

American myths of linguistic assimilation: A sociolinguistic rebuttal

DANIEL ERKER^a AND RICARDO OTHEGUY^b

^a*Boston University, USA*

^b*The Graduate Center, CUNY, USA*

ABSTRACT

This study examines the behavior of 331 Spanish speakers, 269 immigrants to the United States and sixty-two native-born individuals, through questionnaires and sociolinguistic interviews. Results show that increased US life experience correlates with expanded use of English in both private and public domains of life. Additionally, greater use of English co-exists with maintenance of fine-grained patterns of structured linguistic variation in Spanish, such that US-born speakers demonstrate remarkable similarity to the immigrant generation in their usage of three variables: (i) subject pronoun presence vs. absence, (ii) grammatical subject position, and (iii) syllable-final /s/. The co-occurrence of increased use of English, on one hand, and intergenerational structural continuity in variable linguistic behavior in Spanish, on the other, challenges two misconceptions about Spanish in the United States: that (a) Spanish-speaking immigrants and their US-born children are unwilling or unable to learn English, and (b) regular use of English entails attrition and/or failed acquisition of Spanish. Neither of these views finds empirical support in our data. (Spanish in the United States, comparative variationist linguistics, subject personal pronouns, grammatical subject position, syllable final /s/, bilingualism)

INTRODUCTION

In 1919, ten years after his presidency ended, Theodore Roosevelt shared his thoughts on immigration in a letter to the American Defense Society: ‘In the first place we should insist that if the immigrant who comes here does in good faith become an American and assimilates himself to us, he shall be treated on an exact equality with every one else, for it is an outrage to discriminate against any such man because of creed or birthplace or origin. But this is predicated upon the man’s becoming in very fact an American and nothing but an American’. If there was any confusion about what becoming ‘nothing but an American’ entailed linguistically, another quote attributed to Roosevelt is unambiguous: ‘This is a nation—not a polyglot boarding house. There is not room in the country for any 50–50 American, nor can there be but one loyalty—to the Stars and Stripes’.

While Roosevelt's imperious patriotism may strike contemporary readers as out-moded, his sentiments are consistent with a national suspicion of multilingualism that is alive and well. Indeed, similar attitudes towards language, immigrants, and assimilation can be found in legislation recently proposed by United States senators Tom Cotton and David Purdue, whose *Raise Act* would revise US immigration policy to include English proficiency tests of potential immigrants. President Trump's decision to announce his support of the senators' bill in front of a portrait of Roosevelt imbued the endorsement, intentionally or not, with a sense of genuine historical continuity. It also underscored the fact that many such efforts to promote English amount to a rejection of other languages as legitimate forms of linguistic behavior in the United States.

At a debate among Republican presidential candidates in September 2015, then candidate Trump scolded Jeb Bush, former Governor of Florida, for using Spanish on the campaign trail. To cheers from the audience, Trump said, "This is a country where we speak English, not Spanish". Of course, Trump was wrong about this, since US citizens speak English and Spanish, as well as a host of other languages. But he was not wrong when qualifying his statement to the debate moderator: "I'm not the first one to say this... We've had many people over the years, for many, many years saying the same thing". The group of people 'saying the same thing' includes not only politicians, but also entrepreneurs and private citizens. On Amazon.com, for example, it costs just \$16.98 to buy a t-shirt printed with the US flag and the phrase 'Welcome to America, now speak English'. This sentiment was on full display in a recent incident in which an NYC attorney was captured, in a video that quickly went viral, berating workers for speaking Spanish in a cafeteria in Midtown Manhattan, among the places in America where one might least expect the use of a language other than English to provoke outrage.

These past and present manifestations of an American immigration narrative that sees successful assimilation as a linguistic zero-sum game resonate with scholarship on the relationship between languages and nation states: 'Seen as a community imagined through language, the nation presents itself as simultaneously open and closed' (Anderson 2006:146). It is open to those who would leave aside their ancestral ways of speaking and acquire those of their adopted home. It is closed to those who won't. This means that at the same time that recent arrivals to the United States are making choices about where to live, work, study, shop, bank, eat, and so on, they are also being asked to make a linguistic choice: acquire English and assimilate or maintain an immigrant language and be judged an assimilation failure.

As bilingual scholars intimately familiar with the US Latino community, we find these popularly held ideas to be at odds with our experience regarding attitudes about, and use of, language in the community. But we want to do more than rely on our experiences as a means for criticizing the popular zero-sum linguistic assimilation narrative, which allows only for successful immigrants who abandon their native languages or failed immigrants who refuse to do so. Instead, in this article

we decompose this narrative into a set of hypotheses that make testable predictions about the linguistic behavior of Latino immigrants and their US-born children.

HYPOTHESIS 1 (H1): Spanish-speaking immigrants and their US-born children maintain Spanish at the expense of English (and are therefore undesirable and probably unsuccessful). **HYPOTHESIS 2 (H2):** Spanish-speaking immigrants and their US-born children acquire English at the expense of Spanish (and are therefore desirable and very likely successful). Note that H1 and H2 exist in tension with each other. H1 claims that a group of people refuses to learn or use the new language (and are punished with rejection and failure), while H2 claims that a group strives to learn and use the new language at the cost of loss of knowledge of the ancestral language (but are rewarded with acceptance and success). Despite their differing stances on the linguistic behavior of the immigrant group, H1 and H2 are unified by the same underlying position, namely that bilingualism is impossible. According to both hypotheses, the gain of one language entails the loss of another.

Let us consider the predictions that follow from these hypotheses, imagining what the world would look like were they correct. H1 predicts that Spanish-speaking immigrants to the US are sequestering themselves in linguistic enclaves, living in a state of sustained intergenerational monolingualism. H2, in contrast, predicts intergenerational attrition of Spanish as speakers assimilate into the monolingual English majority. Given these predictions, what is the best way to assess their respective (in)accuracy in the real world?

At minimum, an appropriate empirical test of H1 and H2 would, for a representative sample of the population of interest, document speakers' habitual language use across various dimensions of life, a central aim in the research tradition of *language maintenance and shift* (Fishman 1964, 2003). Additionally, we would want to understand the nature of speakers' linguistic behavior across different contexts. In other words, beyond documenting the distribution of language choice across social domains—that is, in addition to answering Fishman's question of 'who speaks what language to whom and when?'—we also want a reliable way of assessing aspects of the linguistic knowledge that individuals bring to those interactions. In other words, we want to be able to answer the question of who speaks what language to whom and when AND HOW? This is essential to understanding whether and to what extent intergenerational shifts in habitual language use are accompanied by changes in linguistic knowledge. To us, the most compelling way of realizing this second component of an empirical test of H1 and H2 is to employ a *comparative variationist methodology* to analyze patterns of linguistic variation occurring in spontaneous speech (Sankoff 2002; Poplack, Walker, & Malcolmson 2006; Poplack & Levey 2010; Torres Cacoullos & Travis 2011).

Before describing our approach further, a word about why we are not choosing what might seem a more obvious path for exploring the hypotheses under consideration, that of the research tradition of *heritage linguistics* (Zapata, Sánchez, & Toribio 2005; Montrul 2008, 2009, 2013; Benmamoun, Montrul, & Polinsky 2010). Like us, scholars working in this paradigm have a central interest in the

linguistic knowledge and behavior of immigrants and their children. However, as we have argued elsewhere (Otheguy 2016), we see limitations with their approach. The problem lies in their proposal that US-born speakers of immigrant languages are in possession of *incompletely acquired* grammars, a claim we see as based on faulty reasoning and suboptimal data. Theoretically, the notion of incomplete acquisition leaves little room for the natural course of language change, which unfolds in an ongoing cycle of replication, variation, and selection. Rather than viewing linguistic differences between related groups as the output of this cycle, inter-group variability is instead seen from this perspective as evidence of the unsuccessful transmission of a grammar from one generation to the next. This strikes us as an especially unproductive way to frame the analysis of US-born users of immigrant languages. To expect them to present as linguistic replicas of the first immigrant generation endorses a narrow and unrealistic view of language acquisition—one that sees it as process by which copies of languages are uniformly deposited into the minds of language users. This logic is reflected in heritage linguistics's explicit theoretical interest in the category of *native speaker*, which we see as particularly unhelpful in testing H1 and H2.

These theoretical issues are related to methodological problems. Many studies claiming to have uncovered evidence of incomplete acquisition compare monolingual home-country experimental controls to US-born Spanish speakers. Not only do the latter regularly use English in addition to Spanish, differentiating them from the former, they also differ by typically having limited experience with Spanish in formal education (Menken & Kley 2010), the very environment in which speakers are likely to develop the meta-linguistic skills and awareness that often serve as diagnostics for (in)completeness of acquisition, for example, judgments of grammaticality, felicity, and logic (Bialystok 1986). When a methodology that guarantees group differences to emerge combines with a theoretical position biased towards interpreting differences as errors of acquisition, conclusions of incompleteness are all but certain, and, therefore, limited in what they can reveal.

A *comparative variationist* approach to H1 and H2 offers advantages precisely in the areas where the search for (in)completeness of acquisition is limited. In the comparative variationist paradigm, spontaneous use of language through speech, instead of meta-linguistic introspection or other measures designed to probe speaker intuitions, constitutes the primary linguistic data. Because US-born users of Spanish are more likely to utilize the language in speech rather than through reading or writing, it makes sense to focus on this type of linguistic behavior for intergenerational comparisons. In addition, because the comparative variationist approach is built on the tradition of *variationist sociolinguistics* (Labov 1972), it is attuned to the inherent variability of language use. It not only expects variation to arise between and within different groups, it also provides a precise way for comparing them. When similar patterns of structured variation arise in the speech of two (or more) people, we can infer that they share knowledge of the norms that guide linguistic interaction in a speech community. From this perspective, the treatment

of linguistic variables by different groups of speakers is viewed, not as a means for assessing the (in)completeness of their acquisition or the (non)nativeness of their relation to a linguistic system, but rather as a way for discerning whether and to what degree the members of the different groups constitute a single community of language users. This leads to the following formulation in the context of testing our hypotheses: To the extent that the immigrant generation and the US-born demonstrate similar sensitivity to the linguistic and social factors that give rise to structured variation in speech, they belong, to a greater or lesser degree, to the same speech community; if they do not share this sensitivity to linguistic and social factors, they don't, but have diverged into different communities. The second element of our investigation of the linguistic zero-sum narrative, then, involves testing whether US immigrants from Latin America and their children belong to the same or different speech communities.

We analyze the linguistic behavior of 331 Spanish speakers in New York City and Boston and find in our data no evidence to support either H1 or H2. With respect to H1, our results refute claims that the linguistic lives of Spanish-speaking immigrants and their US-born descendants are marked by sustained Spanish monolingualism. Instead, participants in the study are highly motivated to learn and use English; among first-generation immigrants, usage of English significantly increases with time spent in the US. Among the US-born, Spanish monolingualism is non-existent. With respect to H2, we find that expanded use of English in daily life does not come at the cost of the knowledge of Spanish that is assessed by patterns of structured variation. In our data, the prevailing trend is one of intergenerational linguistic continuity and stability. The facts suggest plainly that for our participants and for the variables we study, H1 and H2 are wrong. Even as the use of English in the community increases across generations, Spanish is maintained and the social and linguistic norms guiding its use are robustly conserved. Contrary to H2, in New York City and Boston, Latino immigrants and their children have neither stopped using Spanish nor given up the norms that govern variability, but rather form a single speech community marked by continuity of norms.

SPEAKERS, DATA, AND METHODOLOGY

Speakers

The study's data¹ come from the Otheguy-Zentella Corpus of Spanish in New York and the Boston Spanish Corpus (BSC). Together these corpora contain audio-recorded sociolinguistic interviews with 331 Spanish-speaking adults—140 New Yorkers and 191 Bostonians. The New York corpus, which for ease of presentation we refer to as the New York Spanish Corpus (NYSC), was collected between 2000 and 2005 by researchers at The CUNY Graduate Center under the direction of the second author and his collaborator, Ana Celia Zentella. The Boston data were collected between 2014 and 2017 by a team at Boston University led by the first

author. Interviews were conducted in Spanish by Spanish-speaking interviewers from the respective research teams. Interviews lasted roughly an hour and contained two parts. The first was an interview intended to produce spontaneous speech for analysis from the perspective of variationist sociolinguistics. Participants could talk about whatever they liked, though some topics dominated—for example, immigration narratives, family histories, and life in the US compared to Latin America. The second component of each interview was a questionnaire that asked participants to describe their social and linguistic backgrounds, behaviors, and attitudes.

Participants included 192 women and 139 men, ranging in age from eighteen to eighty years old, with a median and mean age of thirty and 33.9 years old, respectively. Participants' ages of arrival to the US ranged from zero to seventy years old, with a median of twenty-two and a mean of 19.9. Time spent living in the US ranged from one month to sixty-three years, with a median of thirteen years and a mean of 13.9. On the basis of participants' ages, ages of arrival, and years spent in the US, two other variables were generated, *Immigration Category* and *Percentage of Life in the US* (PLUS). Previous analysis of the corpora has found these measures capable of illuminating patterns obscured by conventional immigrant-generation categories (Otheguy & Zentella 2012; Erker & Bruso 2017). In particular, these measures make it possible to investigate whether the linguistic behavior of first-generation immigrants—foreign-born individuals who migrated to the US—varies with increased experience in the United States.

In the study, *Newcomers* were individuals who arrived in the US after their sixteenth birthday and who had spent less than six years in the US. Participants born in the US or who arrived before their third birthday were categorized as *USBorn*. All other participants were categorized as *Established Immigrants*. The study includes 108 Newcomers, 161 Established Immigrants, and sixty-two USBorn participants. PLUS was calculated by dividing years in the US by age, for example, a twenty-year-old who arrived to the US at age ten has a PLUS of fifty.

One reason for defining the participant categories as we have is that they provide tightly focused snapshots at the poles of US life experience, where evidence of potential empirical support for H1 and H2 is likeliest to emerge. The study's Newcomers arrived as linguistically mature adults and have relatively little life experience in a predominantly Anglophone society, a profile presumably favorable to H1. The USBorn, in contrast, were either literally born in the country or arrived well before linguistic maturation, and setting the maximum age of arrival for membership in this group at three years old greatly reduces the chances that these individuals received schooling in Spanish, arguably giving H2 a strong chance to be borne out. Another reason for adopting these participant categories is that they permit direct comparison across the two corpora. Creation of the NYSC preceded that of the BSC by roughly a decade. In the interim, extensive research was carried out on the NYC data using these participant categories, which we build on here. Further subdivision of our intermediate group, the Established Immigrants, would very likely be fruitful, especially if carried out with an eye towards individual

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

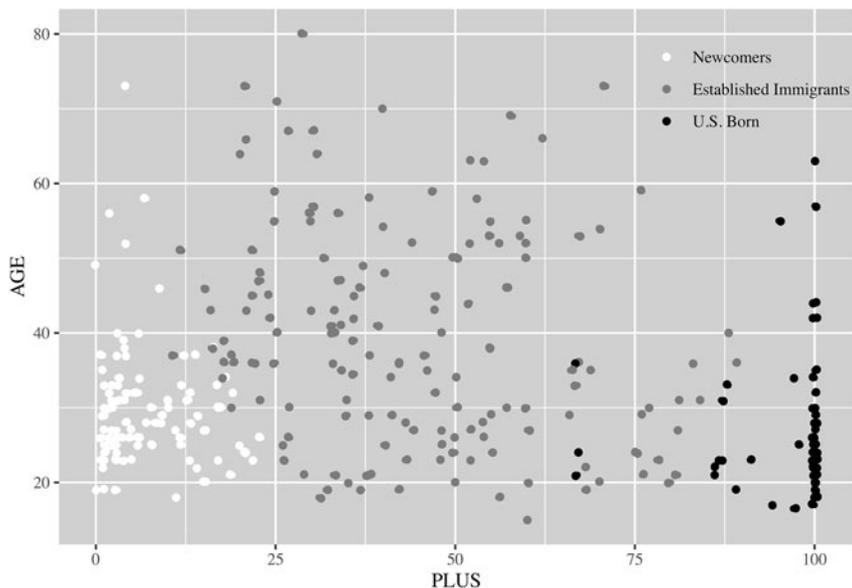


FIGURE 1. Study participants by age, PLUS, and Immigration Category.

differences in educational experience, though we leave this for future research. [Figure 1](#) plots participants according to age, Immigration Category, and PLUS.

In terms of national origin and ancestry, participants represented twenty locales. In order of decreasing numbers of participants these included Mexico (forty-seven speakers), Colombia (forty-two), Dominican Republic (forty-two), Puerto Rico (thirty-nine), Ecuador (thirty), Cuba (twenty-three), Venezuela (twenty-three), El Salvador (twenty), Peru (nineteen), Spain (eighteen), Guatemala (seven), Honduras (three), Argentina (two), Chile (two), Bolivia (one), Costa Rica (one), Nicaragua (one), Panama (one), Paraguay (one), and Uruguay (one). An additional eight participants reported mixed origins, with parents from some combination of the aforementioned locales. [Table 1](#) provides a national breakdown of participants by sex and Immigration Category.

Questionnaire

The questionnaire asked participants about their usage of Spanish and English with particular interlocutors and across specific domains of life. The present analysis focuses on answers to the following question.

- (1) ¿Cuál(es) idioma(s) habla [o hablaba] con su(s): papá, mamá, hermanos, hijos menores, hijos mayores, amigos, jefe, compañeros de trabajo, compañeros de escuela, esposa/o o novia/o?

TABLE 1. *Speakers in the study.*

Origin	# of participants	Sex		Immigration Category		
		Women	Men	Newcomers	Est. Imm.	USBorn
Mexico	47	28	19	18	18	11
Colombia	42	25	17	14	18	10
DominicanRep	42	20	22	9	27	6
PuertoRico	39	19	20	8	16	14
Ecuador	30	21	9	10	14	6
Cuba	23	13	10	7	13	4
Venezuela	23	15	8	13	9	1
ElSalvador	20	12	8	4	13	3
Peru	19	10	9	8	11	0
Spain	18	9	9	8	9	1
Mixed	8	5	3	3	1	4
Guatemala	7	3	4	1	6	0
Honduras	3	3	0	0	3	0
Argentina	2	2	0	2	0	0
Chile	2	2	0	2	0	0
Bolivia	1	1	0	1	0	0
Costa Rica	1	0	1	0	1	0
Nicaragua	1	1	0	0	0	1
Panama	1	1	0	0	1	0
Paraguay	1	1	0	0	1	0
Uruguay	1	1	0	0	0	1
Total	331	192	139	108	161	62

‘Which language(s) do [or did] you speak with your: father, mother, siblings, younger children, older children, friends, boss, coworkers, classmates, spouse or boy/girlfriend?’

Participants were asked to answer *español* ‘Spanish’, *inglés* ‘English’, or *ambos* ‘both’ to this question. Participants were also asked to rate their proficiency in Spanish and English, choosing from the options *excelente* ‘excellent’, *muy bien* ‘very good’, *pasable* ‘passable’, and *pobre* ‘poor’.

Though self-reported data on language usage and proficiency has known limitations (Bourhis 1983; Martin-Jones 1991; Marian, Blumenfeld, & Kaushanskaya 2007), the inclusion of the option ‘both’ for question (1)—which many participants selected for several interlocutors—was meant to capture the fact that in multilingual settings ‘language use does not necessarily fall into neat little patterns of complementary distribution across domains’ (Shameem 1998:87). Further confidence in the validity of responses regarding language usage comes from the fact that participants’ answers to question (1) are largely consistent with those that they gave for related items in the questionnaire that, for reasons of space, we do not discuss here, that is, descriptions of their habitual language use at home, at school, in

social activities, when reading, watching television, and listening to music. Finally, we do not use proficiency self-ratings as rigorous metrics for assessing participants' linguistic knowledge. Rather, we see them primarily as reflecting language attitudes and linguistic (in)security. The speech data under investigation are a much more direct source of insight into the nature of participants' linguistic knowledge.

Sociolinguistic variables

Speech data were examined with respect to three widely studied variable phenomena in Spanish: (i) presence versus absence of subject personal pronouns (e.g. *yo canto* ~ *canto*, both 'I sing'), (ii) position of grammatical subjects (e.g. *Juan llegó* vs. *Llegó Juan*, both 'John arrived'), and (iii) syllable-final or coda /s/ (e.g. the /s/ of *más* 'more', *mismo* 'same' and *hablas* 'you speak'), which is realized along a continuum of frication, ranging from a voiceless alveolar fricative, [s], to the complete absence of frication. While all 331 speakers are included in our analysis of habitual language use and self-rated proficiency, the sociolinguistic analyses are, due to their labor-intensive nature, carried out on subsets of speakers. In each case, they are based on large numbers of tokens as well as a sufficient number of speakers to explore patterns of variation across immigration categories. The analysis of subject pronouns was based on 69,933 finite verbs collected from interviews with 174 speakers (140 from the NYSC and thirty-four from the BSC). Analysis of variation in subject position included 5,600 grammatical subjects collected from the interviews of forty speakers (fourteen from the NYSC and twenty-six from the BSC). Finally, the analysis of coda /s/ was based on 7,054 tokens collected from sixty-two speakers (twenty from the NYSC and forty-two from the BSC). While there are, of course, other variable phenomena that could be put into service of our research questions, the three features we have chosen are attractive not only because they each have extensive research literatures behind them—meaning that we are well-positioned to identify factors that constrain their variability—but they also represent variation in Spanish across different grammatical domains, that is, phonology and (morpho)syntax.

TESTING H1: SPEAKERS MAINTAIN SPANISH AT THE EXPENSE OF ENGLISH

The claim that Spanish-speaking immigrants and their US-born children fail to linguistically assimilate has a long history in America (see Chavez 2008 for an overview). Within academia, H1 has its most well-known articulation in the work of Harvard political scientist Samuel Huntington. In 'The Hispanic challenge', Huntington (2004) argues that Latinos are unmotivated, unwilling, or simply unable to acquire English: 'Unlike past immigrant groups, Mexicans and other Latinos have not assimilated into mainstream U.S. culture, forming instead their own political and linguistic enclaves. By ignoring this question, Americans acquiesce to

their eventual transformation into two peoples with two cultures (Anglo and Hispanic) and two languages (English and Spanish)'. According to Huntington, the supposed political and linguistic separatism of US Latinos constitutes nothing short of an existential threat to the United States. The present study's data make it possible to conduct a strong test of Huntington's claims, especially in light of the fact that our participants are, demonstrably, routine users of Spanish. That is, a prerequisite of participating in the study was being willing and able to participate in an hour-long interview in Spanish. Our participants thus represent a group of people who could largely restrict their linguistic interactions to Spanish, should they desire to do so.

Usage of Spanish and English with specific interlocutors

As a first step, consider participants' reported language use with interlocutors. Table 2 lists rates of *Spanish Only*, *Both Spanish and English*, and *English Only* for each interlocutor in the questionnaire (in order of decreasing use of *Spanish Only*) for the sample overall.²

While exclusive Spanish use is higher with family members, comparing across Immigration Category reveals a steady decline in exclusive Spanish use even within the family. In Figure 2 the left column of the bar plot for each interlocutor represents Newcomers, the middle column corresponds to Established Immigrants, and the right represents USBorn participants. Each column corresponds to one hundred percent of the answers given by each group for that interlocutor. Note the diminished amount of exclusive Spanish use, represented in black, when scanning from left to right. Also note the simultaneous expansion of the use of English—either exclusively (in grey)—or in combination with Spanish (in white). Chi-square tests returned statistically significant differences in language use by Immigration Category for each familial interlocutor.

TABLE 2. *Reported language use across interlocutors.*

Interlocutor	Spanish Only		Both Span/Eng		English Only	
	Percent	N	Percent	N	Percent	N
Mother	88.0	272	8.7	27	3.2	10
Father	85.6	245	10.8	31	3.5	10
Siblings	57.3	172	27.0	81	15.6	47
(own) Older children	52.9	18	38.2	13	8.8	3
(own) Younger children	51.1	22	37.2	16	11.6	5
Significant other	40.5	82	21.7	44	37.6	76
Friends	23.7	76	56.3	181	19.9	64
Classmates	23.7	56	41.1	97	35.1	83
Coworkers	18.9	50	31.8	84	49.2	130
Boss	15.3	42	14.2	39	70.3	192

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

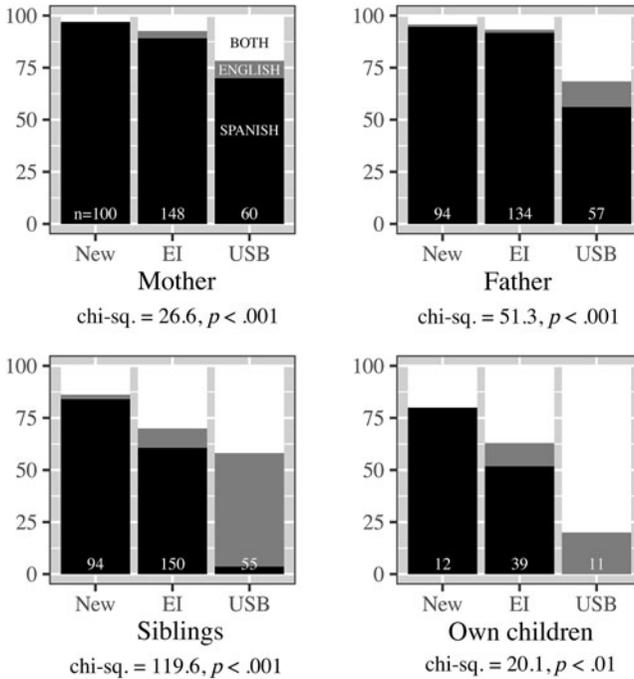


FIGURE 2. Reported language use with family members by Immigration Category. In figures throughout, New = Newcomers, EI = Established Immigrants, USB = USBorn.

An even more dramatic shift in language usage occurs outside of family relationships. In Figures 3 and 4, observe the expansion of exclusive English use (grey bars) across *significant other, friends, classmates, coworkers, and boss*. Once more, chi-square tests returned significant differences in language use with interlocutors across immigration categories, with one exception, *boss*, for which English Only was the strong preference in each group. Further exploration of the data showed that these trends cut across (i) *City of residence*—they emerge in NYC and Boston, (ii) *Sex*—men and women do not differ from each other, (iii) *Regional origin*—individuals of Mexican origin, for example, do not differ from those with origins in Peru, El Salvador, or the Dominican Republic, and so on, and (iv) *Socioeconomic status*—working, middle, and upper class individuals all participate in these trends.

Usage of Spanish and English with interlocutors overall

Now consider participants' reported use of Spanish and English with interlocutors in the aggregate along lines of PLUS and Immigration Category. Because some participants did not provide answers for all interlocutors (when, for example, they did

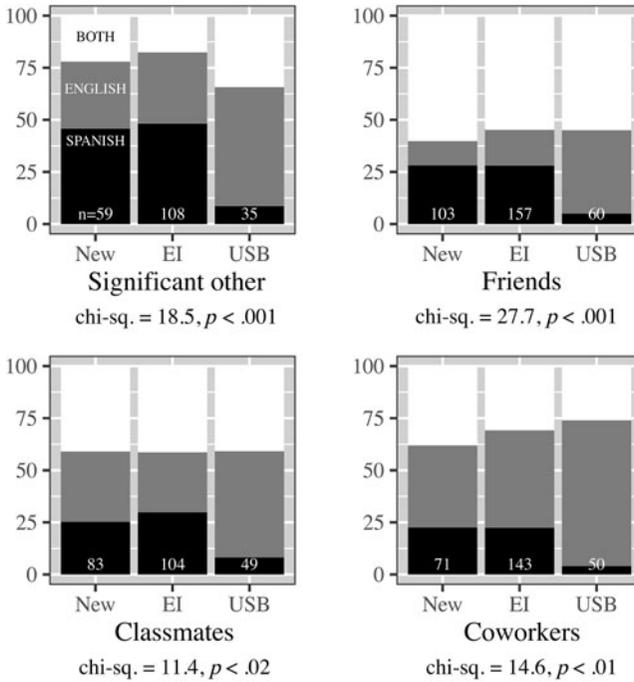


FIGURE 3. Reported language use with nonfamilial interlocutors by Immigration Category.

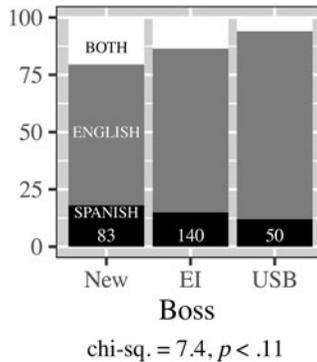


FIGURE 4. Reported language use with boss by Immigration Category.

not have siblings or a significant other), responses were converted to percentages. For example, speaker 37VZ gave answers for seven interlocutors (no answer for children or coworkers). She answered *español* ‘Spanish’ for four (*father*,

mother, siblings, and boyfriend), *ambos* ‘both’ for one (*friends*), and *inglés* ‘English’ for two (*boss* and *classmates*). These correspond to fifty-seven, fourteen, and twenty-nine percent, respectively, for the variables *Percentage of interlocutors* with whom the participant speaks (a) *Spanish only*, (b) *Both Spanish and English*, and (c) *English only*.

Clear relationships emerged between the aggregate measures and US life experience. Beginning with *Spanish Only*, results showed that as PLUS increases, exclusive Spanish usage significantly decreases. A parallel result emerges across immigration categories: rates of exclusive Spanish use were highest among Newcomers, lower for Established Immigrants, and lowest for USBorn participants. With respect to the measures *Both English and Spanish* and *English Only*, results showed significant increases with PLUS as well as significant differences between immigration categories: rates for both measures were lowest among Newcomers, higher among Established Immigrants, and highest for USBorn participants. Figures 5–7 summarize the results, providing the Pearson *r* coefficient for each correlation (below the left square of each figure) as well as the F-statistic associated with respective ANOVAs and groups means for the immigrant categories (the right square). The number of participants, which is the same across the three figures, is given in Figure 5.

Summary of results for H1

Among participants, longer life experience in the US corresponds to a significant decline in exclusive Spanish usage and a significant increase in exclusive English usage. These trends are not restricted to life outside the home, with results revealing a steady decline in exclusive Spanish use also within family relationships. Our participants, therefore, are not failing or refusing to learn and use

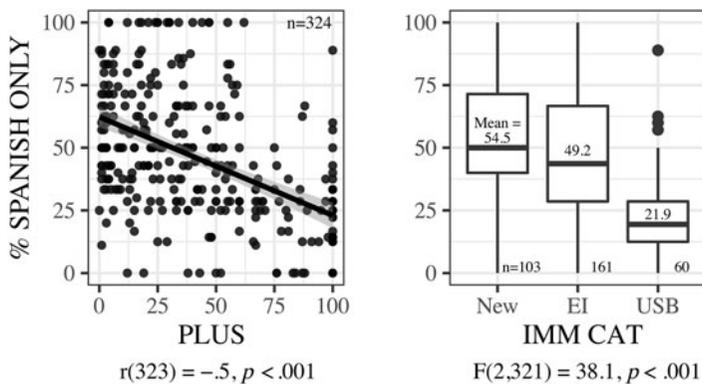


FIGURE 5. Exclusive Spanish use by PLUS and across Immigration Category. In figures throughout, PLUS = Percent life in the US, IMMCAT = Immigration Category.

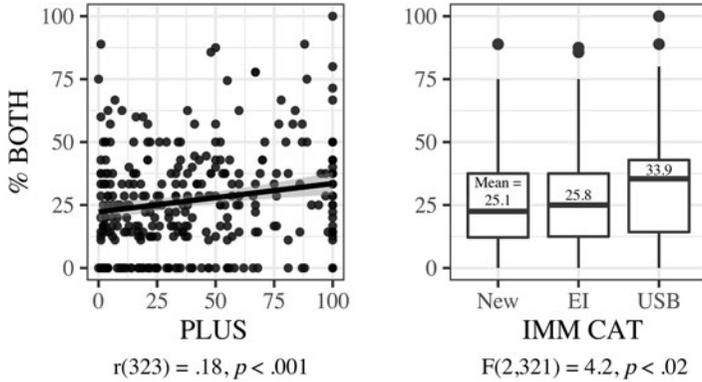


FIGURE 6. Use of both Spanish and English by PLUS and across Immigration Category.

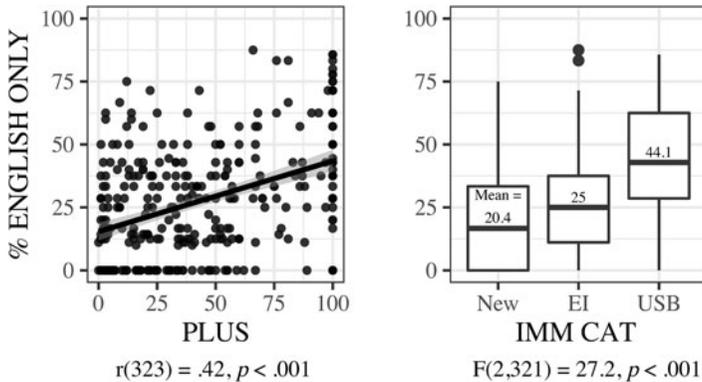


FIGURE 7. Exclusive English use by PLUS and across Immigration Category.

English. To the contrary, the data reveal a robust pattern of increased English usage in both public and private life, supporting a rejection of H1 and linking the corpora with other studies of habitual language use, including work outside of academia, that have reported similar findings (Veltman 1983; Gynan 1997; Hurtado & Vega 2004; Rumbaut, Massey, & Bean 2006; Chavez 2008; Krogstad, Stepler, & Lopez 2015).

In light of these results, we caution against hasty inference that our rejection of H1 amounts to de facto support for H2. Specifically, it is a mistake to interpret the inverse correlation between exclusive use of English and exclusive use of Spanish as evidence in favor of the zero-sum view of bilingualism. Along this erroneous line of thinking, one might conclude that while our results may counter the claim that Spanish speakers are reluctant to learn English, they do show that English's gain

is Spanish's loss. But this interpretation would be problematic for three reasons. First, it fails to appreciate the significant intergenerational INCREASE in bilingual interactions observed here. That is, the proportion of interlocutors with whom participants speak both Spanish and English GOES UP with PLUS and across immigrant generation. At the aggregate level, we see this in the significant positive correlation and significant ANOVA results in Figure 6. Evidence for increased bilingual interaction is also visible at the level of individual interlocutors. Note, for instance, the steady intergenerational increase in the size of the white bars in Figure 2, showing Established Immigrants and USBorn increasingly using both Spanish and English with their mothers, fathers, siblings, and children. If participants were really on a path towards English monolingualism, one would not expect them to choose to use Spanish at all in interactions with speakers for whom the use of English is an option. In other words, frequency of bilingual interactions would be expected to decline intergenerationally, not increase.

A second error of interpretation lies in failing to see the decrease in exclusive Spanish as reflecting the transition from a largely monolingual to a largely bilingual community. Indeed, that the QUANTITY of exclusive Spanish use should negatively correlate with US life experience is unsurprising, given what being bilingual means. Failure to expect a reconfiguration of the participants' exclusive Spanish usage in the direction of bilingualism is to miss the fact that their use of English has to take place somewhere, and that the number of Spanish-only domains will necessarily contract. A third error is to conflate the reduction in quantity of exclusive Spanish use with a change in its QUALITY. This amounts to saying that because participants with more US life experience are using Spanish less, they must also be using it differently. This does not follow. Qualitative differences must be independently demonstrated, and their appropriate characterization would require additional work. As things stand, the results presented here permit no conclusions about qualitative differences, much less claims that such differences constitute a particular state of affairs such as perhaps incomplete acquisition or intergenerational shift in sociolinguistic norms.

TESTING H2: SPEAKERS ACQUIRE ENGLISH AT THE EXPENSE OF SPANISH

With the preceding picture of habitual language use in mind, let us turn to the questions raised by H2: Does the expanded use of English described above co-occur with a pattern of Spanish attrition among Established Immigrants and/or limited acquisition of Spanish among the USBorn? That is, does the pattern of language shift that we observe come at the cost of speakers' knowledge of Spanish? Is English's gain really Spanish's loss? To answer these questions, we now turn to (i) an examination of participants' self-assessment of their Spanish and English skills, as reflected in their answers to the questionnaire item described above, and (ii) patterns of linguistic variation in their spontaneous speech.

Self-reported proficiency

In the overall sample, self-reported proficiency is higher for Spanish than for English. The great majority of participants, 91.4 percent, reported Spanish proficiency of at least *very good*. The corresponding percentage for English was 64.8 percent. Broadly speaking, more life experience in the US corresponded to an increase in reported English proficiency and a decrease in reported Spanish proficiency. Chi-square tests showed that differences in proficiency across immigration categories were statistically significant, such that Newcomers had significantly greater Spanish proficiency and lower English proficiency than Established Immigrants. In contrast to both immigrant groups, a large majority of USBorn participants reported being more proficient in English than in Spanish. In addition, there were significant differences in PLUS across proficiency ratings. Participants who rated their Spanish as *excellent* tended to have lower PLUS than those with ratings of *very good*, who themselves had lower PLUS than those with ratings of *passable* and *poor*. Results for English proficiency and PLUS are a mirror image of the Spanish results. Table 3 and Figure 8 summarize these findings.

Once more, we signpost results with a word of caution. The picture presented by participants' self-assessment of their language proficiency would at first glance seem to support the predictions associated with H2. We see that as US life experience increases, reported proficiency in Spanish significantly declines while reported proficiency in English goes up. We also know that these ratings come from a sample of individuals whose language habits have been intergenerationally reconfigured towards diminished exclusive use of Spanish. In other words, self-ratings for proficiency seem to align with the zero-sum view of linguistic assimilation (and with H2 in particular): The results might appear to show that as speakers undertake the process of assimilating linguistically, Spanish usage contracts and proficiency declines.

TABLE 3. *Self-reported Spanish and English proficiency across Immigration Category.*

	Newcomer (n = 161)	Est. Imm. (105)	USBorn (60)
Spanish			
Poor	0.0	1.0	3.3
Passable	1.9	6.7	20.0
Very good	20.0	31.0	40.0
Excellent	78.1	61.4	36.6
English			
Poor	11.9	14.6	0.0
Passable	45.5	18.5	3.3
Very good	29.7	28.6	13.3
Excellent	12.8	38.2	83.3

chi-sq. = 37.7, $p < .001$ for Spanish, 93.4, $p < .001$ for English

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

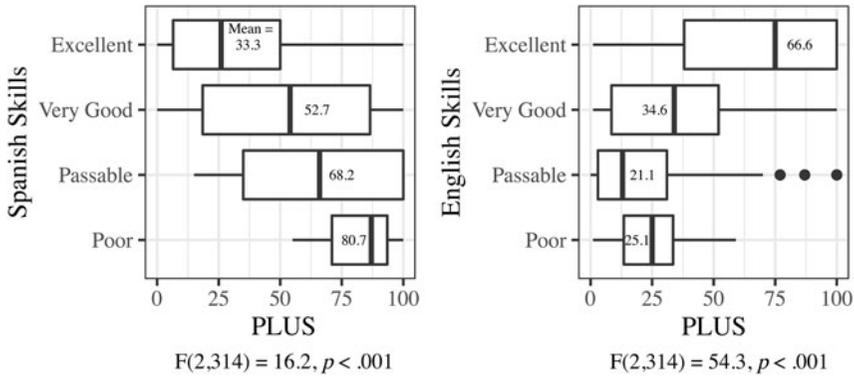


FIGURE 8. Self-reported Spanish and English proficiency by PLUS.

There are, however, reasons to doubt this interpretation. The limitations of self-reported language proficiency can be exacerbated in multilingual settings where different languages are perceived to vary in their respective utility and prestige (Shameem 1998). Of additional importance is the well-known link between speakers’ increased apprehension about using a language and their negative self-perception of their proficiency in it (MacIntyre, Noels, & Clément 1997). This link is consistent with our experiences conducting interviews, during which we have often observed some degree of anxiousness around the topic of proficiency, especially among individuals born in the US—those who are least likely to have formal education in Spanish. So instead of accepting speaker self-reports as reliable and direct metrics of language skills, we are instead inclined to view them primarily as indicators of linguistic (in)security and confidence about using Spanish and English. Our approach makes space for the possibility that someone who uses Spanish less frequently than others do, who is linguistically insecure about her Spanish skills, and/or that someone who describes her proficiency as weak (perhaps because of apprehension related to limited formal education in the language), is nonetheless in possession of linguistic knowledge that is indistinguishable from that of a more frequent and self-confident user of Spanish.

While it may not be immediately obvious that patterns of variation can be used to explore this possibility, numerous scholars have shown that structured variability in speech can be employed as a robust diagnostic of intergenerational linguistic (dis)continuity (Poplack & Levey 2010; Torres Cacoullós & Travis 2018). Indeed, one of the hallmarks of a speech community is that its members share norms for the systematic treatment of linguistic variables. Consequently, determining the degree to which a group of language users shares sensitivity to linguistic and social factors underlying structured linguistic variation is a powerful way of assessing whether and to what extent they should be viewed as members of the same or different linguistic communities. And because sensitivity to such factors is most clearly

displayed via speech (Labov 1972) our interview data represent a site where genuinely direct comparisons can be built. That is, unlike intergenerational comparisons of language skills that are primarily honed in the context of formal education, which can easily lead to overestimation of the degree of linguistic difference between groups of people and be confounded by wide variation in linguistic self-confidence, our approach is based on data that closely approximates how participants naturally use language, namely, through the spontaneous conversations they have with people they encounter in daily life.

Before diving into the analysis, a final bit of context is needed. In line with variationist sociolinguistic methodology, our treatment of the three variables of interest includes comparative analysis at two levels, *rates* and *constraints*. The first of these is a simple comparison of the rates at which different groups use a feature in their speech, in our case the percentage of verbs found with subject pronouns compared to the percentage found without them, the percentage of preverbal versus postverbal grammatical subjects, and the percentage of deleted coda /s/ tokens. The second type of analysis, that of constraints, relies on multivariate regression to assess the effects of the factors that shape variation. In statistical terms, constraints are the independent variables in a multivariate regression model. The regression results are expressed in *constraint hierarchies*. These are lists of independent variables rank-ordered according to their greater or lesser effect on a dependent variable. In addition to being interested in the relative strength of constraints, we also pay close attention the nature of their respective effects, highlighting how the values associated with each constraint affect the likelihood that speakers' choose one or another variant of the dependent variables in question. Comparing the results of different regressions is our method for assessing whether Newcomers, Established Immigrants, and USBorn speakers are similarly sensitive to the same constraints; that is, whether they have shared norms for the use of a particular feature in their respective linguistic repertoires.

Of the two levels of analysis, rates and constraints, the latter is more important. This is because constraint hierarchies are understood to reflect 'the rules which govern variation in the speech community' (Weinreich, Labov, & Herzog 1968:188). As such, in assessing inter-group similarity and difference, rates are less meaningful evidence: 'Differences in rates of variant selection must be used with caution... The key diagnostic in assessing the relationship and provenance of forms is the constraint hierarchy, the configuration of environmental factors affecting the probability that a given variant form will be selected, along with the direction of their effects. We construe these as a portion of the grammar underlying variability' (Poplack & Levey 2010:400).

We, then, compare and provide commentary on rates, but focus on comparing constraint hierarchies. Specifically, we are on the lookout for evidence that would support the prediction that follows from H2: that increased US life experience (and the shifts in habitual language use that it entails) coincides with a breakdown in structured linguistic variation. That is, evidence in support of H2 would

take the shape of inter-group erosion of norms for the treatment of linguistic variables. What we will see in the following intergenerational comparative analysis of subject pronouns, subject position, and coda /s/ is that such evidence largely fails to present itself. Instead, the prevailing trend is one of remarkable similarity across groups, indicative of structural continuity and maintenance of community norms even as patterns of habitual language use are reconfigured, that is, even as exclusive use of Spanish decreases and use of English increases.

Subject pronoun presence vs. absence

Variation in the use of subject personal pronouns with finite verbs (e.g. the alternation between *yo nado* and *nado*, both ‘I swim’) is widely studied in Hispanic linguistics (Bayley & Pease-Alvarez 1997; Flores-Ferrán 2004; Carvalho, Orozco, & Shin 2015; inter alia). Of the studies that have examined pronoun use in the US, Otheguy is responsible for one of the largest (Otheguy & Zentella 2012). While that research was not explicitly framed with H1 and H2 in mind, it nonetheless produced results that amount to a serious challenge to H2, showing that among the present study’s 140 New Yorkers—those in the NYSC—norms for the use of subject pronouns are highly intergenerationally conserved. In other words, a substantial fraction of the participants we have been describing here have been analyzed in prior research with respect to this variable feature, and their usage displays clear structural continuity across immigration categories. This is, therefore, a good place to start in illustrating how the comparative variationist approach works. After examining the New York data with respect to subject pronoun variation, we then move on to analyzing a subset of the BSC, which has heretofore not been studied for this feature.

The first regression results that we examine show how seven linguistic constraints shape pronoun use in the speech of Newcomers, Established Immigrants, and USBorn among our New York participants. The constraints included in the models are (i) the *Person and number* of host verbs, (ii) whether a verb constitutes a *Switch in discourse referent*, (iii) a verb’s *Tense-mood-aspect*, (iv) the *Clause type* in which a verb occurs, (v) the *Lexical content* of a verb—whether it is estimative, or describes a mental or external activity, (vi) a verb’s *Reflexivity*, and (vii) the *Definiteness* of a verb’s referent. For each constraint, we are interested in knowing whether it is significantly predictive, how strongly it influences pronoun use relative to the other constraints, and the nature of its effects, for example, whether a verb is first or third person, in subjunctive or indicative mood, and so on. Table 4 gives the *constraint hierarchy* for each immigration category in the NYSC. It is a hierarchy because the constraints are ranked according to a statistic, the Wald value, that quantifies their relative predictive power in the regression. The bigger the Wald statistic, the more robustly predictive the constraint is. In Table 4 we see that the hierarchies for the three immigration categories are identically ordered for the first four constraints. In all three groups *Person and number* of the verb has the strongest

TABLE 4. *Constraint hierarchies for pronoun presence vs. absence across groups in the NYSC.*

Rank	n tokens = 16,735 $R^2 = .18$		n = 33,514 $R^2 = .19$		n = 10,976 $R^2 = .2$	
	Newcomers (n part. = 39)	Wald	Established Immigrants (75)	Wald	USBorn (26)	Wald
1	Person and number	826.3		2041.5		830.4
2	Switch reference	440.1		918.7		203.4
3	TMA	285.3		526.3		190.6
4	Clause type	98.5		225.9		32.3
5	Lexical type	59.3	Definiteness	219.8	Lexical type	26.4
6	Reflexivity	34.7	Reflexivity	53.9	Reflexivity	19.2
7	Definiteness (nonsig)	1.7	Lexical type	10.9	Definiteness	8.2

Table reproduced from Otheguy & Zentella (2012:160), n part. = number of participants.

effect on pronoun use, followed by *Switch reference*, *TMA*, and *Clause type*. The last three constraints vary in order among the groups, and *Definiteness* does not significantly influence pronoun use among the Newcomers. The three regressions are also similar in overall explanatory power, which is indicated by the R-squared statistic. In each case, the constraints account for roughly twenty percent of the variance or error in the model. Overall, the table tells us that the three groups of speakers are similarly attuned to the same linguistic constraints when choosing whether or not to use a subject pronoun with a finite verb.

In addition to having similarly configured hierarchies, the immigrant groups in the NYSC are also similar with respect to the nature and direction of the effects of the constraints. While a full account of the effects of the individual values within each constraint is beyond the scope of the present discussion (see Otheguy & Zentella 2012:178–95 for details), we can summarize them as follows: Among all three groups in the NYSC, (i) singular subjects increased the probability of pronoun presence while plural subjects decreased it; (ii) switches in discourse referent increased the likelihood of pronoun presence while continuity of referent decreased it; (iii) imperfect indicative verbs favored pronoun presence compared to preterit verbs, for example, *cantaba* vs. *canté*, both ‘I sang’; and (iv) nonreflexive verbs favored pronoun presence compared to reflexive verbs.

Before shifting our attention to the Boston data, which has not been previously analyzed with respect to pronoun use, let us briefly discuss pronoun rates in the New York data. Alongside the preceding picture of intergenerational similarity in constraint hierarchies, Otheguy and Zentella also observed a link between US life experience and rates of pronoun use. Specifically, increased life experience in NYC positively correlates with pronoun rates, a finding consistent with the idea that regular use of Spanish and English has the potential to promote structural convergence between parallel linguistic features. That is, because users of English

produce subject pronouns with finite verbs at very high rates, increased use of English can promote parallel reconfiguration of analogous structures in Spanish. But recall the relative importance of rates within the comparative variationist framework: Changes in rates of use, while potentially illuminating, are not alone considered to represent major shifts in community norms. It is well-established that differences in data structure have the potential to produce misleading results in comparative analysis of rates, especially if the speech of the groups being compared significantly differs in terms of the number of environments that strongly (dis)favor the use of a particular structure, such as a subject pronoun. This is especially problematic when comparative speech data are drawn from different settings and genres—that is, interviews designed to elicit narratives vs. interactive conversations. Different speech genres tend to promote the use of different structures—frequent use of first-person subjects, continuity in verbal tense-mood-aspect, and so on—which can, with respect to subject pronouns in Spanish, lead to different rates of use across genres (Travis 2007:30). Such differences may, in turn, be mistaken for evidence of genuine language change. These facts, along with findings that groups with very different overall rates of pronoun use can be remarkably similar in their sensitivity to conditioning factors (Cameron 1995), are further reason why scholars of contact-induced change emphasize the importance of constraint-based analysis over that of comparisons of rates.

The NYSC (as well as the BSC) was collected with these concerns in mind, and interviewers were uniformly trained to limit their own speech during interviews. They are thus relatively similar in genre in that they were structured to promote narratives. In addition, the very large sample sizes of the data in question reduce the impact of any single participant whose interview may be highly asymmetrical with respect to a particular conditioning factor. While we are inclined, therefore, to interpret the intergenerational increase in pronoun rates in the NYC data as evidence of genuine change, we reiterate that, in keeping with the best practices of the field, rates provide less robust evidence of linguistic innovation than constraints. And in terms of constraints the prevailing trend in the NYSC is the robust cross-generational conservation of norms.

With this as context, now consider a similar analysis of a subset of the Bostonians in the study. Here, the data consist of 11,037 finite verbs produced by thirty-four speakers. Beginning this time at the level of rates, we see that speakers in the BSC display a weak, nonsignificant increase in pronoun use alongside increased PLUS. Similarly, while mean pronoun rates increased across immigration category, differences were not significant. Figure 9 summarizes these results, and compares pronoun rates alongside PLUS and across immigration category for each corpus separately.

To examine pronominal variation in the BSC at the level of constraints, we once again ran three identical regressions, one for each immigration category. The statistical particulars of these regressions are somewhat different³ from those run on the NYSC, but results are highly comparable nonetheless. The main difference is

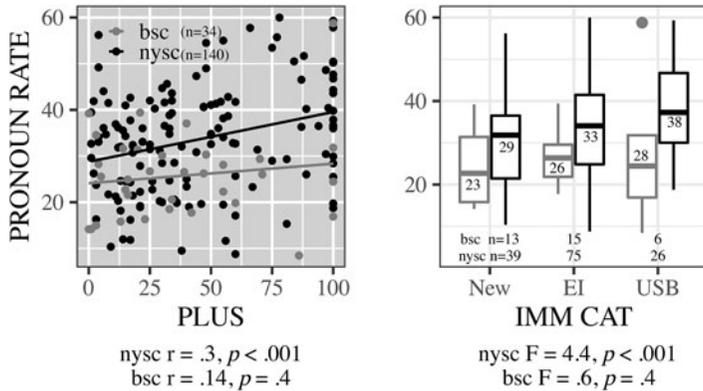


FIGURE 9. Pronoun rates in the two corpora by PLUS and Immigration Category.

that instead of using a Wald statistic to rank constraints, we rely on an analogous measure known as *AIC-increase*, which quantifies how much less explanatory the regression would be if the constraint in question were removed. Additionally, the regressions run on the BSC included only the constraints that had the strongest influence on pronoun use in the NYSC data: *Switch reference*, *Person and number of the verb*, *TMA*, and *Reflexivity of the verb*. Table 5 shows the results, listing predictors in decreasing order of strength, as indicated by their AIC-increase values. Though Established Immigrants have a different ordering of constraints than the other two groups, the rankings for Newcomers and USBorn are identical, paralleling Otheguy & Zentella’s (2012) results. Additionally, the direction of the effects of constraints are very similar across all three groups. Table 6 gives the rbrul weights associated with each constraint-internal value. Weights above .5 indicate that the value favors pronoun presence while those below .5 disfavor it. The further the weight is from .5, the stronger the effect. Among all three groups, pronouns are

TABLE 5. Constraint hierarchies for pronoun presence vs. absence across groups in BSC.

Rank	n tokens = 5,158		n = 3,866		n = 2,013	
	R^2 fixed = .11 total = .29		R^2 fixed = .2 total = .32		R^2 fixed = .21 total = .45	
	Newcomers (n part. = 13)	+AIC	Established Immigrants (15)	+AIC	USBorn (6)	+AIC
1	Switch reference	115.8	Person and number	193.4	Switch reference	30.1
2	TMA	59.1	Switch reference	120.9	TMA	13.6
3	Person and number	43.9	TMA	55.8	Person and number	12.4
4	Reflexivity	3.5	Reflexivity ($p < .08$)	0.9	Reflexivity	1.5

Random factors: Speaker, Verb

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

TABLE 6. *Constraint-internal factor weights for pronoun presence across groups in the BSC.*

Newcomers			Established Immigrants			USBorn		
factor	value	n	factor	value	n	factor	value	n
Switch reference								
different	0.6	2879	different	0.62	2261	different	0.59	1259
same	0.39	2279	same	0.38	1605	same	0.41	754
TMA								
past.subj	0.68	28	imp.ind	0.7	434	conditional	0.73	36
present.ind	0.65	3370	conditional	0.66	65	imp.ind	0.68	208
imp.ind	0.64	419	present.ind	0.62	2382	pret.ind	0.66	217
pret.ind	0.58	620	past.subj	0.59	40	past.subj	0.6	15
conditional	0.57	106	present.subj	0.57	97	present.ind	0.59	1199
periphrastic	0.54	122	periphrastic	0.52	80	present.subj	0.57	36
perfect	0.52	216	pret.ind	0.51	541	perfect	0.52	63
present.subj	0.47	160	perfect	0.48	140	periphrastic	0.48	62
imperative	0.07	117	imperative	0.07	87	imperative	0.03	177
Person and number								
3SG	0.61	668	3SG	0.78	600	2SG	0.66	139
2SG	0.6	744	1SG	0.65	1648	3SG	0.57	348
1SG	0.49	2478	2SG	0.54	568	1SG	0.46	1014
1PL	0.45	418	3PL	0.32	731	1PL	0.42	156
3PL	0.36	850	1PL	0.23	319	3PL	0.39	356
Reflexive								
no	0.55	4785	no	0.54	3572	no	0.59	1912
yes	0.45	373	yes	0.46	294	yes	0.41	101

avored with (i) verbs that constitute switches in referent, (ii) nonreflexive verbs, and (iii) singular as opposed to plural subjects. While there is greater inter-group variability with respect to the variable TMA, *imperfect indicative* verbs forms favor pronoun use compared to preterit verb forms for all three groups. These results mirror what was observed for the same factors among speakers in the NYSC.

To summarize, at the level of rates there is evidence in both corpora of increased frequency of pronoun use with greater US life experience. But this trend is statistically significant only in the NYSC. At the more important level of constraint analysis, the overall trend in both corpora is one of structural continuity. Not only are US-born speakers and the immigrant generations comparably attuned to the same conditioning factors—that is, an identical set of constraints does similarly predictive statistical work across different groups—the way that the various constraints serve to guide speakers’ behavior is largely uniform. These results constitute, in the domain of morphosyntactic variation, strong evidence against H2’s prediction of intergenerational discontinuity. Though increased life experience in the US may promote increased English usage as well as negative self-assessment of Spanish skills, it does not interfere with intergenerational transmission and maintenance

of the norms that underly fine-grained patterns of variable pronoun use in the speech community.

Grammatical subject position

Variability in the position of Spanish grammatical subjects (i.e. the alternation between preverbal and postverbal subjects, *Carlos llegó vs. Llegó Carlos*, both ‘Carlos arrived’) is another widely studied phenomenon (Bolinger 1954; Silva-Corvalán 1982; Delbecque 1988; Ocampo 2009; Mayoral Hernández 2014; Silva-Corvalán 2014; inter alia). Once more, our analysis consists of modeling and comparing variation in different groups. Methodologically, we follow the work of Erker, Ho-Fernández, Otheguy, & Shin (2017), who analyzed a subset of the data examined here. Tokens consisted of subjects co-occurring with finite verbs. Gerundive and infinitival verbs were excluded, as were verbs that appeared without a subject. Data for the present analysis consist of 5,600 grammatical subjects collected from forty participants, fourteen from the NYSC and twenty-six from the BSC. Because of the relatively smaller number of participants examined for this variable, especially compared to the larger sample for pronoun presence vs. absence, we present only a combined analysis and save for future research a comparison of the NYSC and BSC. Grammatical subjects were coded for the following constraints: (i) *Subject type*—personal pronoun, lexical noun phrase, demonstrative pronoun, quantifier, or clause, (ii) *Verb type*—copulative (forms of *ser* or *estar*), experiencer-object/presentative verbs (e.g. *encantar* ‘to make happy’, *gustar* ‘to please’, *parecer* ‘to seem’, *aparecer* ‘to appear’), or occurrence (external and mental activity verbs, e.g. *trabajar* ‘to work’, *creer* ‘to think’), (iii) *Subject referent*—human vs. not human, (iv) *Sentence type*—interrogative vs. declarative, and (v) *Clause type*—main vs. subordinate.

At the level of rates, results indicate that the general preference among Spanish speakers for preverbal subjects is even stronger for those with greater US life experience: There is a significant negative correlation between rates of postverbal subjects and PLUS, and an ANOVA comparing rates across immigration categories returned significant results, showing that US-born speakers used subjects postverbally at a significantly lower rate than Newcomers and Established Immigrants. Figure 10 illustrates these trends and provides the statistical details. These results bear a likeness to those for rates of subject pronoun use, where increased US life experience co-occurred with an increased preference for subject pronoun presence. In both cases, intergenerational shifts are consistent with interpretations of structural convergence, given English speakers’ very strong preferences for (i) explicit nominal subjects and (ii) preverbal placement of subjects.

Shifting to constraint analysis, results show a moderate amount of intergroup variability in the ranking of constraints. However, all of the constraints emerged as significantly predictive of subject position for each group, with the exception of *Clause type*, which was significant only for the USBorn. In

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

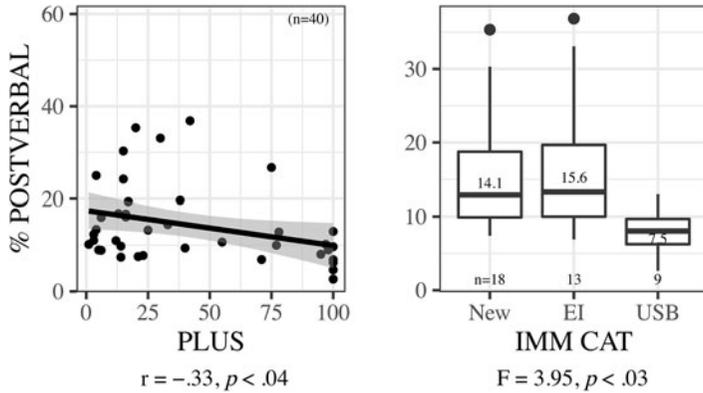


FIGURE 10. Rates of postverbal subjects across PLUS and immigration categories (data from NYC and Boston combined).

addition, the nature of the effects of the constraints was highly similar across the groups: For each group, the probability that a grammatical subject will occur postverbally significantly increases with (i) *experiencer presentative* verbs vs. *occurrence* and *copulative* verbs, (ii) with *nonhuman* vs. *human* referents, (iii) with *clausal* and *lexical NP* subjects vs. *pronominal* subjects, and (iv) in *interrogative* vs. *declarative* sentences. Tables 7 and 8 summarize the regression results.

Overall, results for variation in grammatical subject position are akin to those for pronoun presence/absence. Though there is evidence that the frequency with which speakers are choosing particular variants has shifted, the underlying linguistic constraints giving rise to structured variation are largely invariant intergenerationally.

TABLE 7. Constraint hierarchies for subject position across groups (data from NYC and Boston combined).

Rank	n tokens = 2,646		n = 2,116		n = 838	
	Newcomers (n part. = 18)	AIC	Est. Imm. (13)	AIC	USBorn (9)	AIC
	R^2 fixed = .26 total = .48		R^2 fixed = .38 total = .61		R^2 fixed = .47 total = .52	
1	Verb type	60.74	Subject type	193.38	Human referent	15.34
2	Human referent	23.98	Verb type	120.92	Subject type	16.18
3	Subject type	24.99	Sentence type	55.77	Sentence type	6.45
4	Sentence type	16.1	Human referent	0.94	Verb type	5.60
5	Clause type	ns	Clause type	ns	Clause type	3.32

Random factors: Speaker, Verb. All variables sig. at $p < .001$ except Clause type $p < .05$.

TABLE 8. *Constraint-internal factor weights for postverbal subjects (data from NYC and Boston combined).*

Newcomers			Established Immigrants			USBorn		
factor	value	n	factor	value	n	factor	value	n
Verb type								
exp.pres	0.92	119	exp.pres	0.92	149	exp.pres	0.81	41
occurrence	0.27	1766	occurrence	0.28	1416	copulative	0.34	191
copulative	0.2	761	copulative	0.19	551	occurrence	0.31	606
Human ref								
nonhuman	0.63	874	nonhuman	0.59	794	nonhuman	0.71	149
human	0.37	1772	human	0.42	1322	human	0.29	689
Subj. type								
clausal	0.72	91	clausal	0.83	87	clausal	0.86	19
lexical.np	0.61	950	lexical.np	0.7	688	lexical.np	0.65	242
quantifier	0.58	110	quantifier	0.6	75	quantifier	NA	NA
dem.pro	0.41	194	dem.pro	0.44	177	dem.pro	0.27	36
pers.pron	0.36	1213	indef.pro	0.23	127	pers.pron	0.2	541
indef.pro	0.32	88	pers.pron	0.21	962	indef.pro	NA	NA
Sentence type								
interrogative	0.71	48	interrogative	0.74	81	interrogative	0.8	18
declarative	0.29	2598	declarative	0.26	2035	declarative	0.2	820

In these two variable features, then, we have observed substantial overlap between the immigrant generations and the USBorn, despite the fact that among the latter group shifting habits of language use and negatively self-assessed Spanish proficiency are at their most intense.

Coda s-deletion

To test H2 in the context of our third and final variable, we relied on a set of constraints known to shape synchronic patterns of *s-deletion*, focusing on those relating to (i) the phonetic context in which /s/ occurs, and (ii) various aspects of speaker identity, including *Sex*, *education level*, and *Regional origin/ancestry*. With respect to contextual factors, *s-deletion*, like many phonetic reduction phenomena, has been shown to occur more frequently in fast speech and in high frequency carrier words (File-Muriel 2009). In addition, /s/ is more likely to be reduced when it (a) occurs in unstressed vs. stressed syllables (Poplack 1981; Brown & Torres Cacoullos 2002), (b) occurs word-finally compared to word-internally, for example, *más* ‘more’ vs. *mismo* ‘same’ (Erker 2012), and (c) precedes consonants rather than vowels and pauses (Ma & Herasimchuk 1971; File-Muriel 2009). In terms of social factors, *s-deletion*, while widespread in the Hispanophone world, is generally viewed as a nonstandard linguistic behavior. Within communities, deletion is typically more frequent among men than women and among speakers with lower levels of education (Fontanella de Weinberg 1974; Alba 1990; Cepeda 1995). Regionally speaking, the highest rates of *s-deletion* have been reported in communities located in the Caribbean and among coastal areas of the Latin American mainland (Canfield 1981; Lipski 1994; inter alia).

For the present analysis, 7,054 tokens of syllable final /s/ were collected from twenty speakers in the NYSC and forty-two from the BSC. Speakers included thirty-three women and twenty-nine men. They were grouped according to three regions: *Caribbean*, *Andean*, and *Central*. Twenty-four participants had origins or ancestry in Caribbean locales, including Cuba (one speaker), Dominican Republic (seven), Puerto Rico (fourteen), and Coastal Venezuela (three). Fourteen had origins in the Andean region of South America, including (noncoastal) Colombia (five), Ecuador (four), Peru (four), and (noncoastal) Venezuela (one). Twenty-four had origins in the Central region, which included El Salvador (fourteen), Guatemala (one), and Mexico (nine). With respect to immigration category, participants included twenty-five Newcomers, twenty-six Established Immigrants, and eleven USBorn. In terms of education, two participants had completed only primary schooling, ten were educated through high school only, forty had attended at least some college, and eight had earned postgraduate degrees. Presence versus absence of frication in the speech stream was determined through inspection of waveform and spectrographic displays in PRAAT (Boersma & Weenink 2019). All tokens for which no evidence of frication was observed were categorized as instances of *s-deletion*.

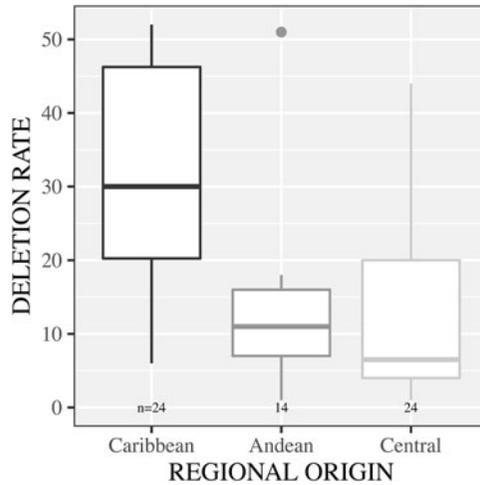


FIGURE 11. s-deletion rates by Region (data from NYC and Boston combined).

Because of the central role of dialectal variation in the study of s-deletion, our analysis of rates begins with an assessment of the influence of regional origin. Figure 11 shows that the distribution of deletion rates falls out along expected lines, with Caribbeans deleting at a higher rate than participants of Andean and Central origin, respectively. A one-way ANOVA comparing deletion rates returned a significant main effect for region ($F(2,59) = 14.55, p < .001$). Bonferroni posthoc tests showed that Caribbeans have a significantly higher deletion rate ($M = 32.42, sd = 15.24$) than Andeans ($p < .001$) and Central speakers ($p < .001$). However, Andeans ($M = 13.21, sd = 11.94$) and Central speakers ($M = 12.42, sd = 11.89$) do not significantly differ from each other in terms of mean deletion rate ($p < .98$).

When speakers were subsequently grouped by immigration category and Region, the same pattern of dialectal variation surfaced in all three groups. While there was some variation in rates across categories—most notably, a relatively low rate of deletion among Caribbean Established Immigrants, ANOVAs comparing rates within each regional group returned no significant differences between immigration categories. See Figure 12 and Table 9.

At the level of constraints, let us once again compare regression results. We see that all six constraints—*Following sound*, *Word position*, *Stress*, *Regional origin*, *Sex*, and *Education level*—are significantly predictive of s-deletion for Newcomers and Established Immigrants. In the regression for USBorn participants, *Stress*, *Education*, and *Sex* were not statistically significant. Despite this difference, which may be due in part to the relatively smaller sample size for the USBorn, the groups are similar in their sensitivity to *Word position*, *Following sound*, and *Region*. Indeed, these three constraints are identically ranked for Newcomers and

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

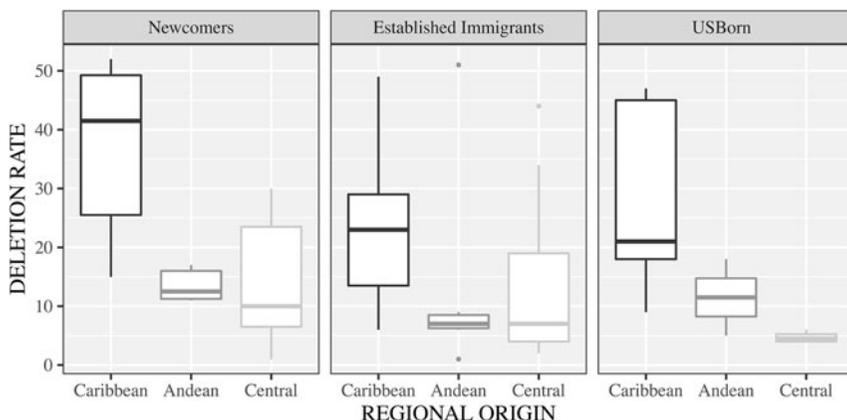


FIGURE 12. s-deletion rates by Region and Immigration Category (data from NYC and Boston combined).

TABLE 9. ANOVA results comparing mean s-deletion rates within regional groups across Immigration Category (data from NYC and Boston combined).

Region	Newcomers	n	Established Immigrants	n	USBorn	n	F	p-value
Caribbean	37.6	12	23.3	7	28.0	5	2.35	0.12
Andean	13.5	6	13.5	6	11.5	2	0.02	0.98
Central	14.4	7	13.7	13	4.8	4	1.01	0.38

USBorn participants. In addition, these three constraints account for more variance in the model for the USBorn, an R-squared of .35, than do the six variables in the regressions for the Newcomers and the Established Immigrants, with R-squared values of .27 and .3, respectively. Thus, while the hierarchy for the USBorn includes a smaller number of significant predictors, they are robustly predictive of structured variation. With respect to the nature of the effects of the variables, results are consistent with the literature: In each group /s/ is more likely to be deleted when it occurs word-finally, precedes a consonant, and is produced by individuals of Caribbean origin. Additionally, for Newcomers and Established Immigrants, /s/ is also significantly more likely to be deleted when it occurs in an unstressed syllable and is produced by men and by speakers with a high school education only. Tables 10 and 11 provide the statistical details of these results.

To summarize, rates of s-deletion are relatively stable intergenerationally and demonstrate qualitatively similar regional differentiation across immigration categories. In addition, analysis of constraints produced hierarchies that were, on balance, more similar to each other than different. The Newcomer and Established

TABLE 10. *Constraint hierarchies for s-deletion across groups (data from NYC and Boston combined).*

Rank	N = 2,879		n = 3,312		n = 863	
	R^2 fixed = .27 total = .38		R^2 fixed = .3 total = .48		R^2 fixed = .35 total = .5	
	Newcomers (n part. = 25)	AIC	Established Immigrants (26)	AIC	USBorn (11)	AIC
1	Word position	76.14	Following sound	110.60	Word position	44.04
2	Following sound	70.87	Word position	105.10	Following sound	8.94
3	Region	27.21	Sex	9.38	Region	2.92
4	Stress	23.88	Stress	8.59	Stress	ns
5	Education	7.82	Region	6.10	Education	ns
6	Sex	1.13	Education	2.69	Sex	ns

Random factors: Speaker, Verb. All variables sig. at $p < .05$ unless indicated.

TABLE 11. *Constraint-internal factor weights for s-deletion (data from NYC and Boston combined).*

Newcomers			Established Immigrants			USBorn		
factor value	weight	n	factor value	weight	n	factor value	weight	n
Word position								
final	0.66	2109	final	0.71	2494	final	0.81	649
internal	0.34	770	internal	0.29	818	internal	0.19	214
Following sound								
consonant	0.65	1880	consonant	0.71	2093	consonant	0.53	543
vowel	0.53	607	vowel	0.29	759	vowel	0.48	199
pause	0.33	392	pause	0.28	460	pause	0.23	121
Region								
caribbean	0.83	1549	caribbean	0.72	1159	caribbean	0.81	300
andean	0.48	687	central	0.41	1461	andean	0.41	265
central	0.18	643	andean	0.36	692	central	0.25	298
Stress								
unstressed	0.59	1631	unstressed	0.56	1966	unstressed	ns	465
stressed	0.41	1248	stressed	0.44	1346	stressed	ns	398
Education								
highschool	0.65	301	primary	0.66	150	highschool	ns	639
postgrad	0.23	657	highschool	0.59	935	postgrad	NA	0
somecollege	0.22	1921	postgrad	0.46	394	somecollege	ns	74
primary	NA	0	somecollege	0.29	1833	primary	ns	150
Sex								
male	0.56	1878	male	0.68	1074	male	ns	307
female	0.44	1001	female	0.32	2238	female	ns	556

Immigrant hierarchies were highly similar, sharing the same set of significant predictors operating in the same ways, but differing in terms of relative strength. While *Region* was a relatively weaker predictor in the hierarchy for Established Immigrants, the general difference between Caribbean and Central/Andean speakers persisted. While the constraint hierarchy of the USBorn differed to a greater degree—neither *Sex*, *Stress*, nor *Education* were significant in the model—regional differences clearly persisted within this group. Furthermore, the two linguistic factors that most robustly constrain deletion among Newcomers and Established Immigrants, *Following sound* and *Word position*, accounted for the most variance in the regression for the USBorn.

SUMMARY AND CONCLUSIONS

Results indicate that among study participants frequency of use of Spanish significantly decreases as US life experience increases. Reduced use of Spanish occurs alongside increased use of English, in both private life as well as in public settings. But if this reduction in Spanish usage is a harbinger of language loss, there is no evidence for it in our findings. In New York and Boston we see Latinos speaking more English and less Spanish, but by no means do we see them, even in the USBorn generation, abandoning Spanish. The changed pattern of language use, more English, less Spanish, correlates with increasingly negative self-assessments of Spanish proficiency among immigrants of longer residence and among the USBorn. However, contrary to participants' own evaluations of their Spanish skills, we have demonstrated that expanded use of English is NOT accompanied by Spanish grammatical erosion with respect to structured linguistic variation in the features examined here. Rather, comparative variationist analysis of spontaneous speech revealed robust patterns of intergenerational continuity in aspects of grammar guiding variation; the factors that Newcomers pay attention to when they use the linguistic features that we studied are largely the same as those that the Established Immigrants and the USBorn pay attention to. What's more, the direction of the effect of these factors is the same across the groups. That is, the groups not only share conditioning factors but also clearly pay the same or similar attention to them.

To the extent that intergenerational differences arose in the data, they were clearest at the level of rates of use of the different forms, where greater US life experience appeared to increasingly align some aspects of speaker's linguistic behavior in Spanish with parallel structures in English. This type of difference, consistent with the notion of structural convergence, was observed in speakers' use of subject pronouns and in their preferences for grammatical subject position, but much less so in the case of *s*-deletion, for which rates were essentially intergenerationally constant, reflecting instead a stable and persistent cross-generational pattern of regionally constrained variation.

Of the three linguistic variables that were analyzed, s-deletion showed significant differences between the USBorn and the immigrant generations even at the level of constraints, in that three variables that were significant in the regressions for the latter were not significantly predictive for the former. But even here we have noted that the variables in question were the weakest three among the immigrant groups, and that the three variables that all three groups had in common accounted for more variance in the regression for the USBorn than did all six variables for the two immigrant groups.

In light of these results, we argue that both hypotheses outlined at the beginning of this article should, for the participants and variables examined here, be rejected. As to H1 (Spanish maintained, English not acquired), its incompatibility with our results will not come as a surprise to anyone familiar with the relevant literature on this topic. It is simply a myth that American immigrant populations resist learning English. Despite the repeated failure of H1 (in its various guises) to withstand empirical scrutiny, it nonetheless persists in shaping attitudes towards language and immigration among nontrivial numbers of Americans. And the impact of the myth embodied in H1 is wide-ranging, asserting itself in federal public policy while also shaping demands for t-shirts and generating viral video content on the internet. We therefore see no choice but to continue engaging this myth, in the hope that the misconceptions about language and immigration that it props up will eventually collapse under the collective empirical weight of this and other related studies.

In addition, we have demonstrated that for our participants H2 (Spanish speakers acquire English at the expense of Spanish) is also a myth, one that not only promotes misunderstanding of language among the general public (e.g. by promoting linguistic insecurity among USBorn speakers of immigrant languages), but that also negatively impacts linguistic inquiry, propounding the mistaken notion of a US Spanish that is on a path to extinction or at best incompletely acquired. But with respect to the component of linguistic knowledge that guides the structuring of variation, which by all accounts is of considerable magnitude, we have shown that in fact knowledge of Spanish erodes but little in the transition from the baseline of the Newcomers to the comparison with Established Immigrants. And, more important, we have shown that when we turn our attention to the USBorn, there is also little evidence of erosion in the factors that guide structured variation in the features examined here.

We do not deny that patterns of language shift can and have led to attrition of immigrant languages in the US. Nor do we claim that our results are exhaustive in terms of the linguistic variables that could be brought to bear on our research questions about Spanish speakers. Indeed, there are many other features that, if subjected to an analysis such as ours, could potentially reveal different results, for example, selection of *ser/estar*, possessive constructions for body parts, periphrastic vs. synthetic future verbal morphology, and so on (though see Bookhamer's 2013 study for results that parallel ours with respect to subjunctive verbal

morphology in the NYSC). Finally, we are not claiming that our results, based as they are on samples drawn from two major urban centers, are either representative of all Spanish-speaking communities in the US or applicable to every child in the US who is born to Spanish-speaking parents. Indeed, some of those children, especially if they have only one Spanish-speaking parent, very rarely use Spanish, and some of them will become adults whose own children will never use it. However, we contend that the inevitability of this outcome has been overstated and that arguments for language loss or incomplete acquisition have been supported by very poor evidence, such as comparisons between monolingual home-country controls and speakers born or raised in the US, using experimental tasks highly encumbered by Spanish-medium schooling, to which the former get plenty of exposure and the latter little or none.

Some researchers have recently defended the notion of incomplete acquisition in response to this type of criticism. Silva-Corvalán (2018:245) writes that ‘incomplete acquisition implies that the bilingual has acquired a language system that is different from that of the providers of language input’. This seems to us disingenuous, as the term *incomplete* is normally read to imply not only that the grammar of the bilingual is DIFFERENT but that it is, along some dimensions, DEFICIENT. The reformulation strikes us as a strained attempt to rebrand an idea that should be rejected. If the term *incomplete acquisition* were actually meant to describe simply differences between grammars, instead of serving as a way to describe a kind of imperfect outcome, then how does the word *incomplete* contribute to the interpretation of that term in any transparently meaningful way? From our perspective, it doesn’t. Instead of pursuing a research agenda under a flawed banner, we suggest that scholars interested in the intergenerational trajectory of Spanish, or other non-English languages spoken by multilinguals in the United States, stay away from the prescriptivism embedded in the search for (in)completeness of acquisition. Furthermore, we encourage such research to take advantage of structured linguistic variation as a means for understanding significantly important aspects of grammatical competence, intergenerational language maintenance, and language change.

Finally, the outcome of our tests of H1 and H2 shows the zero-sum approach to the lives of languages in the US to be simplistic and naive (in addition, of course, to being personally deleterious and politically dangerous). The key missing part in the zero-sum narrative embodied in H1 and H2 is an awareness of the possibility and reality of bilingualism. In many parts of the world, multi-generational families of bilinguals are common, as indeed our data shows them to be also in the US. To be sure, in the process of intergenerational acquisition, and before that in the process of adapting to a new setting, languages may turn out to be different in some structured ways in bilinguals than in monolinguals, without any inference warranted, however, that these differences are terminal or a manifestation of incompleteness. And if our analysis is correct, these differences between immigrant groups and immigrant generations are much smaller than generally believed when examined through the lens of variationist research. In our case, variationist

analysis of three prominent grammatical features of Spanish has shown structural variation across several generations of US Latinos to be guided by the same or very similar factors operating in the same or very similar ways, a fact reflecting the nearly seamless continuity of grammar for these features and the considerable strength of individual and community bilingualism.

NOTES

¹The authors gratefully acknowledge the support of the National Science Foundation (BCS-1423840, BCS-0004133, BCS-0919029).

²Because not all participants provided a response for every interlocutor, Ns across rows do not add up to the total number of participants.

³These analyses, along with all of the remaining regressions in the article, consist of mixed effects logistic regressions carried out using the *rbrul* package for R (Johnson 2009). *Speaker* and *Token* were included as random effects in each regression, reflecting recent trends in variationist analysis (and quantitatively driven linguistics in general). Otheguy & Zentella's (2012) results were produced by fixed effects regression using the SPSS software package. Because of the very large samples in their analyses, these regression results are robust against the types of concerns that have motivated the recent implementation of mixed effects models.

REFERENCES

- Alba, Orlando (1990). *Variación fonética y diversidad social en el español dominicano de Santiago*. Santiago: Pontificia Universidad Católica Madre y Maestra.
- Anderson, Benedict (2006). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.
- Bayley, Robert, & Lucinda Pease-Alvarez (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9(3):349–71.
- Benmamoun, Elabbas; Silvina Montrul; & Maria Polinsky (2010). *Prolegomena to heritage linguistics*. Urbana-Champaign: University of Illinois at Urbana-Champaign, and Cambridge, MA: Harvard University, ms.
- Bialystok, Ellen (1986). Factors in the growth of linguistic awareness. *Child Development* 57:498–510.
- Boersma, Paul, & David Weenink (2019). Praat: Doing phonetics by computer. Online: <http://www.praat.org>.
- Bolinger, Dwight L. (1954). Meaningful word order in Spanish. *Boletín de Filología* 8:45–56.
- Bookhamer, Kevin (2013). *The variable grammar of the Spanish subjunctive in second-generation bilinguals in New York City*. New York: Graduate Center City University of New York dissertation.
- Bourhis, Richard Y. (1983). Language attitudes and self-reports of French-English language usage in Quebec. *Journal of Multilingual and Multicultural Development* 4(2–3):163–79.
- Brown, Esther, & Rena Torres Cacoullos (2002). *¿Qué le vamos aher?* Taking the syllable out of Spanish /s/ reduction. *University of Pennsylvania Working Papers in Linguistics* 8(3):17–31.
- Cameron, Richard (1995). The scope and limits of switch reference as a constraint on pronominal subject expression. *Hispanic Linguistics* 6–7:1–27.
- Canfield, D. Lincoln (1981). *Spanish pronunciation in the Americas*. Chicago: University of Chicago Press.
- Carvalho, Ana M.; Rafael Orozco; & Naomi Lapidus Shin (2015). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press.
- Cepeda, Gladys (1995). Retention and deletion of word-final /s/ in Valdivian Spanish (Chile). *Hispanic Linguistics* 6–7:329–53.

- Chavez, Leo (2008). *The Latino threat: Constructing immigrants, citizens, and the nation*. Stanford, CA: Stanford University Press.
- Delbecq, Nicole (1988). On subject position in Spanish: A variable rule analysis of constraints at the level of the subject NP and of the VP. *Literary and Linguistic Computing* 3(3):185–201.
- Erker, Daniel (2012). *An acoustically based sociolinguistic analysis of variable coda /s/ production in the Spanish of New York City*. New York: New York University dissertation.
- , & Joanna Bruso (2017). Uh, bueno, em: Filled pauses as a site of contact-induced change in Boston Spanish. *Language Variation and Change* 29(2):205–44.
- ; Eduardo Ho-Fernández, ; Ricardo Otheguy; & Naomi Lapidus Shin (2017). Continuity and change in Spanish among Cubans in New York: A study of placement of subjects of finite verbs. In Alejandro Cuza (ed.), *Cuban Spanish dialectology: Variation, contact and change*, 63–82. Washington, DC: Georgetown University Press.
- File-Muriel, Richard J. (2009). The role of lexical frequency in the weakening of syllable-final lexical /s/ in the Spanish of Barranquilla, Colombia. *Hispania* 92:348–60.
- Fishman, Joshua A. (1964). Language maintenance and language shift as a field of inquiry: A definition of the field and suggestions for its further development. *Linguistics* 2(9):32–70.
- (2003). Who speaks what language to whom and when? In Wei Li (ed.), *The bilingualism reader*, 95–111. London: Routledge.
- Flores-Ferrán, Nydia (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change* 16(1):49–73.
- Fontanella de Weinberg, María Beatriz (1974). Aspectos sociolingüísticos del uso de -s en el español bonaerense. *Orbis* 23:85–98.
- Gynan, Shaw N. (1997). Hispanic immigration and Spanish maintenance as indirect measures of ethnicity: Reality and perceptions. *Ethnic Studies Review* 20:41–57.
- Huntington, Samuel P. (2004). The Hispanic challenge. *Foreign Policy* March/April:30–45.
- Hurtado, Aída, & Luis A. Vega (2004). Shift happens: Spanish and English transmission between parents and their children. *Journal of Social Issues* 60(1):137–55.
- Johnson, Daniel Ezra (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3(1):359–83.
- Krogstad, Jens Manuel; Renee Stepler; & Mark Hugo Lopez (2015). *English proficiency on the rise among Latinos: U.S. born driving language changes*. Washington, DC: Pew Research Center.
- Labov, William (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Lipski, John M. (1994). *Latin American Spanish*. New York: Longman Linguistics Library.
- Ma, Roxana, & Eleanor Herasimchuk (1971). The linguistic dimensions of a bilingual neighborhood. In Joshua Fishman (ed.), *Bilingualism in the barrio*, 349–464. Bloomington: Indiana University Press.
- MacIntyre, Peter D.; Kimberly A. Noels; & Richard Clément (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning* 47(2):265–87.
- Marian, Viorica; Henrike K. Blumenfeld; & Margarita Kaushanskaya (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50:940–67.
- Martin-Jones, Marilyn (1991). Sociolinguistic surveys as a source of evidence in the study of bilingualism: A critical assessment of survey work conducted among linguistic minorities in three British cities. *International Journal of the Sociology of Language* 90(1):37–56.
- Mayoral Hernández, Roberto (2014). Subject position in Spanish: A study of factor interactions with prototypical verbs. In Andrés Enrique-Arias, Manuel J. Gutiérrez, Alazne Landa, & Francisco Ocampo (eds.), *Perspectives in the study of Spanish language variation: Papers in honor of Carmen Silva-Corvalán*, 112–31. Santiago de Compostela: Universidade de Santiago de Compostela.
- Menken, Kate, & Tatyana Kleyn (2010). The long-term impact of subtractive schooling in the educational experiences of secondary English language learners. *International Journal of Bilingual Education and Bilingualism* 13(4):399–417.

- Montrul, Silvina (2008). Incomplete acquisition in Spanish heritage speakers: Chronological age or interfaces vulnerability? In Harvey Chan, Heather Jacob, & Enkeleida Kapia (eds.), *Proceedings of the 32nd annual Boston University Conference on Language Development*, 299–310. Somerville: Cascadilla.
- (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *International Journal of Bilingualism* 13(2):239–69.
- (2013). Bilingualism and the heritage language speaker. In Tej K. Bathia & William C. Ritchie (eds.), *The handbook of bilingualism and multilingualism*, 168–89. Malden, MA: Wiley-Blackwell.
- Ocampo, Francisco A. (2009). El orden de palabras en el español hablado. La construcción sujeto verbo objeto directo. In Veyrat Rigat Montserrat & Enrique Serra Alegre (eds.), *La lingüística como reto epistemológico y como acción social: Estudios dedicados al profesor Ángel López García*, 501–11. Madrid: Arco/Libros.
- Otheguy, Ricardo (2016). The linguistic competence of second-generation bilinguals: A critique of incomplete acquisition. In Christina Tortora, Marcel den Dikken, Ignacio Montoya, & Teresa O’Neill (eds.), *Romance linguistics 2013*, 301–20. Amsterdam: John Benjamins.
- , & Ana Celia Zentella (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. New York: Oxford University Press.
- Poplack, Shana (1981). Mortal phonemes as plural morphemes. In David Sankoff & Henrietta Cedergren (eds.), *Variation omnibus*, 59–71. Edmonton: Linguistic Research.
- , & Stephen Levey (2010). Contact-induced grammatical change: A cautionary tale. In Peter Auer & Jürgen Erich Schmidt (eds.), *Language and space: An international handbook of linguistic variation, vol. 1: Theories and methods*, 391–419. Berlin: Mouton de Gruyter.
- ; James A. Walker; & Rebecca Malcolmsen (2006). An English like no other: Language contact and change in Quebec. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 51(2–3):185–213.
- Rumbaut, Rubén G.; Douglas S. Massey; & Frank D. Bean (2006). Linguistic life expectancies: Immigrant language retention in Southern California. *Population and Development Review* 32(3):447–60.
- Sankoff, Gillian (2002). Linguistic outcomes of language contact. In Jack Chambers, Peter Trudgill, & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 638–68. Oxford: Blackwell.
- Shameem, Nikhat (1998). Validating self-reported language proficiency by testing performance in an immigrant community: The Wellington Indo-Fijians. *Language Testing* 15(1):86–108.
- Silva-Corvalán, Carmen (1982). Subject expression and placement in Mexican-American Spanish. In Jon Amastae & Lucía Elías Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*, 93–120. New York: Cambridge University Press.
- (2014). *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge: Cambridge University Press.
- (2018). Bilingual acquisition: Difference or incompleteness? In Naomi Lapidus Shin & Daniel Erker (eds.), *Questioning theoretical primitives in linguistic inquiry*, 245–68. Amsterdam: John Benjamins.
- Torres Cacoullous, Rena, & Catherine E. Travis (2011). Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism* 15(3):241–67.
- , & ——— (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge: Cambridge University Press.
- Travis, Catherine E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2):101–35.
- Veltman, Calvin J. (1983). *Language shift in the United States*. Berlin: Mouton.
- Weinreich, Uriel; William Labov; & Marvin Herzog (1968). Empirical foundations for a theory of language change. In Winfred Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–188. Austin: University of Texas Press.

AMERICAN MYTHS OF LINGUISTIC ASSIMILATION

Zapata, Gabriela C; Liliana Sánchez; & Almeida Jacqueline Toribio (2005). Contact and contracting Spanish. *International Journal of Bilingualism* 9(3–4):377–95.

(Received 16 October 2018; revision received 20 July 2019; accepted 27 August 2019; final revision received 28 October 2019)

Address for correspondence:

Daniel Erker
Boston University
Department of Linguistics
621 Commonwealth Avenue
Boston, MA 02215, USA
danny.erker@gmail.com