EFFECTS OF CODA VOICING ON VOWEL PHONATION

Chelsea Sanker

Brown University
chelsea_sanker@brown.edu

ABSTRACT

This paper presents an acoustic analysis of how coda obstruents influence vowel phonation; this sheds some light on the articulatory mechanisms involved, as well as illuminating secondary cues that listeners may use to identify coda voicing. Analyses are based on 12 American English speakers producing words in isolation and in a frame sentence, with measurements tracked across each vowel.

Vowels before voiced obstruents exhibit a greater harmonics-to-noise ratio (HNR), lower spectral tilt, less jitter and shimmer, and slightly lower F0 than vowels before voiceless obstruents. The differences are most apparent in the final quarter of the vowels. Some patterns depend on the particular contrast; voicing-conditioned differences in spectral tilt are absent with alveolar stops.

Keywords: Speech production, acoustics, voicing contrasts, phonation

1. INTRODUCTION

Vowel characteristics are influenced in a variety of ways by coda voicing. The most thoroughly described effect is the greater duration of vowels before voiced consonants than before voiceless consonants, which is observed in many languages [2, 14].

Lisker summarizes many of the vowel characteristics influenced by following consonant voicing, in addition to other characteristics that distinguish voicing in stops; these vowel features include transition duration, F1, F0, and intensity decay time [13]. Subsequent work confirms most of these qualities (e.g. [16, 7, 9]), though the effect on F0 has not been consistently replicated (e.g. [5]).

Spectral tilt has also sometimes been observed to be higher in vowels next to voiceless obstruents than next to voiced obstruents, both onsets [10] and codas [4], at least for some places and manners of articulation. On the other hand, American English coda /p/ and /t/ often have glottal constriction that produces creakiness in the vowel, reflected in lower spectral tilt and several other characteristics, including more jitter [3]. Effects induced by consonant voicing can interact with other characteristics, so it is important to consider the particular consonantal contrasts.

Characteristics of vowel phonation are likely to be influenced by the articulation of voiced and voiceless codas, though there are relatively few studies that examine such effects. The greater laryngeal opening in anticipation of voiceless codas is likely to lower the harmonics-to-noise ratio (HNR) and increase spectral tilt; greater vocal fold tension should increase F0. The transition out of regular voicing is likely to increase jitter and shimmer.

This study reports the influence of coda obstruent voicing on spectral tilt, jitter, shimmer, HNR, and F0 within vowels. There are several significant phonation differences, which may serve as secondary cues to voicing. Parallel effects of duration suggest that some of the phonation effects of voicing could contribute to the duration effects, as changes in duration could enhance the existing differences in phonation.

2. METHODS

Recordings were made of 12 American English speakers (6 male; mean age = 25.8) reading English words, elicited in randomized order with PsychoPy [15] and recorded in a sound attenuated booth with a stand-mounted Blue Yeti microphone at a 44.1 kHz sampling rate. Each word was produced twice in isolation, in succession, and once in the frame sentence "she said _____ like them," providing three different contexts of production.

The set of words included 47 minimal (or near-minimal) pairs contrasting in the voicing of the final obstruent, e.g. *bad, bat* and *ridge, rich*, roughly balanced across six combinations of place and manner of articulation (/p-b, t-d, k-g, f-v, s-z, tʃ-dʒ/) and six vowel qualities (/i, ɪ, ɛ, æ, ɑ, ʌ/).

Each vowel was segmented into quarters, to measure change across the vowel. Spectral tilt was measured as H1-H2. Jitter and shimmer were measured as local percentages, as given by the Voice Report function in Praat. Harmonicity was calculated by forward cross-correlation. F0 mean is also reported.

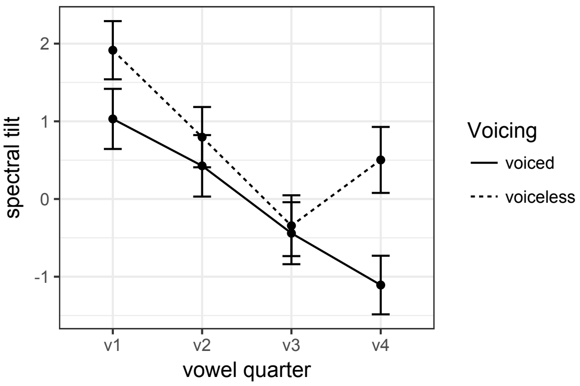
Regression models for factors influencing each of these characteristics were calculated using the lme4 package in R [1]; p-values were calculated by the lmerTest package [11].

3. RESULTS

3.1. Main effects of coda voicing

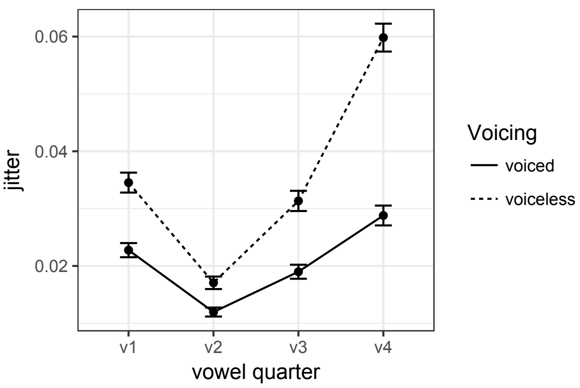
Spectral tilt decreased within vowels, consistent with the falling F0 and previous work showing that spectral tilt and F0 tend to be positively correlated [17]. However, in vowels preceding voiceless codas, spectral tilt increased in the final quarter to be significantly above the spectral tilt in vowels before voiced codas. Fig. 1 illustrates these effects. Error bars indicate the 95% confidence interval.

Figure 1: Spectral tilt by coda voicing.



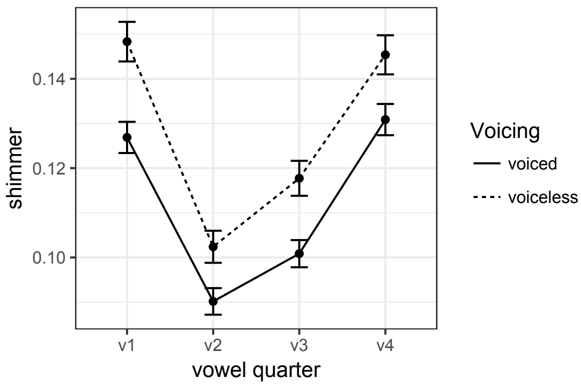
There was more jitter in vowels before voiceless codas than before voiced codas; the difference increased towards the end of the vowel, as illustrated in Fig. 2. Irregularity induced by neighboring consonants covers a larger percentage of shorter vowels, which likely contributes to the difference. However, voicing is a predictor of jitter even when duration is also included as a predictor (see Section 3.2).

Figure 2: Jitter (local) by coda voicing.



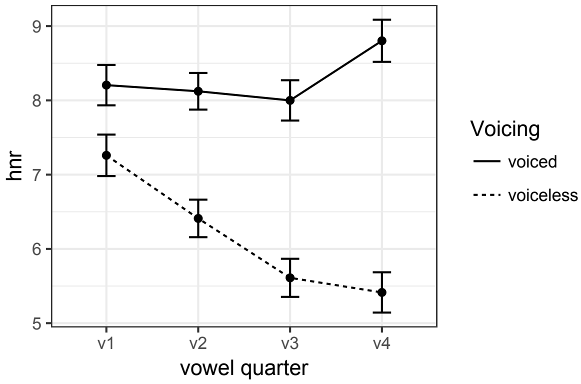
There was more shimmer throughout vowels before voiceless codas than those before voiced codas, as illustrated in Fig. 3. However, shimmer was primarily affected by vowel boundaries, likely due to shifting intensity caused by neighboring consonants.

Figure 3: Shimmer (local) by coda voicing.



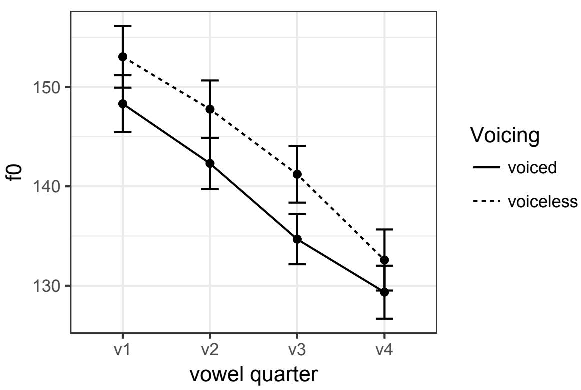
There was a substantially higher harmonics-to-noise ratio (HNR) in vowels before voiced codas than before voiceless codas, indicating less aperiodic noise in this environment. The difference increased throughout the vowel, primarily due to decreasing HNR before voiceless codas. Effects are illustrated in Fig. 4.

Figure 4: HNR by coda voicing.



F0 was slightly higher in vowels before voiceless codas than in vowels before voiced codas, but the distributions overlapped substantially. There was an overall decrease in F0 in both environments. These effects are illustrated in Fig. 5.

Figure 5: F0 by coda voicing.



3.2. Effects of duration vs. voicing

To test whether effects are due to voicing itself, or are indirect results of voicing-conditioned vowel duration, linear mixed effects models including both coda voicing (voiced; voiceless) and continuous vowel duration as fixed effects were calculated for each measure in the final vowel quarter. In each case, the random effects were participant and word. The reference value for coda voicing was Voiced.

Table 1 presents a regression model for spectral tilt. Spectral tilt was significantly greater before voiceless codas than before voiced codas, and lower in longer vowels.

Table 1: Regression model for spectral tilt.

	β	SE	t	p value
(Intercept)	-0.2	1.4	-0.15	0.88
Duration	-0.02	0.0099	-2.0	0.046
Coda-voiceless	1.2	0.51	2.3	0.023

Table 2 presents a regression model for jitter. There was more jitter before voiceless codas. Vowel duration was also a significant predictor, with less jitter in longer vowels.

Table 2: Regression model for jitter.

	β	SE	t	p value
(Intercept)	0.045	0.0076	5.9	$<$0.001
Duration	$-2.4\cdot10^{-4}$	$4.9\cdot10^{-5}$	-4.9	$<$0.001
Coda-voiceless	0.026	0.0022	11.7	$<$0.001

Table 3 presents a regression model for shimmer. There was significantly more shimmer before voiceless codas and a trend towards less in longer vowels.

Table 3: Regression model for shimmer.

	β	SE	t	p value
(Intercept)	0.14	0.0084	17.0	$<$0.001
Duration	$-1.7\cdot10^{-4}$	$1\cdot10^{-4}$	-1.7	0.086
Coda-voiceless	0.011	0.0041	2.6	0.01

Table 4 presents a regression model for HNR. Voiced codas were a significant predictor of greater HNR, as was longer vowel duration.

Table 4: Regression model for HNR.

	β	SE	t	p value
(Intercept)	6.5	1.1	5.9	$<$0.001
Duration	0.035	0.0063	5.6	$<$0.001
Coda-voiceless	-2.6	0.52	-5.0	$<$0.001

Table 5 presents a regression model for F0. Coda voicing was not a significant predictor of F0. However, longer vowel duration was a significant predictor of lower F0.

Table 5: Regression model for F0.

	β	SE	t	p value
(Intercept)	139.3	16.0	8.7	$<$0.001
Duration	-0.15	0.038	-4.0	$<$0.001
Coda-voiceless	-0.14	1.6	-0.088	0.93

Most characteristics were similarly affected by coda voicing and vowel duration, though the two factors were sufficiently independent to both be significant predictors. The notable exception was F0, for which there was an effect of duration but no effect of coda voicing after accounting for duration.

3.3. Consonant-specific effects

Some of the effects of coda voicing depended on the particular contrast, rather than simply voicing.

Table 6: Vowel duration by coda voicing and particular coda.

	voiced	voiceless
Labial stop	250 ms	170 ms
Alveolar stop	263 ms	183 ms
Velar stop	248 ms	161 ms
Labiodental fricative	288 ms	181 ms
Alveolar fricative	296 ms	182 ms
Post-alveolar affricate	235 ms	161 ms

Duration was affected by voicing for all of the coda pairs similarly, with vowels about 50% longer before the voiced item, though there was some variation in overall vowel duration due to coda place and manner of articulation, as illustrated in Table 6.

Table 7: Spectral tilt in the final vowel quarter by coda voicing and particular coda.

	voiced	voiceless
Labial stop	-0.42	0.77
Alveolar stop	-1.8	-1.4
Velar stop	-0.73	2.6
Labiodental fricative	-0.89	1.8
Alveolar fricative	-1.2	0.54
Post-alveolar affricate	-0.55	1.6

Spectral tilt was influenced differently by some of the coda voicing contrasts, as illustrated in Table 7. Higher spectral tilt before voiceless codas was apparent for most contrasts, but the difference was somewhat weaker for labial stops and extremely weak for alveolar stops. Spectral tilt before both alveolar stops was lower than in other environments, consistent with frequent creakiness.

Table 8: Jitter in the final vowel quarter by coda voicing and particular coda.

	voiced	voiceless
Labial stop	3.1	6.3
Alveolar stop	2.7	6.8
Velar stop	3.1	6.3
Labiodental fricative	2.6	4.7
Alveolar fricative	2.8	4.6
Post-alveolar affricate	3.3	6.8

The effect of coda voicing on jitter in the vowel was apparent across places and manners of articulation, though slightly weaker before fricatives than before stops and affricates, as illustrated in Table 8.

Table 9: HNR in the final vowel quarter by coda voicing and particular coda.

	voiced	voiceless
Labial stop	7.8	4.2
Alveolar stop	8.4	4.9
Velar stop	8.1	5.0
Labiodental fricative	10.3	7.3
Alveolar fricative	10.8	7.0
Post-alveolar affricate	6.9	3.8

Overall HNR was higher before fricatives than before stops or affricates. However, the effect of coda voicing was similar across places and manners of articulation, with an only slightly smaller effect for fricatives, as illustrated in Table 9.

Voicing effects on F0 and shimmer are not divided by consonant, as they were weak overall.

While there are also effects of vowel quality on some of these characteristics, they exhibit little evidence for interacting with effects of voicing, and will not be presented here.

3.4. Effects of production environment

There was no consistent effect of production environment across characteristics. Table 10 presents the mean qualities of the final vowel quarter from each production environment with each coda voicing.

For some measures, the effect was stronger in one environment. Differences in spectral tilt based on coda voicing were larger in the frame sentence than in isolated words.

In contrast, the voicing effect on HNR was larger for words in isolation, as was the effect on duration.

Table 10: Characteristics of the final vowel quarter by coda voicing and word environment.

	isolation 1		isolation 2		frame	
Voicing	vc	vcls	vc	vcls	vc	vcls
Duration	287	183	289	188	221	156
Spectral tilt	-0.59	0.38	-1.4	-0.15	-1.3	1.3
Jitter, %	2.5	5.7	3.3	6.4	2.8	5.8
Shimmer, %	12.3	14.6	13.3	14.6	13.7	14.4
HNR	9.2	5.2	9.1	5.4	8.1	5.7
F0	134	138	122	126	132	134

The weak effect of voicing on shimmer was clearer for words in isolation than for words in the frame sentence.

Word environment had no apparent effects on F0 or jitter, either overall or interacting with voicing.

4. DISCUSSION AND CONCLUSIONS

Voiceless codas increase spectral tilt in vowels. This likely results from increased tension in the vocal folds and separation between them, which makes the vibration cycle more sinusoidal and amplifies lower harmonics [8]. This effect of voicelessness sometimes competes with glottalization, which decreases spectral tilt in /t/ [3]; the low spectral tilt observed before /d/ suggests that vowels are also being glottalized in this environment in American English.

Voiceless codas decrease HNR and increase jitter in vowels. There is a similar but weaker relationship with shimmer. Movement to open the glottis begins before voicing ends [6]; the decrease in HNR could reflect breathiness corresponding to this widening of the glottis, and the increase in jitter could reflect irregularity due to the vocal folds shifting out of the orientation for regular voicing. These effects are paralleled by similar relationships between vowel duration and these phonation measures; longer vowels allow a more stable laryngeal posture, producing greater HNR and less jitter.

F0 is only slightly higher in vowels before voiceless codas, but the difference decreases, if anything, across the vowel. Duration is a stronger predictor; longer vowels have a lower final F0, consistent with broader declination tendencies [12].

Listeners may use some of these phonation differences as acoustic cues to voicing; these characteristics may become production targets as a result, and be exaggerated to help enhance voicing contrasts. The parallel effects of coda voicing and vowel duration, particularly for jitter and HNR, may suggest that creating vowel duration differences could strengthen existing phonation differences.

5. REFERENCES

[1] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.

[2] Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.

[3] Chong, A. J., Garellek, M. 2018. Online perception of glottalized coda stops in American English. *Laboratory Phonology* 9(1), Article 4.

[4] Coleman, J. 2003. Discovering the acoustic correlates of phonological contrasts. *Journal of Phonetics* 31, 351–372.

[5] Gruenenfelder, T. M., Pisoni, D. B. 1980. Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from speech perception and production. *Perception and Psychophysics* 28(6), 514–520.

[6] Hirose, H., Gay, T. 1972. The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica* 25, 140–164.

[7] de Jong, K. 1991. An articulatory study of consonant-induced vowel duration changes in English. *Phonetica* 48, 1–17.

[8] Klatt, D. H., Klatt, L. C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87(2), 820–857.

[9] Kohler, K. J. 1982. F0 in the production of lenis and fortis plosives. *Phonetica* 39, 199–218.

[10] Kong, E. J., Beckman, M. E., Edwards, J. 2012. Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese. *Journal of Phonetics* 40, 725–744.

[11] Kuznetsova, A., Bruun Brockhoff, P., Haubo Bojesen Christensen, R. 2015. *lmerTest: Tests in Linear Mixed Effects Models*. R package version 2.0-29.

[12] Ladd, D. R. 1984. Declination: A review and some hypotheses. *Phonology* 1, 53–74.

[13] Lisker, L. 1986. "voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29(1), 2–11.

[14] Maddieson, I. 1977. Further studies on vowel length before aspirated consonants. *UCLA Working Papers in Phonetics* 38, 82–90.

[15] Pierce, J. W. 2007. PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods* 162, 8–13.

[16] Summers, W. V. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82(3), 847–863.

[17] Swerts, M., Veldhuis, R. 2001. The effect of speech melody on voice quality. *Speech Communication* 33, 297–303.