2017. *Proc Ling Soc Amer* 2, 24:1-15. https://doi.org/10.3765/plsa.v2i0.4071.

# Word-final velar place assimilation in English

Lisa Lipani[*]

**Abstract**. In English, word-final alveolar consonants assimilate in place. Additionally, there is recent evidence that assimilation can occur in word-final nasals at all places of articulation (Coleman et al. 2016). Some anecdotal evidence exists that word-final velars can assimilate (Barry 1985), but this has not been substantiated. This study uses the Santa Barbara Corpus of American English (DuBois et al. 2000–2005) to examine word-final velar consonant variation, which was measured by the F2 transitions in the preceding vowel. Given the present data, word-final velars do not seem to undergo categorical assimilation or gradient coarticulation processes.

**Keywords**. speech production; assimilation; coarticulation; phonological models; English

**1. Introduction**. The variability of word-final consonants has been investigated among coronals by a number of researchers (Gow and Hussami 1999; Gow 2002; Gow and McMurray 2007; Dilley and Pitt 2007, among others), and word-final alveolar consonants /t, d, n, s, z/ are said to assimilate in place to the following consonant, as shown in example (1):

(1) *line perfectly* /laɪ**n** pɛɹfɛktli/ → [laɪ**m** pɛɹfɛktli]
*mat belonged* /mæ**t** bilɔŋd/ → [mæ**p** bilɔŋd] (Gow 2002)

In addition to word-final coronal variability, word-final nasal variability has been recently documented by Coleman et al. (2016), who provide evidence from formant values that place assimilation is not limited to coronals; place assimilation can occur in labial and velar nasals, i.e. [m, ŋ], albeit rarely. If there is evidence that word-final nasals and coronal consonants can assimilate in place, the next question is: "Can velar stops undergo the same assimilation process?" Anecdotal evidence of velars assimilating in place to the following consonant exists, shown in example (2) for one speaker of Received Pronunciation (RP) English.

(2) *like that* /laɪ**k** ðæt/ → [laɪ**t** ðæt] (Barry 1985)

Barry (1985) comments that this is likely relatively uncommon, but phonological theory should be able to account for this process. Additionally, researchers acknowledge that velars are susceptible to phonetic change. For example, /k/ is palatalized before and after a high vowel, such as /i/, yielding [kʲi] or [ikʲ]. These are "phonetic variations within the place category DORSAL" (Coleman et al. 2016: 454), and "it is not clear what the 'canonical' place of articulation of the /k/ is" (Ohala 1993: 157).

Two processes could be affecting the pronunciation of word-final velars: assimilation, a categorical phonological process, or coarticulation, a gradient phonetic process[1], and these

---

[*] I would like to thank my major professor, Dr. Margaret Renwick, the rest of my committee: Drs. Keith Langston, Bill Hollingsworth, and Chad Howe, and everyone who provided me with feedback on this research. I'd also like to thank my family: Don Lipani, Marcia Lipani, and Sara Lipani, and my fellow graduate students at the University of Georgia. Author: Lisa Lipani, University of Georgia (llipani@uga.edu).

[1] Ohala (1993) uses the terms "assimilation" and "coarticulation" interchangeably, but does not believe the palatalization of /k/ should be considered coarticulation. Hardcastle and Hewlett (2006) argue the opposite. The definition of coarticulation is contested and determining a definition is outside the scope of this paper.

processes would be reflected differently in the data. This study uses definitions of processes similar to what Schuppler et al. (2012) and Coleman et al. (2016) outline: 1. a categorical phonological process in which one pronunciation variant is selected from the lexicon and 2. a gradient phonetic process that occurs due to the movement of articulators. If the phonological process of assimilation were occurring, data centered on two modes (one larger mode representing productions where assimilation does not occur, and a smaller mode where assimilation does occur) is expected. If the phonetic process of coarticulation were occurring, data that has a wider distribution than a corresponding distribution that lacks coarticulation is expected.

Variation of word-final velars has potential implications for phonological theory. Coronal place assimilation has been explained in prior phonological theory by place underspecification (Avery and Rice 1989). However, this account is not satisfactory, as there is evidence that velar and labial nasals can also assimilate. Coleman et al. (2016) suggest a probabilistic underspecification model that captures the likelihood of the value of the second formant, an acoustic cue for place. This model describes the assimilation of any word-final segment more elegantly than current rule-based models, especially because it encompasses the fact that assimilation processes are not equally likely (as coronal place assimilation is more likely than nasal place assimilation, and perhaps velar place assimilation).

The results in this paper suggest that word-final velars are not subject to assimilatory processes or coarticulatory influences from the following word-initial consonant, as the test distributions did not vary substantially from the control distribution. These results, combined with previous research, can give us more information about how to build a model that encompasses all variation (or lack thereof) in word-final consonants.

## 2. Methods and Procedures.

2.1 SANTA BARBARA CORPUS OF SPOKEN AMERICAN ENGLISH. The Santa Barbara Corpus of Spoken American English (DuBois et al. 2000–2005) was used because it contains audio data and orthographic transcriptions. Additionally, this corpus was analyzed because it contains spontaneous speech, where variation is more likely to occur.

The Santa Barbara Corpus is a part of the International Corpus of English and contains spontaneous speech collected from 216 speakers from 2000 to 2005. Age, sex, hometown, home state, current state of residence, education, years of education, occupation, and ethnicity of the speakers were recorded, if possible. The corpus, which contains approximately 249,000 words, consists of primarily face-to-face conversation, and each audio file is accompanied by an orthographic transcription with a brief description of the context of the conversation. Metadata for the sex of four speakers are absent, and in this study, decisions about sex were made by listening to the audio.

2.2 METHODS. Pairs of words in three conditions were investigated and are shown in Table 1. The last two conditions, velar # labial and velar # coronal, are potentially susceptible to assimilation or coarticulation from the word-initial consonants. The first condition, velar # velar, functions as a control.

To find the place of articulation of the word-final consonant of the word pair's first word (i.e. the place of articulation of /k/ of "back" in the word pair "back to"), a major cue to place of articulation was measured: the second formant transition from the steady state of the vowel to just before the consonant closure (Stevens 1998). This measurement of change in Hertz is called F2 difference (cf. Dilley and Pitt 2007). In this study, F2 difference is defined as F2 at the midpoint minus F2 at 90% of the vowel, which means there is an inverse relationship between F2 movement and F2 difference. Since F2 is affected by the vowel quality, separate analyses are conducted by vowel type: high front /i, ɪ/, low front /æ, ɛ/, and back /u, ʊ, ɔ, ɑ/. If the place of articulation of the word-final consonant had changed (in the test conditions velar # labial or velar # coronal), there would be a difference from the control (velar # velar). For example, if the previously mentioned example “back to” was pronounced like “ba[t] to”, an F2 difference that varies from the F2 difference of an unassimilated token of “bac[k] to” would be expected; the vowel preceding the velar in the test case, velar # coronal, would be expected to exhibit formant transitions like those before a coronal. These second formant measurements were also examined in the context of the acoustics of the word-final velar, and the presence or absence of the release of the stops was recorded. This was determined by the stop burst for velar oral stops /k, g/ and if a silence was present between the nasal stop /ŋ/ and the following stop.

| Word pair | Word 1 | Word 2 | Example |
|---|---|---|---|
| velar # velar | ends with /k, g, ŋ/ | begins with /k, g/ | *big kid* [bɪ**g** **k**ɪd]<br>*doing good* [duɪ**ŋ** **g**ʊd] |
| velar # labial | ends with /k, g, ŋ/ | begins with /p, b/ | *walk by* [wɑ**k** **b**aɪ]<br>*being put* [biɪ**ŋ** **p**ʊt] |
| velar # coronal | ends with /k, g, ŋ/ | begins with /t, d/ | *back to* [bæ**k** **t**u]<br>*sing to* [sɪ**ŋ** **t**u][2] |

Table 1: Test and control condition examples

Praat (Boersma and Weenink 2015) was used to segment sound files into smaller pieces in order for a forced automatic aligner, SPPAS (Bigi 2012), to be used. SPPAS outputs text grids segmented by phone, which were checked by hand. Monophthongal vowels before consonants /k, g/ were segmented based on a decrease in amplitude in the speech wave, and decreasing amplitude in the signal and lightening of formants in spectrograms were the primary means used to segment between the vowel and the velar nasal /ŋ/. Sounds that were of poor quality, contained overlap, contained too much background noise, or were unintelligible were removed. Praat was used to automatically extract formant values from the midpoint and the 90% point of the vowel. Formant settings used by Praat were based on sex; for females, Praat’s automatic formant tracker found five formants up to 5500 Hz and for males, Praat found five formants up to 5000 Hz. R statistical software (R Core Team 2015) was used to analyze and visualize data.

2.3 HYPOTHESES. This section defines F2 formant transition values typical to the control condition (velar # velar) and those typical to canonical labial or alveolar stops. Delattre et al. found that the locus of the bilabial stop /b/ is “some point low on the frequency scale” (1955: 769). Furthermore, in transitions from vowels to bilabial stops /p, b/, the F2 movement downwards seems to be more important than moving to a specific locus (Olive et al. 1993). Transitions to alveolar consonants depend on vowel quality. Before alveolar stops /t, d/, front vowels with a higher F2 (i.e. front vowels) show F2 movement downwards, and vowels with lower F2 values (i.e. back vowels) show movement upwards. The vowel transitions move towards the locus of the alveolar, which has been defined as about 1800 Hz (Liberman et al. 1954; Delattre et al. 1955). For pre-velar vowels, F2 always increases (Olive et al. 1993). In

---

[2] The liquids /l/ and /ɹ/ are known to have a long-distance effect on F3 (Local and Kelly 1985; West 1999) and can have an effect on F2 (Tunley 1999), so orthographic patterns containing /l/ or /ɹ/ were excluded from the analysis.

Table 2, the expected effect of place of articulation on the preceding velar is outlined and grouped by the following place of articulation in addition to the preceding vowel.

If a categorical process of assimilation were occurring, word-final velars are predicted to feature a non-symmetric bimodal distribution of F2 differences, where there is a larger mode of unassimilated tokens and a smaller mode of assimilated tokens. Having this mode that exhibits deviation from the expected F2 difference of the control (velar # velar) and is similar to the F2 difference for vowel preceding a canonical labial or coronal may indicate assimilation. If a gradient coarticulation process were occurring, differing distributions between the control condition (velar # velar) and the test condition (velar # coronal or velar # labial) is expected; a distribution featuring coarticulation wider than the control condition. When comparing unreleased and released velars, an unreleased velar is expected to be more susceptible to coarticulatory or assimilatory processes.

| Preceding vowel quality[3] | Following place of articulation | Expected effect of place of articulation on preceding velar |
|---|---|---|
| any | bilabial /p, b/ | positive F2 difference |
| | | |
| high front /i, ɪ/ | alveolar /t, d/ | positive F2 difference |
| back /u, ʌ, ɔ, ɑ/ | alveolar /t, d/ | negative F2 difference |
| low front /æ, ɛ/ | alveolar /t, d/ | no F2 difference |

Table 2. Hypothesized F2 differences

To summarize, the variable being investigating is F2 difference ($F2_{50} – F2_{90}$). The study examines the data for three patterns in the distributions: 1. a bimodal distribution of the test conditions (which might indicate the phonological and categorical process of assimilation), 2. a wider distribution in the test conditions than in the control condition (which might indicate the phonetic and gradient process of coarticulation), and 3. a difference in the distributions of unreleased and released velars.

**3. Results**. Overall, the data consist of 269 tokens, and the subsets of the data are shown in Table 3. Columns show the place of articulation of the word-initial segment, and the rows show the word-final segment. These data are further subset by vowel type; the data were subset by front high vowels [ɪ], low front vowels [ɛ, æ], and back vowels [ʌ, ɔ, ɑ]. There were no instances of [u] or [i] in the data.

| | Males | | | Females | | |
|---|---|---|---|---|---|---|
| | Labial | Coronal | Velar | Labial | Coronal | Velar |
| Nasal velar | 14 | 60 | 7 | 7 | 56 | 10 |
| Nonnasal velar | 5 | 48 | 4 | 5 | 52 | 1 |

Table 3. Number of tokens

3.1 SECOND FORMANT MEASUREMENTS: ASSIMILATION. This section presents results examining bimodality, which might indicate assimilation, in distributions of the data subsets.

---

[3] Preceding vowel quality would also have an effect on a velar consonant, as discussed previously, such as in the examples *key* [$k^j$i] and *coo* [ku]. The effect of this fronting, only expected to affect /i, ɪ/, is F2 being slightly lower at the end of the transition to the velar. However, I would still expect this to show movement upwards among these high front vowels (and have a negative F2 difference), as opposed to showing movement downward as hypothesized (and have a positive F2 difference).

3.1.1 NASALS. In the velar nasal data, only the front high vowels have sufficient data to be analyzed. There are only two non-front, non-high vowels for the female group (one velar and one labial), and two non-front, non-high vowels (one velar and one coronal) for males, and further analysis was not done.

3.1.1.1 FRONT HIGH VOWELS: MALES. In the subset of male data of front high vowels before nasal velars, there are 59 (out of a total 60 of instances of nasal velars followed by a coronal — see Table 3) coronals, 14 (out of 14) labials, and 6 (out of 7) velars. In other words, only 2 other tokens in this subset feature any vowel besides [ɪ].

The data were plotted, and outliers above 350 Hz or below −500 Hz were examined. Three outliers occurred due to Praat incorrectly tracking formants, and they were corrected by hand. Figure 1 shows a density plot of F2 difference of high vowels preceding velars separated by place of articulation of the word-initial stop for males. For example, “type = coronal” is an instance of ŋ#t or ŋ#d. If assimilation were not occurring, F2 difference is expected to be negative and to look similar to the dashed line in Figure 1 (the control condition: velar # velar), and if assimilation were occurring, a mode centered around a positive F2 difference (assimilated tokens) in addition to a mode near the mode of the control, i.e. a bimodal distribution that is indicative of a categorical process, would be expected. The mode of all the three places of articulation is roughly equal, and there is not a bimodal distribution; consequentially, there is little evidence of assimilation. While Figure 1 shows how the distributions overlap, and are unlikely to be bimodal, the data are limited and may possibly be over-smoothed.

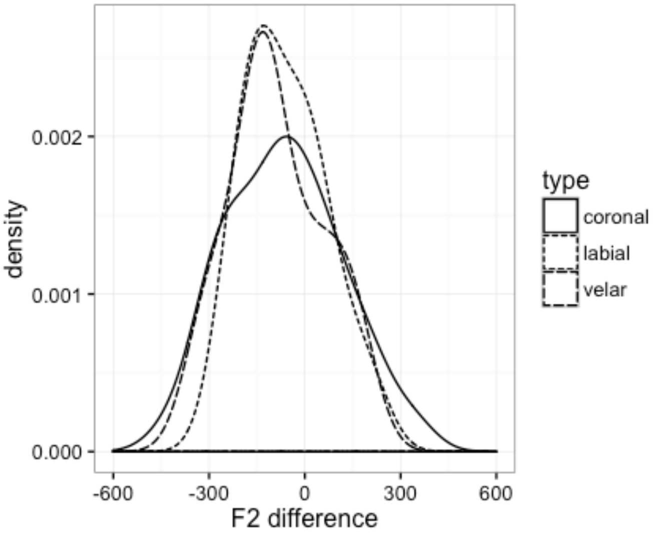

Figure 1. Density curve for F2 difference in front high vowels for males

3.1.1.2 FRONT HIGH VOWELS: FEMALES. In the subset of female velar nasals following a front high vowel for female speakers, there are 56 (out of 56) coronals, 6 (out of 7) labials, and 9 (out of 10) velars.

Six data points under −500 Hz or over 350 Hz were examined and corrected by hand. Figure 2 shows a density plot of F2 difference of high vowels preceding velars separated by place of articulation of the stop after the velar. Similarly to the male data, there is substantial overlap of the distributions. The velar data does have two peaks, but this is likely due to the limited velar data points; a Hartigans’ dip test (Hartigan and Hartigan 1985) does not give a *p*-value low enough to reject the null hypothesis that the data is unimodal (*p* = 0.2498).

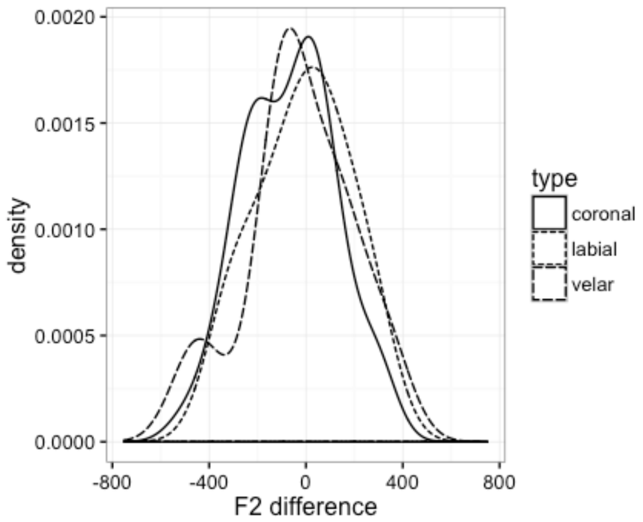

Figure 2. Density curve for F2 difference in front high vowels for females

3.1.2 NONNASALS. In the velar oral stop data, only the front low vowel and male back vowel subsets have sufficient data to be analyzed. In the subset of female back vowels, there are only six tokens. For high vowels, there are only ten data points (three velar, five coronal, and two labial) in the male subset, and six data points (five coronals and one labial) in the female subset. The male subset of data has one data point, which was measured by hand. These sparse data were not plotted.

3.1.2.1 FRONT LOW VOWELS: MALES. In this subset, there are 27 (out of 48) coronals, 2 (out of 5) labials, and 0 (out of 4) velars. The 2 labials are not shown in the density plot in Figure 3.

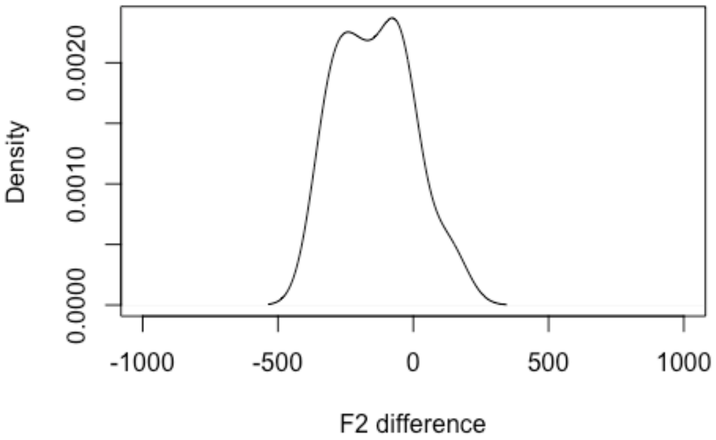

Figure 3. Density curve for F2 difference in front low vowels for males.

The data are centered on one mode; a Hartigans' dip test does not give a *p*-value low enough ($p = 0.7647$) to reject the null hypothesis of unimodality. Most of the data are negative, indicating that assimilation is not occurring.

3.1.2.2 FRONT LOW VOWELS: FEMALES. In this subset, there are 33 (out of 52) coronals, 2 (out of 5) labials and 0 velars (out of 1). Labials and velars are not shown in Figure 4.

The central mode of the data does seem to encompass 0 Hz, which is what would be expected for assimilation, but most of the data fall in the negative F2 difference range, which is what would be expected if assimilation were not occurring. The data are not bimodal and no categorical assimilation process is occurring, which can be supported with a *p*-value ($p = 0.6236$) from a Hartigans' dip test. This density curve is not smooth and has three peaks, but this is due to the limited data being plotted.

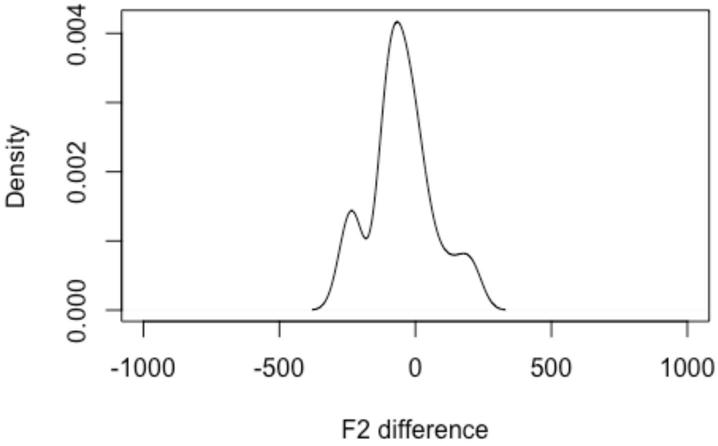

Figure 4. Density curve for F2 difference in front low vowels for females

3.1.2.3 BACK VOWELS: MALES. In this subset, there are 21 data points containing back vowels preceding velars that are followed by coronals. One data point was corrected by hand. As shown in Figure 5, the central mode of the data does seem to encompass 0 Hz, but most of the data fall in the negative F2 difference range. Again, the smaller peaks in this density curve are due to limited data.

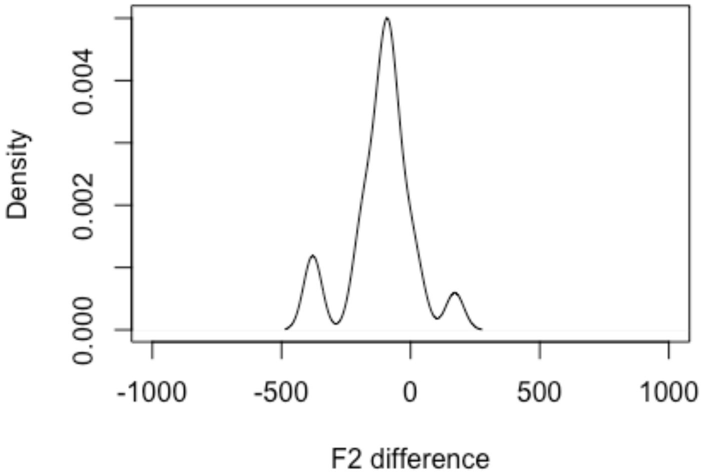

Figure 5. Density curve for F2 difference in back vowels for males

3.1.3 HARTIGANS' DIP TESTS. Table 4 gives *p*-values for Hartigans' dip tests of every subset of the data. A few subsets have too few data points to have an interpretable *p*-value, and those are indicated by a dash. In summary, the visualizations of the density plots combined with Hartigans' dip test results (which show the null hypothesis that the data are unimodal cannot be rejected) suggest that there is not likely to be bimodality in the data for which there are sufficient data points, and therefore, it is unlikely that categorical assimilation is occurring.

| | | Males | | | Females | | |
|---|---|---|---|---|---|---|---|
| | | Bilabial | Coronal | Velar | Bilabial | Coronal | Velar |
| Nasal | High front | 0.6533 | 0.8721 | 0.8906 | 0.964 | 0.7295 | 0.2498 |
| | Low front | – | – | – | – | – | – |
| | Back | – | – | – | – | – | – |
| Nonnasal | High front | – | 0.5995 | – | – | 0.8165 | – |
| | Low front | – | 0.7647 | – | – | 0.6236 | – |
| | Back | – | 0.9681 | – | – | 0.9947 | – |

Table 4. *p*-values for Hartigans' dip test

3.1.4 MODELING ASSIMILATION. These Hartigans' dip tests may not be able to detect bimodality if a mode is sufficiently small, and this may be the case in rare assimilatory processes: Dilley and Pitt note that assimilation is an infrequent phenomenon: "assimilation [of word-final coronals] occurred in only 9% of all possible [4349] assimilable environments" (2007: 2350). To examine the sensitivity of the Hartigans' dip test, an artificial bimodal distribution can be created that is a rough approximation of what might be expected from a bimodal distribution indicating assimilation. For example, the control condition, [ɪg], is expected to exhibit an F2 difference of approximately −300 Hz[4], and [ɪd] (what the velar might assimilate to, if the word-final velar were to precede a word-initial coronal) is expected to exhibit an F2 difference of approximately 400 Hz. A bimodal distribution was created using R, consisting of 100 normally-distributed observations centered around a mean of −300 Hz with a standard deviation of 100 Hz and 9 normally-distributed observations (an educated guess of assimilation rates based on 9% assimilation rate found in Dilley and Pitt (2007) and 8.3% assimilation rate of velar nasals found in Coleman et al. (2016)), centered around a mean of 400 Hz with a standard deviation of 100 Hz. The Hartigans' dip test gives a *p*-value of 0.8739, which would not be enough to reject the null hypothesis that the data are unimodal. However, visualization in the form of a quantile-quantile plot suggests a bimodal distribution, as shown in Figure 6. Including 500 normally-distributed observations centered around a mean of -300 Hz with a standard deviation of 100 Hz and 45 normally-distributed observations centered around a mean of 400 Hz with a standard deviation of 100 Hz does give a significant Hartigans' dip test value ($p = 0.01297$), and the quantile-quantile and density plots show this bimodality as well, as shown in Figure 6.

More data points will strengthen the results from the Hartigans' dip test, but visualization of the data can also help show any non-unimodality. For comparison, Figure 7 shows quantile-quantile plots for a subset of the actual nasal velar data for males and females. These quantile-quantile plots show no bimodality.

---

[4] Olive et al. (1993) report a spectrogram of /ig/ where F2 moves up approximately 400 Hz, and since the F2 of [ɪ] tends to be slightly lower, the value 300 Hz was used. The same assumption was made for /ɪd/ based on /id/.

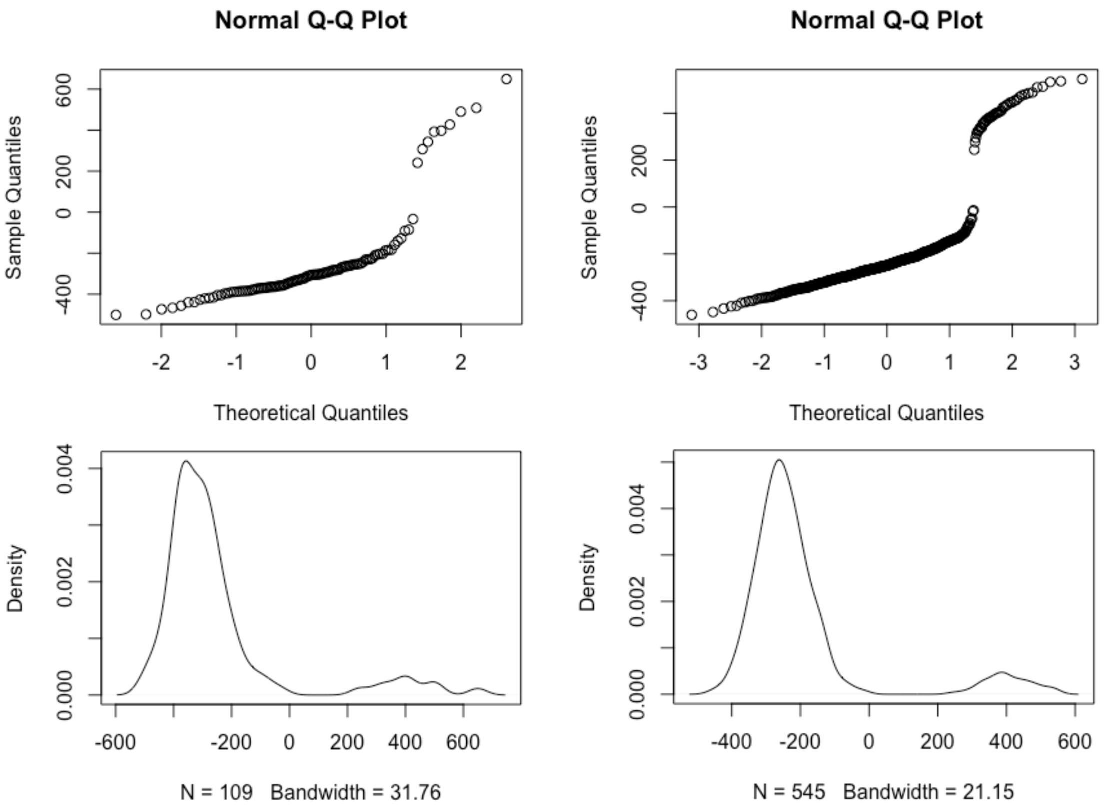

Figure 6. Quantile-quantile (above) and density plots (below) for theoretical distributions, left: 109 observations, right: 545 observations

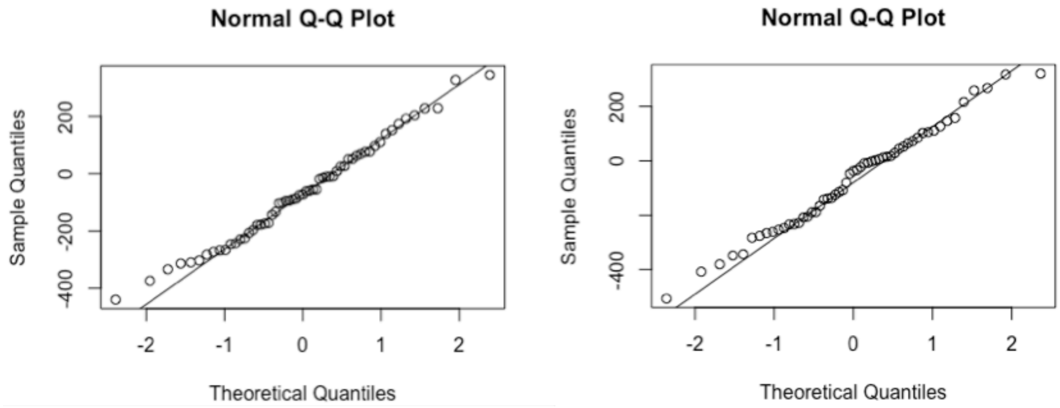

Figure 7. Quantile-quantile plots for high vowels preceding a word-final velar preceding a coronal stop, e.g. going to, left = males, right = females

3.2 SECOND FORMANT MEASUREMENTS: COARTICULATION. If assimilation, a categorical process, is occurring, the visualizations might show a bimodal distribution. However, a more continuous phonetic process that might not feature two modes (i.e., coarticulation) may be occurring. Coarticulation may cause the distribution in the density plot to be wider than the control condition. A two-sample Kolmogorov-Smirnov test was used to compare each of the test condition (labial and coronal) to the control condition (velar), and *p*-values for these are given in Table 5. However, many subsets do not have enough data to be compared; these are marked by a dash in Table 5. As Table 5 shows, no subset containing enough data has *p*-values low enough to reject the null hypothesis that these samples come from the same distribution.

| | | Males | | Females | |
|---|---|---|---|---|---|
| | | Bilabial | Coronal | Bilabial | Coronal |
| Nasal | High front | 0.9983 | 0.771 | 0.8531 | 0.4979 |
| | Low front | − | − | − | − |
| | Back | − | − | − | − |
| Nonnasal | High front | 1 | 0.9643 | − | − |
| | Low front | − | − | − | − |
| | Back | − | − | 0.6667 | 0.25 |

Table 5. *p*-values for Kolmogorov-Smirnov tests

3.2.1 MODELING COARTICULATION. The Kolmogorov-Smirnov tests may not be sufficient to detect a low percentage of coarticulation. To simulate a skewed distribution to test the sensitivity of the Kolmogorov-Smirnov test, a normal distribution was created with 100 observations centered around a mean of −300 Hz with a standard deviation of 100 Hz combined with 9 random numbers from −100 Hz to 0 Hz representing coarticulated tokens (which are moving towards a positive F2 difference). This was compared to a different simulated normal distribution that did not contain the additional observations representing coarticulation (100 observations centered around a mean of −300 Hz with a standard deviation of 100 Hz). The Kolmogorov-Smirnov test gives a *p*-value of 0.6005. Creating a distribution in the same way as above with 400 tokens in the normal distribution tokens plus 36 random numbers and comparing this to a distribution with 400 tokens yields a significant *p*-value ($p = 0.03803$). Thus, it is possible that the Kolmogorov-Smirnov test may be insufficiently sensitive to detect coarticulation in the current sample, and more data are needed. However, as shown in Figure 8, visualization of the data can also help show the skewedness of a distribution. Figure 8 compares a theoretical distribution lacking coarticulation (represented by a solid line) with a theoretical distribution including coarticulation (represented by a dashed line).

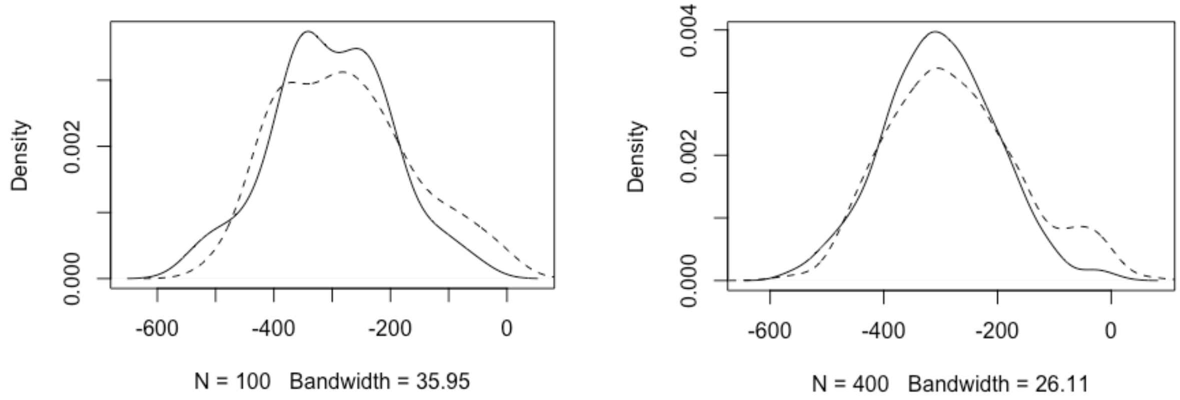

Figure 8. Density plots comparing theoretical distributions

3.3 MIXED-EFFECTS MODELS. Mixed-effects models can also test if place of articulation (coronal, labial, or velar) is a predictor of F2 difference. Two nested mixed models (Baayen et al. 2008) were made using the R package 'lme4' (Bates et al. 2015) and *p*-values[5] were produced using the

---

[5] However, these *p*-values should be interpreted cautiously according to Baayen (2008).

R package ‘lmerTest’ (Kuznetsova et al. 2016). The first is a full mixed model, the results of which are shown in Table 6, with the predictors sex, nasality, place of articulation, vowel, an interaction between nasality and place of articulation, and random effects for both speaker and word[6].

| | Estimate | Std. Error | t value | *p*-value |
|---|---|---|---|---|
| (Intercept) | −88.786 | 44.205 | −2.009 | 0.081 |
| Sex = male | −10.567 | 21.838 | −0.484 | 0.633 |
| Nasal = TRUE | −1.917 | 49.141 | −0.039 | 0.969 |
| Prec. vowel = ɑ | 0.233 | 109.229 | 0.002 | 0.998 |
| Prec. vowel = ɛ | 166.357 | 123.560 | 1.346 | 0.18 |
| Prec. vowel = ɪ | 55.009 | 60.115 | 0.915 | 0.371 |
| Prec. vowel = ɔ | 0.030 | 60.188 | 0 | 1 |
| Prec. vowel = ʌ | 52.499 | 92.427 | 0.568 | 0.571 |
| Following POA = labial | 0.287 | 59.619 | 0.005 | 0.996 |
| Following POA = velar | −6.518 | 83.989 | −0.078 | 0.938 |
| Nasal = TRUE + Following POA = labial | −20.761 | 72.550 | −0.286 | 0.775 |
| Nasal = TRUE + Following POA = velar | 12.112 | 96.033 | 0.126 | 0.9 |

Table 6. Results of a mixed-model with place of articulation as a predictor

The second reduced mixed model (Table 7) contains the same predictors as the first except for place of articulation. These models contain no significant predictors, including the place of articulation of the word-initial consonant. Another way to investigate the significance of a predictor variable is to compare the nested models using an ANOVA, which showed that the models are not statistically significantly different ($p = 0.9901$). The place of articulation predictor variable has no effect on the model.

| | Estimate | Std. Error | t value | *p*-value |
|---|---|---|---|---|
| (Intercept) | −88.000 | 42.739 | −2.059 | 0.073 |
| Sex = male | −12.456 | 21.343 | −0.584 | 0.566 |
| Nasal = TRUE | −3.185 | 42.325 | −0.075 | 0.940 |
| Prec. vowel = ɑ | −2.193 | 108.510 | −0.02 | 0.984 |
| Prec. vowel = ɛ | 166.961 | 123.020 | 1.357 | 0.177 |
| Prec. vowel = ɪ | 53.419 | 57.415 | 0.93 | 0.365 |
| Prec. vowel = ɔ | −0.681 | 59.223 | −0.011 | 0.991 |
| Prec. vowel = u | 109.007 | 94.187 | 1.157 | 0.252 |
| Prec. vowel = ʌ | 51.943 | 86.068 | 0.604 | 0.548 |

Table 7. Results of a mixed-model without place of articulation

---

[6] Initially, this model included an interaction between place of articulation and vowel. However, this model had high correlation between predictors (i.e. the model displayed collinearity, making some of the predictors essentially redundant), so the model was fit without this interaction.

A larger data set would be beneficial in obtaining significant predictors, and ultimately, it is difficult to ascertain the usefulness of these models.

3.4 SECOND FORMANT MEASUREMENTS: ACOUSTICS OF THE VELAR. Finally, Kolmogorov-Smirnov tests were used to compare two distributions: 1. released word-final consonants: stops /k, ɡ/ containing a burst before another stop or a velar nasal /ŋ/ with a silence after it before the following stop and 2. unreleased word-final consonants: unreleased stops /k, ɡ/ before another stop or a velar nasal /ŋ/ that was not released. Sixty-six of 266 tokens either feature a burst (in [k, ɡ]) or a silence after the [ŋ]. Out of 152 nasals, 44 feature a silence, and out of 114 stops, only 10 feature a burst. Due to sparse data, the nonnasal subset of this data was not analyzed. In the nasal subset of this data, only high front vowels have sufficient data to be compared with the Kolmogorov-Smirnov test, and the *p*-values of this test are given in Table 8. In Table 8, no subset has a small enough *p*-value to reject the null hypothesis that these samples come from the same distribution.

| | Males | | Females | |
|---|---|---|---|---|
| | Bilabial | Coronal | Bilabial | Coronal |
| High front nasal | 0.6264 | 0.9003 | 0.9333 | 0.4913 |

Table 8. *p*-values for Kolmogorov-Smirnov tests: unreleased vs. released

**4. Discussion**. In addition to the above analysis, I listened to the data. Impressionistically, no tokens sounded assimilated to my ears. Overall, the results show no strong evidence of assimilation or coarticulation. However, assimilation may be infrequent; for example, Dilley and Pitt found that "assimilation occurred in only 9% of all possible [4349] assimilable environments" (Dilley and Pitt 2007: 2350) in the Buckeye Corpus, and assimilation of velars is expected to be rare (Coleman et al. 2016). Additionally, as discussed in the sections above, statistical tests might not also be sensitive enough to detect a relatively rare process in a limited data set, and due to the small data set, density plots are potentially over-smoothed. Nothing suggesting categorical assimilation, gradient coarticulation, or place of articulation having any effect on F2 difference was found. Since the present results do not show evidence of velar assimilation or coarticulation thus far, I will proceed under the assumption that velars do not undergo these processes.

The fact that velars do not seem to assimilate suggests velars are specified for place with the [DORSAL] feature; however, Coleman et al. (2016)'s evidence from nasals show that these rules that account for place underspecification are unsatisfactory. Allowable variation can be modeled probabilistically following Coleman et al. (2016) based on Blackburn and Young (2000) and Keating's window model (1990), as shown in Figure 9. Probability distributions can be modeled based on density curves like those created previously in Figures 1–5 and incorporated into the probability distributions (shown by arrows) in the model (right). This probabilistic underspecification model can account for assimilation by using bimodal distributions, but it can also account for this data that does not show assimilation with a narrower Gaussian model that does not allow as much variation as Gaussian models for segments that do assimilate.

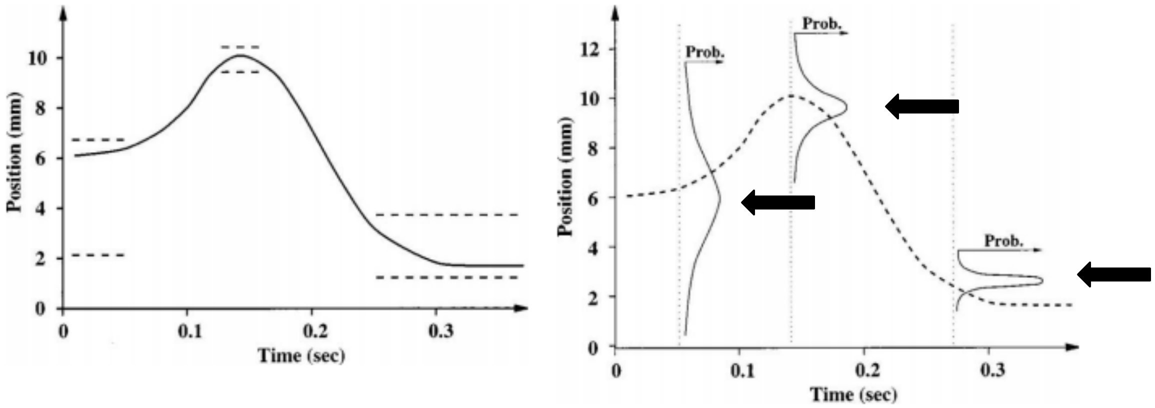

Figure 9. Partially reproduced from Blackburn and Young (2000): Keating's (1990) window model (left) and Blackburn and Young's (2000) adaptation that incorporates probability (right)

5. **Conclusion**. This type of study of phonetic detail (including assimilation and coarticulation) can have an enormous impact on the fields of speech synthesis and automatic speech recognition. A large part of the unnaturalness in speech synthesis technology today can be attributed to a lack of fine phonetic detail, which is argued by Hawkins (2003) to contribute to perceptual coherence (also see Remez et al. 1994), a part of what makes speech understandable and natural. Duffy and Pisoni (1992) also comment that synthetic speech fails to make use of multiple acoustic cues to transmit information, and that this has led to less intelligible, less natural sounding speech. Duffy and Pisoni (1992) also argue that a lack of acoustic detail leads to increased mental processing costs, and this leaves fewer resources for other necessary comprehension processes (see Myers and Duffy 1990). For example, Hidden Markov Models (HMMs) can be used for both speech synthesis and speech recognition, and because HMMs use probability density functions for both these tasks, probabilistic modeling can be used to improve these.

The above study needs more data, but there is no current evidence that velars assimilate in place to a following stop consonant. This study examined possible assimilatory contexts in the Santa Barbara Corpus of Spoken American English, and specifically investigated F2 difference of a vowel before a velar followed by three different places of articulation (coronal, labial, and velar stops), where velars preceding velars served as a control. Through a combination of visualizing the data with density plots and a comparison of mixed models, the evidence is not sufficient to make the claim that velars assimilate.

Future research will use a larger corpus such as the Audio British National Corpus (Coleman et al. 2012) or the Nationwide Speech Project (Clopper and Pisoni 2006), which all have transcriptions and audio data. Additionally, looking at word-internal instances where assimilation could occur will expand data, for example, examining the F2 difference in the [ɛ] in *technically* /tɛ**kn**ɪkli/, which could be realized as [tɛ**tn**ɪkli].

## References

Avery, Peter & Keren Rice. 1989. Segment Structure and Coronal Underspecification. *Phonology* 6 (2). 179–200. https://dx.doi.org/10.1017/S0952675700001007

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language* 59 (4). 390–412. https://dx.doi.org/10.1016/j.jml.2007.12.005

Barry, Martin. 1985. A Palatographic Study of Connected Speech Processes. *Cambridge Papers in Phonetics and Experimental Linguistics* 4. 1–16. http://dx.doi.org/10.1006/jmla.1993.1011

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67 (1). 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bigi, Brigitte. 2012. SPPAS: A tool for the phonetic segmentations of speech. *The Eighth International Conference on Language Resources and Evaluation*. 1748–1755. ISBN 978-2-9517408-7-7

Boersma, Paul & David Weenink. 2015. *Praat: Doing Phonetics by Computer* (version 5.4.22). http://www.fon.hum.uva.nl/praat/

Clopper, Cynthia G & David B. Pisoni. 2006. The Nationwide Speech Project: A New Corpus of American English Dialects. *Speech Communication* 48 (6). 633–644. https://dx.doi.org/10.1016%2Fj.specom.2005.09.010

Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. Audio BNC: The Audio Edition of the Spoken British National Corpus. *Phonetics Laboratory, University of Oxford*. http://www.phon.ox.ac.uk/AudioBNC

Coleman, John S., Margaret E.L. Renwick & Rosalind Temple. 2016. Probabilistic Phonetic Underspecification in Nasal Place Assimilation. *Phonology* 33 (3). 425–458. https://dx.doi.org/10.1017/S0952675716000208

Delattre, Pierre C., Alvin M. Liberman & Franklin S. Cooper. 1955. Acoustic Loci and Transitional Cues for Consonants. *Journal of the Acoustical Society of America* 27 (4). 769–773. http://dx.doi.org/10.1121/1.1908024

Dilley, Laura & Mark Pitt. 2007. A Study of Regressive Place Assimilation in Spontaneous Speech and Its Implications for Spoken Word Recognition. *Journal of the Acoustical Society of America* 122: 2340–2353. http://dx.doi.org/10.1121/1.2772226

DuBois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson & Nii Martey. 2000–2005. *Santa Barbara Corpus of Spoken American English, Parts 1–4*. Philadelphia: Linguistic Data Consortium.

Duffy, Susan A. & David B. Pisoni. 1992. Comprehension of Synthetic Speech Produced by Rule: A Review and Theoretical Interpretation. *Language and Speech* 35 (4.: 351–389.

Gow, David W., Jr. 2002. Does English Coronal Place Assimilation Create Lexical Ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 28 (1). 163–179. http://dx.doi.org/10.1037/0096-1523.28.1.163

Gow, David W., Jr. & Peter Hussami. 1999. Acoustic Modification in English Place Assimilation. *Journal of the Acoustical Society of America* 106 (4). 2243–2243. http://dx.doi.org/10.1121/1.427646

Gow, David W., Jr. & Bob McMurray. 2007. Word Recognition and Phonology: The Case of English Coronal Place Assimilation. In Cole, Jennifer and José Ignacio Hualde (eds.) *Papers in Laboratory Phonology,* vol. 9, 173–200. Berlin: Mouton de Gruyter.

Hardcastle, William J. & Nigel Hewlett. 2006. *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press.

Hartigan, John A. & P. M. Hartigan. 1985. The Dip Test of Unimodality. *The Annals of Statistics* 13 (1). 70–84. http://dx.doi.org/10.1214/aos/1176346577

Hawkins, Sarah. 2003. Roles and Representations of Systematic Fine Phonetic Detail in Speech Understanding. *Journal of Phonetics* 31 (3). 373–405. https://dx.doi.org/10.1016/j.wocn.2003.09.006

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2016. *lmerTest: Tests in Linear Mixed Effects Models.* (version R package version 2.0-32). https://CRAN.R-project.org/package=lmerTest

Liberman, Alvin M., Pierre C. Delattre, Franklin S. Cooper & Louis J. Gerstman. 1954. The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants. *Psychological Monographs: General and Applied* 68 (8). 1-13. http://dx.doi.org/10.1037/h0093673

Local, John & John Kelly. 1985. Notes on Phonetic Detail and Conversational Structure. *Belfast Working Papers in Language & Linguistics* 7. 1–15.

Myers, Jerome L. & Susan A. Duffy. 1990. Causal Inferences and Text Memory. *Psychology of Learning and Motivation* 25. 159–173. https://dx.doi.org/10.1016/s0079-7421(08)60254-6

Ohala, John J. 1993. Coarticulation and Phonology. *Language and Speech* 36 (2–3). 155–170.

Olive, Joseph P., Alice Greenwood & John Coleman. 1993. *Acoustics of American English Speech: A Dynamic Approach*. New York: Springer-Verlag New York.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Remez, Robert E., Philip E. Rubin, Stefanie M. Berns, Jennifer S. Pardo & Jessica M. Lang. 1994. On the Perceptual Organization of Speech. *Psychological Review* 101 (1). 129–156. http://dx.doi.org/10.1037/0033-295X.101.1.129

Stevens, Ken. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT.

Tunley, Alison. 1999. Coarticulatory Influences of Liquids on Vowels in English. Cambridge, UK: University of Cambridge dissertation. http://www.ling.cam.ac.uk/people/sarah/docs/AliTunleyThesis.pdf

West, Paula. 1999. The Extent of Coarticulation of English Liquids: An Acoustic and Articulatory Study. *Proceedings of the International Conference of Phonetic Sciences*. 1901–1904.