

A sociophonetic account of gradient /z/ devoicing among Chicax high schoolers

Nicole Holliday* and Franny D. Brogan

Pomona College, USA

*Corresponding author. Email: nicole.holliday@pomona.edu

Abstract

This paper examines final /z/ devoicing among Chicax teens in Southern California to investigate the degree to which devoiced final /z/ neutralizes with final /s/ in this dialect. Results indicate on the one hand that devoiced /z/ remains distinct from /s/: as expected, devoiced /z/ is significantly less voiceless than /s/ and has a significantly lower center of gravity (COG). However, unexpectedly, devoiced /z/ has a significantly longer fricative duration and a significantly shorter preceding vowel duration than /s/, a pair of results that run counter to general tendencies for voiced fricatives to be shorter and have longer preceding vowels than their voiceless counterparts. We propose that these durational findings may explain, at least in part, the salience of final /z/ devoicing in Latinx Englishes despite its ubiquity among speakers of mainstream US English. In this first instrumental sociophonetic account of final /z/ devoicing in Latinx Englishes, we also find that, counter to existing segmental accounts, the morphological status of /z/ is no longer a significant predictor of devoicing. Moreover, while both following segment and speaker gender are significant predictors of devoicing, they do not condition devoicing in the expected ways.

Keywords: Obstruent devoicing; Chicano English; ethnolectal variation; youth language; sociophonetics

Final obstruent devoicing is prevalent crosslinguistically. While many languages phonologically neutralize final obstruents (e.g., German, Catalan, Russian), others do so variably; in English, complete and partial devoicing of final obstruents is relatively common (e.g., Docherty, 1992 and references therein). The prevalence of this phenomenon can be attributed to the aerodynamics of voiced obstruents, which require airflow through the glottis yet prevent supraglottal airflow from being readily released. Thus, “active oral tract expansion (for example, by tongue root advancement or larynx lowering) is necessary to maintain airflow in an obstruent” but “cannot be continued indefinitely or controlled tightly” (Hayes & Steriade, 2004:8). While these aerodynamic difficulties are exacerbated in prepausal position, obstruent devoicing in English also variably occurs word- and phrase-medially and initially.

Despite its ubiquity in varieties of Mainstream U.S. English (MUSE), final /z/ devoicing has been consistently cited as a salient feature of Latinx¹ Englishes (e.g., Frazer, 1996; Thomas & Van Hofwegen, 2019; Thompson, 1975), that is, varieties spoken by individuals and communities with an ethnic background and descent

from Mexico, Central, or South America. For these varieties, final /z/ devoicing has been attributed to substrate effects: Spanish lacks phonemic /z/. However, it is not clear from the existing literature that speakers of Latinx Englishes devoice final /z/ at higher rates than speakers of MUSE; the strong connection listeners make between this feature and Latinx speakers remains unexplained. The primary goal of the present study is to explore whether the acoustic characteristics of devoiced final /z/ might provide clues as to why devoicing is so much more salient in these language varieties.

In addition to exploring the acoustic characteristics of devoiced final /z/ in Latinx English, a secondary goal of the present study is to evaluate the potential benefits of analyzing this phenomenon as continuous rather than categorical. That is, while it is well established that voicing is a gradient phonetic process, existing work on this variable has conceived of it segmentally ([z] vs. [s], [z] vs. [z̥], etc.). This methodological approach requires researchers to draw arbitrary lines between variants, relying on dichotomous auditory judgments and hindering replicability. We thus aim to provide the first instrumental study of final /z/ devoicing in Latinx English in which we analyze this phenomenon gradiently via the continuous measure of *percent voicelessness* and subsequently compare our findings to existing segmental accounts.

Through an acoustic analysis of the naturalistic speech of eighteen Latinx high schoolers in Southern California, the paper endeavors to address two distinct research questions:

1. In other dialects of English, research has shown that devoiced /z/ retains various acoustic characteristics that prevent it from fully neutralizing with /s/. Is there evidence for more complete neutralization that might explain the feature's increased salience in Latinx dialects?
2. When /z/ devoicing is examined gradiently rather than categorically, do the following segment, morphological status of /z/, and gender identity of the speaker—the most robust predictors of /z/ devoicing in previous accounts of the phenomenon—still condition variation in the expected ways?

The sections that follow review the literature that serves as a foundation for these research questions.

(De)voicing as a gradient property

Previous studies of /z/ devoicing in both Latinx and other dialects of English have largely treated this variable as categorical, coding variants in either a bipartite or tripartite manner. Yet crosslinguistic studies on the inverse phenomenon of /s/ voicing have made compelling cases, both theoretical and practical, for treating voicing as gradient.

In a paper examining /s/ voicing assimilation in a non-/s/ weakening dialect of Peninsular Spanish, Campos-Astorkiza (2014) found ample variation in the degree of voicing before a following voiced obstruent, an environment thought to trigger phonological voicing assimilation across dialects of Spanish. The author concluded that “[s/ voicing] is not a categorical process, but rather gradient and incomplete

in many cases” and attributed this gradience to “increased gestural overlap between two adjacent and contradictory glottal gestures,” which “results in gestural blending and, consequently, gradient surface assimilation.”² The author further emphasized that “the question is not whether assimilation takes place or not, but rather to what extent the involved laryngeal gestures overlap; there could be no or minimal overlap, or total overlap” (Campos-Astorkiza, 2014:31). Earlier work on other dialects of Spanish (e.g., Romero, 1999; Schmidt & Willis, 2011) has made similar claims.

Investigations of /s/ voicing in other languages such as Greek have yielded parallel results and led scholars to arrive at similar conclusions about the gradient nature of voicing. Pelekanou and Arvaniti (2001), for example, examined /s/ voicing sandhi in two regional dialects of Greek and found that, in /s/+sonorant environments, degree of voicing is unsystematic. In line with Campos-Astorkiza (2014), the authors argued that the fact that identical phonological environments engendered different degrees of voicing showed that “/s/-voicing should best be treated as a gradient phenomenon, i.e., as the result of gestural overlap between the gestures of the vocal folds” (Pelekanou & Arvaniti, 2001:73). Baltazani (2006) later reaffirmed Pelekanou and Arvaniti’s results in a similar paper and cautiously situated her findings within Articulatory Phonology (Browman & Goldstein, 1986, 1989, 1992). Akin to explanations offered by scholars aforementioned in this section, Baltazani (2006:9) framed gradient voicing assimilation as “variable reduction in the amplitude of the opening gesture of the glottis, which is responsible for voicelessness.”

/z/ devoicing as a feature of Latinx Englishes

Because /z/ devoicing is common across dialects of English, the question of why it is such a salient feature of Latinx Englishes is one that scholars have been interested in investigating. Early accounts of /z/ devoicing posited that frequency may be the answer. In one of the first studies of this feature, for example, Thompson (1975) examined the speech of forty second-generation Mexican American men in Austin, Texas, in an auditory coding task and found that twenty-three of his participants devoiced /z/ between 10% and 15% of the time (typically after /ɪ/, as in ‘cars’ /kɑ.ɪz/, also a feature of Anglo speech in Austin), while the other seventeen devoiced at a rate of 35-75%. He concluded that only this latter group showed increased devoicing attributable to contact with Spanish and thus proposed a threshold of 25% for Spanish-influenced /z/ devoicing.

Some more recent work has shown that Mexican American speakers do, indeed, devoice /z/ at rates at or above this threshold. For example, in a study of twenty-four Mexican American teens in San Antonio, Texas, Bayley and Messing (2008) found that final /z/ devoicing occurred in 39.1% of the tokens. In the same community a handful of years later, Bayley and Holland (2014) examined final /z/ devoicing among thirteen novel speakers and found that the rate of devoicing was almost identical to that obtained by Bayley and Messing.

That said, not all findings have been so consistent. Frazer (1996), for example, found that Mexican American speakers in Sterling and Rock Falls, Illinois, devoiced /z/ either very infrequently or not at all, with individual rates ranging from 0% to 16%, all well under Thompson’s (1975) threshold. Similarly, in a comprehensive

study of Chicana English speakers in Austin, Texas, Galindo (1987) found that only 10% of tokens of /z/ were devoiced to [s]. While differences in methodology³—specifically, what constitutes devoicing—and the social characteristics of participants may have contributed to these disparate findings, it is nevertheless difficult to conclude from this body of literature that speakers of Latinx Englishes devoice /z/ at higher rates than speakers of other dialects.

Studies of /z/ devoicing in MUSE further complicate the story that frequency explains the salience of this feature in Latinx Englishes. In an examination of English /z/ devoicing via electroglottography (EGG) measures, for example, Smith (1997) found substantial variation in devoicing among four Midwestern and Western U.S. English-speaking participants. Binning tokens into three groups—“voiced” (tokens in which more than 90% of the fricative showed vocal fold vibration), “partially devoiced” (tokens in which 25–90% of the fricative was voiced), and “devoiced” (tokens in which less than 25% of the fricative was voiced)—Smith found that speakers produced “devoiced” tokens at rates of approximately 24–67% and “partially devoiced” tokens at rates of approximately 17–45%.⁴ All participants devoiced or partially devoiced /z/ more than 50% of the time, a rate higher than those found in various studies of /z/ devoicing in Latinx Englishes. Other studies of /z/ devoicing in MUSE have yielded similar results. For example, José (2010) found that White speakers in Northwestern Indiana devoiced postsonorant /z/ at a rate of 34% overall, well above Thompson’s (1975) threshold.

Just as the frequency of /z/ devoicing does not suffice to explain its salience in Latinx Englishes, nor do the contexts in which the phenomenon occurs: the phonological environments and morphological classes that favor devoicing have been found to be almost identical in Latinx and non-Latinx dialects of English. (Findings for Latinx Englishes are reviewed in the following section; for other dialects of English, see, for example, Holmes [1996], José [2010], and Smith [1997]). We thus turn to possible acoustic explanations.

Accounts of /z/ devoicing in non-Latinx dialects of English have examined the acoustic cues that potentially differentiate devoiced /z/ from /s/. Smith (1997), for example, found that, even when fully or partially devoiced, /z/ retains many of the acoustic cues that typically differentiate voiced fricatives from their voiceless counterparts; that is, the duration of /z/ is significantly shorter than that of /s/, the duration of the vowel preceding /z/ is significantly longer, and airflow for /z/ is significantly lower. Smith concluded that /z/ and /s/ resist complete neutralization, even when the former is produced without vocal fold vibration. As no similar analyses (to our knowledge) have been conducted with speakers of Latinx Englishes, we hypothesize that more complete neutralization of devoiced /z/ and /s/, a loss of contrast that would make /z/ devoicing more perceptually salient, might acoustically distinguish Latinx English /z/ devoicing.

Existing sociophonetic accounts of Latinx English /z/ devoicing

The voicing and manner of the following segment and the morphological class of /z/ have consistently been identified as the best predictors of devoicing in Latinx Englishes. Unsurprisingly, given expected regressive assimilatory effects, Bayley and

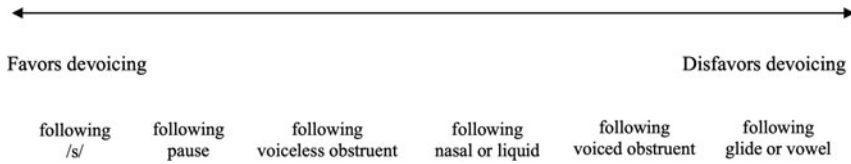


Figure 1. Attested relationship between following segment and /z/ devoicing.

Messing (2008), Bayley and Holland (2014), Doviak and Hudson-Edwards (1980), and others have identified the voicing of the following segment as the most important determinant of /z/ devoicing. Most recently, Bayley and Holland (2014) found that /z/ was devoiced most frequently when followed by an /s/ (e.g., /dagzslip/ ‘dogs sleep’), a pause (e.g., /dagz/ ‘dogs’), or another voiceless consonant (e.g., /dagzklɔdəl/ ‘dogs cuddle’). While voiced following segments generally disfavor devoicing, manner matters: devoicing is most prevalent preceding nasals and liquids (e.g., /dagznɔv/ ‘dogs know’) and least prevalent preceding glides and vowels (e.g., /dagzit/ ‘dogs eat’). These results echo, for the most part, those of Galindo (1987) and are summarized in Figure 1.

With respect to morphology, Bayley and Messing (2008) found that inflectional /-z/ (e.g., /dag-z/ ‘dog-s’) favored devoicing (except in cases of the third person singular, e.g., /dɔg-z/ ‘jog-s’), while /z/ that occurred as part of a monomorpheme (e.g., /lenz/ ‘lens’) was less likely to be affected. Bayley and Holland (2014) further extended this result, showing that plurals, possessives (e.g., /dag-z/ ‘dog’s’), and contracted copulas (e.g., /hiz/ ‘he’s’) all favored devoicing, while third person singular was neutral, and monomorphemes disfavored it. The authors noted that, while this finding “is the opposite of what we find in well-studied variables, such as English coronal stop deletion where inflected forms tend to resist deletion (Bayley, 1994; Labov, 1989),” the case of /z/ devoicing is different in that “there is no loss of information when an affix is devoiced. That is, from a functional perspective, plurals, possessives, and contracted copulas retain the same information whether the affix is realized as [s] or as [z]” (396-98).

Finally, gender has been consistently identified as a predictor of /z/ devoicing in Latinx Englishes, with scholars, regardless of when and where research is conducted, finding that women devoice /z/ at higher rates than their male counterparts (e.g., Bayley & Holland, 2014; Bayley & Messing, 2008; Frazer, 1996; Galindo, 1987; among others). It is worth noting that this finding is not unique to Latinx English. Rather, gender seems to be the most robust social predictor of /z/ devoicing in other dialects of English as well (e.g., José, 2010; Verhoeven, Hiron, & Basavaraj, 2011).

Methods

Participants

The speakers in this study are eighteen self-identified Chicax high school students (15–17 years old) residing in Los Angeles, San Bernardino, and Riverside Counties of Southern California. Eleven identify as female, and seven identify as male. All students report English and Spanish as their first languages and indicate that they use both English and Spanish on a daily basis. Figure 2 shows the geographic locations of the various communities from which the participants hail,⁵ with the Latinx population

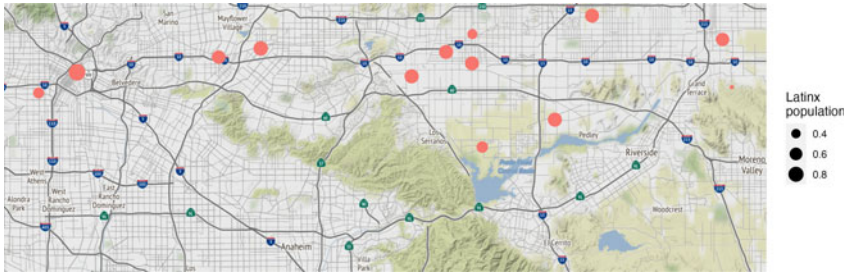


Figure 2. Participants' cities of origin by Latinx population in Los Angeles, San Bernardino, and Riverside Counties (Southern California).

(which ranges from 28.2% [Loma Linda] to 96.2% [East Los Angeles])—represented by dot size, based on data from the U.S. Census (U.S. Census, 2020). While the Census data reflect those who identify as any type of Latinx, the students in the current study all have Mexican American heritage and do identify as Chicax (Fought, 2003).

Speakers were interviewed during June and July of 2019 while participating in a Summer Enrichment Program (SEP) sponsored by Pomona College that provides English and Math classes for high-achieving, low-income/first-generation college students living in Southern California. All interviews, lasting approximately thirty minutes each, were conducted by six high school seniors also participating in the SEP, under the direction of a research team led by the first author and two undergraduate students. The SEP provides enrichment opportunities such as supplementary English and math classes for low-income/first-generation high schoolers in Southern California with the goal of assisting them in gaining admission to selective colleges and universities. One hundred percent of the students participating in the SEP identified as students of color, with 52% identifying as Hispanic/Latinx. The SEP student researchers involved in the data collection had participated in the program since their first year of high school and, as incoming seniors, they had the opportunity to collaborate with faculty and undergraduate students on a research project. The students were trained in research ethics and sociolinguistic interview procedures before embarking on the interview process.

Given the original research team's interest in the relationship between the use of nonstandard linguistic forms and school experience, interviews focused on topics related to language use, discipline and punishment, and race/ethnicity. As the interviewers were peers of the participants speaking in a private setting, students may have felt more comfortable using their authentic speech style and openly discussing their experiences at school (Rickford & McNair-Knox, 1994). Three of the interviewers identified as Chicax, two as Asian American, and one as African American. Each student conducted 3-5 one-on-one peer interviews as part of the study. Data collection met the ethical standards required by Pomona College and was approved through its Institutional Review Board (IRB).

Analysis

All interviews were orthographically transcribed and annotated by the first author and the undergraduate researchers on the research team using ELAN (ELAN,

2019) and were subsequently forced aligned using DARLA (Reddy & Stanford, 2015). For each recording, all instances of expected word-final [z]⁶ were identified and coded for preceding segment, following segment, and word on an interval tier in *Praat* (Boersma & Weenink, 2020). If the segment preceding /z/ was a vowel, then that vowel was also segmented. We extracted TextGrid information as well as measures of percent voicelessness, fricative duration, preceding vowel duration (if applicable), and center of gravity (COG) for 2,382 tokens of /z/ using a *Praat* script (Brown, 2014). Each participant contributed an average (mean) of 132 tokens of /z/ to the final dataset. While no prior studies (to our knowledge) have examined the role of COG in maintaining perceptual contrast between [z̥] and [s], we thought it prudent to do so given that voicing often pulls down COG values even when a low pass-band filter has been applied (Lindhout, 2016; Niebuhr, Lancia, & Meunier, 2008), giving voiceless fricatives reliably higher COGs than their voiced counterparts.

Once tokens and their corresponding acoustic and temporal measures had been extracted, the research team coded them for morphological class: third person singular /-z/ (e.g., /gōv-z/ 'goes'), other inflectional /-z/ (e.g., /dag-z/ 'dogs'), or lexical /z/ (e.g., /lɛnz/ 'lens'). Before conducting statistical analyses, we converted COG measures to z-scores using the *scale()* function in *R* (R Core Team, 2020) in an effort to account for variance in vocal tract size. All statistical analyses were conducted in *R*, and all data visualizations were produced using the *ggplot2* package (Wickham, 2016).

Results

Almost half of the /z/s analyzed (48%) were categorically voiced (37.3%) or voiceless (10.4%); the other 52% were intermediately devoiced. Figure 3 summarizes the distribution of data in a density plot.

While previous studies have shown that Latinx English speakers tend to devoice /z/ at rates of more than 25%, a gradient analysis reveals that the proportion of categorically voiceless tokens among our speakers is substantially lower. That said, almost two-thirds of tokens are devoiced to some degree, highlighting the potential issues posed by treating this phenomenon segmentally and the necessity of examining voicing in more precise detail.

Acoustic/temporal cues differentiating devoiced /z/ from /s/

Given that, in other dialects of English, research has shown that devoiced /z/ retains various acoustic characteristics that prevent it from fully neutralizing with /s/, we first

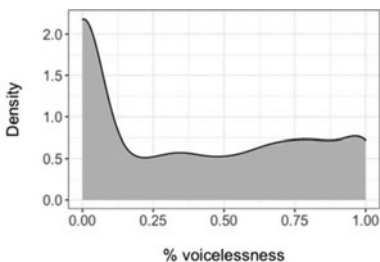


Figure 3. Distribution of % voicelessness for tokens of /z/ (n = 2382).

investigated whether there was evidence for more complete neutralization in our data that might explain the feature's increased salience in this dialect. For all tokens of final /z/ and final /s/ for which percent voicelessness was greater than zero ($n = 2100$), we compared four acoustic or temporal measures that typically distinguish phonemic /z/ and /s/: percent voicelessness, COG, fricative duration, and the duration of the preceding vowel (when applicable; $n = 1645$). In the case of incomplete neutralization, we would expect all of these measures to be significantly higher for /s/ except for the duration of the preceding vowel, which tends to be longer before /z/ (and voiced obstruents in general).

Using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R, we constructed four linear mixed-effects models. Each model examined the effect of phoneme (/s/ versus /z/) on one of the four acoustic/temporal measures and included "speaker" as a random effect. This model structure was selected due to the fact that, when the model is run with random effects of both "speaker" and "word," R yields a "singular fit" warning, which indicates model overfitting (see Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). It was also clear that the slope of the effect of "word" in this model was close to zero, indicating that "word" has very limited explanatory power and thus does not add to the reliability of the model.

The first model, the output of which appears in Table 1, examined percent voicelessness and revealed that, as expected, /s/ is significantly more voiceless than partially- and fully- devoiced /z/.

The second model, the output of which appears in Table 2, examined normalized COG and revealed that, as expected, tokens of devoiced /z/ have a significantly lower normalized COG than those of /s/.

For ease of comparison, differences in *raw* (unnormalized) COG for /s/ and devoiced /z/ are shown in Figure 4.

While the effects of phoneme on percent voicelessness and COG are as expected, the third model revealed no such parallel for fricative duration. That is, while voiced fricatives are thought to be generally shorter than their voiceless counterparts (Kaiser, 1997), our speakers' devoiced /z/s are significantly *longer* than their /s/s. The output of this model is presented in Table 3.

Table 1. Effect of phoneme on % voicelessness

	Tokens	Mean % voicelessness	Results of linear mixed model; DV = percent voicelessness ($n = 2986$)		
			β (coefficient)	Standard Error	<i>p</i> -value
Intercept			78.45	3.60	<.001***
Phoneme					<.001***
/s/	1492	78.73	Ref	–	–
/z/ (devoiced)	1494	61.55	–21.20	0.85	<.001***

Significance codes: * $p < .05$; ** $p < .01$; *** $p < .001$. Dashes indicate reference level of variable.

Table 2. Effect of phoneme on normalized COG

	Tokens	Mean COG (normalized)	Results of linear mixed model; DV = COG (normalized) (n = 2986)		
			β (coefficient)	Standard Error	p-value
Intercept			0.16	0.16	0.34
Phoneme					<.001***
/s/	1492	0.09	Ref	–	–
/z/ (devoiced)	1494	–0.09	–0.26	0.03	<.001***

Significance codes: * $p < .05$; ** $p < .01$; *** $p < .001$. Dashes indicate reference level of variable.

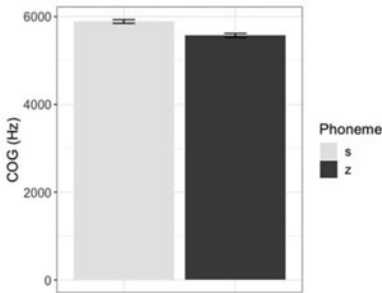


Figure 4. COG in Hz by phoneme.

Table 3. Effect of phoneme on fricative duration

	Tokens	Mean fricative duration (ms)	Results of linear mixed model; DV = fricative duration (ms) (n = 2986)		
			β (coefficient)	Standard Error	p-value
Intercept			102.84	4.55	<.001***
Phoneme					<.001***
/s/	1492	101.88	Ref	–	–
/z/ (devoiced)	1494	106.04	6.00	2.56	.02**

Significance codes: * $p < .05$; ** $p < .01$; *** $p < .001$. Dashes indicate reference level of variable.

To ensure that the observed differences are not simply an artifact of the speakers producing longer-than-expected /z/s even in the absence of devoicing, we compared duration across three categories: realizations of final /s/, devoiced realizations of final /z/, and fully voiced realizations of final /z/ (the reader will recall that this third category was excluded from the analyses above). This comparison, presented in [Figure 5](#), reveals that fully voiced final /z/s are much shorter than their devoiced counterparts, confirming that this lengthening is, indeed, unique to the latter variant.

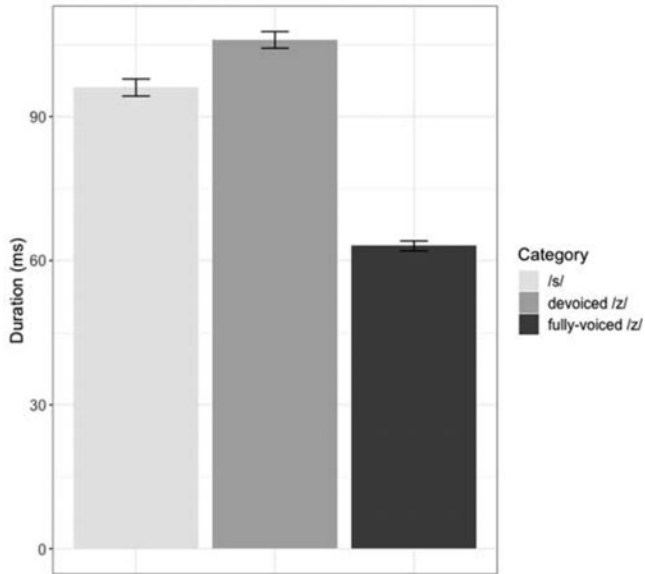


Figure 5. Duration of final /s/, devoiced final /z/, and fully voiced final /z/ in milliseconds.

The fourth and final model, the output of which appears in Table 4, examined the duration of the vowel preceding /s/ and devoiced /z/ (where applicable) and revealed that, again unexpectedly, vowels preceding /z/ are significantly shorter than those preceding /s/.

In summary, results show that in some respects devoiced /z/ remains distinct from /s/: as expected, devoiced /z/ is significantly less voiceless than /s/ and has a significantly lower COG. However, unexpectedly, devoiced /z/ has a significantly longer fricative duration and a significantly shorter preceding vowel duration than /s/, a pair of results that run counter to general tendencies for voiced fricatives to be shorter and have longer preceding vowels than their voiceless counterparts.

Table 4. Effect of phoneme on preceding vowel length

	Tokens	Mean preceding vowel length (ms)	Results of linear mixed model; DV = preceding vowel length (ms) (n = 1645)		
			β (coefficient)	Standard Error	p-value
Intercept			120.72	5.64	<.001***
Phoneme					<.001***
/s/	619	120.16	Ref	-	-
/z/ (devoiced)	1026	101.09	-19.00	3.71	<.001***

Significance codes: *p < .05; **p < .01; ***p < .001. Dashes indicate reference level of variable.

Linguistic and social constraints

The second goal of the present study is to ascertain whether, when examined gradiently, following segment, morphological status of /z/, and gender identity of the speaker still condition variation in the expected ways. To address this question, we tested the effects of these three variables on percent voicelessness. Following Chappell and García (2017) and Bolyanatz and Brogan (2021), percent voicelessness was modeled using zero-one inflated beta regression in the *gamlss* package (Rigby & Stasinopoulos, 2005). Zero-one inflated beta regression is preferable to standard linear regression when the values of the dependent variable are bounded $0 \leq y \leq 1$ and the data are nonnormally distributed.

The initial model included three fixed effects: following segment (with eight levels reflecting both voicing and manner distinctions), morphological status of /z/, and gender, as well as a random effect for “speaker.” The final model includes the fixed effects of following segment and gender; the morphological status of /z/ was removed because it was not statistically significant, a finding that stands in contrast with segmental work on /z/ devoicing. Table 5 summarizes the model output.

Akin to previous work on this phenomenon, gender serves as a significant predictor of /z/ devoicing. However, it does so in an unexpected direction: men produce significantly more voiceless /z/s than their female counterparts. This finding, as well the effect of the following segment, are discussed in more detail in the following section.

Table 5. Social and linguistic factors predicting % voicelessness of /z/

	Tokens	Mean % voicelessness	Results of inflated beta regression with <i>mu</i> values reported; DV = percent voicelessness (<i>n</i> = 2382)		
			β (coefficient)	Standard Error	<i>p</i> -value
Intercept			0.58	0.15	<.001***
Gender					
Female	1375	35.32	Ref	–	–
Male	1007	43.09	0.09	0.03	<.01**
Following segment					
[s]	65	44.74	Ref	–	–
Pause	637	60.79	–0.04	0.15	0.80
Voiceless obstruent	250	48.87	–0.16	0.16	0.32
Nasal	118	24.34	–0.19	0.19	0.32
Glide	151	33.43	–0.29	0.17	0.09
Vowel	590	27.30	–0.31	0.16	<0.05*
Voiced obstruent	304	26.33	–0.41	0.16	<0.05*
Liquid	267	22.76	–0.44	0.16	<0.01**

Significance codes: **p* < .05; ***p* < .01; ****p* < .001. Dashes indicate reference level of variable.

Discussion

Shades of neutralization

In our effort to ascertain what acoustic distinctions, if any, are retained between devoiced /z/ and /s/ in this dialect, we found that devoiced /z/ was significantly less voiceless and had a significantly lower COG than /s/. While these results are expected based on the existing literature, the findings with respect to fricative and preceding vowel duration are not. In naturalistic speech, voiced fricatives are reliably shorter than their voiceless counterparts while the vowels preceding voiced fricatives are reliably longer. Smith (1997) confirmed that devoiced /z/ remained significantly shorter than /s/ while the preceding vowel remained significantly longer among her MUSE speakers. However, our speakers' devoiced /z/s were, in fact, significantly longer than their /s/s, while the vowel preceding /z/ was significantly shorter than that preceding /s/. These findings are notable as previous work has found that, the more advanced a neutralization process driven by substrate effects is in an ethnolect, the more stigmatized the feature (Chand, 2009).

While the limited amount of instrumental research on this variable makes us cautious in our interpretation of these asymmetries, we propose that, because fricative and preceding vowel duration are strong perceptual cues for voicing, the unexpected length of devoiced /z/ and its preceding vowel may help explain the particularly high salience of this feature in Latinx Englishes. Of course, perceptual research is needed to confirm this hypothesis, and we encourage other scholars to pursue this line of inquiry.

/z/ devoicing: categories versus continua

The data presented in Figure 3 highlight the inherently gradient nature of final /z/ devoicing. When coding the data categorically, where does one draw the line? If “devoiced” means 100% voiceless, the frequency of devoicing in this data is about 10%; if “devoiced” means not fully voiced, the frequency of devoicing in our data is greater than 50%. As the body of literature on Latinx /z/ devoicing continues to grow, it would behoove scholars to either treat this feature as the gradient phenomenon it is or bin segmental categories (i.e., [s]/[z̥]/[z]) according to continuous acoustics to aid in replicability.

Moreover, our sociophonetic analysis reveals that, when the dependent variable is treated as gradient, the linguistic predictors of /z/ devoicing shift in turn. Particularly of note is the absence of a significant effect for morphological status of /z/, which has reliably predicted /z/ devoicing in segmental accounts. Results with respect to following segment are more nuanced. Figure 6 juxtaposes our findings with those of existing segmental accounts; key differences are marked in *italics*.

As seen in Figure 6, while a following /s/, pause, and non-/s/ obstruent affect devoicing as in segmental accounts, we find differences vis-à-vis where following sonorants fall on the devoicing continuum. In previous accounts, less sonorous sonorants (nasals and liquids) tended to favor devoicing more than more sonorous segments (glides and vowels). However, in our model, vowels and glides favored devoicing more than liquids, which favored devoicing least of any following sound type. Upon examining the distribution of values for percent voicelessness by

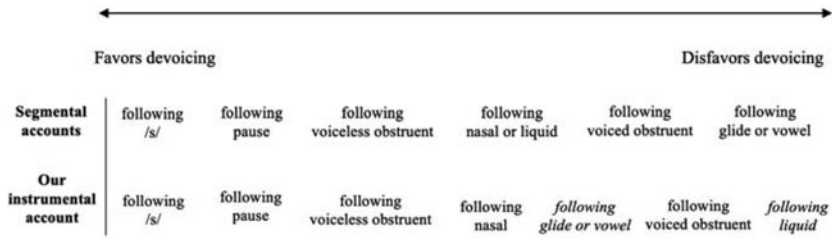


Figure 6. Effect of following segment on final /z/ devoicing.

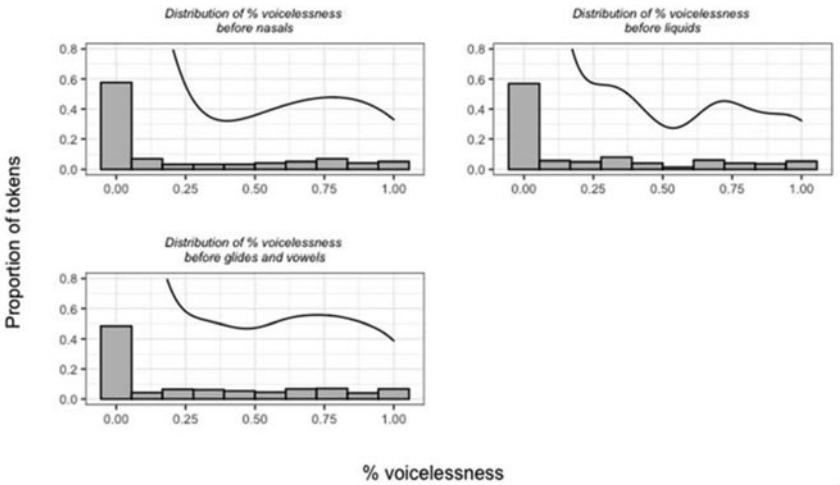


Figure 7. Distribution of percent voicelessness by following segment type.

following segment in Figure 7, we see that nasals and vowels have a higher concentration of values on the right (more voiceless) end of the distribution as compared to liquids, which have more values clustered below 50% voicelessness.

Crucially, while nasals and liquids look almost identical at the poles of the distribution, potentially explaining why they tend to cluster together in segmental accounts, the shape of the distribution of values for nasals is much more akin to that of glides and vowels. A gradient analysis of final /z/ devoicing takes these underlying patterns into account.

A novel gender finding

As described in earlier sections of this paper, much previous work on final /z/ devoicing in Latinx Englishes has found that women devoice final /z/ at higher rates than their male counterparts (e.g., Bayley & Holland, 2014; Bayley & Messing, 2008). While this may be surprising given the general tendency for men to produce nonstandard variants at higher rates than women (Labov, 2001:293), it aligns with

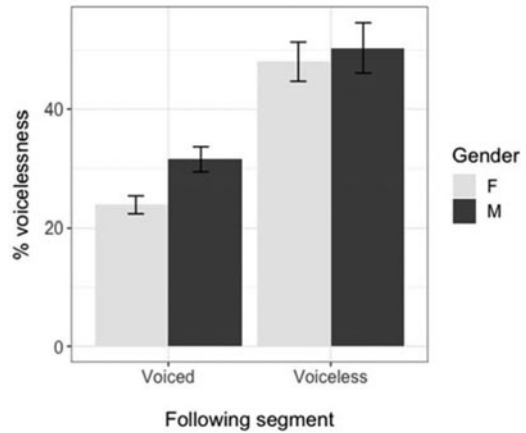


Figure 8. Percent voicelessness by gender and voicing of the following segment.

physiological expectations: men tend to have longer vocal tracts and thicker vocal folds (Beck, 1999), both of which can make glottal adduction harder to control.⁷

Our finding that male participants devoiced /z/ at significantly higher rates than their female counterparts is inconsistent with the existing body of work and runs counter to what we might expect if variation were driven purely by physiological differences. Thus, we propose that, in the speech community in question, the characteristics, qualities, and so on indexed through use of the nonstandard variant (devoiced [z]) may play a crucial role in explaining patterns of variation. With this proposal in mind, it is worth noting that our male speakers also produce significantly more voiceless /z/s preceding voiced consonants as compared to their female counterparts, as shown in Figure 8.

This finding is rather curious, as the prevoiced environment is not only the most phonetically unnatural context for devoicing but is also the only context in which Spanish phonologically voices /s/, in words such as /besbol/ ['bez.bol] 'baseball.' We argue that this finding adds credence to our argument that /z/ devoicing carries social meaning for speakers, as scholars have frequently cited use of an already-salient variable in phonetically unnatural contexts as a means for emphasizing in-group solidarity. Schilling-Estes (2000:159), for example, in her account of /ay/ monophthongization among Lumbee speakers in North Carolina, noted that the "social salience [of /ay/ monophthongization] may have led speakers [...] to heighten their usage levels for the variant in phonetically less favored as well as highly favored contexts in order to increase the prominence of an already highly noticeable feature." Future research should investigate the potential social meaning of final /z/ devoicing among Chicax youth in Southern California.

Conclusion

The goals of the present study were twofold. First, we endeavored to explore the acoustic nature of final /z/ devoicing among speakers of Chicax English. We compared realizations of devoiced /z/ with those of /s/ across multiple acoustic and

temporal measures. We found that, for measures of percent voicelessness and COG, devoiced /z/ and /s/ remained sufficiently differentiated. However, while previous studies of /z/ devoicing in other dialects of English show that devoiced /z/ retained a shorter fricative duration and longer preceding vowel as compared to /s/, we found that these contrasts were not just lost but reversed: our speakers' devoiced /z/s were significantly longer than their /s/s, and vowels preceding devoiced /z/ were significantly shorter than those preceding /s/. These novel findings led us to wonder if these durational asymmetries might explain the elevated salience of this variable in this dialect.

The present study also aimed to provide a sociophonetic account of final /z/ devoicing in Chicano English and championed the use of acoustic measures to do so. Previous work on this variable has treated it categorically, which is problematic both theoretically and practically. First, given the robust evidence that voicing is a gradient process, our analytical methods should reflect as much. Furthermore, because of the gradient nature of this variable, binning variants into artificial categories is necessarily subjective as "the relationship between phonological representation and phonetic realization becomes complex and no clear boundary between /z/ and /s/ can be established" (Thomas & Van Hofwegen, 2019:63). In this paper, we showed that the vast distribution of voicelessness across tokens of /z/ necessitates that the variable either be treated as continuous or binned acoustically to aid in replicability.

Moreover, when comparing our sociophonetic model to those presented in previous studies, we found that the linguistic predictors of /z/ devoicing differed slightly when the outcome variable was treated as continuous. We also found that male speakers devoiced final /z/ significantly more than their female counterparts (and did so in more phonetically unnatural environments), a finding that runs counter to much of the existing literature. Cautiously, this provides some evidence that this variable carries social meaning for these male speakers, who are physiologically predisposed to producing less voicelessness, not more. With the recognition that social meaning is locally constructed, we believe that additional in-depth, qualitative research is required to unearth the meaning of this variable in this particular community.

Finally, while durational measures were not the focus of this paper, one limitation of this methodology is that we did not normalize duration by local speech rate, which could potentially affect duration measures. Additionally, we did not conduct a full prosodic analysis examining pause length and boundaries, which could also potentially impact variation at the individual level. Future work that aims to corroborate our finding with respect to fricative length should consider integrating speech rate into its analyses to avoid this shortcoming.

Acknowledgments. We are grateful to our participants as well as the student interviewers for sharing their time and stories with the research team, and to Pomona College's Department of Linguistics and Cognitive Science for financial support. No external funding was obtained in conducting this research.

Competing interests. The authors declare none.

Notes

1. Though the terms "Latino" and "Latina" have been used historically for individuals with a familial and ethnic background from Latin America, the term "Latinx" has increased in popularity due to its inclusion of

people of all genders. The recent volume by Salinas and Lozano (2021) contains several chapters related to the history of the term, but the authors ultimately advocate for “Latinx” (and Chicanx) as the preferred terms in academic writing, due to their broader inclusivity.

2. See Browman and Goldstein (1989) for a comprehensive discussion of articulatory gestures.
3. Tokens for the studies discussed here used primarily impressionistic transcription and coded variants within a binary system (voiced versus voiceless).
4. Smith (1997) did not provide exact percentages; these estimates are based on Figure 2, found on page 481.
5. There are eighteen participants but only thirteen communities as some participants came from the same communities.
6. That is, we only included tokens of /z/ that “should” surface as [z] according to the phonological rules of Mainstream U.S. English (MUSE). For the remainder of this paper, we will refer to the qualifying tokens as “/z/” with the understanding that those /z/s that are phonologically devoiced in MUSE were excluded.
7. See Chappell and García (2017:7–10) for a comprehensive review of literature examining the effect of physiological factors on the ability to control voicing cessation.

References

- Baltazani, Mary. (2006). On /s/-voicing in Greek. *Proceedings of the 7th International Conference on Greek Linguistics*. York: University of York.
- Bates, Douglas, Martin Maechler, Ben Bolker, & Steve Walker. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67:1–48. doi:10.18637/jss.v067.i01
- Bayley, Robert. (1994). Consonant cluster reduction in Tejano English. *Language Variation and Change* 6:303–26. doi:10.1017/S0954394500001708
- Bayley, Robert, & Holland, Cory. (2014). Variation in Chicano English: The Case of Final (z) Devoicing. *American Speech* 89:385–407. doi:10.1215/00031283-2908200
- Bayley, Robert, & Messing, Cory L. (2008). Final (z)-devoicing in Chicano English. In M. Grosvald & D. Soares (eds.), *Proceedings of the 38th Western Conference on Linguistics* 19:1–11. Davis: University of California.
- Beck, Janet Mackenzie (1999). Organic variation of the vocal apparatus. In W. J. Hardcastle & J. Laver (eds.), *The handbook of phonetic sciences*. 256–89. Oxford: Blackwell.
- Boersma, Paul, & Weenink, David. (2020). Praat: Doing phonetics by computer [Computer software]. Amsterdam, Netherlands. Retrieved from <http://www.praat.org>
- Bolyanatz, Mariška, & Brogan, Franny D. (2021). Acoustic Differences between Chilean and Salvadoran Spanish /s/. *The Journal of the Acoustical Society of America* 150(4):2446–60. doi:10.1121/10.0006443
- Browman, Catherine P., & Goldstein, Louis M. (1986). Towards an articulatory phonology. *Phonology Yearbook* 3:219–52. doi:10.1017/S0952675700000658
- Browman, Catherine P., & Goldstein, Louis M. (1989). Articulatory Gestures as Phonological Units. *Phonology* 6:201–51. doi:10.1017/S0952675700001019
- Browman, Catherine P., & Goldstein, Louis M. (1992). Articulatory phonology: An overview. *Phonetica* 49:155–80. doi:10.1159/000261913
- Brown, Earl. (2014). “Extract info about /s/” [Praat Script]. Retrieved from <http://www-personal.k-state.edu/~ekbrown/#>.
- Campos-Astorkiza, Rebeka. (2014). Sibilant Voicing Assimilation in Peninsular Spanish as Gestural Blending. In M.-H. Côte, É. Mathieu, & S. Poplack (eds.), *Variation within and across romance languages*. 17–38. Amsterdam: John Benjamins.
- Chand, Vineeta. (2009). [v]at is Going On? Local and Global Ideologies about Indian English. *Language in Society* 38:393–419. doi:10.1017/S0047404509990200
- Chappell, Whitney, & García, Christina. (2017). Variable Production and Indexical Social Meaning: On the Potential Physiological Origin of Intervocalic /s/ Voicing in Costa Rican Spanish. *Studies in Hispanic and Lusophone Linguistics* 10:1–37. doi:10.1515/shll-2017-0001
- Docherty, Gerard J. (1992). *The timing of voicing in British English obstruents*. Berlin: De Gruyter Mouton.
- Doviak, Martin J., & Hudson-Edwards, Allison. (1980). Phonological Variation in Chicano English: Word-final (z)-devoicing. In E. L. Blansitt & R. V. Teschner (eds.), *A Festschrift for Jacob Ornstein: Studies in General Linguistics and Sociolinguistics*. Rowley: Newbury House. 82–95.

- ELAN (Version 5.7). (2019). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive [Software]. Retrieved from <https://archive.mpi.nl/tla/elan>
- Fought, Carmen R. (2003). *Chicano English in Context*. New York: Palgrave MacMillan.
- Frazier, Timothy C. (1996). Chicano English and Spanish Interference in the Midwestern United States. *American Speech* 71:72–85. doi:10.2307/455470
- Galindo, Delma Letticia. (1987). Linguistic Influence and Variation on the English of Chicano Adolescents in Austin, Texas. Doctoral dissertation, University of Texas.
- Hayes, Bruce, & Steriade, Donca. (2004). Introduction: The Phonetic Bases of Phonological Markedness. In B. Hayes, R. Kirchner, & D. Steriade (eds.), *Phonetically based phonology*. Cambridge: Cambridge University Press. 1–33.
- Holmes, Janet. (1996). Losing voice: Is Final /z/ Devoicing a Feature of Maori English? *World Englishes* 15:193–205. doi:10.1111/j.1467-971X.1996.tb00105.x
- José, Brian. (2010). The Apparent-time Construct and Stable Variation: Final /z/ Devoicing in Northwestern Indiana. *Journal of Sociolinguistics* 14:34–59. doi:10.1111/j.1467-9841.2009.00434.x
- Kaiser, Ed. (1997). *Strong Fricatives*. Amsterdam: Phonetic Sciences. Retrieved from <https://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/speech.bme.ogi.edu/tutordemos/SpectrogramReading/cse551.html/cse551/node32.html>
- Labov, William. (1989). The Child as Linguistic Historian. *Language Variation and Change* 1:85–97. doi:10.1017/S0954394500000120
- Labov, William. (2001). *Principles of linguistic change: social factors (Vol. 2)*. Oxford: Blackwell.
- Lindhout, Maartje. (2016). *The acquisition of the “voicing” distinction in Dutch fricatives*. Master’s thesis, Leiden University.
- Matuschek, Hannes, Kliegl, Reinhold, Vasishth, Shravan, Baayen, Harald, & Bates, Douglas. (2017). Balancing Type I Error and Power in Linear Mixed Models. *Journal of Memory and Language* 49:305–15. doi:10.1016/j.jml.2017.01.001
- Niebuhr, Oliver, Lancia, Leonardo, & Meunier, Christine. (2008). On Place Assimilation in French Sibilant Sequences. In R. Sock, S. Fuchs, & Y. Laprie (eds.), *8th International Seminar on Speech Production*. Strasbourg: INRIA. 221–24.
- Peleanou, Theodora, & Arvaniti, Amalia. (2001). Postlexical Rules and Gestural Overlap in a Greek Spoken Corpus. In C. Clairis (ed.), *Proceedings of the 5th International Conference on Greek Linguistics*. Paris: L’Harmattan. 71–74.
- R Core Team. (2020). R: A language and environment for statistical computing [Software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Reddy, Sravana, & Stanford, James. (2015). DARLA: A Web Application for Automated Dialect Analysis. In M. Gerber, C. Havasi, & F. Lacatusu (eds.), *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. Denver: Association for Computational Linguistics. 71–75.
- Rickford, John R., & McNair-Knox, Faye. (1994). Addressee- and Topic-influenced Style Shift: A Quantitative Sociolinguistic Study. In D. Biber & E. Finegan (eds.), *Sociolinguistic perspectives on register*. New York: Oxford University Press. 235–76.
- Rigby, Robert A., & Stasinopoulos, D. Mikis. (2005). Generalized Additive Models for Location, Scale and Shape (with Discussion). *Applied Statistics* 54:507–54.
- Romero, Joaquín. (1999). The Effect of Voicing Assimilation on Gestural Coordination. In J. J. Ohala (ed.), *Proceedings of the 14th International Congress of Phonetic Sciences*. Berkeley: University of California. 1793–96.
- Salinas, Cristobal, & Lozano, Adele. (2021). The History and Evolution of the Term Latinx. In E.G. Murillo, Jr., D. Delgado Bernal, S. Morales, L. Urrieta, Jr., E. Ruiz Bybee, J. Sánchez Muñoz, V.B. Saenz, D. Villanueva, M. Machado-Casas, & K. Espinoza (eds.), *Handbook of Latinos and education*. New York: Routledge. 249–63.
- Schilling-Estes, Natalie. (2000). Investigating Intra-Ethnic Differentiation: /Ay/ in Lumbee Native American English. *Language Variation and Change* 12:141–74. doi:10.1017/S0954394500122021
- Schmidt, Lauren B., & Willis, Erik W. (2011). Systematic Investigation of Voicing Assimilation of Spanish /s/ in Mexico City. In S. M. Alvord (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*. Somerville: Cascadilla Proceedings Project. 1–20.

- Smith, Caroline L. (1997). The Devoicing of /z/ in American English: Effects of Local and Prosodic Context. *Journal of Phonetics* 25:471–500. doi:10.1006/jpho.1997.0053
- Thomas, Erik R., & Van Hofwegen, Janneke. (2019). Consonantal Variables Correlated with Ethnicity. In E. R. Thomas (ed.), *Mexican American English*. Cambridge: Cambridge University Press. 61–92.
- Thompson, Roger M. (1975). Mexican-American English: Social Correlates of Regional Pronunciation. *American Speech* 50:18–24. doi:10.2307/3087859
- US Census. (2020). United States Census of 2020. <https://www.census.gov/2020census>. Accessed on 06/28/2022.
- Verhoeven, Jo, Hirson, Allen, & Basavaraj, Kavya. (2011). The Devoicing of Fricatives in Southern British English. In W. Lee & E. Zee (eds.), *Proceedings of the 17th International Conference on Phonetic Sciences*. Hong Kong: City University of Hong Kong. 2070–73.
- Wickham, Hadley. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Cite this article: Holliday N, Brogan FD (2022). A sociophonetic account of gradient /z/ devoicing among Chicana high schoolers. *Language Variation and Change* 34, 165–182. <https://doi.org/10.1017/S0954394522000114>