

`Structure` increases for any change in a structural input in the direction predicted to promote change to $\{1,2\}$. (`Coda2` is divided by 3 because it takes on values between 0 and 3 in the data, while `Coda1` and `Coda2.td` are 0 or 1.)

The model was fitted in R using `lmer`, from the `lme4` package for mixed-effects models (Bates and Maechler 2009), with a logistic link. Fixed effects were included for `LogFreq`, `Structure`, and their interaction, and a random effect of `Prefix` on the intercept. `LogFreq` and `Structure` were centered, and there was little collinearity after centering (condition index = 1.31).

The estimated variance for the random effect term is $s^2 = 3.0$. This term’s contribution to model likelihood was nearly significant ($\chi^2(1) = 3.8$, $p = 0.052$), giving some support to the assumption that different prefix classes have different intrinsic probabilities of change.

The estimated fixed effects are summarized in Table 2. The effect of structure and its interaction with frequency are qualitatively the same as in Model 1: the probability of change increases with the amount of hypothesized change-promoting structure, as does the slope of `LogFreq`. An important difference from Model 1 is the significant main effect of frequency: on average, change becomes less likely as frequency increases. This comes from controlling for prefix class, not from using a composite predictor for structure: the main effect of frequency is marginal ($p =$

prefix class, and adjusting them to follow a normal distribution; the key fact is that the smaller the prefix class, the more its intercept can be adjusted.

¹⁹Bearing out this intuition, the resulting model does not validate well: some fitted coefficients are not significantly different from 0 using bootstrap samples. (See Footnote 16.)

Testing for frequency and structure

	<i>Frequency quartile</i>					<i>Structure quartile</i>			
	Q1	Q2	Q3	Q4		Q1	Q2	Q3	Q4
$\rho_{structure}$	-0.21	0.35	0.39	0.21	ρ_{freq}	-0.40	-0.31	0.051	0.72
p	0.22	0.044	0.026	0.25	p	0.004	0.07	0.77	0.008

Table 3: Left: Spearman rank correlations between Structure and changed for words in each frequency quartile. Right: Correlations between LogFreq and changed for words in each Structure quartile. Bold values are significant ($p < 0.05$).

0.10) if Model 2 is fitted without the random effect term.²⁰

4.4 Regression models: interpretation

Models 1 and 2 give significant insight into the roles of frequency and structure in conditioning the DSS.

Structure In both models, there were significant main effects for structural inputs in the hypothesized directions. To our knowledge, this is the first demonstration that stress shift in N/V pairs is conditioned by phonological structure.

Model 1 is useful for understanding the relative role of different structural variables. More σ_1 coda consonants, fewer σ_2 coda consonants, and the presence of final t/d all increased the likelihood of change. It would be nice to know the relative importance of the three structural inputs to how likely a word is to change; this comparison is complicated by the presence of interaction terms. One approach is to compare the *average predictive difference per unit* (APDU) of each input: the average difference in probability predicted by the model, across all data points, when a given input is varied by one unit, with other inputs held at their observed values.²¹ The APDU of Coda1 is 0.24, compared to 0.17 for Coda2.td and -0.13 for Coda2. By this measure, whether a word has a σ_1 coda is somewhat more important in determining its probability of changing than the length of its σ_2 coda, or whether it ends in /t/ or /d/.

Model 2 is useful for a broader understanding of the role of structure. Because we centered LogFreq, the main effect of Structure is to be interpreted with frequency held at its mean value: on average, more hypothesized change-promoting structure makes change more likely. But because of the large interaction with frequency, the effect of structure is predicted to diminish (slope nearer

²⁰Model 2 validates well: all coefficients are different from 0 in bootstrapped samples with $p < 0.05$. (See Footnote 16.)

²¹See Gelman and Hill (2007:§5.7, §21.4). For Coda1 and Coda2.td we used Eqn. 5.9; for Coda2 we divided Eqn. 5.9 by 3 (the range of Coda2).

to 0) as frequency is decreased. This prediction is reflected somewhat in the data (Table 3), in that the correlation between *Structure* and *changed* is lower for words in the lowest frequency quartile ($\rho = -0.21$) than for words in other quartiles ($0.21 \leq \rho \leq 0.39$). However, ρ for the highest and lowest quartiles is not significantly different from zero, making any conclusion about the effect of frequency on the slope for *Structure* tentative. The interaction is more important for the effect of *Structure* on the slope for frequency, discussed shortly.

Frequency In both models, the main effect of frequency was in the direction predicted by Phillips (lower frequency makes change more likely), but it was only significant when prefix class was controlled for (Model 2). This agrees with Phillips' view (1984, 2006) that it is within prefix classes that lower frequency promotes change, rather than across all words.

However, the role of frequency is affected by structure. In Model 2, the negative main effect of frequency means that on average (*Structure* held at its mean), the slope of frequency is negative. But due to the interaction with *Structure*, this slope is predicted to diminish (become null), then reverse (become positive) as *Structure* is increased. This pattern is seen in the empirical data (Table 3): the correlation between frequency and *changed* is significantly negative for words in *Structure* quartile 1 (Q1), significantly positive for words in *Structure* Q4, and increases steadily from Q1 to Q4. While lower frequency does promote change on average, the size of this effect diminishes as the amount of change-promoting structure increases, and reverses given enough change-promoting structure.

5 Discussion

The observed interaction between frequency and structure is intriguing in light of a central hypothesis in the literature on frequency and lexical diffusion, that changes can be classified by whether it lower or higher frequency that promotes change, and that which frequency effect is seen reflects the underlying mechanism of change: low frequency-first changes involve "lexical analysis" and tend to be analogical changes, while high frequency-first changes involve only the surface phonetic realization of words (less lexical analysis) and tend to be "physiologically motivated" (Phillips 1998, 2006:56). This classification does not, however, make predictions for changes with multiple motivations, or where the effect of frequency is modulated by other conditioning factors.

The DSS shows *both* types of frequency effect, perhaps reflecting two underlying sources of change. We do not have a compelling explanation for the observed interaction between frequency and structure, in particular how structure influences the effect of frequency. A heuristic account is that different mechanisms are at play for words with different amounts of change-promoting structure. Words with less

Testing for frequency and structure

structure are being treated “analogically”: lower frequency words are more likely to have their pronunciation modeled on similar words (i.e., disyllabic nouns). Words with more structure are being treated “phonetically”: for high-frequency words, there is a higher probability of stress placement being influenced by syllable weight and final segment identity, in the same vein as the correlation between word frequency and the degree of vowel reduction (see references in Bell et al. 2009). Why structure would condition frequency effects in this way is not clear.

Aspects of the frequency/structure interaction are compatible with recent studies of change in progress. Both types of frequency effect within a single change have also been observed by Clark and Trousdale (2009), in the context of consonant fronting in a variety of Scottish English. An interaction between frequency and other factors (though not a reversal in the direction of the frequency effect) was found by De Schryver et al. (2008) for Dutch fricative devoicing. Along with the DSS, these cases suggest that the two-way typology of frequency effects requires refinement to account for the more general case of frequency effects in changes with several motivations.

6 Conclusion

This paper has examined the contributions of frequency and phonological structure to conditioning the DSS, using multiple logistic regression models. Both frequency and structural factors are important for predicting whether an N/V pair has undergone the DSS, and there is a substantial interaction between the effects of frequency and structure. In Models 1 and 2, we found a significant effect of structure on the likelihood that an N/V pair has undergone the DSS. In Model 2, we controlled for the effect of prefix class, and observed a significant frequency effect. There were no significant effects of structure or frequency in simple observed proportions (i.e., the mean frequency of changed vs. unchanged words), showing that understanding the interaction between frequency and structure is important for understanding the relative contribution of each, and illustrating the utility of regression models incorporating both.

References

- Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald, Richard Piepenbrock, and Léon Gulikers. 1996. *CELEX2 (CD-ROM)*. Philadelphia: Linguistic Data Consortium.
- Baker, Robert and Philip Smith. 1976. A Psycholinguistic Study of English Stress Assignment Rules. *Language and Speech* 19(1):9–27.

Morgan Sonderegger

- Bates, Douglas and Martin Maechler. 2009. *lme4: Linear Mixed-Effects Models Using S4 Classes*. R package version 0.999375-32.
- Bell, Alan, Jason Brenier, Michelle Gregory, Cynthia Girand, and Dan Jurafsky. 2009. Predictability Effects on Durations of Content and Function Words in Conversational English. *Journal of Memory and Language* 60(1):92–111.
- Boyer, Abel. 1700. *The Royal Dictionary: In Two Parts...* London.
- Carroll, John, Peter Davies, and Barry Richman. 1971. *The American Heritage Word Frequency Book*. Boston: Houghton Mifflin.
- Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Clark, Lynn and Graeme Trousdale. 2009. Exploring the Role of Token Frequency in Phonological Change: Evidence from TH-Fronting in East-Central Scotland. *English Language and Linguistics* 13(1):33–55.
- De Schryver, Johan, Anneke Neijt, Pol Ghesquière, and Mirjam Ernestus. 2008. Analogy, Frequency, and Sound Change. The Case of Dutch Devoicing. *Journal of Germanic Linguistics* 20(2):159–195.
- Fudge, Erik. 1984. *English Word-Stress*. London: Allen & Unwin.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Giegerich, Heinz. 1992. *English Phonology*. Cambridge: Cambridge University Press.
- Guion, Susan, J.J. Clark, Tetsuo Harada, and Ratreë Wayland. 2003. Factors Affecting Stress Placement for English Nonwords Include Syllabic Structure, Lexical Class, and Stress Patterns of Phonologically Similar Words. *Language and Speech* 46(4):403–427.
- Harrell, Frank, Jr. 2001. *Regression Modeling Strategies*. New York: Springer.
- Hayes, Bruce. 1982. Extrametricality and English Stress. *Linguistic Inquiry* 13(2):227–276.
- Janda, Richard and Brian Joseph. 2003. On Language, Change, and Language Change – Or, of History, Linguistics, and Historical Linguistics. In B. Joseph and R. Janda, eds., *Handbook of Historical Linguistics*, 3–180, Oxford: Blackwell.

Testing for frequency and structure

- Kelly, Michael. 1989. Rhythm and Language Change in English. *Journal of Memory & Language* 28:690–710.
- Kiparsky, Paul. 1995. The Phonological Basis of Sound Change. In J. Goldsmith, ed., *The Handbook of Phonological Theory*, 640–670, Cambridge, MA: Blackwell.
- Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal Factors*. Oxford: Blackwell.
- Leech, Geoffrey, Paul Rayson, and Andrew Wilson. 2001. *Word Frequencies in Written and Spoken English: Based on the British National Corpus*. Harlow: Longman.
- Minkova, Donka. 1997. Constraint Ranking in Middle English Stress-Shifting. *English Language and Linguistics* 1(1):135–175.
- Phillips, Betty. 1984. Word Frequency and the Actuation of Sound Change. *Language* 60(2):320–342.
- Phillips, Betty. 1998. Word Frequency and Lexical Diffusion in English Stress Shifts. In R. Hogg and L. van Bergen, eds., *Historical Linguistics 1995. Vol. 2: Germanic Linguistics*, 223–232, Amsterdam: Benjamins.
- Phillips, Betty. 2006. *Word Frequency and Lexical Diffusion*. New York: Palgrave.
- Poldauf, Ivan. 1984. *English Word Stress*. Oxford: Pergamon.
- Ross, John R. 1973. Leftward, Ho! In S.R. Anderson and P. Kiparsky, eds., *Festschrift for Morris Halle*, 166–173, New York: Holt, Rinehart and Winston.
- Sherman, Donald. 1975. Noun-Verb Stress Alternation: An Example of the Lexical Diffusion of Sound Change in English. *Linguistics* 159:43–71.
- Sonderegger, Morgan and Partha Niyogi. 2010. Variation and Change in English Noun/Verb Pair Stress: Data, Dynamical Systems Models, and Their Interaction. Ms., to appear in A.C.L. Yu, ed., *Origins of Sound Patterns: Approaches to Phonologization*, Oxford: Oxford University Press.

Morgan Sonderegger
University of Chicago
Department of Computer Science
1100 East 58th Street
Chicago, IL 60637

morgan@cs.uchicago.edu