Language Variation and Change, **29** (2017), 205–244.

doi:10.1017/S0954394517000102

Uh, bueno, em … : Filled pauses as a site of contact-induced change in Boston Spanish

DANIEL ERKER AND JOANNA BRUSO
Boston University

ABSTRACT

There is mounting evidence that the filled pauses that pervade spontaneous speech constitute a rich site of linguistic inquiry. The present study uses a comparative variationist method to explore possible effects of language contact on pause behavior, examining 3810 filled pauses produced by 24 Spanish-speaking residents of Boston, Massachusetts. Interspeaker differences in pause behavior correlate with intensity of contact. Participants who have lived in the United States for a larger fraction of their lives, who use English more frequently, and who do so more proficiently fill pauses differently when speaking Spanish than do those who have spent less time in the contact setting and whose English skills and usage are more restricted. Results show that a greater degree of contact corresponds to increased use of centralized vowels in phonologically filled pauses (i.e., more frequent use of [a(m)] and [ə(m)] at the expense of [e(m)]). This pattern is interpreted as evidence of contact-induced change.

A large and growing body of research indicates that nonsilent hesitations, or *filled pauses* (FPs), are an intrinsic component of natural speech. Instead of representing clumsy interruptions in the flow of conservation, FPs actually facilitate both the production and perception of speech (Bell, Jurafsky, Fosler-Lussier, Girand, Gregory, & Gildea, 2003; Swerts, 1988; Watanabe, Hirose, Den, & Minematsu, 2008). For speakers, FPs provide time to plan utterances and a way to hold the conversational floor. For listeners, they provide cues to discourse structure and help to signal conversational turn-taking. Additionally, the distribution of FPs appears to be sensitive to the same types of cognitive, linguistic, and social factors that constrain variation in a broad range of decidedly linguistic phenomena (Corley, MacGregor, Lucy, & Donaldson, 2007; Liberman, 2014; Tottie, 2015). As FPs become more integrated into linguistic inquiry, a number

The authors gratefully acknowledge the support of the National Science Foundation (BCS-1423840). We also wish to thank Byron Ahn, Charles Chang, Gregory Guy, Naomi Shin, Neil Myler, and Ricardo Otheguy for valuable commentary on the present study. We are grateful to the three referees for their insightful reviews. Lastly, we wish to acknowledge the enthusiastic research team of the Spanish in Boston Project: Gretchen Anderson, Corey Angers, Yadirys Collado, E. J. Conlon, Gabriela Diaz-Quiñones, Kendra Dickinson, Daniela Estrella, Lily Fawcett, Lauren Gerrish, Morganne Goddard, Karina Hernández, Madeline Reffel, Emma Temkin, Meda Sandu, Deb Waughtal, and Abbey Tolentino-Winter.

of new questions arise. The current study asks whether filled pauses can shed light on the outcomes of language contact.

In situations of contact, structural differences between linguistic systems are potential inflection points for language change. Indeed, it is in the ways that languages differ that their influence on each other (or lack thereof) can be most clearly assessed in contact settings. Because FPs have occupied a relatively peripheral position in linguistic research, structural descriptions of most languages tend not to include a systematic accounting of the units that speakers regularly use to fill pauses. However, because they are pervasive in speech, it is relatively easy to identify the most common FPs of a language. It is also easy to see that speakers of different languages use different pause fillers. For example, while [ə] and [əm] are favored FPs of English speakers, Spanish speakers frequently fill pauses with [e] or [em]. In other words, FPs are language-specific and constitute sites for cross-linguistic comparison and contrast.

The study of Spanish in the United States has generated a substantial amount of research on the outcomes of contact between speakers of Spanish and speakers of English, who, given high levels of bilingualism, are often one and the same. Of special interest is whether contact promotes *linguistic convergence*, "a type of contact-induced change whereby two languages become more structurally similar over time through the adoption of features from one language (usually the socially dominant language) into the other (usually the socially subordinate language)" (Benevento & Dietrich, 2015:407). While several studies of the linguistic behavior of U.S. Latinos indicate that contact may indeed promote convergence between the norms of some features of English and Spanish (Erker & Otheguy, 2016; Lipski 1996; Otheguy & Zentella, 2012; Silva-Corvalán, 1994), other research has failed to find such evidence, casting doubt on the inevitability of convergence as a contact outcome (Benevento & Dietrich, 2015; Flores-Ferrán, 2004; Torres Cacoullos & Travis, 2010, 2016). These studies, together with additional work that takes a sociolinguistic approach to contact linguistics (Poplack & Levey, 2010; Poplack, Walker, & Malcolmson, 2006; Poplack, Zentz, & Dion, 2012; Sankoff, 2002), motivate and frame the present study of variation in FPs. The central question is whether contact with English may be promoting linguistic convergence in the filled pause behavior of Spanish speakers. To explore this possibility, the present paper examines a set of 3810 FPs that occurred in the spontaneous speech of 24 residents of Boston, Massachusetts.

In much of the literature just cited, the linguistic features used to investigate outcomes of contact are independently well understood. For example, the work of Benevento and Dietrich, Flores-Ferrán, Lipski, Otheguy and Zentella, and Torres Cacoullos and Travis focuses on variation in Spanish subject pronoun use. This variable has been extensively researched in noncontact communities, and there is broad consensus regarding factors that condition its variation. Furthermore, regional differences in pronoun use across the Hispanophone world are well documented, meaning that scholars of Spanish in the United States are well positioned to explore the outcomes of dialectal contact in addition to the

potential effects of Spanish-English contact. By comparison, variationist research on FPs is scant, and it is simply unknown whether FPs represent a site of dialectal variation in the Hispanophone world. The present study is thus necessarily exploratory in scope, simultaneously offering a novel study of Spanish FPs within a variationist framework while also pursuing answers to questions of language contact.

We take a broad view of what may constitute an FP, including both lexical and nonlexical phenomena in the analysis. Lexical[1] FPs included tokens of *bueno* 'good', *como* 'so', *este* 'this', *osea* 'that is', and *y* 'and', as well as the English item *so*. Nonlexical, or phonological, FPs included tokens of [e(m)], [a(m)], and [ə(m)]. Variation in the use of these forms was investigated in relation to several measures of contact experience, including (a) self-reported English and Spanish proficiency, (b) reported use of Spanish and English with selected interlocutors (e.g., with parents, siblings, friends, coworkers), (c) reported use of Spanish and English in selected domains (e.g., listening to the radio, watching TV, reading for pleasure), and (d) percentage of life spent living in the United States.

While variation in lexical FPs remains difficult to interpret in light of the study's results, a clear relationship emerged between intensity of contact and variation in phonological FPs. Speakers with a higher degree of contact—corresponding to more life experience in the contact setting, as well as greater proficiency in and more extensive use of English—fill phonological FPs differently than do speakers with a lower degree of contact. Speakers with less experience in the contact setting, lower English proficiency, and more restricted use of English demonstrate a strong preference for [e(m)]. However, as contact intensity increases, this preference gives way to higher rates of use of both [a(m)] and [ə(m)]. This shift—a pattern of centralization in the preferred vowels used to fill phonological FPs in Spanish—emerges at both the perceptual/impressionistic and acoustic levels of analysis.

The study's results reinforce a view of FPs as decidedly linguistic phenomena, the use of which is systematically variable and capable of illuminating the dynamics of language contact. The study also raises a number of additional questions, some of which are methodological and others that are analytical. Methodologically, the study presents questions regarding the optimal application of variationist methods to the study of FPs. How, exactly, should the variable context itself be operationalized? What are the relevant variants, and what is their shared meaning? In terms of analysis, we will argue that the study's results are best interpreted as evidence of contact-induced linguistic convergence. However, the nature of the mechanism underlying this change remains largely in the domain of speculation. That is, it is not entirely clear *why* this change is taking place. While we offer answers to some of these questions in the paper, its primary (and more modest) goals are to increase interest in the sociolinguistic study of filled pauses by offering a variationist analysis of their use in a corpus of spontaneous Spanish speech, and to make a contribution to the field's understanding of the outcomes of language contact.

PREVIOUS RESEARCH ON FPS

Among the first researchers to systematically analyze FPs were psychoanalysts, who were primarily interested in understanding the relationship between "speech disturbances" and their patients' emotional states. In a series of studies, Mahl (1956) developed a taxonomy of such disturbances, of which FPs were a single category (under the label[2] *ah*). Also included were what he called *sentence corrections*, *sentence incompletions*, *repetition of words*, *stutters*, *intruding incoherent sounds*, *tongue slips*, and *omission* of words or parts of words. Mahl found that increased anxiety corresponded to higher rates of "non-ah" speech disturbances. Mahl also observed variation within and between individuals in the frequency of occurrence of different disturbance types as well as in patients' willingness to tolerate silence in therapy sessions.

Informed by Mahl's work, MacClay and Osgood (1959) outlined a smaller set of hesitation phenomena that included a more specific enumeration of phonological FPs. Their categorization included (a) *repeats*, which they defined as "non-semantically significant repetitions," (b) *false starts*, or "incomplete or self-interrupted utterances," (c) *filled pauses*, defined as "all occurrences of the English hesitation devices [ɛ, æ, ɹ̩, ə]," and (d) *unfilled pauses*, defined as "silence of unusual length" (24). MacClay and Osgood found that false starts were more likely to occur within lexical words while repeats were more likely to occur with function words. The overwhelming majority of pauses—both filled and unfilled—occurred at word boundaries rather than word-internally, and while both types of pauses occurred more frequently before lexical words than function words, FPs occurred more frequently at phrase boundaries. Unfilled pauses occurred more frequently within phrases. Among FPs, [ə] was by far the most frequent.

In addition to differences in the distribution of hesitation types, MacClay and Osgood, like Mahl, observed variation within and between speakers. Some speakers produced hesitation phenomena with relatively high frequency while others produced very few at all. Among those who demonstrated high rates of hesitation phenomena, some produced them at a relatively constant rate, while others' rates varied widely within their own speech. With respect to the function of FPs, MacClay and Osgood (1959:42) proposed the following:

> Let us assume that the speaker is motivated to keep control of the conversational ball until he has achieved some sense of completion. He has learned that unfilled intervals of sufficient length are the points at which he has usually lost his control—someone else has leapt into the gap. Therefore, if he pauses long enough to receive the cue of his own silence, he will produce some kind of signal (m, schwa) or perhaps a repetition of the immediately preceding unit, which says in effect: 'I'm still in control—don't interrupt me'.

The view that filled and unfilled pauses are functionally linked, with the former serving as a kind of conversational failsafe for when the latter last too long, is

examined in the work of Goldman-Eisler (1961:18), who further narrowed the study of hesitation phenomena by focusing exclusively on pauses. While she also found that pauses (both filled and unfilled) are more likely to occur before lexical words than function words, she argued that FPs and silence "reflect different internal processes." In more recent decades, researchers have continued to explore FPs[3] along these and other lines, examining how FPs affect speech processing (Bailey & Ferreira, 2003; Bell et al. 2003; Corley et al. 2007; Fox Tree, 2001; Swerts, 1998; Watanabe, Den, Hirose, & Minematsu 2004; Watanabe et al. 2008), how they may be assessed semantically and lexically (Clark & Fox Tree, 2002; Corley & Stewart, 2008; O'Connell & Kowal, 2005), and the ways in which their distribution is shaped by social factors (Fruehwald, 2016; Liberman, 2014; Tottie, 2014, 2015).

With respect to speech processing, several studies indicate that FPs help listeners predict the relative complexity and probability of forthcoming utterances. Swerts (1998:487) found a relationship between the distribution of FPs and topic shifts. Phrases that marked major discourse boundaries were significantly more likely than not to begin with an FP, suggesting that "discourse structure can to some extent be predicted from characteristics of FPs." Relatedly, Fox Tree (2001) examined the role of *uh* and *um* in the processing of English and Dutch. Fox Tree argued that *uh* and *um* differ with respect to speech comprehension. While the former increases speakers' ability to recognize forthcoming words, the latter has no effect. In a more recent study, Bailey and Ferreira (2003) examined the effect of FPs on the processing of *garden path* sentences by English speakers. They found that the presence of an FP after a syntactically ambiguous noun increases the likelihood that a sentence will be rated as ungrammatical. That is, the occurrence of a filled pause after the noun *dog* in a sentence such as 'While the boy scratched the dog, *uh*, barked' increases the chances that the sentence will be judged as ill-formed.

A number of psycholinguistic approaches to FPs appear to reinforce Bailey and Ferreira's (2003:196) claim that "disfluencies significantly influence the operation of the parser." In an event-related potentials experiment, Corley et al. (2007) found that the presence of the hesitation *er* (produced as [ɛː]) reduced an *N400* effect corresponding to the processing of less predictable words. That is, when preceded by *er*, a less predictable word was made easier to process. Furthermore, the results of a memory test showed that words preceded by *er* were better remembered, leading the authors to propose that "hesitations heighten listeners' immediate attention to upcoming speech" (Corley et al., 2007:659). Additional evidence that FPs aid in speech processing is reported by Watanabe et al. (2008), who examined response times in an image identification task in Japanese. Participants heard descriptions of both simple and complex shapes and were asked to press a button when they had identified the shape being described. Sentences describing the objects were preceded by either an FP (the unit *eto*), a silent pause of the same duration, or no pause at all. Response times were shorter when descriptions were preceded by FPs, and this effect was more robust

for more complex descriptions, indicating that FPs cue listeners to the complexity of forthcoming utterances.

This view is reinforced by production studies. Watanabe et al. (2004) found that the average number of lexical items in Japanese clauses was higher for those preceded by FPs. Moreover, in a study of phonetic reduction in English function words, Bell et al. (2003) found that these units were less likely to be reduced in the presence of FPs. Arguing that words are lenited when they are more predictable or probable, they view FPs as signals that speakers are planning more complex and less predictable utterances. Together, these studies highlight the dual role of FPs in the production and perception of speech. They are both an indication of the status of the speaker—that (s)he is planning a more complex utterance—as well as a signal to the listener to heighten attention and prepare for less predictable speech.

With respect to the potential meaning and lexical status of FPs, Clark and Fox Tree (2002) argued that *uh* and *um* are conventional English words. Specifically, they argued that *uh* and *um* should be classified as interjections. Their proposal is based on analysis of four corpora of English in which they found that *um* was significantly more likely than *uh* to be followed by an unfilled pause before speech resumed. Clark and Fox Tree argued that the basic meanings of *uh* and *um* therefore differ in that *uh* announces a forthcoming short delay in speech while *um* announces that a longer delay is about to occur. O'Connell and Kowal (2005) challenged this proposal, supporting their skepticism with an analysis of *uh* and *um* in the speech of Hillary Clinton and six television personalities who interviewed her. While O'Connell and Kowal did find that silent pauses following *uh* are indeed shorter than those following *um*, they reported that only 24 percent of all instances of *uh* and *um* in their data are in fact followed by a silent pause. In the other cases, *uh* and *um* are either unflanked by a pause, both preceded *and* followed by a pause, or preceded but not followed by a pause. A similar finding is reported by Tottie (2015), whose analysis of American and British English corpora showed that a large proportion of instances of *uh* and *um* did not occur with a period of silence at all, and that nearly half of all instances of *uh* and *um* ended rather than initiated delays.

Regarding the sociolinguistic status of FPs, several studies indicate that variation in FP behavior is constrained by social factors such as age, sex, class, and region of origin, as well as contextual factors such as speech setting and interlocutor identity. Tottie's (2014) study compares rates of use of *uh* and *um* (which are treated together as UHM) in the *Santa Barbara Corpus of Spoken American English* (Du Bois, Chafe, Meyer, Thompson, Englebreston, & Martey, 2000–2005) to those found in the *British National Corpus*. Results show that (a) British speakers use UHM more frequently than American speakers do (14.5 and 7.5 times per 1000 words, respectively), (b) older speakers in both corpora use UHM at higher rates than younger speakers do, (c) upper class speakers in the *British National Corpus* have higher rates of UHM, and (d) speech produced in more private settings (in speakers' homes, for instance) contained fewer UHMs than speech produced in nonprivate settings (e.g., courtrooms, offices). In an

examination of FPs in the *Fisher Corpus*, Liberman (2014) observed differences between men and women. Men prefer *uh* to *um* while women show the opposite pattern. Liberman's analysis also revealed interlocutor accommodation effects such that male speakers' rates of *uh* are lower when speaking to women than when speaking to other men. Women, conversely, produce higher rates of *uh* when speaking to men than when speaking to women. There is also an age effect in the data such that older speakers prefer *uh* to *um* while younger speakers prefer *um*, supporting observations that increased rates of *um* may represent a site of ongoing generational change (Fruehwald, 2016).

Compared with research on FPs in English, the study of Spanish FPs has been limited. Johnson, O'Connell, and Sabin (1979) compared hesitation phenomena in English and Spanish to assess cross-linguistic differences in the temporal organization of elicited speech. Johnson et al. proposed a distributional and functional equivalence between *uh* and *um* in English and the Spanish lexical items *este* 'this', *pues* 'so', and *bueno* 'good'. The possibility that Spanish lexical items may do some of the same discursive work as phonological FPs appears in more recent studies as well. Galué (2002) investigated the distribution and function of discourse markers in the Spanish of 15 residents of Caracas, Venezuela. Galué suggested that in addition to its role as a demonstrative pronoun, *este* has an independent discourse function similar to that of [e(m)]. Both units, she argued, facilitate turn-taking and give speakers time to plan utterances. Graham (2013) took Galue's analysis further by proposing that *este* and [e(m)] are in free variation with each other. Graham supported his proposal with an analysis of *este* and [e(m)] as they occurred in interviews with 18 residents of San Juan, Puerto Rico. The most robust finding of that study is an age effect whereby younger speakers significantly prefer *este* to [e(m)].

Additional insight into the inventory of Spanish FPs comes from the field of speech synthesis. Adell, Escudero, and Bonafonte (2012:460) attempted to incorporate FPs into synthesized speech on the hypothesis that doing so will increase its likeness to natural speech: "Because humans use filled pauses, machines will potentially sound more human-like if they do so." Restricting their attention to phonological FPs, Adell et al. modeled their speech synthesis system to include, in decreasing frequency of occurrence, [e], [a], and [em] as Spanish FPs. Results were consistent with their hypothesis, showing that users rated synthesized speech as more natural when it included FPs.

In studies of Spanish in the United States, there is little explicit analysis of FPs. Two relevant studies are Lipski (2005) and Torres and Potowski (2008). In his study, Lipski is interested in Spanish-English code-switching as a source of language change. He suggested that English origin units in monolingual Spanish discourse (*so*, *but*, *anyway*, *you know*, and *I mean*) can serve as a window into "the factors that facilitate the seamless introduction of L2 functional elements into an expanse of L1 discourse" (3). Lipski's analysis focused principally on insertion of *so* into spontaneous Spanish speech. He observed that *so* frequently occurs phrase-peripherally and often with adjoining unfilled pauses. The use of *so* by Spanish speakers is also of interest to Torres and Potowski (2008), who

examined the speech of Mexicans, Puerto Ricans, and speakers of mixed Mexican and Puerto Rican heritage in the greater Chicago area. Their results indicated that *so* is widely incorporated into Spanish discourse by all groups, though, among speakers with lower Spanish proficiency, *so* appears to be gaining ground at the expense of *entonces* 'so/then' (but see Aaron [2004] for slightly different results).

Together, these studies provide a solid foundation for an analysis of FPs in a situation of language contact. They make the case that FPs are an integral part of the production and processing of speech cross-linguistically, serving important functions for speakers and hearers alike. They also show that language-internal variation in the distribution of FPs is constrained by cognitive, communicative, and social factors. Finally, and of central importance to questions of contact, previous research on FPs shows that their form is dependent on language-specific phonological and lexical properties. As such, they raise the possibility that in situations of contact, FPs may constitute a site of linguistic convergence.

DATA AND METHODOLOGY

Speakers and contact measures

To assess the potential link between language contact and variation in FP behavior, 24 speakers were selected from the *Boston Spanish Corpus*, a collection of sociolinguistic interviews with 185 Spanish-speaking residents of the Greater Boston area. All interviews in this corpus were conducted in Spanish[4] and carried out in the same quiet room using a Zoom h4n digital recorder and an SM93 lavalier microphone. Recordings were made at a sampling rate of 44.1 kHz and were subsequently transcribed. Interviewers asked open-ended questions intended to elicit long turns of spontaneous speech. Frequent topics of conversation included immigration stories, personal histories, future plans, and comparisons of Boston to Latin America. All participants answered a questionnaire designed to collect sociodemographic data as well as information about language use and attitudes. Most interviews in the corpus are roughly 1 hr in length. The shortest is 40 min and the longest is nearly 3 hr. All participants were compensated for their time.

The 24 speakers selected for the present study vary along a number of social dimensions that routinely correlate with patterns of linguistic variation (e.g., sex, age, region of origin, level of education). Thirteen are women and 11 are men. They range in age from 19 to 73 years old. Together they represent seven different locales in the Spanish-speaking world, including Colombia, Dominican Republic, El Salvador, Mexico, Puerto Rico, Peru, and Venezuela. Speakers were either born in these places or their parents emigrated from them. While the majority (21 speakers) have at least some university level education, 3 do not. Of these, two were educated through high school, while the other was educated through elementary school only. The speakers also vary in terms of their length of residence in the contact setting and their use of and proficiency in Spanish

and English, respectively. Variation in these dimensions was assessed using the following measures:

1. Percentage of life in the United States (PLUS)
2. Self-reported language use with selected interlocutors
3. Self-reported proficiency in English and Spanish
4. Self-reported Spanish use in selected domains

The first of these measures, *PLUS*, is the fraction of a speaker's life that has been spent living in the United States. It is calculated by dividing years spent in the United States by age. The second contact measure is based on speakers' responses to the following question, which was administered as part of the sociolinguistic questionnaire: *¿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?* '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?' For each of the 10 interlocutors, speakers were asked to answer *español* 'Spanish', *inglés* 'English', or *ambos* 'both'. Because some speakers did not provide an answer for every interlocutor (e.g., if they were without siblings, children, a significant other), responses were converted to percentages to make cross-speaker comparison possible. To illustrate, speaker 37VZ provided answers for seven interlocutors (she did not give an answer for *younger/older children* or *coworkers*). Of the seven interlocutors for which she provided a response, this speaker answered *inglés* for two (*boss* and *classmates*), *español* for four (*father*, *mother*, *siblings*, and *boyfriend*), and *ambos* for one (*friends*). These values correspond to 29, 57, and 14 percent, respectively, for the variables *Percentage of interlocutors with whom participant speaks* (a) *English only*, (b) *Spanish only*, and (c) *Both Spanish and English*.

The third contact measure in the study, language proficiency, was based on participants' self-reported description of their skills in both English and Spanish. For each language, participants were asked to choose from the options *excelente* 'excellent', *muy bien* 'very good', *pasable* 'passable', and *pobre* 'poor'. For the final contact measure, use of Spanish in selected domains, speakers were asked to answer either *mucho* 'a lot' or *poco* 'little' in response to the following question about their use of Spanish: *¿Cuánto español usa Ud. en: casa, la escuela, el trabajo, actividades sociales, al leer, al escuchar la radio, al mirar la television?* 'How much Spanish do you use: at home, at school, at work, in social activities, when reading, when listening to the radio, when watching TV?' Responses regarding the use of Spanish in these domains were, like those for language use with selected interlocutors, converted to percentages to allow for cross-speaker comparison (e.g., some participants were not students or reported that they did not regularly listen to the radio at all). For example, speaker 28PU provided answers for six domains of use (he did not provide an answer for *at school*). He answered *mucho* for four (*at home*, *in social activities*, *listening to*

the radio, and *watching tv*) and 'poco' for two domains (*work* and *reading*). These responses correspond to 67 and 33 percent (4/6 and 2/6), respectively, for the variables *Percentage of domains in which speaker uses Spanish* (a) *a lot* and (b) *little*.

Table 1 describes the study's speakers in terms of the preceding contact measures. It also lists their sex and age as well as their regional origin, which is indicated by the speaker's study-internal identification number (country abbreviations are as follows: CO = Colombia, DR = Dominican Republic, EL = El Salvador, ME = Mexico, PR = Puerto Rico, PU = Peru, VZ = Venezuela).

Figures 1–4, which are a series of histograms, summarize the frequency distributions of the various contact measures. Figure 1, which illustrates the distribution for *PLUS*, shows that 5 speakers have a *PLUS* of 5 or less, that 3 have a plus of 100, and that the remaining 16 speakers fall in between those two poles. A large majority of the study's participants (18 speakers) have spent less than half of their lives in the United States. The mean and standard deviation for *PLUS* are 33.8 and 32.9, respectively.

Starting with *Spanish only*, Figure 2 shows that most participants (16 speakers) report using only Spanish with 25–75 percent of the study's selected interlocutors. Five speakers fall below this range and three exceed it. Of the three below the 25-percent mark, two participants failed to report speaking only Spanish with any interlocutor. That is, they have a value of 0 for the variable *Percentage of interlocutors with whom participant speaks Spanish only*. There is one speaker who reports using only Spanish with each of the study's selected interlocutors. The mean and standard deviation for this variable are 46.4 and 24.8. With respect to *English only*, Figure 2 shows that a large majority of study participants (all but three) use only English with less than half of the study's selected interlocutors. The mean and standard deviation for this variable are 25 and 20.9. A similar distribution is observed for the variable *Percentage of interlocutors with whom participant speaks Spanish and English*. Only two speakers reported *ambos* 'both' (Spanish and English) for more than half of the study's interlocutors. The mean value for this measure is 28.6, with a standard deviation of 18. The preceding histograms, together with the central tendency measures, indicate that although there is variation among study participants, as a group they use more Spanish than they do English with the interlocutors specified in the questionnaire. This is perhaps unsurprising in light of the asymmetry in reported language skills among study participants, which is illustrated in Figure 3.

This figure shows that the overwhelming majority of participants (21 speakers) described their Spanish skills as *excelente*. By comparison, there is considerably greater variation in speakers' self-evaluations of their English skills, which 10 described as *excelente*, 7 as *muy bien,* 4 as *pasable*, and 3 as *pobre*. Finally, Figure 4 illustrates the reported use of Spanish in selected domains.

Means for the measures *Percentage of selected domains in which speakers use Spanish a lot* and *little*, are, respectively, 51.4 and 48.6, each with a standard deviation of 28.2. Three speakers report using Spanish 'a lot' in 100 percent of

TABLE 1. *PLUS, language use, and proficiency measures for the 24 speakers in the study, ordered in terms of increasing PLUS*

		% Interlocutor with Whom Speaker Uses			Self-Reported Prociency		% Domains Use Spanish			
Speaker ID	PLUS	English Only	Spanish Only	English and Spanish	English	Spanish	'Mucho'	'Poco'	Sex	Age
37VZ	1	29	57	14	Excellent	Excellent	71	29	F	19
32DR	3	0	50	50	Very good	Excellent	67	33	M	33
45PU	3	17	50	33	Passable	Excellent	33	66	F	31
25PR	4	0	71	29	Poor	Excellent	60	40	M	37
75CO	5	0	43	57	Very good	Excellent	29	71	M	32
26VZ	13	40	20	40	Excellent	Excellent	42	58	M	31
63ME	14	60	40	0	Excellent	Excellent	14	86	F	22
01PR	15	33	17	50	Very good	Excellent	50	50	F	20
11DR	16	10	30	60	Passable	Excellent	33	67	F	38
40PU	16	0	83	17	Poor	Excellent	100	0	F	43
17EL	20	14	57	29	Very good	Excellent	66	34	M	25
65CO	21	29	71	0	Passable	Excellent	50	50	F	66
35PU	23	33	50	17	Excellent	Excellent	14	86	M	26
62VZ	25	29	43	28	Very good	Excellent	33	67	M	36
13DR	30	0	100	0	Poor	Passable	100	0	F	43
31ME	33	22	56	22	Very Good	Excellent	83	17	F	43
28PU	40	10	50	40	Very Good	Excellent	67	33	M	70
49PR	42	33	33	34	Excellent	Excellent	20	80	M	36
66CO	55	16	84	0	Passable	Excellent	100	0	M	38
42DR	71	25	25	50	Excellent	Excellent	50	50	M	73
07ME	86	33	33	33	Excellent	Very Good	28	72	M	22
16EL	100	62.5	0	37.5	Excellent	Excellent	14	86	M	20
34EL	100	25	50	25	Excellent	Excellent	16	84	F	23
38PR	100	80	0	20	Excellent	Passable	25	75	F	24

Note: Country abbreviations were defined in text.

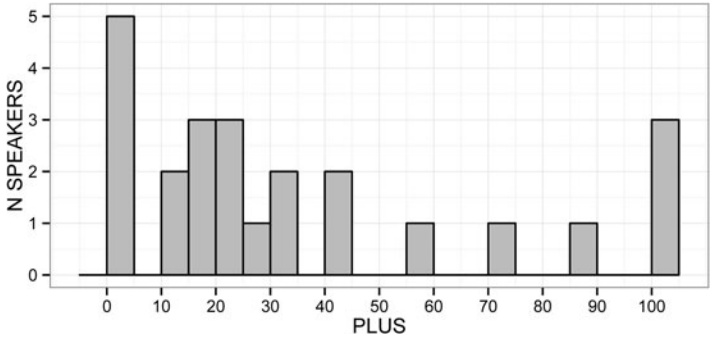

FIGURE 1. Histogram of *PLUS*.

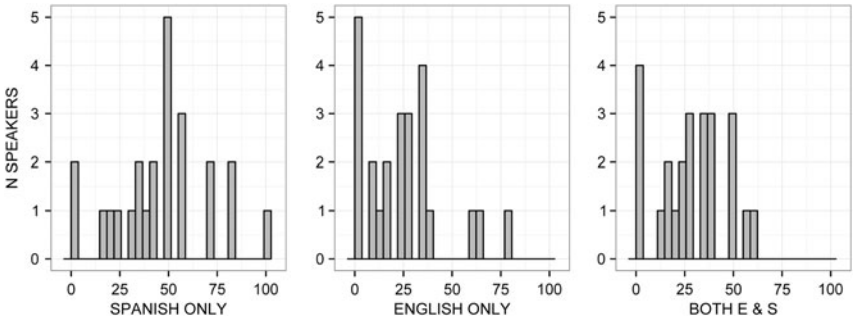

FIGURE 2. Histogram of language use. Percentage of interlocutors with whom participants report speaking *Spanish*, *English*, and *Both English and Spanish*.

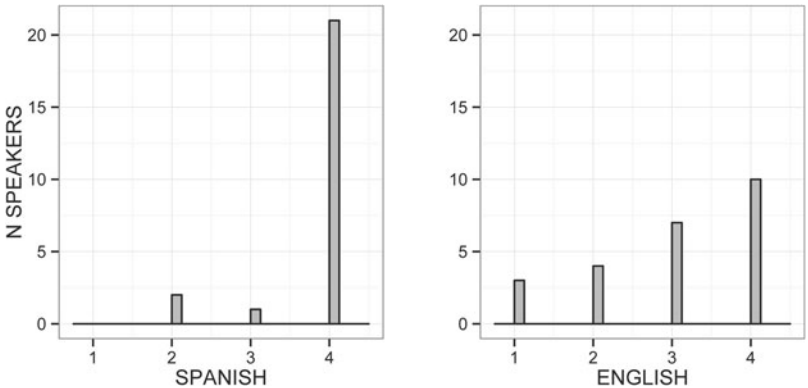

FIGURE 3. Histogram of self-reported proficiency in Spanish and English (numerical values from 1 to 4 correspond, respectively, to *pobre*, *pasable*, *muy bien*, and *excelente*).

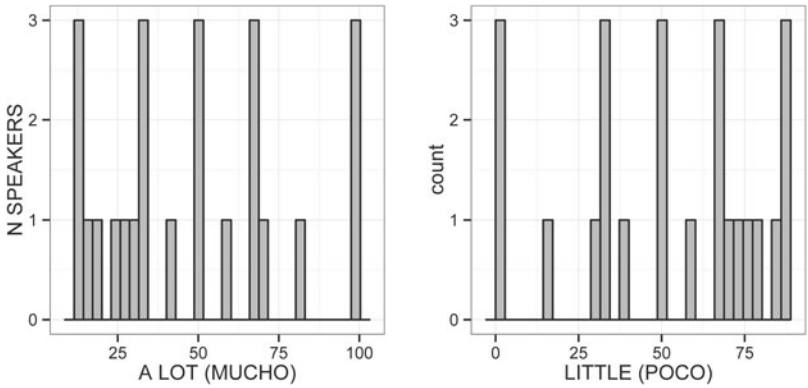

FIGURE 4. Histogram of *Use of Spanish in selected domains*.

the domains specified in the questionnaire (these are the same three speakers who report using 'little' Spanish in no domains, i.e., in all selected domains they use Spanish a lot). All speakers report using a lot of Spanish in at least one domain, which also means that no speaker reported using little Spanish in all selected domains.

Taken together, the study's contact measures reveal a substantial amount of variation among participants. However, in terms of reported usage and proficiency, they are, on the whole, best characterized as a group of Spanish speakers who vary in terms of their English proficiency and use. In this regard, they are not a group that presents as undergoing (or having undergone) language shift, but rather one of varied bilingualism.

Data collection and the variable context

Recall from the literature review that several researchers have suggested that the use of certain Spanish lexical items, some of which have been analyzed elsewhere as discourse markers, is functionally equivalent to that of phonological FPs (Galué, 2002; Graham, 2013; Johnson et al., 1979; Lipski, 2005). For example, Johnson et al. (1979:348) link *este*, *pues*, and *bueno* to *uh* and *um* in English, remarking that, "in context [these forms] in Spanish serve as a filler, just as 'uh' does in English." Lipski (2005:7) made a similar claim: "*este* is a pause-filler roughly equivalent to *uh/um*." Galué draws a Spanish-internal comparison between *este* and [e(m)], and Graham (2013:68), who went as far as proposing that these two forms are in free variation, remarked that they "both translate to English uh or um and are used in the same fashion." These remarks, especially when viewed in light of the relative dearth of variationist research on Spanish FPs, motivate a broadly defined variable context for the present study, and, in particular, one that includes both lexical and phonological FPs. This decision, while well-motivated, presents methodological challenges, especially with respect to identifying the variants of the variable under investigation as well as their shared meaning.

Six[5] lexical and three phonological FPs are included in the study: *bueno*, *como*, *este*, *osea*, *so*, *y*, [e(m)], [a(m)], and [ə(m)]. It would not make sense to suggest that these forms are phonetic variants of a single underlying representation. However, they nonetheless constitute "variable ways of saying the same thing" (Labov, 1978:6) when considered from a functional perspective. When used to fill pauses, they are different ways for the speaker to communicate the message *Please wait, I'm thinking*. Or, to quote MacClay and Osgood's (1959:42) more assertively phrased function: "I'm still in control—don't interrupt me."

The functional equivalence between these forms is marked by shared formal properties. When used to fill a pause, these units are semantically and syntactically independent from the surrounding utterance. They also tend to contain at least one vowel that is relatively long in duration and spectrally stable. Indeed, the spectral stability and lengthy duration of vocalic elements in FPs has been exploited by engineers as a mechanism for filtering them out of automatic speech-to-text transcription (Shriberg, 1999; Stouten, Duchateau, Martens, & Wambacq, 2006). Consider Figures 5 and 6, which illustrate these properties.

Figure 5 shows the spectrogram and waveform for utterances produced by one of the study's participants. In them, *como* occurs twice, once at the end of the phrase *En las vacaciones me las pasaba porque* ***como*** ... 'On vacation I spent them because, uh' and again in the phrase *Es* ***como*** *te dije que soy* ... 'It's like I told you that I am …'. The first *como* is an FP. It is structurally independent from the preceding phrase and makes no obvious formal contribution to the propositional content of the utterance. It is 671 msec long, and the bulk of its duration (406 msec) consists of the vocalic nucleus of the second syllable, [o], which is "spectrally stable during a relatively long period of time" (Swerts, 1998:494). In contrast, the second *como* is structurally integral to the utterance in which it occurs. Without it the phrase would be ungrammatical. At 265 msec, it is less than half as long as its FP counterpart.

In Figure 6, three tokens of [e] are segmented. The first occurs in the final syllable of *entonces* 'then', the second is an FP, and the third occurs in the second syllable of the word *operación* 'operation'. At 241 msec, the [e] filling a pause is much longer than the tokens that occur within lexical items, which are 85 and 65 msec, respectively.

As mentioned, some of the properties outlined as functional and formal attributes of FPs have also been ascribed to *discourse markers*. However, while these elements, like FPs, contribute to discourse structure and facilitate conversational interaction, they also appear to serve more complex discursive functions. Travis (2006:2), for instance, defined discourse markers as "the heterogeneous group of linguistic items that act on (or 'mark') segments of discourse, and function to indicate how those segments are to be understood in the context of surrounding discourse." This function of discourse markers is nicely illustrated in Hernández and Baldazo's (2013) analysis of the variable use of Spanish *verdad* 'truth'. These researchers propose that in its capacity as a discourse marker this form "*incrementa la cohesión entre enunciados y codifica el significado modal entre el hablante y el enunciado*" '[*Verdad*] increases

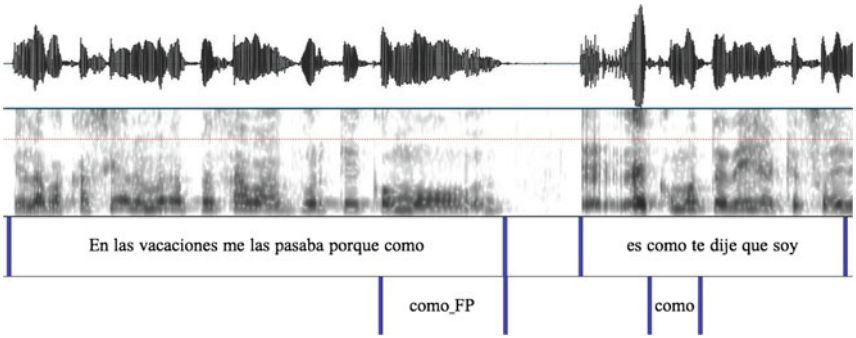

FIGURE 5. *Como* as an FP.

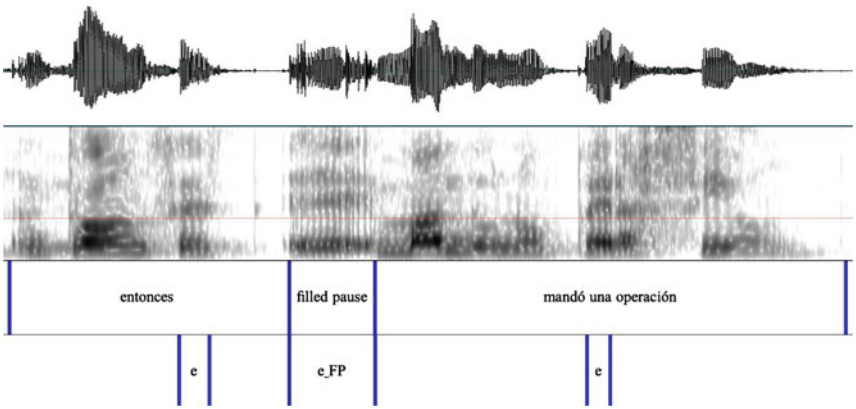

FIGURE 6. [e] as an FP.

cohesion between utterances and encodes a particular meaning between the speaker and the utterance' (Hernández & Baldazo, 2013:115). It also "*añade mayor veracidad a la carga proposicional*" 'increases the veracity of propositional content' (Hernández & Baldazo, 2013:116). That is, the discourse marker *verdad* provides a kind of commentary that relates speaker and utterance, indexing the former's assertion of the accuracy of the latter. This interpretation is similar to that of Galué (2002:27), who suggested that discourse markers "support the illocutionary force of speech acts."

While the research literature reviewed previously does not cite this as a potential function of FPs, our data collection experience suggests that some elements that serve this kind of discursive function can themselves simultaneously function as FPs. In other words, we suggest that the potential overlap between FPs and discourse markers is asymmetrical in nature. A pause might be filled with a turn-extending element that at once signals "processing difficulty or the need for planning time" while also providing semantic-pragmatic meta-commentary on an utterance (Tottie, 2015:53). The upshot of this line of thinking is that some of

the items under analysis as FPs in the current study may, in addition to their role in communicating the message *Please wait I'm thinking*, also be serving to relate speakers and their utterances. That is, our methodology allows for the possibility that some FPs are also discourse markers. We leave for future research, however, the task of analyzing the second of these functions. Ultimately, the potential overlap between these two classes of elements will be somewhat peripheral to the matter of interpreting the study's findings with respect to language contact.

Up to 200 FPs per speaker were collected in the manner illustrated in Figures 5 and 6. In addition to being segmented in *Praat* text grids (Boersma & Weenink, 2016), all phonological FPs were impressionistically coded as tokens of [e(m)], [a(m)], and [ə(m)] or *cannot tell* (these were FPs containing a vowel whose quality could not be confidently coded perceptually). F1 and F2 measurements were taken at the midpoint of the vocalic portion of phonological FPs, excluding any nasal segments. A script was used to extract duration and formant measurements. Silent pauses were excluded from the present analysis on the basis of the contested nature of their relationship with FPs. While some researchers have argued that the two phenomena are linked (Clark & Fox Tree, 2002; Goldman-Eisler, 1961; MacClay & Osgood, 1959), the more recent research of O'Connell and Kowal (2005) and Tottie (2015) raised substantial questions about such a link, suggesting that unfilled pauses constitute an independently variable phenomenon.

Tokens were coded for the following linguistic factors: (a) *Duration* of the FP, in milliseconds. (b) *Repetition:* Whether the pause was (i) a direct repeat of the immediately preceding pause token, in other words, was the FP observed prior to an instance of [e(m)] also a token of [e(m)]?, or (ii) a repeat of the preceding pause type, in other words, was the FP observed prior to a lexical FP also a lexical FP? (c) *Phonetic context:* Tokens were coded for the identity of the segments immediately preceding and following them, as well as whether the preceding and following syllables were stress-bearing or not. (d) *Interviewer FPs:* To assess the possibility that interviewers' FP behavior influenced that of the study's participants, 20 FP tokens were collected from each interviewer. Each interviewer was coded for his/her (i) most frequently used FP, (ii) preference for lexical or phonological FPs, and (iii) mean F1 and F2 in phonological FPs.

RESULTS

The 24 interviews yielded a total of 3810 FPs: 1595 tokens, or 42 percent of the data, were phonological FPs, and the remaining 2215 were lexical. Figure 7 illustrates the observed frequency distributions of the FP types included in the study (ct refers to cannot tells, and schwa is used here and in other figures to refer to [ə], as the ggplot2 package used to create the figures (Wickham, 2009) does not easily interface with data frames containing International Phonetic Alphabet symbols).

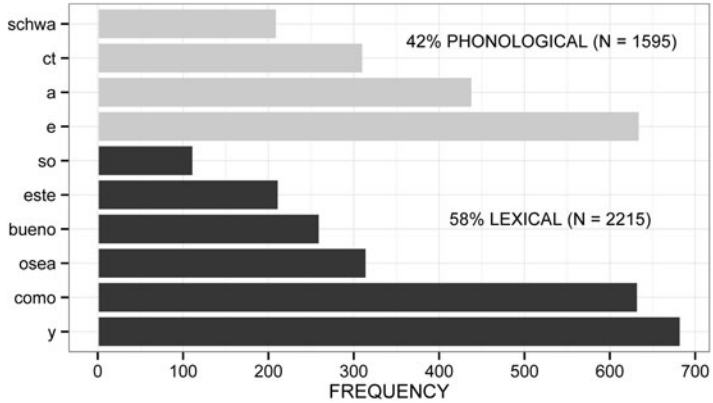

FIGURE 7. Frequency of FP types (phonological FPs in gray, lexical FPs in black).

Analysis of the 3810 FPs is presented in several steps. First, we adopt and expand on the approach of Graham (2013). In that study, the dependent variable was formulated as a choice between a lexical and phonological FP, *este* and [e(m)], respectively. Here we generalize this approach, attempting to ascertain which, if any, linguistic and social factors in the study shape speakers' choices between using lexical or phonological FPs. Following this analysis, lexical and phonological FPs are analyzed separately, with an emphasis on the potential relationship between the study's contact measures and patterns of variation within each of these classes.

Lexical versus phonological FPs

As seen in Figure 7, there is a modest overall preference for lexical FPs in the data. There are a number of factors that modulate this preference. Consider Figure 8, which is a bar graph showing proportions of phonological and lexical FPs (relative to the total set of FPs) for each individual in the study. Speakers are presented in order of increasing preference for lexically filled pauses. The dotted horizontal line set at .58 marks the overall proportion of lexical FPs in the data. At the speaker level, the observed ratios differ significantly from what is expected on the basis of the entire sample ($\chi^2 = 536.4$, $df = 23$, $p < .001$). While no speaker used exclusively lexical or phonological FPs, some strongly prefer the former to the latter, and vice versa.

Speakers' preferences for phonological or lexical FPs are unrelated to how frequently they fill pauses. While speakers who prefer phonological FPs are slightly more frequent pause fillers in general (3.2 FPs per minute compared to 2.9, see Figure 9), this difference is not significant ($t = -.749$, $p = .46$).

Several other factors do produce significant differences in the choice between phonological and lexical FPs. These include (a) pause duration, (b) repetition, (c) speaker age, (d) sex, (e) geographic origin, and (f) *PLUS*. The variables

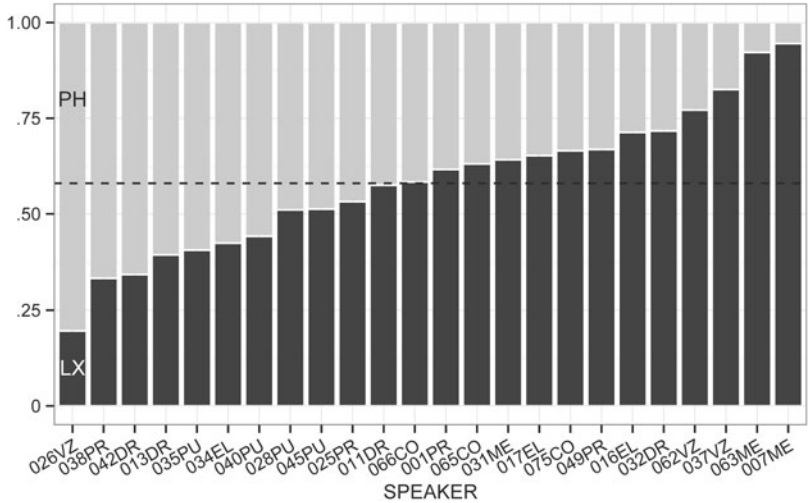

FIGURE 8. Proportions of phonological (PH) and lexical (LX) FPs by speaker (dotted line indicates overall proportion of lexical FPs).

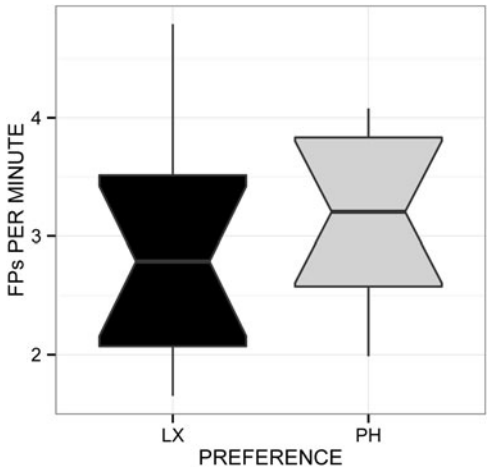

FIGURE 9. Rate of FP use per minute, by phonological (PH) to lexical (LX) preference.

related to phonetic context and interviewer FPs did not return any significant results. Nor did the other contact measures, in other words, those related to language skills, language choice with interlocutors, and Spanish use in selected domains. Mean duration for phonological FPs is 434 msec, which is significantly longer than that of lexical FPs (332 msec, $t = -13.9$, $p < .001$), showing that when speakers fill a shorter pause, they are more likely to do so with a lexical FP. See Figure 10.

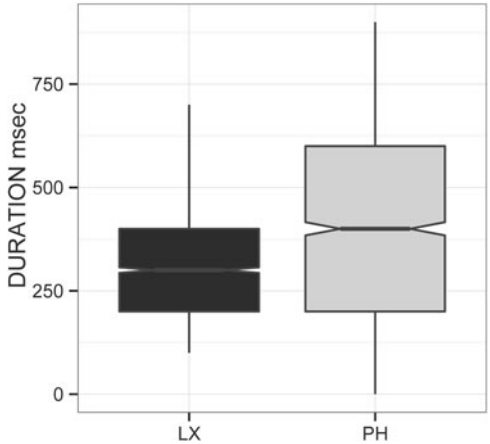

FIGURE 10. Duration of lexical versus phonological FPs.

In addition to being preferred in shorter pauses, lexical FPs are also more likely to be direct repeats of the immediately preceding pause filler in the discourse ($\chi^2 = 13.6$, $df = 1$, $p < .001$). They are also more likely to be used by speakers who are younger ($\chi^2 = 22.7$, $df = 2$, $p < .001$), male ($\chi^2 = 5.4$, $df = 1$, $p < .02$), and of Mexican origin ($\chi^2 = 160.2$, $df = 6$, $p < .001$). Finally, in a result that is difficult to interpret, differences emerge with respect to *PLUS*, such that speakers who have lived between half and three quarters of their life in the United States disfavor lexical FPs compared with the rest of the sample ($\chi^2 = 14.9$, $df = 3$, $p < .001$). Figure 11 illustrates these results with a series of proportional bar graphs. Again, the dotted horizontal line indicates the overall distribution of phonological to lexical FPs in the data.

In a mixed-effects logistic regression[6] with FP type (lexical vs. phonological) as the dependent variable, *Speaker* as a random factor, and six fixed effects (all except *Region*), significant effects persist, in decreasing order of predictive power, for *Duration* (Wald $\chi^2 = 186.9$, $p < .001$), *Country* (Wald $\chi^2 = 14.39$, $p < .03$), and *Repetition* (Wald $\chi^2 = 3.65$, $p < .05$). To summarize, results indicate that choices between phonological and lexical FPs are widely variable between speakers but independent of their overall FP rates. Among the factors included in the study, what primarily constrains the choice between a lexical and phonological FP is duration. The shorter the FP, the more likely it is to be lexical as opposed to phonological. Also, FPs that are direct repeats of immediately preceding tokens are more likely to be lexical than phonological. Finally, significant differences emerge along lines of country; Mexican speakers have the strongest preference for lexical FPs and Peruvians have the strongest preference for phonological FPs. Overall, while the linguistic and social factors included in the study shed some light on variation in the choice between lexical or phonological FPs, their relevance to questions of contact-induced change is not immediately apparent.

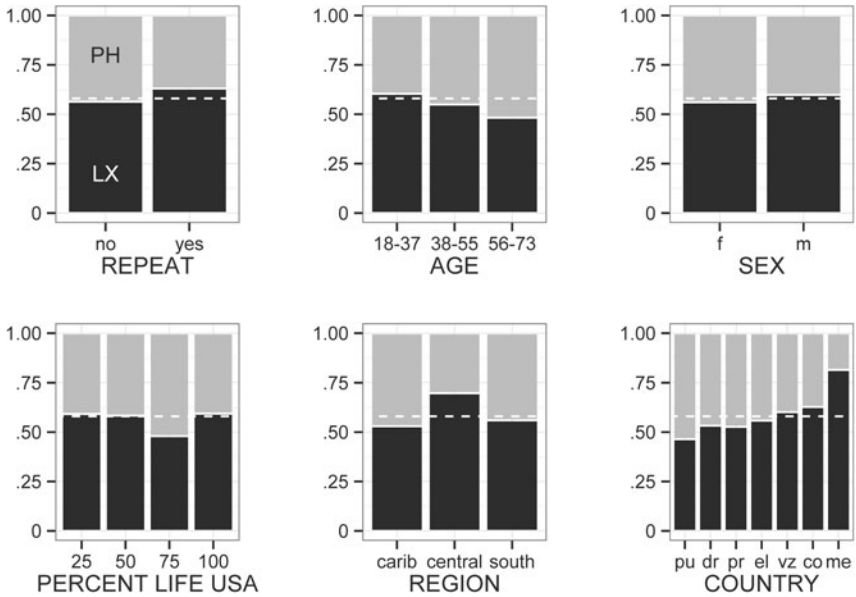

FIGURE 11. Lexical (LX) versus phonological (PH) FPs by six factors (country-to-region groupings are as follows: Caribbean [PR and DR], Central [EL and ME], and South [CO, PU, VZ]).

The next two sections explore the possibility that contact effects are better assessed by examining lexical and phonological FPs separately.

Lexical FPs and the study's contact measures

Within the lexical data, speakers vary with respect to both their FP inventories as well as their frequency of use of particular items. Consider Figure 12, which illustrates lexical FP inventories by speaker. The absence of a square in a given row indicates that this participant did not use a particular item as an FP during his or her interview. All speakers used *y* and *como* as pause fillers. All but one used *bueno* and only four failed to use *osea*. Fifteen speakers also employed *este* and 12 speakers used *so*. Although use of *so* would be expected to correlate with more intense language contact, results are inconclusive in this regard. For example, Figure 13, a proportional bar graph, shows that *so* is not only used across a spectrum of *PLUS*, but that the group using it with the second highest frequency consists of speakers who have spent less than a quarter of their lives in the United States.

Figures 14–16, also proportional bar graphs, illustrate the use of lexical FPs across the study's other contact measures. In Figure 14, two patterns are of potential interest. First, the use of *so* appears to roughly track with the contact measures in the expected directions. Its use is extremely limited among participants who report using *Spanish only* with more than 50 percent of the

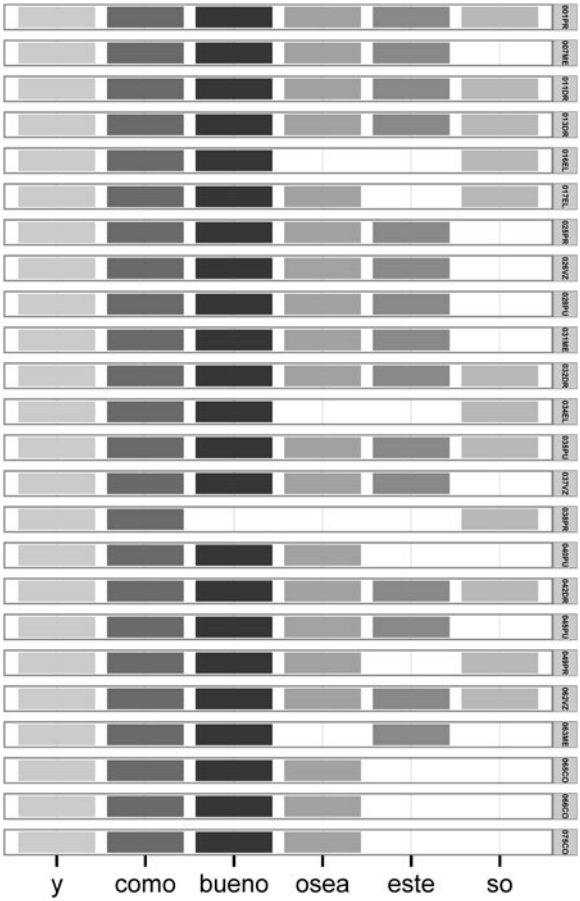

FIGURE 12. Lexical FP inventories by speaker.

study's selected interlocutors. By contrast, the use of *so* is highest among those in the rightmost columns of the *English only* and *E & S* images in the figure—that is, the 76th to 100th percentile for each of these measures. Together, these plots suggest that greater use of English corresponds to greater use of *so* as a lexical FP. However, statistical evidence to support this claim is lacking: Significant correlations fail to emerge between rates of use of *so* and any of the contact measures in the study.

A second pattern of interest emerges in the center image of Figure 14, in the data for the *English only* measure. There is an apparent difference in lexical FP inventory size as well as patterns of use between participants who speak only English with more and less than 50 percent of their interlocutors, respectively. While all six lexical FPs appear in the first and second columns, only four appear in the third and fourth. The smaller inventory in columns 3 and 4 is driven by expanded use of *como*. While there is no correlation between rates of

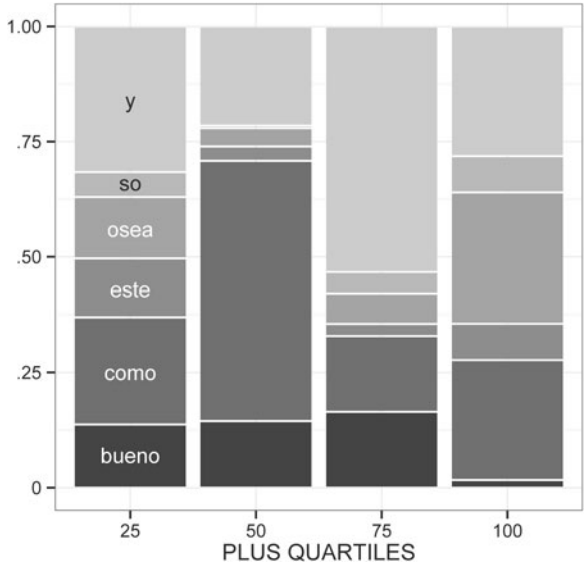

FIGURE 13. Proportion of each lexical FP item according to *PLUS* quartile.

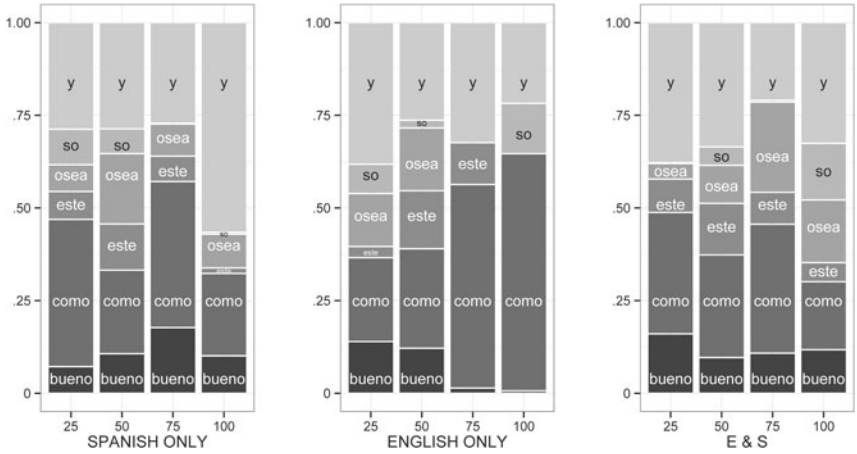

FIGURE 14. Lexical FP use and *Percentage of interlocutors with whom speakers use Spanish, English,* and *both English and Spanish* (by quartiles).

use of *como* and any of the study's contact measures, a significant negative correlation exists between the lexical FP inventory size and rates of *English only* ($r = -.41$, $p < .05$). As exclusive English use increases, the number of different lexical FPs that speakers use decreases. This finding is complemented by a nearly significant correlation between lexical FP inventory size and *PLUS* ($r = -.38$, $p < .06$).

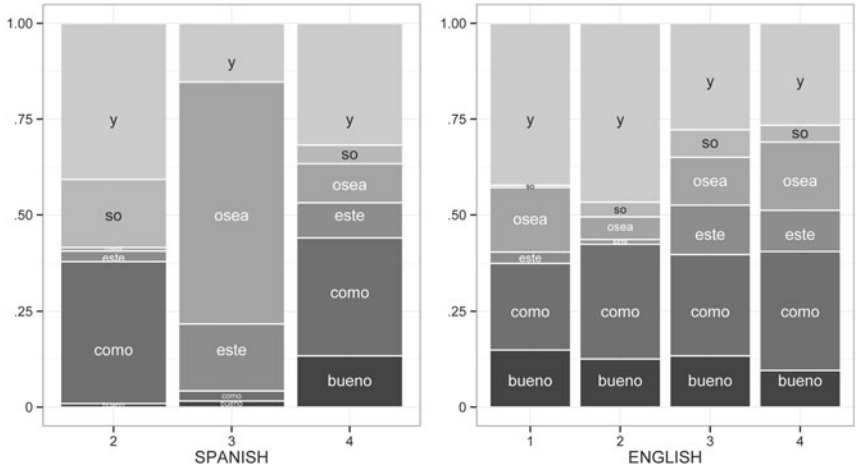

FIGURE 15. Lexical FP use and self-reported proficiency in Spanish and English (numerical values from 1 to 4 correspond, respectively, to *pobre*, *pasable*, *muy bien*, and *excelente*).

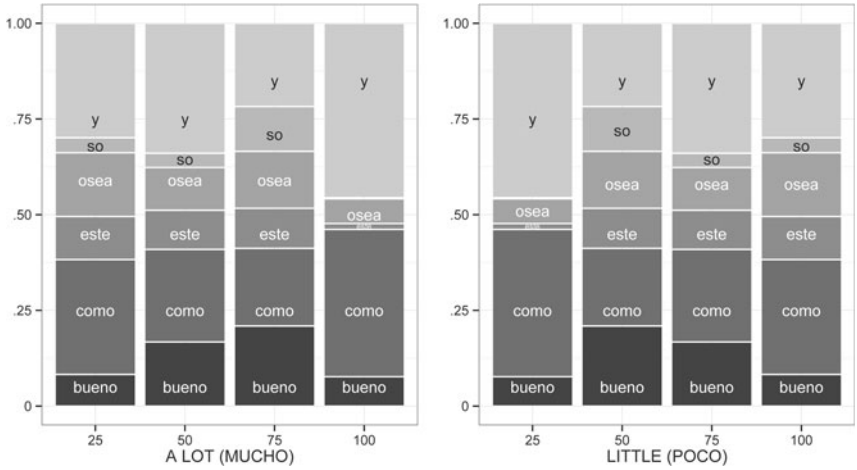

FIGURE 16. Lexical FP use and self-reported use of Spanish in selected domains.

Figures 15 and 16 round out the study's contact measures, illustrating lexical FP use across different levels of Spanish and English proficiency as well as use of Spanish in selected domains. While the figures are suggestive of a relationship between use of *so* and the contact measures—there is frequent use of *so* among the two speakers who report *passable* Spanish and infrequent use of *so* among speakers who report using *a lot* of Spanish in 100 percent of selected domains—there is no statistical evidence for such a relationship. Nor is there any relationship between lexical FP inventory size and either of these last two measures.

Phonological FPs and the study's contact measures

Recall that phonological FPs were perceptually coded for vowel quality and also measured for F1 and F2 at their vocalic midpoints. A total of 635 tokens, or 39.8 percent of the data, were coded as containing [e]. Of these, 85 tokens, or 13.3 percent, also contained a postvocalic nasal segment. The other 86.7 percent were entirely vocalic. The next most frequent vowel category was [a], with 439 tokens, or 27.5 percent of the total number of phonological FPs. Among this group, there was a nearly even split between [a] and [am], with 223 and 216 tokens, respectively. A total of 210 tokens, or 13.1 percent of the data, were coded as cases of [ə(m)]. Within this subset, [ə] was more frequent than [əm], with 132 cases of the former and 78 of the latter. Finally, 311 tokens, or 19.4 percent of the data, were coded as *cannot tell*. For these cases, it was not possible to confidently code the quality of the vocalic portion of the FP. Figure 17 illustrates the distribution of phonological FPs by speaker, focusing on vowel choice. Tokens of *cannot tell* are excluded. There is a significant amount of variance, across speakers, from the overall mean distribution ($\chi^2 = 791.6$, $p < .001$).

As was the case with the lexical FPs, significant differences between individuals are in and of themselves unsurprising. However, in contrast to variation in lexical FPs, which was not greatly illuminated by the study's contact measures, a clear pattern emerges in the use of phonological FPs. As contact intensity increases, speakers' preferences for particular vowels shift: rates of use of [e] decrease, while those of [a] and [ə] increase. Table 2 summarizes the results of correlation analyses assessing the relationship between each contact measure and rates of use of [e], [a], and [ə] in phonological FPs.

Figures 18–23 illustrate the results from Table 2, showing how these trends emerge across the study's speakers. Only measures that significantly correlate with the use of at least one vowel category are shown (i.e., *Both English and Spanish* with interlocutors and *Spanish proficiency* are excluded). Each point in the figures corresponds to an individual participant, the *x*-axis plots the contact measures, and the *y*-axis plots speakers' rates of use of the given vowel category.

To summarize, speakers who have spent a larger portion of their life in the United States, who use more English, and who do so with greater proficiency disfavor [e] when producing phonological FPs. Instead they prefer [a], and, to a lesser extent, [ə]. A useful method for examining the relationship between the contact measures and rates of use of particular phonological FPs is to plot a conditional inference tree (see Appendix). The tree in Figure 24 was generated by the binary partitioning algorithm included in the *party* package for R (Hothorn, Hornik, & Zeilis, 2006). It shows that *PLUS* and rates of *English only* provide the best model for predicting the use of [e] versus [a]/[ə] in phonological FPs. Here, the response variable is formulated in terms of centralization (with tokens of [a] and [ə] assigned a value of 'yes'). The predictor variables included in the model are all of the contact measures illustrated in Figures 18–23. The

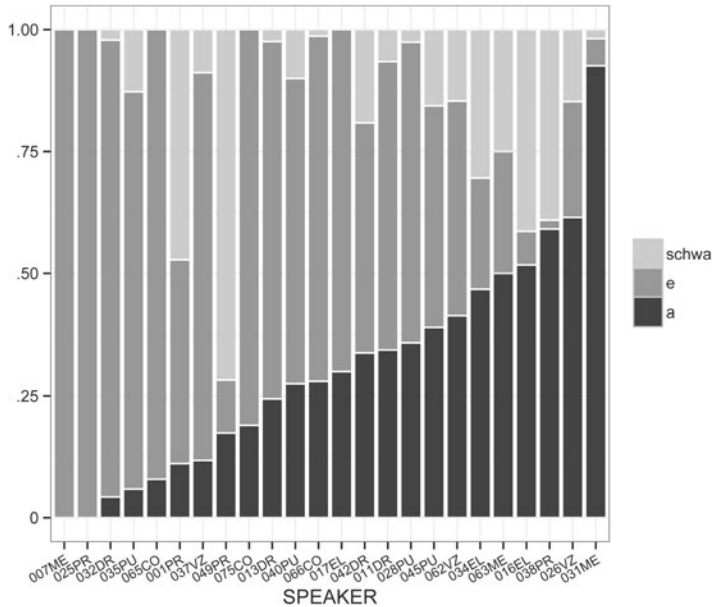

FIGURE 17. Phonological FP rates by type by speaker ('cannot tell' excluded).

cannot tell tokens are excluded. Rate of *English only* use (abbreviated in the tree as 'PERC_INTL_E_ONLY) and *PLUS* emerge as the only significant predictors.

The first split in the tree partitions the data in terms of *English only*, separating into the rightmost node in the figure the FPs produced by participants who speak only English with more than 36.5 percent of their interlocutors. Of the 265 phonological FPs produced by this group, 85 percent contain a centralized vowel (i.e., either [a] or [ə]). At node 2, the 1019 FPs produced by speakers with an *English only* value of less than 36.5 are partitioned in terms of *PLUS*. The leftmost node in the figure (node 3) contains the 636 phonological FPs produced by participants with an *English only* rate of less than 36.5 and a *PLUS* of less than 31.4. Of these FPs, only a quarter contains a centralized vowel. The final split in the tree is at node 4, where the data are once again partitioned in terms of *English only*. Particularly noteworthy is that the FPs binned in nodes 5 and 6 are all produced by speakers who have lived at least a third of their lives in the United States (i.e., with a *PLUS* of 31.4 or higher). However, those who use English less—that is, those with an *English only* rate of less than 19—have only a slightly higher rate of centralization than the participants with a *PLUS* of less than 31.4 (in node 3). In contrast, those whose *English only* rate is higher than 19 (but still lower than 36.5) produce centralized FPs at a rate of 75 percent, still lower than but comparable to the rate in observed in node 7.

To distill the tree's results, consider the following set of questions that one would want to ask in order to then predict the behavior of a speaker in the study. First and most importantly, what proportion of your family, friends, and

TABLE 2. *Correlations between the study's contact measures and rates of use of [e], [a], and [ə] in phonological FPs*

	[e]	[a]	[ə]
PLUS	**−.43***	.31	.33
% Interlocutors with whom participants speak			
English only	**−.64*****	**.43***	**.55****
Spanish only	**.56****	−.32	**−.57****
Both English and Spanish	−.04	−.06	.14
Self-reported language proficiency			
English	**−.44***	.23	**.46***
Spanish	.05	−.11	.04
% of domains in which Spanish used			
'Mucho' (a lot)	.34	−.07	**−.51****
'Poco' (a little)	−.34	.07	**.50****

Note: $*p \leq .05$, $**p \leq .01$, $***p \leq .001$. Bold values are statistically significant.

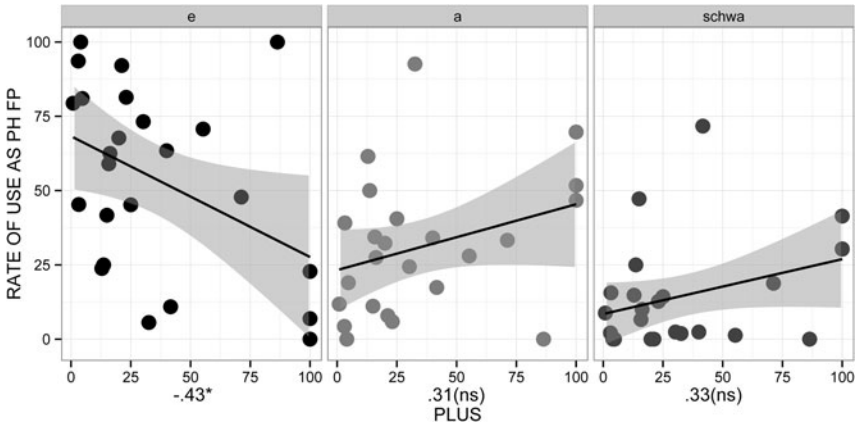

FIGURE 18. *PLUS* and rates of use of of [e], [a], and [ə] in phonological (PH) FPs. $*p < .05$. ns, not significant.

coworkers do you speak only English with? If you use only English with more than a third of these interlocutors, you are very likely to have a massive preference for centralized vowels and will rarely use [e] in your phonological FPs. However, if you use exclusively English with *less* than a third of your interlocutors, then it becomes important to know how much of your life have you lived in the United States. If it is less than a third of your life, you will strongly disfavor centralized vowels and will very frequently use [e] in phonological FPs. If, on the other hand, you have lived *more* than a third of your life in the United States, then how you use English once again becomes relevant. If you use English with more than a fifth of your interlocutors, you will strongly prefer centralized vowels in your FPs. Less than a fifth, and you will strongly disfavor using them.

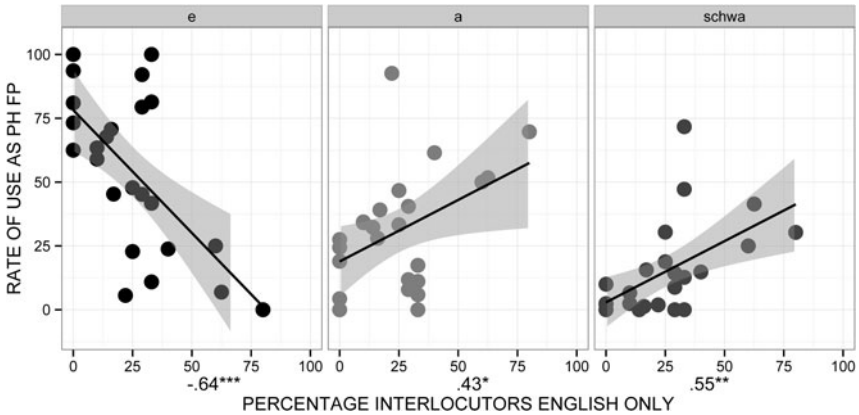

FIGURE 19. *Percentage of interlocutors with whom speaker uses English only* and rates of use of [e], [a], and [ə] in phonological (PH) FPs. $*p < .05$, $**p < .01$, $***p < .001$.

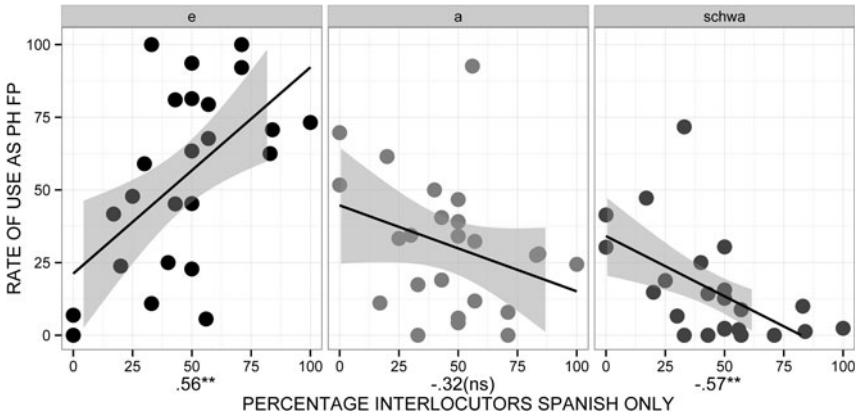

FIGURE 20. *Percentage of interlocutors with whom speaker uses Spanish only* and rates of use of [e], [a], and [ə] in phonological (PH) FPs. $**p < .01$. ns, not significant.

The acoustic data complement and enrich this picture. First, they indicate that *cannot tell* tokens are spectrally intermediate between [e] and the two central vowel categories. As a group, *cannot tell* tokens are higher in the vowel space than [a] and more fronted than [ə]. Consider Figure 25, which presents box plots for F1, F2, and duration across phonological FP types. Formant values were normalized across speakers using the norm.lobanov algorithm in the Vowels package for R (Kendall & Thomas, 2015). One-way analyses of variance show (unsurprisingly) that the vocalic portions of phonological FPs differ spectrally. Significant differences between types are observed for both F1 ($F(3, 1591) = 430.3$ $p < .001$) and F2 ($F(3, 1591) = 370.7$, $p < .001$.). Recall that the most robust predictor of the choice between lexical and phonological FPs was

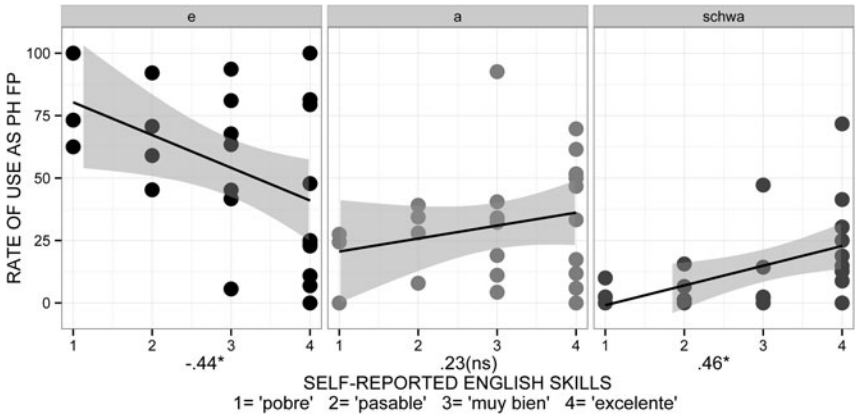

FIGURE 21. English skills and rates of use of [e], [a], and [ə] in phonological (PH) FPs. $*p < .05$. ns, not significant.

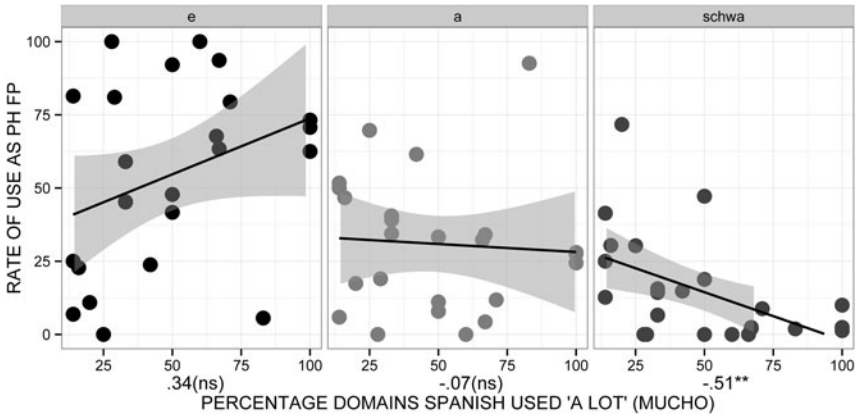

FIGURE 22. Spanish use in selected domains (MUCHO) and rates of use of [e], [a], and [ə] in phonological (PH) FPs. $**p < .01$. ns, not significant.

duration. Longer pauses favored the latter. Within phonological FPs, duration differences between types are not significant ($F(3, 1591) = 2.2$, $p = .09$).

At the acoustic level, a general pattern of centralization emerges with increasing *PLUS* and rates of *English only* that are consistent with the results we have presented. As *PLUS* increases, the F2 of the vocalic portion of phonological FPs significantly decreases ($r = -.271$, $p < .001$) while F1 increases ($r = .24$, $p < .001$). The same trends emerge for increased *English only;* F1 increases ($r = .41$, $p < .001$) and F2 decreases ($r = -.28$, $p < .001$). Figures 26 and 27 illustrate these trends at the token and speaker levels, respectively. At the token level, the darkness of each point indicates the *PLUS* and rate of *English only* of the speaker who produced it. At the speaker level, individual points represent

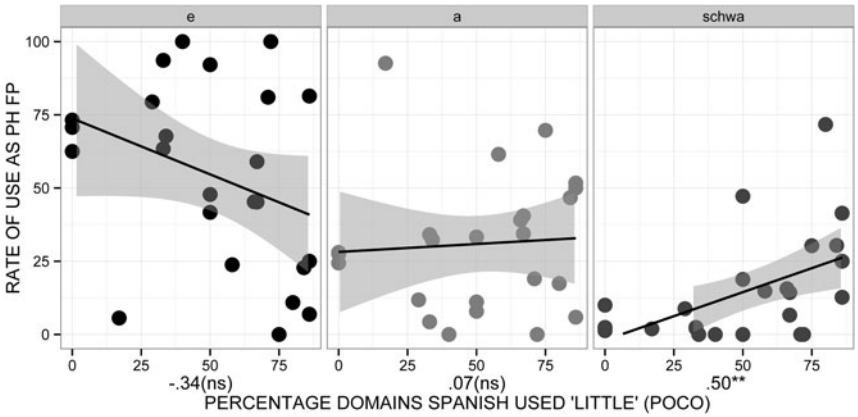

FIGURE 23. Spanish use in selected domains (POCO) and rates of use of [e], [a], and [ə] in phonological (PH) FPs. $^{**}p < .01$. ns, not significant.

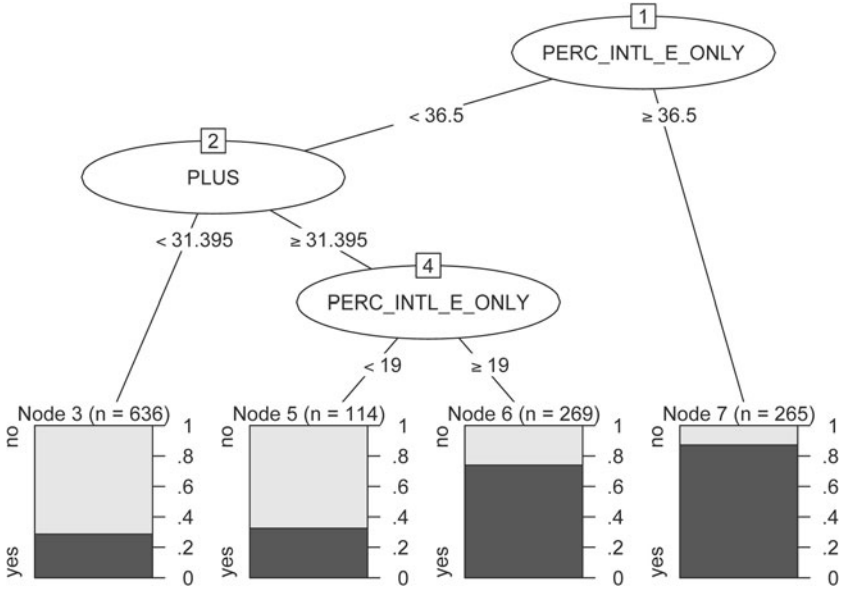

FIGURE 24. Conditional inference tree for phonological FP type: centralized versus [e] ('no').

speakers, and they are plotted according to the mean midpoint F1 and F2 of the vocalic portions of their phonological FPs.

In summary, a clear relationship exists between contact intensity and variation in phonological FPs. The strong preference for [e(m)] seen among those with more limited contact diminishes as contact intensity increases. Among the study's contact measures, *PLUS* and *English only* are the most robust predictors of

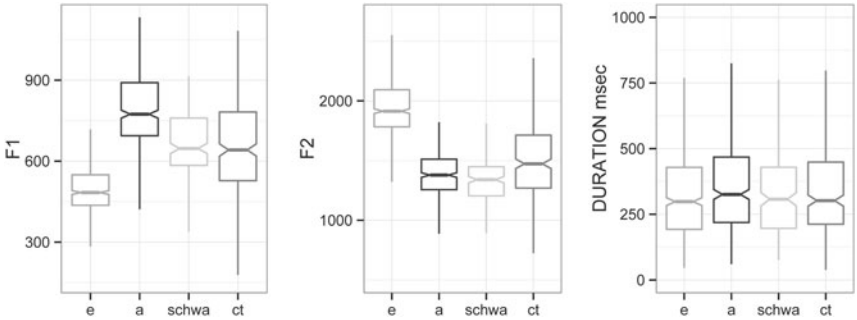

FIGURE 25. F1, F2 in hertz and *Duration* in milliseconds of phonological FPs.

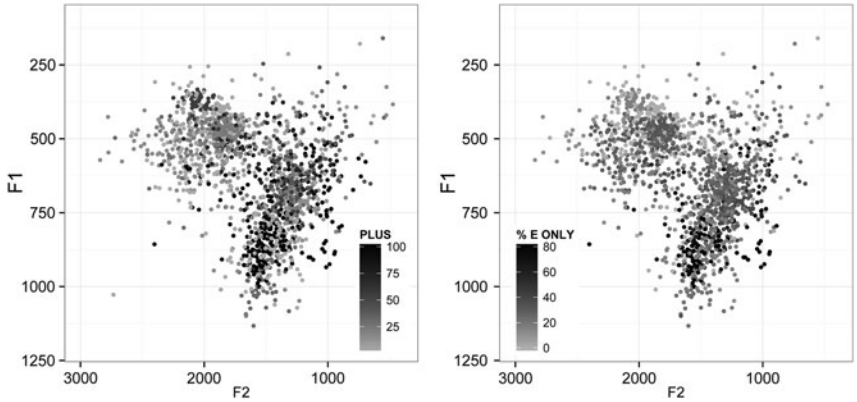

FIGURE 26. F1 and F2 (in hertz) of all phonological FP tokens by *PLUS* and *Percentage of interlocutors with whom speaker uses English only*.

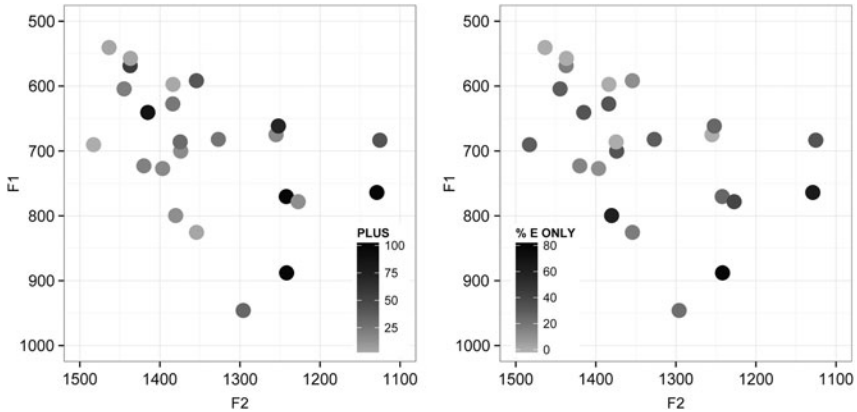

FIGURE 27. Mean F1 and F2 (in hertz) of phonological FPs by speaker, according to *PLUS* and *Percentage of interlocutors with whom speaker uses English only*.

variation in phonological FPs. As values for these two measures increase, so do speakers' preferences for centralized vowels in phonological FPs.

DISCUSSION

In recent years, several scholars (Poplack et al., 2006; Poplack & Levey, 2010; Sankoff, 2002; Torres Cacoullos & Travis, 2010, 2016) have expressed skepticism toward what they view as two prevalent assumptions about language contact. The first is that contact inevitably leads to language change. The second is that, given sufficient social and linguistic pressure, the potential nature of such change is essentially unconstrained. Of particular importance to critiques of these assumptions are the phenomena of lexical borrowing and code-switching. Both phenomena, these scholars argued, have been misunderstood as evidence of change and have been wrongly viewed as agents that promote structural convergence in contact situations. For instance, in their study of the use of English in Canada, Poplack et al. (2006:186) remarked that the notion that "lexical manifestations of contact function as agents of structural change, [is] an idea with no basis in scientific fact." Similarly, in their ongoing research into the linguistic behavior of Spanish-English bilinguals in New Mexico, Torres Cacoullos and Travis (2010:243) concluded, "Thus, despite prolonged contact, the data do not support Spanish convergence with English in this variety, nor code-switching as a mechanism of language change."

These scholars have suggested that arguments to the contrary—which is to say, claims in support of the inevitability of contact-induced change, and of the view that lexical borrowing and code-switching are catalysts for increased structural similarity between languages—are likely to emerge when any of a series of analytical mistakes are made. These include drawing conclusions on the basis of limited data and insufficiently empirical methods, mistaking patterns of variation for evidence of change, wrongly identifying the contact situation as the source of an actual change, and missing potential language-internal causes for purported change.

In addition to identifying such potential missteps, these researchers offer a way to avoid them by formulating and exemplifying an approach to contact linguistics designed to provide a rigorous method for establishing the presence of contact-induced change. This method is perhaps best characterized as "an approach to language contact grounded in sociolinguistics" (Sankoff, 2002:640). Indeed, as an empirical approach to the investigation of contact-induced change, Poplack et al. (2006) offered what is essentially a particular instantiation of the variationist paradigm (Labov, 1972), namely, one that compares pre- and postcontact varieties. For the comparison to be valid, however, several steps must be taken. The first is the identification of an appropriate contact community, one that is likely to be conducive to contact-induced change on the basis of several considerations, including intensity of contact, length of contact, status of the languages in the community, and size of the speaker population.

The next step is to identify a specific linguistic candidate for potential convergent change and systematically examine variation in its use, ideally as it occurs in the spontaneous speech of community members under relatively natural conditions. Then, the analyst must quantitatively model variation in the candidate feature, identifying the sets of linguistic and social factors that constrain its use. The final step is to compare constraint hierarchies across the relevant varieties (e.g., pre-, post-, and noncontact). Confidently inferring contact-induced change ultimately hinges upon the following determination: "Only when a candidate for change in a contact variety features a constraint hierarchy different from that of its pre-contact predecessor, but parallel to that of its presumed source, can we conclude in favor of contact-induced change" (Poplack & Levey, 2010:401).

The present study of FPs was designed with this comparative variationist approach to language contact in mind. However, several aspects of the study present obstacles to making a wholly satisfactory determination about contact-induced change. The most serious challenge is presented by the linguistic variable itself. As a candidate for potential convergence, Spanish FPs are problematic in that they are, at present, under-researched in general. This introduces uncertainty into the study's design in two ways. First, it complicates the operationalization of the variable context, making the question of what should count as variants of the variable part of the study itself. The field's limited understanding of variation in FPs also introduces uncertainty into the task of constructing constraint hierarchies. Because noncontact baselines for variation in Spanish FPs are not well established, the present study is in the position of actively searching for the potential effects of linguistic and social conditioning factors, rather than assessing the sensitivity of pre- and postcontact varieties to the influence of well-known constraints.

The study's results indicate that additional research will be necessary in both of these domains. With respect to the variable context, additional consideration of the relationship between lexical and phonological FPs will be required. Despite claims in the literature of functional equivalence between lexical and phonological FPs, and despite the fact that the lexical and phonological forms included in the study do share functional as well as formal properties, what shapes speakers' choices between them remains largely unclear. Results indicate that duration plays a role, such that phonological FPs are significantly longer than lexical FPs. There is also evidence that speakers are significantly more likely to repeat an immediately preceding FP if it is lexical rather than phonological. However, there is little evidence of a relationship between the choice of a lexical or phonological FP and the other factors in the study (e.g., those related to phonetic context and interlocutor FP use as well as the study's social factors and contact measures).

One possibility is that we have failed to identify the operative constraints on speakers' choices between lexical and phonological FPs. Some of the research discussed indicates that FPs occur with increased frequency before utterances that are more complex and novel. Perhaps these factors may also shape the choice between lexical and phonological FPs. Future analysis might therefore

include as potential constraining factors those that characterize additional aspects of the utterances that follow FPs, including their syntactic complexity and the relative frequency of their constituent lexical items. Another possibility is that lexical and phonological FPs are better viewed as separate phenomena, and that as such, they should not be assessed as part of a singular variable. This is certainly what is suggested by the separate analyses of the two subsets of data in relation to the study's contact measures. While there is clear evidence that variation in phonological FPs is shaped by these measures, such is not the case for lexical FPs. The closest that the latter come toward demonstrating a trend consistent with increased contact intensity is a reduction in FP inventory size.

The preceding concerns naturally create problems for the final step in the approach to contact linguistics that we have outlined, which is the construction of comparative constraint hierarchies. In the absence of a clear sense of the factors that shape variation, it is impossible to compare pre- and postcontact varieties in this way. Ultimately, what this study's results amount to is evidence of a relationship between contact intensity and variation in rates of use of a subset of the variants under analysis (i.e., the phonological FPs). Assessing contact outcomes on the basis of rates alone, rather than rates in tandem with constraints, is a weaker position from which to diagnose change and to determine whether or not it is contact-induced. That said, several aspects of the study's results support a view that interspeaker differences in the use of phonological FPs observed here constitute an instance of linguistic change, one that has the contact situation as its likeliest catalyst.

First, the data set is more than robust enough to support reliable quantitative analysis, and the methods for collecting and describing the data are standard practices in variationist research, in other words, data are drawn from spontaneous speech produced in sociolinguistic interviews, and vocalic variants were both perceptually coded as well as instrumentally measured for a number of acoustic properties. Furthermore, the study relies on a set of contact measures that assess critical dimensions of variation between members of the contact community, including language proficiency and usage as well as time spent in the contact setting. In other words, the study's results rest on a strong empirical base.

Second, the differences between speakers observed here are large in magnitude. Indeed, the significant correlations between contact measures and rates of use of [e(m)], [a(m)], and [ə(m)] are moderate to strong in nature, with coefficients ranging from .43 to .64. Furthermore, the conditional inference tree that emerges when the contact measures are considered simultaneously reveals very large differences in rates of use of central vowels. Speakers who report using only English with more than a third of their interlocutors use central vowels in 85 percent of their phonological FPs. In contrast, those who use only English with less than a third of interlocutors and who have also spent less than a third of their lives in the contact setting use central vowels in only 25 percent of their phonological FPs.

Third, the trends observed here are similar to others that have been interpreted as cases of contact-induced change in Spanish in the United States. These include studies that have shown greater contact with English to correlate with (a) higher rates of subject pronoun use and higher rates of pre- as opposed to postverbal subjects in general (Erker & Otheguy, 2016; Otheguy & Zentella, 2012,), (b) a stronger preference for subject-verb-object constituent order as well as reduced word order flexibility overall (Barrera-Tobón, 2013; Raña Risso, 2013), and (c) increased rates of use of indicative verbal morphology at the expense of the subjunctive (Bookhammer, 2013; Silva-Corvalán, 1994). Connecting these studies is their shared observation of what Poplack and Levey (2010:393) called an "overextension of recipient-language options into new contexts." However, while some of these patterns have been interpreted as cases of language-internal change *accelerated* by the contact situation, the shifts in phonological FP behavior observed here are unlikely to have a Spanish-internal source. Indeed, the most recent arrivals to Boston in the study overwhelmingly prefer [e] to [a], greatly weakening the chances that the shift toward frequent use of the latter represents an acceleration or intensification of an incipient Spanish-internal change. Furthermore, while the increased rates of use of [a] do indeed represent a minor use pattern developing into a major use pattern, we also observe the addition of a new category, namely, [ə]. Not only is this vowel typically absent from characterizations of the vocalic inventory of Spanish—that is, it is not a pre-existing recipient-language option—it also happens to be a sound that is routinely used to fill pauses in the contact setting's majority language.

In light of these aspects of the study's results, we conclude that the shifts in phonological FP behavior that correlate with increased contact intensity are best viewed as evidence of contact-induced change, and that this change amounts to an increase in similarity between Spanish and English along this dimension. That is, as participants experience contact more intensely, they become (in an apparent time sense) users of Spanish who, like users of English, prefer to use central vowels in phonological FPs. If this is ultimately the right analysis, then yet another consideration arises: what explains these results? That is, why should FPs constitute a site of linguistic convergence? And what is the specific mechanism driving this trend? These are questions that we can only briefly speculate on here.

One potential explanation is suggested by the *cognitive-load hypothesis*, which features prominently in Silva-Corvalán's (1994:6) research on Spanish speakers in Los Angeles: "In language contact situations, bilinguals develop strategies aimed at lightening the cognitive load of having to remember and use two different linguistic systems" (though see Dumont [2010] for a critique of this claim). With this in mind, the overextension of [a] at the expense of [e] might be viewed as a way of converging on a more generalized FP strategy, one that could be employed whether an individual was speaking in either Spanish or English, and thus be less cognitively burdensome. Of the three FP vowels, [a] represents a site of direct overlap in the phonological inventories of Spanish and English. Schwa, in contrast, is absent from noncontact varieties of Spanish, and the use of

monopthongal [e] by speakers of American English is restricted to particular regions in the United States, none of which include the Greater Boston Area. This formulation of the cognitive-load hypothesis would also predict that individuals with higher rates of [a] in FPs in Spanish would be more likely to use this sound in English FPs as well, perhaps at the expense of schwa. Ultimately, testing this idea would require an analysis of the FP behavior of study participants when speaking English, which is beyond the scope of the present investigation.

Another possible explanation of the study's results is suggested by usage-based approaches to linguistic inquiry. Bybee (2010:1) proposed that the cognitive processes underlying the structure of natural language "are called into play in every instance of language use; it is the repetitive use of these processes that has an impact on the cognitive representation of language and thus on language as it is manifested overtly." With this in mind, increased use of central vowels in Spanish FPs might be interpreted as a by-product of using English more frequently. By virtue of regularly using a language that recruits schwa as the primary vowel for phonological FPs, study participants might establish connections between exemplars of this category and those of the vowel category that is most like it acoustically and that is used to similar effect when speaking in Spanish, namely, [a] (and not [e]). In other words, it might be that increased experience with schwa, in the form of using and hearing it as an FP in English, might indirectly strengthen the representation and probability of use of [a(m)] as an FP when speaking Spanish. A potential source of support for this speculation may be the strength of the variable *English only* relative to that of *PLUS*. Not only is it the stronger of the two significant constraints, it also seems to modulate the effect of *PLUS*. Consider again the difference observed in rates of centralization among speakers who have spent more than a third of their lives in the United States (see Figure 24) but who have an *English only* score of less than 36.5. Among these speakers, those who use English less frequently (i.e., with less than a fifth of their interlocutors) have a centralization rate that is very similar to that of individuals who have lived less than a third of their life in the contact setting. Their infrequent use of English effectively neutralizes the effect of the duration of their experience in the contact setting. By comparison, those who use English more frequently (with more than a fifth but still less than a third their interlocutors) have a centralization rate of roughly 75 percent, which is similar to that of speakers who use only English with more than a third of their interlocutors. In other words, the study's results suggest that frequency of English use is of singular importance in shaping shifts in Spanish FPs.

CONCLUSION

This study reinforces the view that FPs constitute a site of linguistic variation that is worthy of inquiry and relevant to central theoretical questions asked by psychologists, cognitive scientists, and linguists alike. This study's results

motivate additional research on FPs and strongly suggest that an accounting of the units used for filling pauses would represent valuable additions to descriptive grammars of all languages as well as of regional and social varieties of languages. In addition, our results contribute to the field of contact linguistics. They show that FPs are sensitive to the pressures of language contact and that intense contact can lead to convergence in pause behavior.

The study also suggests a number of avenues for future research. These include a concerted effort to catalog regional variation in FPs across the Spanish-speaking world. Until a description of dialectal differences in FP behavior exists, an understanding of the outcomes of Spanish dialectal contact on FP behavior will remain out of reach. Additionally, the study suggests a number of potential experimental investigations. For instance, not only could the perception tasks discussed in the literature review be replicated for Spanish, it seems likely that investigating how bilinguals respond to different types of FPs could be valuable. That is, it could be fruitful to explore whether the attention of Spanish-English bilinguals is equivalently modulated when hearing [e], [a], or [ə]. Lastly, in highlighting the value of the comparative variationist approach to contact linguistics, the study strongly encourages its continued use in the ongoing study of Spanish in the United States.

NOTES

1. Some of the study's lexical FPs have been treated elsewhere as *discourse markers* (Aaron, 2004; Galué, 2002; Travis, 2005, 2006). As seen in the literature review, one of the primary functions of FPs is to signal discourse structure, suggesting that these categories may overlap. Additional commentary on the similarities and differences between FPs and discourse markers is offered in the data and methodology section.

2. Studies vary in how they represent FPs in prose, sometimes using standard orthography, International Phonetic Alphabet symbols, or a combination of both. We recount previous research using the labeling choices of the respective authors.

3. Linguistic research has also continued to progress on the related topic of conversational interaction (Couper-Kuhlen & Ono, 2007; Fox, Hayashi, & Jasperson, 1996). While such studies tend to focus on the broader issue of conversational repair (rather than on FPs exclusively), their results strongly support the view that hesitation phenomena are shaped by the structural properties of linguistic systems. They also illustrate how structural differences (e.g., syntactic left- or right-headedness, agglutinating or fusional morphology) can give rise to cross-linguistic variation in turn-taking and conversational interaction strategies.

4. This is not to say that the corpus lacks either instances of code-switching or the use of English-origin lexical items. Both occur in the corpus. However, interviewers were instructed to speak in Spanish, and, with the exception of an English-language reading passage, all of the study materials (project description, consent forms, questionnaires, etc.) were written and presented in Spanish. In this regard, the *Boston Spanish Corpus* is more akin to the *Otheguy and Zentella Corpus of Spanish in NYC* (Otheguy & Zentella, 2012) than it is to the *New Mexico Spanish-English Bilingual Corpus* of Torres Cacoullos and Travis (2010), which is a Spanish-English bilingual corpus by design.

5. Three other forms—*like*, *well*, and *pues*—were also considered as potential FPs but were ultimately excluded from the analysis due to low frequency. Also excluded were inflected verb forms: *tú sabes*, *mira*, *vea*, *I mean*, *you know*.

6. The regression was run using the lme4 package for R (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2014).

REFERENCES

Aaron, Jessi Elana. (2004). *So* and *entonces* in New Mexican bilingual discourse. *Spanish in Context* 1 (2):161–179.

Adell, Jordi, Escudero, David, & Bonafonte, Antonio. (2012). Production of filled pauses in concatenative speech synthesis based on the underlying fluent sentence. *Speech Communication* 54:459–476.

Bailey, Karl G.D., & Ferreira, Fernanda. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language* 49:183–200.

Barrera-Tobón, Carolina. (2013). *Contact-induced changes in word order and into nation in the Spanish of New York City bilinguals*. PhD dissertation, The Graduate Center of the City University of New York.

Bates, Douglas, Maechler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Bell, Alan, Jurafsky, Daniel, Fosler-Lussier, Eric, Girand, Cynthia, Gregory, Michelle, & Gildea, Daniel. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113(2):1001–1024.

Benevento, Nicole, & Dietrich, Amelia. (2015). I think, therefore *digo yo*: Variable position of the 1sg subject pronoun in New Mexican Spanish-English code-switching. *International Journal of Bilingualism* 19(4):407–422.

Boersma, Paul, & Weenink, David. (2016). Praat: Doing phonetics by computer [Computer program]. Version 6.0.14. Available at: http://www.praat.org/. Accessed February 11, 2016.

Bookhammer, Kevin. (2013). *The variable grammar of the Spanish subjunctive in second-generation bilinguals in New York City*. PhD dissertation, The Graduate Center of the City University of New York.

Bybee, Joan. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Clark, Herbert H., & Fox Tree, Jean E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition* 84 (1):73–111.

Corley, Martin, MacGregor, Lucy J., & Donaldson, David I. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105(3):658–668.

Corley, Martin, & Stewart, Oliver W. (2008). Hesitation disinfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass* 2(4):589–602.

Couper-Kuhlen, Elizabeth, & Tsuyoshi Ono. (2007). "Incrementing" in conversation: A comparison of practices in English, German and Japanese. *Pragmatics* 17(4):513–552.

Du Bois, John W., Chafe, Wallace L., Meyer, Charles, Thompson, Sandra A., Englebretson, Robert, & Martey, Nii. (2000–2005). *Santa Barbara corpus of spoken American English, parts 1–4*. Philadelphia: Linguistic Data Consortium.

Dumont, Jenny. (2010). Testing the Cognitive Load Hypothesis: Repair rates and usage in a bilingual community. *Studies in Hispanic and Lusophone Linguistics* 3(2):329–352.

Erker, Daniel, & Otheguy, Ricardo. (2016). Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish. *Lingua* 172:131–146.

Flores-Ferrán, Nydia. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case for English contact? *Language Variation and Change* 16(1):49–73.

Fox, Barbara, Hayashi, Makoto, & Jasperson, Robert. (1996). Resources and repair: A crosslinguistic study of syntax and repair. In E. Ochs, E. A. Schegloff & S. A. Thompson (eds.), *Interaction and grammar*. Cambridge: Cambridge University Press. 185–237.

Fox Tree, Jean. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition* 29:320–326.

Fruehwald, Josef. (2016). Filled pause choice as a sociolinguistic variable. *University of Pennsylvania Working Papers in Linguistics* 22(2):39–49.

Galué, Dexy. (2002). Marcadores conversacionales: un analisis pragmático. *Boletín de Lingüística* 18:27–48.

Goldman-Eisler, Frieda. (1961). A comparative study of two hesitation phenomena. *Language and Speech* 4(1):18–26.

Graham, Lamar A. (2013). Comparing hesitation markers in Sanjuanero Spanish. *Diálago de la Lengua* 5:66–77.

Hernández, José Esteban, & Baldazo, Rubén. (2013). Distribución variable de un marcador del discurso: No decir nada es como decir *verdad*. *La Revista Española de Lingüística* 43(2):113–150.

Hothorn, Torsten, Hornik, Kurt, & Zeileis, Achim, (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3):651–674.

Johnson, Teresa H., O'Connell, Daniel C., & Sabin, Edward J. (1979). Temporal analysis of English and Spanish narratives. *Bulletin of the Psychonomic Society* 13(6):347–350.

Kendall, Tyler, & Thomas, Erik R. (2015). Vowels: Vowel manipulation, normalization, and plotting package. Available at: http://lingtools.uoregon.edu/norm/. Accessed April 30, 2017.

Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

_____. (1978). Where does the linguistic variable stop? A response to Beatriz Lavandera. *Working Papers in Sociolinguistics* 44:4–22.

Liaw, Andy, & Wiener, Matthew. (2002). Classification and regression by randomForest. *R News* 2 (3):18–22.

Liberman, Mark. (2014). More on *um* and *uh*. Language log. August 3.Available at: http://languagelog.ldc.upenn.edu/. Accessed April 30, 2017.

Lipski, John. (1996). Patterns of pronominal evolution in Cuban-American bilinguals. In A. Roca, & J. B. Jensen (eds.), *Spanish in contact: Issues in bilingualism*. Somerville: Cascadilla. 159–186.

_____. (2005). Code-switching or borrowing? No sé so no puedo decir, you know. In L. Sayahi and M. Westmoreland (eds.), *Selected proceedings of the second workshop on Spanish sociolinguistics*. Somerville: Cascadilla Proceedings Project. 1–15.

MacClay, Howard, & Osgood, Charles E. (1959). Hesitation phenomena in spontaneous English speech. *Word* 15(1):19–44.

Mahl, George. (1956). Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology* 42:3–32.

O'Connell, Daniel C., & Kowal, Sabine. (2005). Uh and Um revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research* 34:555–576.

Otheguy, Ricardo, & Zentella, Ana Celia. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford: Oxford University Press.

Poplack, Shana, & Levey, Stephen. (2010). Contact-induced grammatical change: A cautionary tale. In P. Auer & J. E. Schmidt (eds.), *Language and space: An international handbook of linguistic variation*. Vol. 1. *Theories and methods*. Berlin: Mouton de Gruyter. 391–419.

Poplack, Shana, Walker, James A., & Malcolmson, Rebecca. (2006). An English "like no other"?: Language contact and change in Quebec. *Canadian Journal of Linguistics/Revue canadienne de Linguistique* 51(2):185–213.

Poplack, Shana, Zentz, Lauren, & Dion, Nathalie. (2012). Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition* 15(2):203–225.

R Core Team. (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Raña Risso, Rocío. (2013). *A corpus-based sociolinguistic study of subject pronoun placement*. PhD dissertation, The Graduate Center of the City University of New York.

Sankoff, Gillian. (2002). Linguistic outcomes of language contact. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 638–668.

Shriberg, Eric. (1999). Phonetic consequences of speech disfluency. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (eds.), *Proceedings of the International Congress of Phonetic Science (ICPhS), Symposium on the Phonetics of Spontaneous Speech (S. Greenberg and P. Keating, organizers), San Francisco, CA, USA* 1:619–622.

Silva-Corvalán, Carmen. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Oxford University Press.

Stouten, Frederik, Duchateau, Jacques, Martens, Jean-Pierre, & Wambacq, Patrick. (2006). Coping with disfluencies in spontaneous speech recognition: Acoustic detection and linguistic context manipulation. *Speech Communication* 48:1590–1606.

Swerts, Marc. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30:485–496.

Tagliamonte, Sali A., & Baayen, R. Harald. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2):135–178.

Torres, Lourdes, & Potowski, Kim. (2008). A comparative study of bilingual discourse markers in Chicago Mexican, Puerto Rican, and MexiRican Spanish. *International Journal of Bilingualism* 12 (4):263–279.

Torres Cacoullos, Rena, & Travis, Catherine E. (2010). Testing convergence via codeswitching: Priming and the structure of variable subject expression. *International Journal of Bilingualism* 15(3):241–267.

_____ (2016). Two languages, one effect: Structural priming in code-switching. *Bilingualism: Language, and Cognition* 19(4): 733–753.

Tottie, Gunnel. (2014). On the use of *uh* and *um* in American English. *Functions of Language* 21 (1):6–29.

_____. (2015). Uh and um in British and American English: Are they words? Evidence from co-occurrence with pauses. In N. Dion, A. Lapierre, & R. T. Cacoullos (eds.), *Linguistic variation: Confronting fact and theory*. New York: Routledge. 38–54.

Travis, Catherine E. (2005). *Discourse markers in Colombian Spanish: A study in polysemy*. Berlin: Walter de Gruyter.

_____ (2006). *The Natural Semantic Metalanguage approach to discourse markers. Approaches to discourse particles*. Oxford: Elsevier.

Watanabe, Michiko, Den, Yasuharu, Hirose, Keikichi, & Minematsu, Nobuaki. (2004). Types of clause boundaries and the frequencies of filled pauses. *Proceedings of the 18th Annual Convention of the Phonetic Society of Japan* 8:65–70.

Watanabe, Michiko, Hirose, Keikichi, Den, Yasuharu, & Minematsu, Nobuaki. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication* 50(2):81–94.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

APPENDIX

In conditional inference trees, splits are made on the basis of statistical hypothesis testing such that the data are partitioned into groups that are maximally internally homogeneous with respect to the response variable. Tagliamonte and Baayen (2012:159) provided a concise description of how partitioning proceeds:

> At any step of the recursive process of building such a tree, for each predictor, a test of independence of that predictor and the response is carried out. If the test indicates independence, then that predictor is useless for predicting the [response variable]. If the null hypothesis of independence is rejected, the predictor is apparently useful. If there are no useful predictors, the algorithm stops. If there is more than one useful predictor, the predictor with the strongest association with the response is selected.

To further probe the relative importance of the variables *English only* and *PLUS* to variation in FP vowel choice, we generated a random forest of 2000 trees (using the randomForest package in R; Liaw & Winer [2002]), with *Centralization* as the response variable and the study's contact measures as predictors. Figure 28 illustrates the results of this procedure. Variables are plotted according to *mean decrease in Gini*, which assesses a variable's utility in maximizing the homogeneity of terminal nodes in the trees. The larger the decrease in Gini, the more impactful the variable is in partitioning the data into purer subsets. The variable importance plot is consistent with the results of the single conditional inference tree presented in Figure 24, showing that *English only* and *PLUS* have the strongest associations with the response variable.

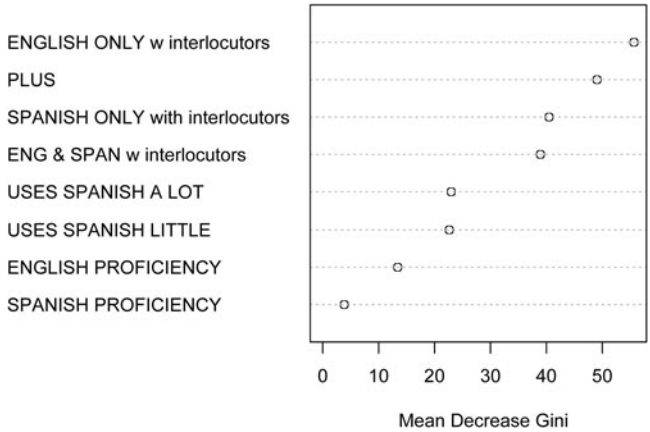

FIGURE 28. Random forest variable importance plot for vowel choice in phonological FPs.