

Bay Area Spanish: regional sound change in contact languages

Annie Helms

University of California, Berkeley
annie_helms@berkeley.edu



Received: 15-03-21
Accepted: 30-09-21
Published: 22-02-22

How to cite: Annie Helms. 2022. Bay Area Spanish: regional sound change in contact languages. RLLT17, eds. Ora Matushansky, Laurent Roussarie, Michela Russo, Elena Soare & Sophie Wauquier. Special issue of *Isogloss Open Journal of Romance Linguistics* 8(2)/17, 1-18.

DOI: <https://doi.org/10.5565/rev/isogloss.116>

Abstract

Compression of the vowel space area and reduced vowel dispersion have been observed as features of the California Vowel Shift (CVS, D’Onofrio, Pratt, & Van Hofwegen 2019), an ongoing process of sound change in white Anglo English spoken in California, United States. The present research focuses on the Bay Area region of California, where there is a large community of Spanish speakers, and provides empirical data to address the gap in the literature regarding the influence of English regional sound change on other languages spoken in the region, namely the influence of the CVS on Spanish spoken in the Bay Area (Bay Area Spanish). Analysis of word list tasks given to bilingual speakers of Bay Area Spanish and monolingual speakers of Mexican Spanish reveals a more compressed vowel space in Bay Area Spanish and patterns that mirror changes in progress in California English. These results suggest that regional sound changes can influence production within contact languages and broaden the scope of regional sound change studies by removing the historically Anglo-centric focus.

Keywords: U.S. Spanish, California Vowel Shift, language contact, vowel space, regional sound change.

1. Introduction

The compression of the vowel space area and reduction of vowel dispersion have been posited as features of California white Anglo English¹ (D’Onofrio, Pratt, and Van Hofwegen 2019) and are identified as changes in progress associated with the California Vowel Shift (CVS). As Eckert (2008) states, “[i]n the dominant discourse of American dialectology, the white Anglo variety is considered a regional dialect, while African American and Latino varieties are considered ethnic dialects,” (p. 27). Accordingly, typical linguistic descriptions of regional patterns of ongoing vowel changes, such as CVS, are based on data collected from mainly white Anglo English speakers. As white Anglo speakers are in the minority of the population of California (United States Census Bureau 2019), it is important to consider whether other speakers in California index regional identity through similar modulations of the vowel system, and whether these modulations are present in regional languages other than English (i.e., contact languages). Though several studies have examined the relationship between ethnicity and participation in CVS (Cheng 2016, Eckert 2008, Fought 1997, 1999, 2003, Godinez and Maddieson 1985, Hall-Lew 2009, A. Wong and Hall-Lew 2014), little work has been done to examine the influence of CVS on contact languages, such as Spanish. Accordingly, this study focuses on the Bay Area of California and investigates whether changes in progress in CVS, namely compression of vowel space area and reduced vowel dispersion, are also present in the vowel space of Spanish in the Bay Area (Bay Area Spanish).

1.1. Spanish in the Bay Area

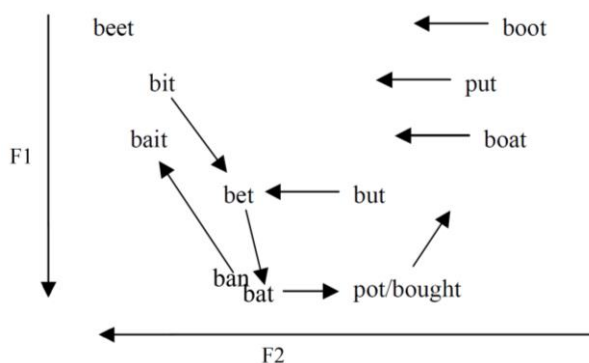
The Bay Area, shown in Figure (1), is a region in California comprised of nine counties surrounding the San Francisco Bay and containing the metropolitan centers of San Francisco, Oakland, and San Jose. The area is both ethnically and linguistically diverse, where 42% of the population is white, 24% is Hispanic (Bay Area Census 2010), and 40% speaks a language other than English at home (United States Census Bureau 2013). Among this portion of the population, 40% speak Spanish at home (United States Census Bureau 2013). Due to the proximity with the Mexican border and California’s history as part of Mexico in the 19th century, Mexican Spanish is the variety that most influences Spanish spoken in California, and 84% of Spanish speakers in California claim Mexican heritage (Pew Hispanic Center 2002). Despite the large Hispanic population in the Bay Area and the prominent presence of Spanish, little to no empirical attention has been given to Spanish in the Bay Area versus Spanish in Los Angeles (Davidson 2019).

¹ Following Eckert (2008), I use the term ‘white Anglo English’ to refer to the variety of English spoken by white-presenting speakers. Similar terms in the literature include ‘Anglo English’ (Fought 1999, Hinton et al. 1987), ‘European English’ (Hall-Lew 2010), and ‘White English’ (Cheng 2016).

Figure 1. The Bay Area of California (highlighted in red)

1.2. The California Vowel Shift

The California Vowel Shift was first described in 1987 by Hinton et al. as the set of ongoing sound changes observed predominately among white Anglo English speakers in California, and predominately among residents of the Bay Area. Since then, further studies have been carried out in Northern California (Cheng 2016, Eckert 2008, Hall-Lew 2009, Hall-Lew et al. 2015), Los Angeles (Fought 1997, 1999, Godinez and Maddieson 1985), and the Central Valley (D’Onofrio, Pratt, and Van Hofwegen 2019, Podesva et al. 2015). Figure (2) demonstrates the shifts that have been observed, namely the fronting of back vowels (higher F2), the raising and backing of the POT-BOUGHT merger (lower F1, lower F2), raising of BAN (lower F1), and backing of BIT, BET, and BAT (lower F2).²

Figure 2. The California Vowel Shift

Source: Adapted from Eckert (2008)

While most studies measure formants of individual vowels to determine vowel height and vowel frontedness, D’Onofrio, Pratt, and Van Hofwegen (2019) examine vowel space area and vowel dispersion as possible metrics for sound change. Whereas measures of vowel height rely solely on F1 and measures of vowel frontedness rely solely on F2, vowel space area and vowel dispersion take into account both formant measures to determine the Euclidean area spanned by the outer points of the vowel system and the average Euclidean distance between outer points and the vowel system

² When speaking of vowel shifts in English, the English words from Figure (2) will be used to describe the lexical sets containing the vowel in question. However, when the discussion focuses on Spanish production, or English production where Spanish may be a substrate, the lexical sets will then be referred to in terms of the corresponding Spanish vowel (i.e., /u/-fronting in Spanish is equivalent to BOOT-fronting in English).

centroid, respectively. Using these two metrics of the vowel system, D’Onofrio and colleagues find that younger speakers have more compressed vowel space areas and reduced vowel dispersion when compared to older generations of speakers. Based on their findings, the authors postulate that compression of the vowel space in and of itself can constitute a sociolinguistic variable that speakers can use as a social index. These findings complement previous studies (Gick et al. 2004, Ramanarayanan et al. 2013, Wilson and Gick 2006) that observe a compression of the vowel space when speakers switch between languages or between speaking styles. D’Onofrio, Pratt, and Van Hofwegen (2019: 212) account for vowel space compression with a distinct articulatory setting of the jaw, which is described as:

slightly lowered, open jaw and a relatively fronted lingual setting . . . [which] accounts not only for the front-back compression of the space, as a lowered jaw reduces the mobility of the tongue in this dimension, but it also accounts for the general lowering of the front vowels observed in the later stages of the shift, as a lowered jaw setting would result in a generally lowered tongue body, leading to lowered vocalic productions.

Their hypothesis is backed by visual evidence from imitations of California English (Pratt and D’Onofrio 2017), particularly of the stigmatized ‘Valley Girl’ and ‘Surfer Dude’ personae. In these imitations, speakers are observed to manipulate their jaw settings, both as a semiotic resource and to derive vowel shifts consistent with CVS. Thus, Pratt and D’Onofrio posit that both jaw setting and vocalic productions are salient and have social meaning in the context of California English.

1.3. Ethnicity and regional sound change

The trend in American dialectology is to define regional varieties of English by the language of white Anglo speakers of English and to classify the varieties of English spoken by other members of society as ethnolects (Eckert 2008, pp. 26-27), for example African American English and Chicano English. This trend not only underrepresents entire communities of speakers in the dialectal designation of ‘regional’, but also fortifies the notion that white Anglo speech is standard and that the speech of others is nonstandard. Additionally, this practice has resulted in little empirical attention paid to regional varieties of African American English and Chicano English. Accordingly, several studies have examined whether ethnic identification prevents or encourages speakers to exhibit features of the ongoing sound change characteristic of white Anglo California English (i.e., CVS). The findings indicate that the relationship between ethnicity and participation in CVS is vowel-dependent, but that the driving force of this sound change is not necessarily ethnicity. Instead, participation in sound change may be more closely related to social networks, for which ethnicity may be a proxy (Cheng, 2016). Particularly relevant to the present paper are studies of the participation of Chinese Americans (Cheng 2016, Hall-Lew 2009, Hall-Lew et al. 2015, A. Wong and Hall-Lew 2014), Korean Americans (Kim and N. Wong 2020), and Chicano English speakers³ (Eckert 2008, Fought 1997, 1999, 2003, Godinez and Maddieson 1985) in

³ In this paper, Chicano English is defined as the variety of English spoken by Americans of Mexican heritage. Speakers of Chicano English have varying levels of Spanish proficiency and fluency, with some not claiming any knowledge of Spanish.

CVS. The results of these studies indicate that these speakers do, in fact, participate in regional sound change previously only attributed to and attested among white Anglo speakers of English. However, the data challenge the idea that ethnicity alone is the factor that is correlated to participation in sound change. Hall-Lew (2009) and A.Wong and Hall-Lew (2014) find that Chinese American participation in the POT-BOUGHT merger mirrors the community-level change in progress, and that this feature is used more so as a regional marker than an ethnic marker. Cheng (2016) observes productions of Asian Americans and white Anglo Americans and finds that some vowel changes associated with CVS vary across ethnicity whereas others do not.

Regional features of Chicano English in the United States have been observed to complement ongoing sound changes in coterritorial white Anglo English (Los Angeles, California: Fought 1997, 1999, 2003, Godinez and Maddieson 1985; Michigan: Roeder 2010; Chicago, Illinois: Konopka and Pierrehumbert 2008; and general Chicano English: Santa Ana and Bayley 2004). In these studies, the influence of the regional variety of white Anglo English has been observed in the sound changes in the coterritorial Chicano English vowel system. For example, /u/-fronting and /o/-fronting have been observed among Chicano English speakers in Los Angeles, where the coterritorial CVS exhibits similar changes (Fought 1997, 1999, 2003, Godinez and Maddieson 1985). However, /u/- and /o/-fronting are not characteristic features of the regional white Anglo English in Chicago, and similar fronting of /u/ and /o/ is not present in Chicago Chicano English (Konopka and Pierrehumbert 2008). Additionally, English BOUGHT has been observed to be produced similarly to Spanish /a/ in Chicano English (Fought 2003, Santa Ana and Bayley 2004) and has been seen to be slightly raised and backed in Los Angeles Chicano English (Godinez and Maddieson 1985), seemingly in line with the shift that takes place in the coterritorial CVS. Therefore, it has been demonstrated that features of CVS are present in the speech of Chicano communities, thereby opening the possibility of the same features spreading to Bay Area Spanish. In order to further justify this possibility, the vowel system of Spanish will be presented in the next subsection, followed by an overview of the literature regarding the influence that sounds in two languages may have upon one another within the phonetic system of a bilingual speaker.

1.4. The Spanish vowel system

The Spanish vowel system consists of five vowels, /a/, /e/, /i/, /o/, and /u/, where each vowel's position in F1/F2 space defines the outer points of the language's vowel system. Compared with corresponding vowels in General American English, Spanish /u/ and /o/ are generally more backed (lower F2) than English BOOT and BOAT, and Spanish /a/ is more backed (lower F2) than English BOUGHT. Whereas Spanish /i/ and English BEET are fairly similar across measures of F1 and F2, Spanish /e/ is not usually associated to English BAIT as this English vowel is usually produced as the diphthong /eɪ/ rather than the monophthong /e/ (Ladefoged and Johnson 2015, p. 98).

1.5. Bidirectional contact effects between L1 and L2

The L1 and L2 are not two separate phonetic systems, rather they co-exist in a mutual phonetic space and can influence one another (Flege 1995, 2002, Flege and Bohn 2021). If an L2 vowel category is perceived to be equivalent to an L1 vowel category, despite auditorily detectable differences, the formation of a new category can be blocked (Flege 1995, Flege and Bohn 2021). Blockage of category formation results in assimilation, by

which the speaker will produce a composite L1-L2 category (Evans and Iverson 2007, Kendall and Fridland 2012). The quality and quantity of input the bilingual receives in their lifetime, in addition to other factors, will determine whether the composite category more closely resembles the L1 or the L2 category (Flege 2002, Yeni-Komshian, Flege, and Liu 2000). In addition to assimilation, additional features of the L1 vowel system such as stress patterns or durational effects, may be applied to the L2 vowel space and vice versa (Amengual 2012, Mack 1982). Therefore, in productions of bilingual Spanish-English speakers in California that receive input from CVS-affected English (whether white Anglo English or Chicano English), L1 and L2 vowel categories perceived as equivalent may assimilate, namely /i/, /u/, /o/, and /a/, in addition to features of the vowel space as a whole, yielding overall compression of the vowel space area and reduced vowel dispersion.

This study provides empirical data that addresses the gap in the literature regarding the influence of American English regional sound change on Spanish as a contact language. Namely, the study observes whether or not vowel space compression is a feature of Bay Area Spanish, a variety in contact with California English, due to the bidirectional relationship between the L1 and L2. I hypothesize that the result of assimilation and phonetic interference will yield a vowel space in Bay Area Spanish that is more compressed than that of Mexican Spanish produced by monolingual speakers. I additionally hypothesize that this trend will vary according to age, as in California English, and relating to a potential change in progress, and according to style, where a compressed vowel space will be more associated with younger speakers than with older speakers and with less-monitored speech styles.

2. Methods

The production data for the present study come from two corpora. The Bay Area Spanish data come from the Corpus of Bay Area Spanish (CBAS, Davidson 2016) and the Mexican Spanish data are a dummy set generated by the author for the present study from the means and standard deviations of Mexican Spanish vowels reported in Grijalva, Piccinini, and Arvaniti (2013). Each corpus will be individually described in the following two subsections.

2.1. CBAS

CBAS consists of speakers that self-identify as speakers of English and Spanish and that have lived in the Bay Area for at least 18 years. The experimental procedure for the corpus consists of a sociodemographic questionnaire, a sociolinguistic interview in Spanish, a word list task in English, a word list task in Spanish, and a translation task from Spanish to English. Data from both the Spanish word list task and the Spanish sociolinguistic interview will be analyzed in the present study. In the word list task, participants were asked to read aloud from a list of isolated words and phrases containing all five Spanish vowels in stressed and unstressed position. The following information was collected from the demographic questionnaires: gender, age, place of birth, self-reported proficiency in each language, relative contact with each language on a daily basis, and age of acquisition of each language. Participants are binned according to age (18-25 years and 40+ years). The sociodemographic profile of each age group is shown in Table (1).

Table 1. Self-reported sociodemographic information across age groups for selected speakers in CBAS, mean responses are out of 1.0 (unless otherwise indicated) and are listed with standard deviations in parentheses.

	Younger (18-25 years)	Older (40+ years)
<i>Speaking Proficiency</i>		
Spanish	0.83 (0.17)	0.61 (0.25)
English	0.95 (0.10)	0.83 (0.28)
<i>Listening Proficiency</i>		
Spanish	0.95 (0.10)	0.83 (0.17)
English	1.00 (0.00)	0.95 (0.10)
<i>Cultural Identity</i>		
Spanish	1.00 (0.00)	0.83 (0.00)
English	0.83 (0.23)	0.78 (0.25)
<i>Age Learned (years)</i>		
Spanish	Birth (0.0)	Birth (0.0)
English	5.3 (2.3)	8.0 (9.8)
<i>Daily Use with Friends</i>		
Spanish	0.17 (0.29)	0.27 (0.46)
English	0.83 (0.29)	0.73 (0.46)
<i>Daily Use with Family</i>		
Spanish	0.50 (0.50)	0.57 (0.51)
English	0.50 (0.50)	0.43 (0.51)
<i>Daily Use at Work</i>		
Spanish	0.40 (0.10)	0.20 (0.26)
English	0.60 (0.10)	0.80 (0.26)

Out of the six speakers, three were categorized as younger (male: n=1; female: n=2) and three were categorized as older (male: n=1; female: n=2). Data for the remaining speakers in CBAS were not included as these speakers would be classified as L1-English L2-Spanish speakers. Though the heritage variety of some speakers in CBAS is distinct from Mexican Spanish (i.e., Salvadoran Spanish, Nicaraguan Spanish, etc.), the vowel space across varieties of monolingual Spanish have previously been classified as fairly stable and invariable (Hualde 2013, Lipski 2019, Tomás 1977). Accordingly, Mexican Spanish was selected as the variety of Spanish for comparison due to the large influence of this variety in the Bay Area (see Section 1.1) and the Mexican Spanish data are described in the following subsection.

2.2. Mexican Spanish

Mexican Spanish was selected as the variety of Spanish for the monolingual control set due to the large influence of this variety in the Bay Area (see Section 1.1). The data for Mexican Spanish come from means and standard deviations of vowel formants reported in Grijalva, Piccinini, and Arvaniti (2013). Grijalva and colleagues recorded five female speakers attending the Universidad Autónoma de Baja California (Mexico) in an elicitation task. Speakers' ages ranged from 19 to 23 with an average of 21 years of age. Each speaker was given a list of test words embedded in carrier phrases in the form of:

Repite ____ con cuidado ‘Say ____ carefully’. Each word was bisyllabic and the vowel of interest was stressed and always in the first syllable. Six recorded repetitions of the list by each speaker resulted in a total of 30 tokens of each Spanish vowel. As data from only female speakers was reported, F1 and F2 measures were not normalized and formant means and standard deviations were reported in raw Hertz. For the purpose of the present study, the author generated a dummy set of 50 instances of each Spanish vowel for 4 speakers (50 repetitions x 4 speakers x 5 vowels = 1,000 instances total), so that the mean and standard deviation for F1 and F2 of each vowel matched those reported in Grijalva et al. (2013).

2.3. Analysis

For the comparison of Bay Area Spanish with Mexican Spanish, data from the word list task in CBAS was compared to the dummy set of Mexican Spanish vowel productions. As the dummy data set contains raw formant values in Hertz from stressed Spanish vowels produced by female speakers in read speech, only stressed vowels from CBAS word list productions by the four female speakers were included in the comparison. Word list productions were transcribed in ELAN (ELAN 2019) and the resulting TextGrid files were submitted to Montreal Forced Aligner for phone-by-phone alignment (McAuliffe et al. 2017). A script (Lennes 2003) was used in Praat (Boersma and Weenink 2018) to extract measurements of F1, F2, and F3 taken at the midpoint of each vowel in order to minimize the effects of co-articulation on formant frequencies. Vowels that were produced in diphthongs and vowels with outlier formants were removed. Outlier productions were defined as having either F1, F2, or F3 measurements that fell outside of the range of boxplots (Dodsworth and Kohn 2012), where the extrema are defined as:

minimum: 1st quartile - (1.5 x Interquartile Range)
 maximum: 3rd quartile + (1.5 x Interquartile Range)

Each vowel production was coded for stress using the *Syltippy* library (nur-ag 2018) in Python and unstressed vowels were omitted from analysis. The formant frequencies from the remaining vowels produced in the CBAS word list task were not normalized so that a comparison could be made between CBAS female productions (n = 490) and Mexican Spanish female productions (n = 1,000).

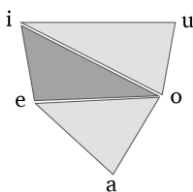
The second analysis was limited to the CBAS data and included both word list and interview data from all six speakers (male: n = 2, female: n = 4). The same procedure detailed above was carried out, except that both stressed and unstressed vowels were included in this second analysis. For the interview data, six-minute portions from the middle of each interview were transcribed and formant measurements were extracted using the same methodology as above. After vowels in diphthongs and vowels with outlier formants were removed, the formant measurements from vowel productions from the word list data (n = 2,006) and the interview data (n = 4,448) were then normalized by ΔF normalization (Johnson 2020).

Following the observation that the California vowel space is undergoing a change in progress towards overall compression (D’Onofrio, Pratt, & Van Hofwegen 2019), measures of vowel space area and vowel dispersion were obtained by speaker, using the methodology from the cited study. For the comparison of the female productions in CBAS and the female productions in the monolingual Mexican Spanish

control, raw formant values were used in the calculations. For the analysis within CBAS, normalized formant values were used. Vowel space area is defined as the Euclidean area of the polygon formed by the outer points of the vowel system in F1/F2 space (Bradlow, Torreta, & Pisoni 1996). As Spanish only has five vowels, all of which are outer points of the vowel system, all five vowels are used in the calculation of vowel space area. The Euclidean area is calculated with Heron's method (D'Onofrio, Pratt, & Van Hofwegen 2019), in which the polygon is divided into three component triangles (as shown in Figure 3) and the areas of the triangles are summed. Vowel dispersion was calculated as the average distance between the centroid of the speaker's vowel space and the centroid of each vowel category (D'Onofrio, Pratt, & Van Hofwegen 2019).

For the comparison with the monolingual control, measures of area and dispersion by speaker were submitted to separate linear fixed effects regression models in R (R Core Team 2018), with VARIETY (Bay Area Spanish, Mexican Spanish) as the sole main effect. For the analysis within CBAS, measures of area and dispersion by speaker were submitted to separate linear mixed effects regression models with *lmerTest* (Kuznetsova et al., 2019), with AGE (younger, older) and STYLE (word list, interview) as an interaction term and SPEAKER as a random intercept. For all regression models, the *r2glmm* package (Jaeger, 2017) was used to calculate marginal R-squared (R^2) effect sizes.

Figure 3. Division of vowel space into three triangles for vowel space area calculation.



3. Results

According to the output of the linear fixed effects regression model in Table (2), VARIETY is approaching significance ($\beta = 125399$, $R^2 = 0.41$, $p = 0.087$) as a predictor of vowel space area, where Mexican Spanish has a larger vowel space area than Bay Area Spanish. The output of the linear fixed effects regression model predicting vowel dispersion across Bay Area Spanish and Mexican Spanish productions in Table (3) reveals a significant compression of the vowel space in Bay Area Spanish as compared to Mexican Spanish, where dispersion in Mexican Spanish is significantly larger than that of the speakers of Bay Area Spanish ($\beta = 188.87$, $R^2 = 0.723$, $p < 0.01$). Figure (4) shows vowel dispersion across the two corpora, where differences in dispersion appear to predominately result from differences in productions of /i/ and /u/ along F2.

Table 2. Regression coefficients for the fixed linear effects model predicting vowel space area. The Intercept coefficient is vowel space area in Bay Area Spanish and the estimates are measures of Hz squared.

	Estimate	Std. Error	t-value	p-value	
<i>(Intercept)</i>	245032	43383	5.648	0.00132	**
<i>Mex. Spanish</i>	125399	61352	2.044	0.08697	.

Figure 4. Spanish vowel space dispersion in Bay Area Spanish and Mexican Spanish.

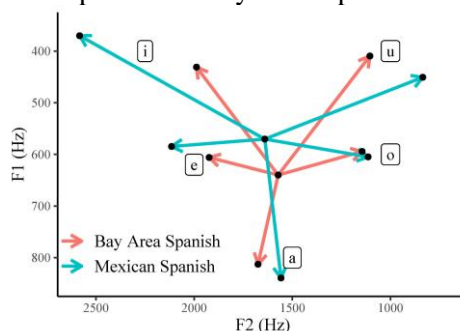


Table 3. Regression coefficients for the fixed linear effects model predicting vowel dispersion. The Intercept coefficient is vowel dispersion in Bay Area Spanish and the estimates are measures in Hz.

	Estimate	Std. Error	t-value	p-value	
<i>(Intercept)</i>	423.66	33.71	12.570	1.55e-05	***
<i>Mex. Spanish</i>	188.87	47.67	3.962	0.00743	**

Within CBAS, vowel productions were analyzed according to STYLE (more-monitored speech: word list task, less-monitored speech: sociolinguistic interview) and AGE (older and younger). The output of the linear mixed effects regression model predicting vowel space area in Table (4) indicates that both STYLE and AGE are significant predictors of vowel space area. In less-monitored speech (sociolinguistic interview) and productions by younger speakers, the vowel space area is reduced (STYLE: $\beta = -0.049$, $R^2 = 0.209$, $p < 0.05$; AGE: $\beta = -0.081$, $R^2 = 0.423$, $p < 0.05$). In Figure (5), vowel space area is visualized, according to age and style. For older speakers, it appears that back vowels (/u/, /o/, and /a/) are more fronted in less-monitored speech, which contributes to a more compressed vowel space area. Among younger speakers, the correlation between vowel space area and individual vowel movement is not as clear but is suggestive of slight centralization across F2 of mid vowels (/e/ and /o/) that may contribute to smaller vowel space area in the less-monitored speech style. Additionally, younger speakers appear to differ from older speakers mainly in the production of the high vowels and mid vowels (/i/, /u/, /e/, and /o/), where younger speakers produce high vowels lower (higher F1) and produce front high and mid vowels (/i/ and /e/) more backed (lower F2). Turning to vowel dispersion, the output of the linear mixed effects regression model predicting dispersion in Table (5) indicates that AGE is a significant correlate of dispersion, where younger speakers have reduced dispersion ($\beta = -0.099$, $R^2 = 0.348$, $p < 0.05$). This relationship is visualized in Figure (5).

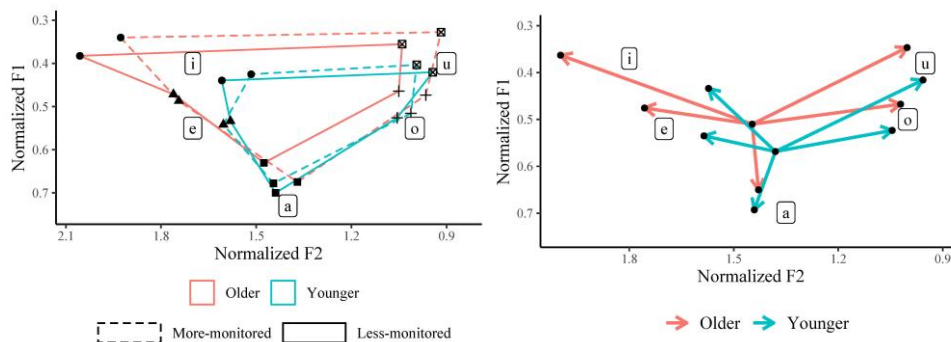
Table 4. Regression coefficients for the mixed linear effects model predicting vowel space area. The Intercept coefficient is vowel space area from word list productions by older speakers and the estimates are squared measures of normalized frequency.

	Estimate	Std. Error	df	t-value	p-value	
<i>(Intercept)</i>	0.19500	0.02023	9.21999	9.637	4.08e-06	***
<i>Interview</i>	-0.04878	0.01921	6.00000	-2.539	0.0442	*
<i>Younger</i>	-0.08128	0.02861	9.21999	-2.841	0.0189	*
<i>Interview: Younger</i>	0.03501	0.02717	6.00000	1.289	0.2450	

Table 5. Regression coefficients for the mixed linear effects model predicting vowel dispersion. The Intercept coefficient is vowel dispersion from word list productions by older speakers and the estimate are measures of normalized frequency.

	Estimate	Std. Error	df	t-value	p-value	
<i>(Intercept)</i>	0.416012	0.028908	9.545884	14.391	8.65e-08	***
<i>Interview</i>	-0.032794	0.028704	6.000005	-1.142	0.2968	
<i>Younger</i>	-0.099073	0.040882	9.545884	-2.423	0.0369	*
<i>Interview: Younger</i>	-0.001205	0.040593	6.000005	-0.030	0.9773	

Figure 5. *Left:* Bay Area Spanish vowel space across age and style; *Right:* Bay Area Spanish vowel space dispersion across age.



4. Discussion

The results of this study indicate that the size of the Bay Area Spanish vowel space is reduced compared to that of Mexican Spanish in terms of vowel dispersion. Whereas previous studies of Spanish vowel variation focus on vowel formants, these results advocate for the use of more holistic measures within Spanish dialectology, such as vowel space area and vowel dispersion. It is possible that the use of such holistic measures may reveal further dialectal variation that has not been previously attested, even among Spanish dialects not in situations of language contact. Though the present investigation complements a small set of studies which demonstrate variability in the Spanish vowel space in situations of bilingualism and language contact (e.g., Alvord and Rogers 2014, O’Rourke 2010, Willis 2005), this study invites additional examinations of vowel space variability across all varieties of Spanish. Furthermore, in this case the compressed vowel space of Bay Area Spanish is consistent with contact effects from California English, a regional variety. This calls for greater awareness in

production and perception studies of speakers' language backgrounds and how certain regional features may impact outcomes of language contact.

The compression of the Bay Area Spanish vowel space is mediated by age and style, where younger speakers and less monitored speech styles are associated with more compressed vowel spaces. The effect of age on the vowel space size is in line with the data on California English gathered by D'Onofrio, Pratt, and Van Hofwegen (2019), where vowel space compression in California English is a change in progress. Younger speakers may also have more contact with and familiarity with the personae of the "Valley Girl" and the "Surfer Dude", making the semiotic resources of the lowered jaw setting and centralization of vowels more salient for these speakers. The effect of style can be attributed to the reduced attention paid to variants in speech, where more non-standard variants are favored in less monitored speech. Additionally, vowel space area has often been correlated with intelligibility, where more dispersed vowel categories yield greater intelligibility (Bradlow, Torretta, and Pisoni 1996, McCloy, Wright, and Souza 2012). Thus, it is to be expected that vowel space size should be correlated with speech style in this way. A potential limitation to the methodology is that duration is not taken into account. It is possible that speech rates across corpora and individual participants also have an effect on vowel space area. Another possible source for the observed compression of the vowel space could be attributed to findings that Heritage Spanish speakers in the United States demonstrate vowel reduction in unstressed vowels, yielding centralization (Ronquest 2013). To comment further on this possibility, stress should be included as a factor in future analyses of both Mexican Spanish and Bay Area Spanish productions. Future work should take into account productions of monolingual speakers of other varieties of Spanish (e.g., Nicaraguan, Peruvian) to better understand the substrates at work in Bay Area Spanish and of these participants, in particular.

It is important to examine how individuals use regional identity markers in their speech, and to broaden the scope of study beyond white Anglo English speakers. Accordingly, studies of regional sound change should strive to include speakers with other ethnic identifications and speakers of other languages spoken in the United States. Within the field of American dialectology, regional varieties are most often defined as the English spoken by white Anglo speakers. As such, little work has been done to study regional variation in English among other populations (i.e., Chicano speakers, African American English speakers) or regional variation in other languages spoken in the United States (such as Spanish). This tendency fortifies the Anglo-centric emphasis in the discipline, furthering the narrative that only white Anglo speakers of English can index regional identity through their speech. The present study addresses this asymmetry by examining Bay Area Spanish as a regional variety in the United States, with features that may be influenced by the coterritorial patterns of sound change in English, namely CVS. The comparison of the vowel space of Bay Area Spanish with that of Mexican Spanish suggests that the compressed vowel space seen in Bay Area Spanish may be a feature unique to this area of the United States. Future studies of Spanish in the United States may reveal Bay Area Spanish to have a more compressed vowel space than the Spanish of other regions, strengthening the possibility that a compressed vowel space is a regional feature of California in Spanish as well as English.

As D'Onofrio et al. (2019) conclude in their study of California English, the features that distinguish California English from General American English and from

which social meaning is derived may not be vowel-specific (i.e., BOOT-fronting). Alternatively, social meaning may be associated with the vowel space as a whole, characterized through holistic measures, such as vowel space area and vowel dispersion, and derived from a lowered jaw setting and fronted lingual position. Due to the high degree of bilingualism of Spanish and English in the Bay Area, it is conceivable that speakers may transfer these articulatory settings from one language to another, thereby producing similar modulations of the vowel space in both languages. Although the present study does not include an analysis of English data that would provide direct evidence for this kind of transfer, the present findings are certainly consistent with this effect⁴. Should vowel space compression be a marker of California identity, the findings suggest that this marker of identity could be present in Spanish, as well as English.

Future studies should continue to explore linguistic factors that influence variation in Bay Area Spanish vowels, such as stress, duration, and linguistic environment. Laboratory studies which analyze the articulation of the jaw during speech are also necessary in order to make the definitive claim that the Bay Area Spanish vowel space is compressing due to a change in articulatory setting, as has been observed in California English. A broader comparative study that includes male as well as female speakers is also needed to further comment on dialectal differences between varieties of Spanish. Additionally, future studies should explore how social factors, such as gender, social network, and regional orientation, may be linked to the presence of regional sound changes in contact languages, which ultimately underscores the mechanisms of bidirectionality in language contact settings.

5. Conclusions

The data presented here highlight the importance of a holistic view of the Spanish vowel system in studies of dialectal variation. These results additionally broaden the scope of regional sound change studies by removing focus from white Anglo speakers of English. The present study suggests that regional sound changes not only may involve speakers of languages other than English but may also influence production within these contact languages. The results advocate for a new, more inclusive focus of American dialectology, where regional varieties can be defined without reference to white Anglo ethnic identity and without sole reference to the English language. The findings pave the way for future studies of speakers of Bay Area Spanish that incorporate English data and can examine the bilingual phonetic repertoire in more depth.

Acknowledgments

I would like to thank Dr. Justin Davidson, Dr. Isaac Bleaman, and Ernesto Gutiérrez Topete of UC Berkeley, as well as my audience at Going Romance 34 and one anonymous reviewer for their feedback on this research and manuscript. All remaining errors are my own.

⁴ As data collection for the only existing corpus of Bay Area Spanish has been on hold due to COVID-19, and as the corpus is limited in its current form, English data is not analyzed in this study, nor is age included as a factor of analysis in the comparison with Mexican Spanish. Future work will use this data to further contribute to this topic.

References

- Alvord, Scott M & Brandon MA Rogers. 2014. Miami-Cuban Spanish vowels in contact. *Sociolinguistic Studies* 8(1): 139-170. <https://doi.org/10.1558/sols.v8i1.139>.
- Amengual, Mark. 2012. Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism* 15(3): 517-530. <https://doi.org/10.1017/s1366728911000460>.
- Bay Area Census. 2010. San Francisco Bay Area. <https://www.bayareacensus.ca.gov/bayarea.htm>.
- Boersma, Paul & David Weenink. 2018. Praat: Doing phonetics by computer [Computer program]. Version 6.0. 37.
- Bradlow, Ann R, Gina M Torretta, & David B Pisoni. 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20(3-4): 255–272. [https://doi.org/10.1016/S0167-6393\(96\)00063-5](https://doi.org/10.1016/S0167-6393(96)00063-5).
- Cheng, Andrew. 2016. A Survey of English Vowel Spaces of Asian American Californians. *UC Berkeley PhonLab Annual Report* 12(1): 348-384.
- D’Onofrio, Annette, Teresa Pratt, & Janneke Van Hofwegen. 2019. Compression in the California Vowel Shift: Tracking generational sound change in California’s Central Valley. *Language Variation and Change* 31(2): 193–217. <https://doi.org/10.1017/s0954394519000085>.
- Davidson, Justin. 2016. Corpus of Bay Area Spanish.
- Davidson, Justin. 2019. La [v]ariabilidad sociofonética en el español de California: Social and Linguistic Underpinnings of the Labiodentalization of /b/ [Conference presentation]. 2019 Hispanic Linguistics Symposium, University of Texas at El Paso.
- Dodsworth, Robin & Mary Kohn. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change* 24(2): 221-245. <https://doi.org/10.1017/S0954394512000105>.
- Eckert, Penelope. 2008. Where do ethnolects stop? *International Journal of Bilingualism* 12(1-2): 25–42. <https://doi.org/10.1177/13670069080120010301>.
- ELAN. 2019. ELAN (Version 5.8) [Computer software]. <https://archive.mpi.nl/tla/elan>.
- Evans, Bronwen G. & Paul Iverson. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the*

- Acoustical Society of America* 121(6): 3814–3826.
<https://doi.org/10.1121/1.2722209>.
- Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research* 92: 233–277.
- Flege, James E. 2002. Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, and A. Rohde (eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Wissenschaftlicher Verlag Trier.
- Flege, James E. & Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). *Second Language Speech Learning: Theoretical and Empirical Progress*. 3–83.
- Fought, Carmen. 1997. The English and Spanish of young adult Chicanos. Ph.D. thesis, University of Pennsylvania, PA.
- Fought, Carmen. 1999. A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics* 3(1): 5–23.
<https://doi.org/10.1111/1467-9481.t01-1-00060>.
- Fought, Carmen. 2003. *Chicano English in Context*. London: Palgrave Macmillan.
<https://doi.org/10.1057/9780230510012>.
- Gick, Bryan, Ian Wilson, Karsten Koch, & Clare Cook. 2004. Language-specific articulatory settings: Evidence from interutterance rest position. *Phonetica* 61(4): 220–233. <https://doi.org/10.1159/000084159>.
- Godinez, Manuel & Ian Maddieson. 1985. Vowel differences between Chicano and general Californian English? *International Journal of the Sociology of Language* 53: 43–58. <https://doi.org/10.1515/ijsl.1985.53.43>.
- Grijalva, Criccely, Page E Piccinini, & Amalia Arvaniti. 2013. The vowel spaces of Southern Californian English and Mexican Spanish as produced by monolinguals and bilinguals. *Proceedings of Meetings on Acoustics ICA2013* 19(1): 060088. <https://doi.org/10.1121/1.4800752>.
- Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in San Francisco English. Ph.D. thesis, Stanford University, CA.
- Hall-Lew, Lauren, Amanda Cardoso, Yova Kemenchedjieva, Kieran Wilson, Rauridh Purse, & Julie Saigusa. 2015. San Francisco English and the California vowel shift. *Proceedings of the 18th International Conference of the Phonetic Sciences*.
- Hinton, Leanne, Birch Moonwomon, Sue Bremner, Herb Luthin, Mary Van Clay, Jean Lerner, & Hazel Corcoran. 1987. It's not just the Valley Girls: A study of

- California English. *Annual Meeting of the Berkeley Linguistics Society* 13: 117–128. <https://doi.org/10.3765/bls.v13i0.1811>.
- Hualde, José Ignacio. 2013. *Los sonidos del español: Spanish Language edition*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511719943>.
- Jaeger, Byron. 2017. *r2glmm: Computes R Squared for Mixed (Multilevel) Models*. R package version 0.1.2. <https://CRAN.R-project.org/package=r2glmm>.
- Johnson, Keith. 2020. The ΔF method of vocal tract length normalization for vowels. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1). <https://doi.org/10.5334/labphon.196>.
- Kendall, Tyler & Valerie Fridland. 2012. Variation in perception and production of mid front vowels in the US Southern Vowel Shift. *Journal of Phonetics* 40(2): 289–306. <https://doi.org/10.1016/j.wocn.2011.12.002>.
- Kim, Ji Young & Nicole Wong. 2020. (Divergent) Participation in the California Vowel Shift by Korean Americans in Southern California. *Languages* 5: 53. <https://doi.org/10.3390/languages5040053>.
- Konopka, Kenneth & Janet Pierrehumbert. 2008. Vowels in contact: Mexican heritage English in Chicago. *Proceedings from The Sixteenth Annual Symposium about Language and Society-Austin* 52: 94-103.
- Kuznetsova, Alexandra, Per B. Brockhoff, & Rune H. B. Christensen. 2017. lmerTestPackage: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13): 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Ladefoged, Peter & Keith Johnson. 2015. *A Course in Phonetics*. Cengage Learning.
- Lenes, Mietta. 2003. Collect formant data from files. Praat script.
- Lipski, John M. 2019. Spanish phonological variation. In S. Colina and F. Martínez-Gil (eds.), *The Routledge Handbook of Spanish Phonology*, 455-469. <https://doi.org/10.4324/9781315228112-26>.
- Mack, Molly. 1982. Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America* 71(1): 173–178. <https://doi.org/10.1121/1.387344>.
- McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, & Morgan Sonderegger. 2017. Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. *Interspeech 2017*, 498–502. <https://doi.org/10.21437/Interspeech.2017-1386>.
- McCloy, Daniel, Richard Wright, & Pamela Souza. 2012. Modeling intrinsic intelligibility variation: Vowel-space size and structure. *Proceedings of Meetings on Acoustics* 164ASA 18(1): 060007. <https://doi.org/10.1121/1.4870070>.

-
- nur-ag. 2018. Syltippy: Spanish word syllabization in Python. <https://github.com/nur-ag/syltippy>.
- O'Rourke, Erin. 2010. Dialect differences and the bilingual vowel space in Peruvian Spanish. *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*. Cascadilla Proceedings Project Somerville, MA: 20–30.
- Pew Hispanic Center. 2002. *2002 National Survey of Latinos*. <https://www.pewresearch.org/hispanic/2004/03/19/latinos-in-california-texas-new-york-florida-and-new-jersey/>
- Podesva, Robert, Annette D'Onofrio, Janneke Van Hofwegen, & Seung Kyung Kim. 2015. Country ideology and the California vowel shift. *Language Variation and Change* 27(2): 157–186. <https://doi.org/10.1017/s095439451500006x>.
- Pratt, Teresa & Annette D'Onofrio. 2017. Jaw setting and the California Vowel Shift in parodic performance. *Language in Society* 46(3): 283–312. <https://doi.org/10.1017/S0047404517000227>.
- R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. <https://www.Rproject.org>.
- Ramanarayanan, Vikram, Louis Goldstein, Dani Byrd, & Shrikanth Narayanan. 2013. An investigation of articulatory setting using realtime magnetic resonance imaging. *Journal of the Acoustical Society of America* 134(1): 510–519. <https://doi.org/10.1121/1.4807639>.
- Roeder, Rebecca V. 2010. Northern Cities Mexican American English: Vowel production and perception. *American Speech* 85(2): 163–184. <https://doi.org/10.1215/00031283-2010-009>.
- Ronquest, Rebecca. 2013. An acoustic examination of unstressed vowel reduction in heritage Spanish. *Selected proceedings of the 15th Hispanic linguistics symposium*. Cascadilla Press, 151–171.
- Santa Ana, Otto & Robert Bayley. 2004. Chicano English: Phonology. In E. W. Schneider (ed.), *A Handbook of Varieties of English*. Vol. 2. Berlin, Germany: Mouton de Gruyter, pp. 219–238. <https://doi.org/10.1515/9783110208405.1.219>.
- Tomás, Navarro. 1977. *Manual de pronunciación española (19th edition)*. Madrid.
- United States Census Bureau. 2013. Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 Years and Over: 2009-2013. <https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>
- United States Census Bureau. 2019. *California Population Estimates*.

<https://www.census.gov/quickfacts/fact/table/CA/PST045219>. Retrieved Aug. 2021.

Willis, Erik W. 2005. An initial examination of Southwest Spanish vowels. *Southwest Journal of Linguistics* 24(1-2): 185–199.

Wilson, Ian & Bryan Gick. 2006. Articulatory settings of French and English monolinguals and bilinguals. *Journal of the Acoustical Society of America* 120(5): 3295-3296. <https://doi.org/10.1121/1.4777977>.

Wong, Amy & Lauren Hall-Lew. 2014. Regional variability and ethnic identity: Chinese Americans in New York City and San Francisco. *Language & Communication* 35: 27-42. <https://doi.org/10.1016/j.langcom.2013.11.003>.

Yeni-Komshian, Grace H, James E Flege, & Serena Liu. 2000. Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition* 3(2): 131-149. <https://doi.org/10.1017/s1366728900000225>