

A variationist analysis of first-person-singular subject expression in Louisiana French

Aarnes Gudmestad* and Katie Carmichael

Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

*Corresponding author. E-mail: agudmest@vt.edu

Abstract

In this study, we investigate first-person-singular subject expression in Louisiana French. This variety is undergoing language death and features extreme variation, with twelve first-person-singular subject forms identified within our corpus. We demonstrate that variationist methods are robust for examining such variation in obsolescing languages, and we provide a model for undertaking such analyses. Examining different aspects of our data, we fit two mixed-effects models, one that analyzes the four most frequent phonological variants of the atonic pronoun *je* ‘I’ and the other that focuses on the tonic pronoun *mon* ‘me.’ Several linguistic and social factors predict the use of these subject forms, supporting the claim that variability in declining languages is systematic, just as variation in healthy languages is. We argue that variationist methodologies have contributions to make to research on obsolescing languages and that variationist examinations of endangered and minority languages can provide methodological and theoretical contributions to the study of language variation and change more broadly.

Keywords: Louisiana French; subject expression; variation; language death; variationist sociolinguistics

Research on endangered languages has generally focused on documenting the speech patterns of remaining speakers (e.g., Austin & Sallabank, 2011; Dorian, 1989); a central objective of this work has been to capture the array of typological diversity among languages of the world. In these accounts, linguistic variation across community members may be described, but it is rarely modeled using variationist methods, with some language death researchers going so far as to say that the variability found in language death contexts is inherently different from that found in healthy language contexts (e.g., Campbell & Muntzel, 1989; Cook, 1989).

This perspective has been challenged, however, by variationist sociolinguists such as Wolfram (2004:780) who have argued that “the variability typifying obsolescing forms is ... no different from the variability that characterizes healthy languages and language varieties,” further advocating for the extension of variationist approaches to language death contexts (777–79). Following up on this point, Kasstan (2019:29) called for “more research at the intersection of variationist sociolinguistics and the study of language obsolescence, so that interactions between linguistic decay, emergent variation and social meaning can be more clearly

integrated into contemporary models of language variation and change.” And Nagy (2017:58) asserted the value of examining smaller minority and endangered languages with a variationist lens, finding “the same sort of orderly heterogeneity frequently observed in large languages.”

In this vein, we aim to demonstrate the value of variationist paradigms—the systematic identification of attested variants and the modeling of their distribution across speakers—in the study of first-person-singular (1sg) subject expression in Louisiana French, which is undergoing gradual language death (Picone, 1997). In the variety of Louisiana French that we examine, there exist numerous phonological variants of the atonic pronoun *je*, pronounced [ʒə], due to allophonic variation between [ʒ] and [h], [s], [z], and [ʃ]. An unexpressed subject, Ø, may also occur, and the tonic pronoun *mon* ‘me’ may be variably appended to the clitic pronoun, or it may be used alone without an overt clitic pronoun (e.g., *mon était à l’école* ‘I was going to school,’ participant OFS2). Although previous research has investigated phonological variation in atonic pronouns (e.g., Dajko, 2009) and the rise of the tonic pronoun *mon* (Dubois, 2001; Rottet, 2005), there exists no research uniting these questions into a multivariate analysis of 1sg subject expression in Louisiana French—perhaps in part due to the wide array of variants attested. In taking on this challenge, our goal is to offer a proof of concept of the valuable methodological and theoretical contributions that this sort of analysis can make to research on language variation and change more broadly.

First, we aim to contribute to methodological discussions in variationist sociolinguistics by investigating a case of heightened variation—a linguistic variable that is made up of numerous variants—and by offering one way in which cases of increased variability can be analyzed quantitatively. Cases of extreme variation are not uncommon in contexts of language death (Dressler, 1972; Harrison & Anderson, 2008; King, 1989; Palosaari & Campbell, 2011; Schmidt, 1985; *inter alia*), though largely such cases have been examined via a typological documentation lens rather than a variationist one. That is, while variants are noted, with a special focus on typological frequency, their patterning and conditioning is not typically modeled. We argue in particular that our discussion of how we coded 1sg subject expression in Louisiana French may provide methodological tools to researchers in healthy language contexts as well, especially since variationist research has traditionally skewed toward recoding or simplifying the variants observed into a binary variable (Johnson, 2009). We thus build on work that demonstrates the value of documenting and modeling more than two variants (e.g., Gudmestad, Edmonds, Donaldson, & Carmichael, 2018; Szmrecsanyi, Biber, Egbert, & Franco, 2016) in order to provide a fuller picture of variation.

Additionally, we aim to demonstrate that variationist investigations of declining languages have the potential to contribute fruitful theoretical insights to research on language variation and change more broadly. We analyze data from interviews with twenty-nine Louisiana French speakers in a corpus including both fluent speakers and semi-speakers (cf., Dorian, 1973, 1977). Crucially, the inclusion of semi-speakers, or nonfluent speakers who never fully acquired the language due to insufficient input or opportunities to speak it, can shed light on the ways variation is acquired, especially in a context with limited input and no formal mechanisms

for acquiring the language. Moreover, in the case of endangered minority languages such as Louisiana French, the evolution of the language is subject to regular forces of language change, as well as external impetuses for change such as language contact. However, often in these contexts (apart from largescale revitalization efforts), there is less standardization pressure in terms of spoken and written language. In the community in which we situate our study, for example, very few speakers are literate in French (Dajko, 2009:72), and it is rare that they have extensive exposure to nonlocal varieties of French. This means, in some ways, such factors are controlled for, making this linguistic situation particularly informative for the ways that language change progresses in the absence of such external standardization pressures. We believe that the analyses we present offer evidence to indicate that the language-internal processes of variation and change in a context of obsolescence are governed by similar forces as healthy languages.

Background

Endangered languages

Language endangerment occurs when intergenerational transmission declines, as speakers of a language (or language variety) shift to another language as their primary mode of communication. In the case of gradual language death, or “[t]he loss of a language due to gradual shift to the dominant language in language-contact situations” (Campbell & Muntzel, 1989:184-85), the reasons for a shift are often social in nature, relating to the social value or prestige of the dominant language in contrast to the obsolescing one. In gradual language death, it is common for language shift to occur first in public domains and then later in private domains (such as conversations among family members in the home). There is also rarely institutional support for such minority languages, as they are not used in schools or other formal settings, and “[t]his weakening of normative pressures may contribute to the hypervariation in phonological, morphological, and syntactic features that appears to accompany, or at least be common in, the gradual dissipation of a speech community that language endangerment entails” (Harrison & Anderson, 2008:247). While language death literature has acknowledged the presence of variability (and even hypervariability) in language contact situations more broadly, it has been less common for researchers to use a variationist sociolinguistic lens on such situations. More recently, however, Blainey (2017), Nagy (2017), and Kasstan (2019) have advanced what we know about sociolinguistic variation in situations of language obsolescence, while acknowledging the challenges of applying variationist methods to such contexts.

Blainey (2017) drew attention to Weinreich’s (1974:3) assertion that investigations of language death must include social variables, advocating for their examination while also outlining the challenges for variationist researchers in language death situations. Of particular relevance to the current study are two issues. First, the specific social factors that are investigated depend on the community (3–4); thus, in some situations, extralinguistic variables that are not typically studied in healthy language communities need to be considered (e.g., exogamy). However, as Blainey noted, in the case of Louisiana French, there is evidence that traditional social factors like age

and gender are important. The second issue is that the multivariate statistical analyses that are a hallmark of variationist linguistics will only be possible if sufficient variation exists in the community (13).

Nagy stressed the importance of increasing the diversity of varieties examined in variationist sociolinguistics to include more investigations of declining languages. She argued that, “In addition to documenting these varieties, [variationist research of endangered languages] contributes to better understanding the processes of language variation and change in lesser-studied varieties, in order to see how well generalizations based on larger and better-documented languages can account for new types of data” (2017:34). In short, variationist studies of obsolescing languages have the potential to make theoretical contributions to sociolinguistics, as the findings that emerge from these investigations may strengthen or refute generalized knowledge that has been built on analyses of healthy varieties.

Kasstan (2019) sought to challenge the notion that endangered languages are monostylistic, an often cited but rarely tested notion within the language death literature. Via a Labovian approach to style-shifting, Kasstan demonstrated that palatalized variants of /l/ in Francoprovençal varied in systematic ways across speech styles (via a targeted elicitation task for speakers’ “best Francoprovençal” versus conversational speech). Supplementing this analysis with an examination of metalinguistic commentary, Kasstan argued that the innovation and spread of palatalized variants across different speaker communities was driven by social and stylistic factors—namely, the iconic (and distinctive from French /l/) nature of these variants, and their valorization by New Speakers within the Arpitan revitalization movement. In the current project, we build on these insights by contributing new empirical knowledge about a particular variable phenomenon in a declining language—1sg subject expression in Louisiana French.

Louisiana French

French came to be spoken in the Louisiana territory in the late 1600s, brought by European colonizers who were in search of an outlet to the Pacific Ocean. A number of varied indigenous groups were living in South Louisiana at that time, including modern-day Terrebonne and Lafourche parishes, where descendants of Houma, Chitimacha, and Biloxi-Chitimacha-Choctaw tribes continue to live along the bayous that represent their ancestral homelands. There is no documentation of the original indigenous languages spoken by these groups, as language shift toward the dominant colonial language, French, supplanted their use centuries ago (Dajko, 2009). These tribes represent some of the last remaining speakers of Louisiana French, as Francophone Louisiana has drastically shifted toward a monolingual English-speaking norm over the past century (e.g., Dubois & Horvath, 2000). At present, within Terrebonne and Lafourche Parishes, most fluent speakers of French are aged sixty and older. However, as is common within language death situations (cf., Dorian, 1977), there are a number of semi-speakers.

Some linguistic features that are common within Louisiana French derive from the regionally marked varieties of French brought by settlers (e.g., *asteur* ‘now’ rather than *maintenant*; *après+infinitive* to express present progressive; *nous-autres/*

vous-autres ‘we/y’all’ rather than *nous/vous*; allophonic variation between /ʒ/ and /h/) and are thus found in other Francophone locales, whereas others are less common in the Francophone world but fit patterns of regularization and simplification that frequently occur in situations of language death (e.g., verb regularization; lack of subjunctive mood) (Rottet, 2001). Regional variation within Louisiana includes a distinction between *quoi* ‘what’ (Western/‘Prairie’ dialects) and *qui* ‘who’ (Eastern/‘Bayou’ dialects); Eastern varieties also feature allophonic alternation between [ʒ] and [h], while Western varieties do not (Dajko & Blainey, 2016). Below we describe the variable forms of 1sg subject expression within the speech community of interest.

Subject expression

Twelve 1sg subject forms have been documented in the variety of Louisiana French spoken in Terrebonne and Lafourche Parishes (Carmichael & Gudmestad, 2019). Phonological variation in the clitic pronoun *je* ‘I’ includes realization as [ʒə] or [ʃə], as in other varieties of French, and less common allophones [sə], [zə], and [hə]. The variable use of the tonic pronoun *mon* ‘me’ as a subject also occurs. It may be used alone (as in *mon a pas changé* ‘I didn’t change, participant OFS7) or in conjunction with an atonic pronoun (as in *mon je veux faire* ‘Me, I [ʒə] want to do it’, participant SS1). Moreover, an unexpressed subject (i.e., the absence of an overt subject form) is attested (Rottet, 1996).¹

The variation in 1sg subject forms have been examined in different ways depending on researcher interests and on the variation present in the community in question. For example, in Terrebonne and Lafourche Parishes, Rottet (1996) noted phonological variants of [ʒ] in the atonic pronoun *je* as [ʃ], [s], [z], and [h], though he did not examine their patterning explicitly. He did, however, examine the patterning of tonic, atonic, and unexpressed subjects (*je*, *mon je*, *mon*, and \emptyset). Working on a variety of Louisiana French without as much phonological variation, Dubois (2001) focused on morphosyntactic variation with the presence or absence of the atonic and tonic pronouns (*je*, *moi je*, and *moi*²). Salmon (2007) and Dajko (2009) examined phonological variation between [ʒ], [h], and [z] more broadly, though each coded for 1sg contexts in particular since this is highest frequency environment for [ʒ]. These examinations have resulted in documentation of some linguistic and social patterning to this variation.

The [h] variant of the atonic pronoun was found to be more frequent when followed by a vowel compared to a consonant (Rottet, 1996) and also more frequent in casual speech than careful speech (Dajko, 2009). In contrast, [z] was found to be less frequent in casual speech by Dajko (2009), who argued that [z] was an identity marker for French-speaking Indians in the community, contrasting themselves with Cajuns. Rottet (1996) found that younger speakers used the tonic pronoun *mon* more often and atonic pronoun *je* less often than older speakers and argued that usage of *mon* alone was a relatively recent innovation in the community of Terrebonne-Lafourche. Rottet (1996) also noted an increase in unexpressed subjects among younger and less fluent speakers in the community. And Rottet (2001) found that men and Indians used *mon* more than women and Cajuns.

While these findings point to systematicity in the variable use of 1sg subject expression, they stem from univariate analyses, which leaves open the question as

to the ways in which multiple factors may work in concert. To our knowledge, Dubois (2001) is the only study to have conducted a multivariate analysis of subject expression in Louisiana French. Dubois investigated a western variety of Louisiana French, spoken in St. Landry Parish. She analyzed interview data gathered from thirty White/Cajun, not Indian, speakers and examined the tonic pronoun *moi*, the atonic pronoun *je*, and the use of both forms together, *moi je*.³ She conducted three separate multivariate analyses. In the analysis that focused on *moi*, for example, the tonic pronoun was compared to the use of *moi je*. She found that verb type influenced the use of the three subject forms. *Moi* alone was favored by regular verbs and the irregular verb *être* ‘to be’ and disfavored by the irregular verb *avoir* ‘to have’ and verbs of opinion and belief. The other two variants (*je* alone and *moi je*) were favored with verbs of opinion and belief (e.g., *moi connais* ‘I know’) and *avoir* and disfavored with regular verbs and *être*. Gender and age did not significantly predict the use of any 1sg subject form. The study also showed that only speakers with a weak degree of exposure to French used *moi* alone, which Dubois interpreted as indicative of their limited exposure to the language.

We seek to build on Dubois’s work by broadening the subject forms investigated to include the full spectrum of variants observed in Terrebonne-Lafourche varieties of Louisiana French (including phonological variants of *je* and the unexpressed subject) and by expanding the independent variables analyzed by looking to other variationist research (namely, subject expression in Spanish), in order to provide a more comprehensive account of the variation observed. We thus address the following research questions:

- (1) What 1sg subject forms does this group of Louisiana French speakers use?
- (2) What factors predict the realization of phonological 1sg subject pronouns in Louisiana French?
- (3) What factors predict the use of morphological 1sg subject pronouns in Louisiana French?

Method

Participants

All participants ($N = 29$) were speakers of Louisiana French, members of the Point-Au-Chien Indian Tribe, and residents of Terrebonne and Lafourche Parishes. Full participant information is provided in the [Appendix](#) and summarized here. Participants ranged in age from twenty-eight to seventy-three ($M = 47.6$, $SD = 12.7$). Their highest level of education ranged from second grade to some college. Fourteen participants were female. They identified their first language as English ($n = 6$), French ($n = 17$), or both languages ($n = 3$); three participants did not provide this information.

While most speakers, young and old, were fluent speakers of Louisiana French, we included in our sample eight semi-speakers, or nonfluent speakers who learned Louisiana French through intergenerational transmission.⁴ These speakers were identified as nonfluent either by themselves or by fellow community members, who characterized their speech as “*baroque*” (‘broken, strange’) or “*manière drôle*” (‘sort of

funny’) or “*pas bien*” (‘not good’). Previous research on semi-speakers has demonstrated that community members are generally able to make this distinction and identify less-proficient speakers of the minority language (Carmichael, 2007, 2017; Rottet, 1996). In fact, part of the definition of a semi-speaker is that their differences from the older fluent norm are noticed and looked upon as mistakes, unlike those of the younger fluent speakers (Dorian, 1977). The semi-speakers in this study span a wide range of ability. The most proficient speakers were able to carry on a full conversation, although leaning heavily on codeswitching into English and occasionally coining neologisms, while the speakers on the low end of proficiency were differentiated from passive bilinguals by the ability to produce novel utterances in French. Only speakers who were able to complete the majority of the interview using French phrases were included in this study. Whereas many studies of endangered languages do not include semi-speakers, others point to the significance of including them in research on language change, noting that the “last generation speakers of endangered languages, that is languages with small and dwindling speaker communities, can and do introduce grammatical and phonological innovations [...] semi-speakers should therefore never be ignored in doing fieldwork in endangered speech communities” (Harrison & Anderson, 2008:266–67).

Data

The data came from sociolinguistic interviews completed in 2007–2008 in which participants were prompted to share memories from childhood, family traditions, and stories about their lives. We coded and analyzed 15–45 minutes of each interview, selected based on the most fluid speech segments and narrative-style stories. The interviews were conducted in French by the second author of the paper and a Cajun (white) community member from the nearby town of Houma who spoke the local dialect fluently.

Data coding and analysis

The envelope of variation is defined functionally (Walker, 2010:13–14) as the subject position of 1sg verbs, and we coded the specific 1sg subject forms that occurred in these contexts in the interview segments ($N = 2,131$). We recognize that the ideal variationist analysis may be one in which all 1sg subject forms are examined in one statistical model because a single analysis would enable us to investigate 1sg subject expression as a unified phenomenon. However, the large number of variants that occurred in our dataset and the low frequency with which the participants used some of these forms (see *Results*)—two characteristics that are common in declining languages (e.g., Harrison & Anderson, 2008; Kasstan, 2019; Palosaari & Campbell, 2011)—meant that a single statistical model that included all variants observed our dataset was not possible. To capture as many variants as possible, we opted to perform two separate analyses, each with a different dependent variable, and distinguished between phonological and morphological pronouns in the analyses. The first dependent variable focused on phonological variants, for which we analyzed the forms that occurred at least fifteen percent of the time in the dataset. The four

variants were unexpressed subject, [ʒ], [h], and [ʃ]. The second dependent variable considered morphological variation and the tonic pronoun *mon*. The two categories were the presence and absence of *mon*. We later suggest that fitting multinomial regression models or fitting multiple regression models that focus on different features of a variable phenomenon (phonological and morphological variants in the case of the current study) might prove fruitful in the case of healthy languages with many variants as well (e.g., Gudmestad et al., 2018).

We tested ten linguistic and social factors as fixed effects in the models. The social factors were age, gender, and fluency. Speaker age and gender have been shown to be connected to aspects of variation that affect 1sg subject expression in Louisiana French—in terms of both allophonic (Dajko, 2009; Salmon, 2007) and morphosyntactic (Rottet, 1996) variation. We investigated age as a continuous factor and gender as binary (women, men). Given the importance of fluency in understanding language practices in situations of language death (Dorian, 1978; Dubois & Noetzel, 2005), we examined possible differences between fluent speakers and semi-speakers, categorized according to the community-based descriptions above.

Concerning the linguistic factors, we drew on two strands of scholarship. One is previous research on subject expression in Louisiana French, from which we identified four factors: verb category, following sound, tonic pronoun, and atonic pronoun. Verb category constitutes a simplified version of the verb type factor investigated by Dubois (2001). The “er” category consists of regular verbs that end in *er* in the infinitive (e.g., *parler* ‘to talk,’ *dancer* ‘to dance’). The ‘other’ verbs in our dataset generally have suppletion in their conjugation (e.g., *aller* ‘to go,’ *pouvoir* ‘to be able,’ *savoir* ‘to know’), though there are exceptions (e.g., *partir* ‘to leave,’ *apprendre* ‘to learn,’ *vivir* ‘to live’). Following sound was also included as a predictor, with the sound coming after the dependent variable coded as a consonant or vowel (Rottet, 1996). Tonic pronoun was included in the phonological model only; it distinguished between the presence and absence of the tonic pronoun *mon*. Similarly, the atonic pronoun factor was analyzed in the morphological model only and differentiated between the presence and absence of an atonic pronoun.

The other area is variationist research on subject expression in Spanish (e.g., Carvalho, Orozco, & Shin, 2015). Although French is not typically a pro-drop language, Spanish is. Two factors found to be important predictors across studies are referent continuity and perseveration. Referent continuity (also called switch reference) concerns the subject form of the preceding tensed verb and whether this form is the same or different from the context being analyzed (see Nagy [2015] for a similar effect on subject expression in other languages). The preceding tense verb could be produced by the participant or an interviewer. Unexpressed subjects are more likely when there is no change in referent and overt subjects are more likely when there is a change in referent (e.g., Otheguy & Zentella, 2012, chapter 8). Perseveration (also called priming and coded differently across investigations) identifies the subject form of the previous mention of the same referent and determines whether the subject form is similar or different from the context being analyzed (e.g., Cameron & Flores-Ferrán, 2004). The final linguistic factor is the frequency of the verb in the interview, for which we analyzed the natural logarithm of the frequency scores as a continuous factor. Frequency of the verb has been found to

have an independent effect on subject expression (e.g., Bayley, Greer, & Holland, 2013) and a mediating effect on other significant constraints (e.g., Erker & Guy, 2012).

Examples of our data and coding are available in (1) and (2). The subject form analyzed in these examples is in bold; the subject form of the preceding tensed verb is in all capital letters (i.e., referent continuity), and the previous mention of the same referent is underlined (i.e., perseveration). It is worth noting that the subject form analyzed for referent continuity and perseveration is the same in example (1).

- (1) OFS 5: *JE crois ça **je** l'ai rencontré une fois mais* 'I believe that I found it once but'
dependent variable: [h]

age: 54

fluency: fluent

gender: female

tonic pronoun: absent

verb category: other

following sound: consonant

perseveration: different ([f])

referent continuity: same (*je*)

log-verb frequency: 2.615

- (2) SS 7: *ouais j'étais trop fou* 'yeah I was very crazy'

Interviewer: *TU faisais le party toujours et* 'you always used to party and'

SS 7 : ***mon** j'ai jamais arrêté ça* so {laugh} 'I never stopped that so {laugh}'=

dependent variable: presence of *mon*

age: 39

fluency: semi-speaker

gender: male

atonic pronoun: present

verb category: other

following sound: vowel

perseveration: different ([z])

referent continuity: switch (*tu*)

log-verb frequency: 2.348

We fit two mixed-effects regression models using SAS® software⁵—one for the phonological dependent variable and one for the morphological dependent variable. Each model included a random effect for participant, enabling us to account for individual variability. In the mixed-effects models, the nominal variables had a reference-point category against which the nonreference point categories were compared. We selected the unexpressed subject as the reference point for the phonological dependent variable because it allowed us to examine differences between the use of overt clitic pronouns against a non-overt subject. For the morphological dependent variable, we modeled the log-odds of the presence of *mon*. The reference points for the nominal independent variables were female (gender), fluent (fluency), *er* verbs (verb category), consonant (following sound), tonic pronoun (absent), atonic pronoun (absent), (same) referent continuity, and (different) perseveration. Age and log-verb frequency

were continuous factors, so they did not have a reference point. In order to determine whether each model fit the observed data well, we computed the proportions of outcomes that were correctly predicted by each model and compared that prediction to a corresponding null model (i.e., a model with only the dependent variable). Moreover, when fitting both mixed-effects models, we employed the Kenward-Roger method for computing the denominator degrees of freedom and the restricted maximum likelihood method to compute the estimated values of the fixed and random effects (e.g., McNeish, 2017). These methods were important for our analyses because they enabled us to reduce the rate of type-I error, reduce bias in the estimation of random effect, and include all participants in the analysis, even when the amount of data they contributed to the dataset was limited.

Results

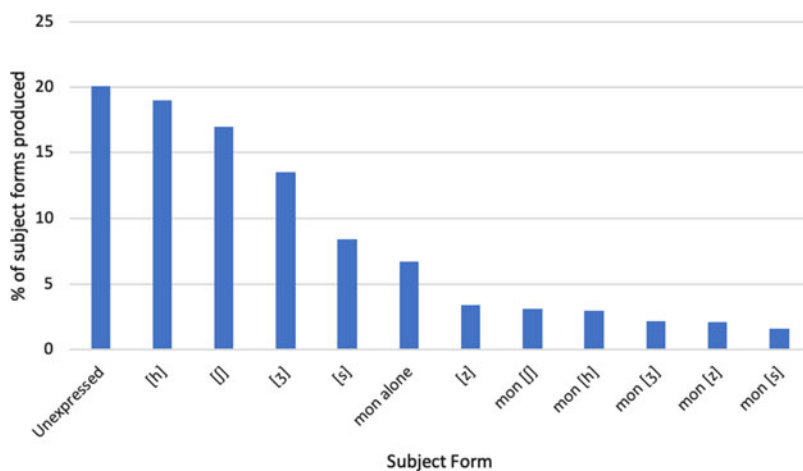
We organize the results around the research questions. We begin with the first question concerning the subject forms produced by the participants in this sample. The repertoire of subject forms includes twelve different forms: unexpressed subjects (\emptyset), five bare clitic forms ([h], [ʃ], [ʒ], [s], [z]), five doubled clitic forms (*mon* [h], *mon* [ʒ], *mon* [z], *mon* [s]), and the tonic pronoun *mon* without an accompanying atonic pronoun. Table 1 provides the frequency of use of these forms and distinguishes between phonological and morphological variants, and Figure 1 illustrates the distribution of all subject forms together. When all forms are considered collectively, the most frequent form was the unexpressed subject, occurring in twenty percent of the 1sg subject contexts that we analyzed. The three other forms that were used at least ten percent of the time were overt atonic subjects: [h], [ʃ], and [ʒ]. When the participants used *mon*, they used it more often with an overt clitic pronoun ($n = 254$) than alone (142 contexts). Overall, however, it was much more common for speakers to not use *mon* ($n = 1735$).

To provide a visual representation of the distribution of some of the more frequent forms among individuals, Figure 2 illustrates the realizations of tonic and atonic pronouns for each speaker. It can be observed that fluent speakers (participant aliases that begin with “OFS” and “YFS”) tended to use overt atonic pronouns more than semi-speakers (participant aliases that begin with “SS”), and although *mon* (alone and with clitic) was found across a number of speakers, it appeared in highest rates among semi-speakers. There are also some semi-speakers who never use the phonological variant [h].

The patterns in Figures 1 and 2 echo some previous findings and challenge others. To begin with, we identified the same twelve forms that occurred in data elicited from a translation task in Carmichael and Gudmestad (2019), suggesting internal validity of this dataset and the particular forms that are in common use in this community regardless of task type. We also see many of the variants identified in previous work on Louisiana French, although the rates at which they appear differ in various ways. In particular Dubois (2001) argued that *mon* alone was the result of limited exposure to the language; since we see its use distributed across fluent and nonfluent speakers alike, our data challenges this interpretation—at least for this particular dialect of Louisiana French. Also noteworthy, the most frequent form in the dataset was

Table 1. Frequency of 1sg subject forms

Phonological variant	<i>n</i>	%	Morphological variant	<i>n</i>	%
∅	429	25	<i>mon</i> alone	142	36
[h]	405	23	<i>mon</i> [ʃ]	65	16
[ʃ]	362	21	<i>mon</i> [h]	63	16
[ʒ]	288	17	<i>mon</i> [ʒ]	47	12
[s]	178	10	<i>mon</i> [z]	45	11
[z]	73	4	<i>mon</i> [s]	34	9
Total	1735	100	Total	396	100

**Figure 1.** Distribution of 1sg subject forms ($N=2131$; see Table 1 for token counts.)

the unexpressed subject, occurring 20% of the time. This is a notable increase in the unexpressed subject compared to Rottet's (1996) findings with data collected in this same community using similar methods 15 years earlier, in which he found an overall rate of 2.7% unexpressed subjects (with his youngest speakers, a mix of fluent and semi-speakers who were under thirty years of age in 1993, featuring the highest rate of unexpressed subjects in his corpus at 7%). This shift toward an increase in unexpressed subjects cannot be explained by contact with English, where the subject is typically expressed, nor can it be explained by limited exposure since it is present in even our oldest fluent speakers. Thus, it appears that these data suggest the advancement of the change in progress originally noted by Rottet (1996), taken up by both fluent and nonfluent speakers.

Next, we turn to the two mixed-effects models, in which each dependent variable and nominal independent variable have a reference point and the other category (or categories) of these variables is (are) compared to the reference point. A positive

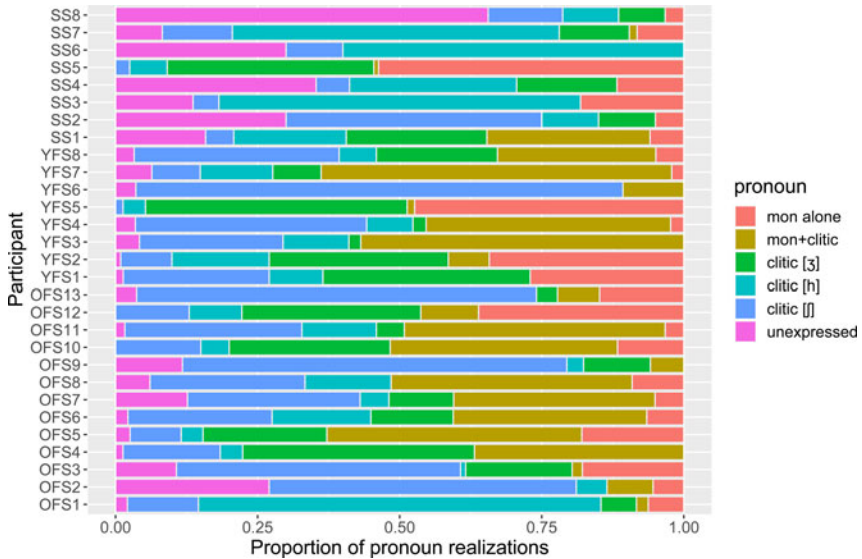


Figure 2. Pronoun realization ($n = 1880$) across participants ($N = 29$).

estimate signifies a higher log-odds of using a nonreference point category (compared to the reference point), whereas a negative estimate denotes a lower log-odds of using a nonreference point category. The p -value shows whether the effect is significant ($\alpha = 0.05$; significance is when $p < \alpha$).

The first mixed-effects model pertained to the second research question and examined the most frequent phonological pronouns: \emptyset , [h], [ʃ], and [ʒ] ($n = 1,801$). This analysis included the use of these subject forms with and without *mon*. The frequency of the four variants was as follows: \emptyset ($n = 571$, 32%), [h] ($n = 468$, 26%), [ʃ] ($n = 427$, 24%), and [ʒ] ($n = 335$, 19%). Thus, this model has a multinomial dependent variable. With this type of regression, the reference point of the dependent variable (\emptyset in this case) is compared to each of the nonreference-point categories. This model revealed that tonic pronoun, verb category, following sound, speaker gender, and fluency predicted the use of the phonological variants (see Table 2 for the distribution of the significant independent variables across the variants). None of the significant factors were strongly correlated. Age, referent continuity, perseveration, and log-verb frequency were not significant. We explored interactions between the significant social factors (gender and fluency). They were strongly correlated with significant fixed effects, so they were removed from the model. The results for the fixed effects are available in Tables 3 through 5 and the details of the random effect are in Table 6. This model accurately predicted 67% of the data, whereas a null model correctly predicted 32% of the data. Thus, the model illustrated in Tables 3 through 6 does a good job of fitting the data.

While there is a separate table for each comparison in the regression (Table 3: \emptyset versus [h], Table 4: \emptyset versus [ʃ], and Table 5: \emptyset versus [ʒ]), it is important to note that these results come from a single mixed-effects model. To facilitate the

Table 2. Frequency of the phonological variants across categories of significant fixed effects

Fixed effect	Category	∅		[h]		[ʃ]		[ʒ]	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Tonic pronoun	Absent	429	29	405	27	362	24	288	19
	Present	142	45	63	20	65	21	47	15
Verb category	<i>er</i>	100	33	55	18	72	24	75	25
	Other	471	31	413	28	355	24	260	17
Following sound	Consonant	328	37	32	4	393	44	144	16
	Vowel	243	27	436	48	34	4	191	21
Gender	Female	250	27	351	38	197	22	120	13
	Male	321	36	117	13	230	26	215	24
Fluency	Fluent	453	32	432	30	314	22	226	16
	Semi-speaker	118	31	36	10	113	30	109	29

Table 3. Details of the fixed effects in the phonological regression model: \emptyset versus [h]

Effect	Estimate	SE	DF	<i>p</i>	CI
Intercept	-1.19	0.44	43.25	0.011	[-2.078, -0.292]
Tonic pronoun [absent]					
Present	-0.50	0.22	1783	0.023	[-0.940, -0.068]
Verb category [er]					
Other	-0.36	0.24	1783	0.125	[-0.827, 0.101]
Following sound [consonant]					
Vowel	3.07	0.21	1783	<0.0001	[2.658, 3.483]
Gender [female]					
Male	-1.90	0.50	25.6	0.001	[-2.929, -0.870]
Fluency [fluent]					
Semi-speaker	-1.85	0.64	35.81	0.007	[-3.143, -0.546]

Note. The reference point for the dependent variable is the unexpressed subject. The reference points for the nominal independent variables are bracketed.

Table 4. Details of the fixed effects in the phonological regression model: \emptyset versus [ʃ]

Effect	Estimate	SE	DF	<i>p</i>	CI
Intercept	-0.44	0.38	36.06	0.259	[-1.212, 0.336]
Tonic pronoun [absent]					
Present	-0.39	0.20	1783	0.048	[-0.770, -0.004]
Verb category [er]					
Other	0.42	0.18	1783	0.019	[0.068, 0.768]
Following sound [consonant]					
Vowel	-1.92	0.20	1783	<0.0001	[-2.303, -1.532]
Gender [female]					
Male	-0.16	0.45	26.14	0.721	[-1.080, 0.757]
Fluency [fluent]					
Semi-speaker	0.43	0.51	28.03	0.409	[-0.618, 1.475]

Note. The reference point for the dependent variable is the unexpressed subject. The reference points for the independent variables are bracketed.

understanding that the results come from one mixed-effects model, we discuss each significant independent variable in turn. Beginning with tonic pronoun, the model revealed that speakers were less likely to use [h] and [ʃ] compared to \emptyset when the tonic pronoun *mon* was present (versus absent). The variant [ʒ] patterned in the same direction as the other overt pronouns but was not significant. The findings

Table 5. Details of the fixed effects in the phonological regression model: \emptyset versus [ʒ]

Effect	Estimate	SE	DF	p	CI
Intercept	-0.44	0.38	36.06	0.259	[-1.212, 0.336]
Tonic pronoun [absent]					
Present	-0.38	0.20	1783	0.061	[-0.766, 0.017]
Verb category [er]					
Other	-0.48	0.19	1783	0.014	[-0.839, -0.096]
Following sound [consonant]					
Vowel	1.02	0.15	1783	<0.0001	[0.719, 1.317]
Gender [female]					
Male	-0.16	0.45	26.14	0.721	[-1.08, 0.757]
Fluency [fluent]					
Semi-speaker	0.39	0.60	28.81	0.521	[-0.842, 1.625]

Note. The reference point for the dependent variable is the unexpressed subject. The reference points for the independent variables are bracketed.

for this factor suggest the use of an overt phonological pronoun discourages the use of *mon*.

For verb category, the findings differed for each comparison. Speakers were more likely to use [ʃ] with verbs other than *er* verbs. The opposite pattern was observed for [ʒ]. They were less likely to use [ʒ] versus an unexpressed subject with other category verbs compared to *er* verbs. This factor was not significant for [h] compared to \emptyset . Thus, despite similarities among the phonological variants with the tonic pronoun variable, the results for verb category suggest that the phonological variants may exhibit different linguistic patterns of use.

The third significant linguistic factor was the following sound, which constrained the use of each phonological variant in the mixed-effects model. Regarding the comparisons between [h] and \emptyset and between [ʒ] and \emptyset , the participants exhibited a higher likelihood of using [h] and [ʒ] when the following sound was a vowel compared to a consonant. They were also less likely to use [ʃ] when the following sound was a vowel compared to a consonant. These findings align with previous documentation of Louisiana French used in this region (Rottet, 2001).

Regarding gender, men were less likely than women to use [h] compared to \emptyset , and there were no significant differences between [ʃ] or [ʒ] and \emptyset . Salmon (2007) and Dajko (2009) also found men to be less likely to use [h] than women in analyses of this variation beyond the 1sg subjects, which they argued was a reflection of women maintaining the French language longer than men due to gendered distributions of labor that meant men worked in (predominantly Anglophone) positions outside the home. Thus, the present investigation adds to the evidence that [h] represents a gendered form of variation within the community, contributing support for the claim that linguistic variables can take on social

Table 6. Results for the random effect in the phonological mixed-effects model

Participant	[h] intercept	[ʃ] intercept	[ʒ] intercept
OFS1	-1.086	-0.123	1.148
OFS2	-1.091	-1.191	-1.083
OFS3	-1.163	-0.348	-0.250
OFS4	0.380	1.492	-0.626
OFS5	0.285	1.068	0.721
OFS6	0.014	0.082	-0.354
OFS7	1.541	-0.358	-0.545
OFS8	0.129	-1.200	-0.402
OFS9	-1.189	-0.619	-1.494
OFS10	0.063	1.228	0.145
OFS11	2.522	-0.739	-0.398
OFS12	0.722	1.154	1.923
OFS13	-0.958	-1.124	0.663
YFS1	-0.831	1.263	1.016
YFS2	-0.726	1.603	2.206
YFS3	0.735	-1.188	-0.927
YFS4	-0.104	-1.486	-1.244
YFS5	-0.039	1.551	3.081
YFS6	-1.009	-1.489	-1.576
YFS7	0.651	0.221	-0.140
YFS8	1.153	0.204	-0.537
SS1	2.430	0.294	0.069
SS2	-0.559	-0.631	-1.155
SS3	-0.278	0.992	0.944
SS4	-0.801	0.162	0.021
SS5	0.396	0.924	2.506
SS6	-0.128	0.658	-0.715
SS7	0.632	0.746	-0.109
SS8	-1.692	-1.161	-1.561

meaning in obsolescing languages (see Kasstan [2019] for additional evidence in Francoprovençal).

The final significant factor was fluency. Semi-speakers were less likely than fluent speakers to use [h] compared to Ø; no significant differences were found for the other comparisons in the model. Thus, although it was reasonable to expect that fluency

modulated the use of 1sg subjects (e.g., Dubois & Noetzel, 2005), we found that its effect on the phonological variants was limited to certain forms, specifically the comparison of [h] and Ø. Notably, this patterning for [h] is not unexpected given prior research. In a broader examination of the [h] allophone that was not limited to its use in subject expression, Carmichael (2017) noted that this variant was exceedingly infrequent in the speech of Point-Au-Chien Indian semi-speakers, which she argued was evidence that this form of variation was not acquired by some speakers as a result of increased exogamy and movement away from the predominantly Francophone environments down the bayou. Our analysis builds on Carmichael's findings. However, to clarify, while this particular form of variation was not acquired by all semi-speakers, several other forms were, highlighting the importance of considering semi-speaker data in language death situations.

Lastly, we fit a mixed-effects model to investigate the use of the tonic pronoun *mon* and to answer the third research question. For this model, we analyzed the entire dataset ($N = 2,131$). The frequency of the two categories of the dependent variable was as follows: presence of *mon* ($n = 396$, 19%) and absence of *mon* ($n = 1735$, 81%). In this model we found that atonic pronoun, following sound, referent continuity, perseveration, and fluency were significant (see Table 7 for the distribution of the significant, independent variables across the categories of the dependent variable). Tables 8 and 9 provide the results for the fixed and random effects, respectively. Verb category, log-verb frequency, gender, and age were not significant.

Concerning the significant linguistic factors, the speakers were less likely to use *mon* when an atonic pronoun was present. This result aligns with the findings from the phonological model, such that when speakers produce an overt subject form, they tend to use either an atonic or tonic subject, not both (thus, the subject form *mon je*—regardless of phonological realization of *je*—was dispreferred). The participants were more likely to use *mon* when the following sound was a vowel. To our knowledge, the present investigation is the first to examine the role of this factor on the use of the *mon*—initially coded for due to our phonological considerations in this study. That is, by expanding our definition of subject expression to include phonological and morphosyntactic variants, we were able to uncover a pattern that was not previously tested for. Participants were also more likely to use *mon* when there was a switch in referent from the preceding tensed verb (referent continuity). Research on Spanish has consistently reported that overt pronouns are more likely than unexpressed subjects in switch-reference contexts (e.g., Carvalho et al., 2015). The current study's findings share similarities with this trend in Spanish, though the results are not identical. Foremost, the phonological model did not reveal a significant effect for referent continuity, meaning that overt atonic pronouns were not more likely to occur than unexpressed subjects in contexts of switch reference. However, the significant results in the tonic model suggest that *mon* was favored in contexts of switch reference. For the final linguistic factor, perseveration, the participants were less likely to use *mon* when the subject form of the previous mention of the same referent was also *mon*, a finding that diverges from findings on Spanish and that invites future research. Lastly, fluency was the sole constraining social factor; semi-speakers were more likely to use *mon* than fluent speakers. That is, while *mon* is not restricted

Table 7. Frequency of the morphological variants across categories of significant fixed effects

Fixed effect	Category	Presence of <i>mon</i>		Absence of <i>mon</i>	
		n	%	n	%
Atonic pronoun	Absent	142	25	428	75
	Present	254	16	1307	84
Following sound	Consonant	191	17	949	83
	Vowel	205	21	786	79
Referent continuity	Same	78	13	536	87
	Switch	318	21	1199	79
Perseveration	Different	299	21	1119	79
	Same	97	14	616	86
Fluency	Fluent	215	13	1442	87
	Semi-speaker	181	38	293	62

Table 8. Details of the fixed effects in the morphological regression model

Effect	Estimate	SE	DF	<i>p</i>	CI
Intercept	-2.11	0.25	81.2	<0.0001	[-2.607, -1.612]
Atonic pronoun [absent]					
Present	-0.65	0.14	2125	<0.0001	[-0.923, -0.367]
Following sound [consonant]					
Vowel	0.50	0.12	2125	<0.0001	[0.257, 0.740]
Referent continuity [same]					
Switch	0.63	0.15	2125	<0.0001	[0.340, 0.912]
Perseveration [different]					
Same	-0.47	0.14	2125	0.001	[-0.738, -0.207]
Fluency [fluent]					
Semi-speaker	1.93	0.36	25.69	<0.0001	[1.200, 2.668]

Note. The model fits the log-odds of using *mon*. The reference points for the independent variables are bracketed.

to nonfluent speakers nor does it appear to be an innovation stemming from semi-speakers' imperfect acquisition of French (evidenced by its attestation in fluent speaker groups across Louisiana), it does appear to be a preferred variant for semi-speakers. This perhaps echoes Rottet's (1996) points about the salience of tonic pronouns in language-acquisition contexts and suggests an important arena where research on language acquisition might provide insights on the processes of

Table 9. Results for the random effect in the morphological mixed-effects model

Participant	Intercept
OFS1	2.242
OFS2	0.681
OFS3	-0.173
OFS4	-0.489
OFS5	-0.515
OFS6	0.553
OFS7	0.206
OFS8	0.455
OFS9	-0.300
OFS10	-0.559
OFS11	-0.006
OFS12	-0.191
OFS13	-0.697
YFS1	-0.106
YFS2	0.574
YFS3	0.013
YFS4	-0.302
YFS5	-0.520
YFS6	-0.663
YFS7	0.153
YFS8	-0.356
SS1	-0.439
SS2	-0.504
SS3	1.082
SS4	-0.088
SS5	-1.433
SS6	0.960
SS7	0.368
SS8	0.054

acquisition in language death situations—especially the acquisition of particular forms of variation.

This model correctly predicted 84% of the observations in the dataset, whereas a null model accurately predicted 81% of the data. Thus, the model illustrated in [Tables 8 and 9](#) does a slightly better job of fitting the data than a null model.

Discussion

The answers to our research questions are as follows:

- (1) The participants produced twelve 1sg subject forms: the unexpressed subject, [h], [ʃ], [ʒ], [s], [z], *mon* [h], *mon* [ʒ], *mon* [z], *mon* [s], and *mon* alone.
- (2) Tonic pronoun, verb category, following sound, speaker gender, and fluency impacted the use of the most frequent phonological variants.
- (3) Atonic pronoun, following sound, referent continuity, perseveration, and fluency predicted the use of the tonic pronoun *mon*.

Thus, these findings suggest that 1sg subject expression in the French of the Point-Au-Chien Indians of Terrebonne and Lafourche Parishes in Louisiana patterns in complex and systematic ways according to both linguistic and social factors, and that this is true for fluent and semi-speakers. This lends support to claims by Wolfram (2004) and others that language variation in language death contexts is similar to that found in healthy language contexts because our analysis provides evidence of the multidimensional systematicity of this variation (i.e., it is conditioned by a range of factors). The fact that patterning was robust for semi-speakers aligns with research in the field of second language acquisition that has demonstrated that learners at different proficiency levels show complex and systematic variability in their additional language (Geeslin & Long, 2014). In other words, it is not a feature that is restricted to highly proficient speakers. The significance of gender as a predictor of 1sg subject expression in Louisiana French may point to potential social significance of this variation, as some other researchers have noted in language death situations (e.g., Kasstan, 2019; Wolfram, 2004:780), though further analysis is needed for such a claim.

We assert that not only were variationist methods suitable and valuable in analyzing this dataset, but that examinations of variation in language death contexts can also provide broader insights to variationists. To begin with, our analysis points to some methodological concerns that are not necessarily specific to obsolescing languages, but this lens of “heightened variation” allowed us to engage with them. By endeavoring to capture all forms of variation in subject expression, not just the phonological variation in atonic pronouns or the morphosyntactic variation in tonic pronouns, we provide a model for applying multivariate analyses to variation with many forms (cf., Gorman & Johnson, 2013). Moreover, while the large number of forms that the participants used for 1sg subjects prevented us from being able to analyze all forms in one mixed-effects model, we explored other ways of conducting a variationist analysis. We decided to fit two models: one consisted of the four most frequent phonological variants and the other centered on the tonic pronoun *mon*. The disadvantages of this analysis are that there are other forms that have not yet been accounted for (e.g., [s]) and that the separation of phonological and morphological forms into two analyses meant that we modeled two phenomena. We attempted to address the latter by including the independent variables of tonic pronoun and atonic pronoun in the phonological and morphological models, respectively. The fact that each of these independent factors was significant allowed us to begin to see the

ways in which the phonological and morphological variants are connected in participants' use of 1sg subject expression. Despite these disadvantages, there are benefits to the decision to fit two models. One strength is specific to Louisiana French: We were able to explain the multifactorial use of more 1sg subject variants than had been investigated previously (Dubois, 2001). Another advantage is that we offered a possible way to extend the variationist approach to a linguistic variable with numerous variants by way of the two mixed-effects models that we fit, including the atonic and tonic pronoun independent variables in order to document the relationship between these forms of variation.

Another issue we faced in this study that is common to language death situations (Blainey, 2017)—but not necessarily *limited* to language death situations—was that of small token counts. Although we found sufficient variation in 1sg subject forms in order to perform multivariate statistical analyses on our data, not all participants contributed enough observations for the models to converge. One solution could have been to eliminate these participants from the model. However, since these participants still contribute knowledge about subject expression, we were motivated to try to keep them in the analysis. Therefore, we used the Kenward-Roger method to compute the denominator degrees of freedom, which leads to a conservative type-I error rate in our model (McNeish, 2017). We also employed the restricted maximum likelihood estimation approach to reduce the bias in the random effect estimates. Such approaches might be relevant to any researcher studying forms of variation with low token counts. Thus, with both methodological issues, we see how research on obsolescing languages can provide insights about the choices we make as researchers because the current study could not be undertaken without finding proper solutions for some of these commonly worked-around issues.

Beyond methodological insights, we contend that variationist research on obsolescing and minority languages can provide a unique lens into the acquisition of variation precisely due to the circumstances of these language situations. For example, in the case of Louisiana French among the Point-Au-Chien Indians, this sample of speakers has minimal contact with mainstream, institutionalized varieties of French and is not literate in French. This means that their vernacular speech in some ways controls for the kinds of hypercorrection that is observed in healthy language contexts (e.g., Labov, 1966). In other words, these factors are controlled for. And yet we still see evidence of both socially and linguistically constrained forms of variation, which we are able to document and explain. Another opportunity we wish to highlight in analyzing variation in language death situations is the light such analyses may shed on the acquisition of variation. Especially in the case of semi-speakers, who have limited input and yet still demonstrate robust acquisition of existing variation, we can circumscribe the limits and reaches of socially meaningful variation even in a context in which domains of use and diversity of interlocutors are greatly diminished. Indeed, we suggest that a key direction for future research on variation in obsolescing languages is to analyze the difference between New Speakers (e.g., Kasstan, 2019), who actively participate in language revitalization often in institutionally supported ways, and semi-speakers who do not have revitalization efforts behind them.

Conclusion

We examined 1sg subject expression in a variety of Louisiana French spoken in Terrebonne and Lafourche Parishes where twelve subject forms exist. Multivariate analyses of a subset of the forms demonstrated that the participants' systematic use of these forms is conditioned by a range of linguistic and social factors. Further research is needed to examine in greater detail all twelve forms that occurred in our dataset and to investigate other communities of Louisiana French speakers to determine how generalizable the findings are. We demonstrated that use of variationist methods is robust for situations of language death and, in doing so, contributed additional evidence to support the claim that variability in obsolescing languages is systematic, just as it is in healthy languages (Kasstan, 2019; Wolfram, 2004).

We furthermore offered examples of how variationist research on obsolescing languages can provide key methodological and theoretical insights on variationists sociolinguistics. First, the solutions we adopted to address the methodological challenges that we encountered may be useful to other researchers who are working with variables that have many variants and for those who are working with variants that have very low frequency for some speakers. Moreover, we proposed that additional research on semi-speakers in language death situations could advance what we know about the acquisition of socially meaningful variation.

Acknowledgments. This work was supported by a grant from the Virginia Tech Institute for Society, Culture and Environment. We also thank the Point-Au-Chien Indian Tribe for welcoming us into their homes and their communities to document their speech, and Rocky McKeon for conducting interviews alongside the second author.

Notes

1. Research on subject expression in Faetar, an endangered Francoprovençal variety, has also employed variationist methods (e.g., Nagy, Iannozzi, & Heap, 2018; Pabst, Konnelly, Wilson, Meslin, & Nagy, 2020). Because it differs in notable ways from our own (e.g., it examines all grammatical persons and does not analyze different phonological variants of the atonic pronoun), we refer readers to this research for thinking more about unexpressed subjects in endangered Romance languages.
2. In the community in which Dubois conducted her research, the tonic pronoun 'me' is pronounced *moi* not *mon*.
3. For additional work on pronoun use in St. Landry, see Girard Lomheim (2017).
4. Semi-speakers are different from New Speakers, who learn an endangered language through revitalization efforts (cf. Kasstan, 2019).
5. The mixed-effects models were generated using SAS software, Version 9.4 of the SAS System for Windows. Copyright © 2018 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

References

- Austin, P. K., & Sallabank, J. (Eds.). (2011). *The Cambridge handbook of endangered languages*. Cambridge: Cambridge University Press.
- Bailey, R., Greer, K., & Holland, C. (2013). Lexical frequency and syntactic variation: A test of a linguistic hypothesis. *University of Pennsylvania Working Papers in Linguistics*, 19(2), 4.
- Blainey, D. (2017). Sociolinguistic research with endangered varieties: The case of Louisiana French. *Canadian Journal of Linguistics* 62:576–95.

- Cameron, R., & Flores-Ferrán, N. (2004). Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context*, 1:41–65.
- Campbell, L., & Muntzel, M. C. (1989). The structural consequences of language death. In N. C. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press. 181–96.
- Carmichael, K. (2007). Gender differences in the substitution of /h/ for /ʒ/ in a formal register of an endangered dialect of Louisiana French. *Southern Journal of Linguistics* 31(2):1–27.
- Carmichael, K. (2017). Stylistic variation and dialect contraction: The case of /ʒ/ and /h/ in Louisiana French. *Fleur de Ling: Tulane University Working Papers, Sociolinguistics* 3(1), 72–89.
- Carmichael, K., & Gudmestad, A. (2019). Language death and subject expression: First-person-singular subjects in a declining dialect of Louisiana French. *Journal of French Language Studies* 29:67–91.
- Carvalho, A. M., Orozco, R., and Shin, N. (Eds.) (2015). *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Washington DC: Georgetown University Press.
- Cook, E.-D. (1989). Is phonology going haywire in dying languages? Phonological variations in Chipewyan and Sarcee. *Language in Society* 18:235–55.
- Dajko, N. (2009). *Ethnic and Geographic Variation in the French of the Lafourche Basin*. Unpublished Tulane University dissertation.
- Dajko, N. & D. Blainey. (2016). Variation in Louisiana French: Prairie Cajuns and Bayou Cajuns. In S. Detey, J. Durand, B. Laks, & C. Lyche (Eds.), *Varieties of Spoken French*. Oxford: Oxford University Press. 478–90.
- Dorian, N. C. (1973). Grammatical Change in a Dying Dialect. *Language* 49(2):413–38.
- Dorian, N. C. (1977). The problem of the semi-speaker in language death. *International Journal of the Sociology of Language* 12:23–32.
- Dorian, N. C. (1978). The fate of morphological complexity in language death. *Language* 54:590–609.
- Dorian, N. C. (Ed.). (1989) *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press.
- Dressler, W. (1972). On the phonology of language death. *Chicago Linguistic Society Papers*, 8, 448–57.
- Dubois, S. (2001). Attrition linguistique ou convergence dialectale: JE, MOI/JE et MOI en français cadien. In A. Donabédian (Ed.), *Langues de diaspora, Langues en contact*. No 18 of Faits de Langue: Revue de Linguistique. Paris: Ophrys. 149–65.
- Dubois, S. & B. Horvath. (2000). “When the Music Changes, You Change Too”: Gender and Language Change in Cajun English. *Language Variation and Change* 11:287–313.
- Dubois, S. & S. Noetzel. (2005). Intergenerational pattern of interference and internally-motivated changes in Cajun French. *Bilingualism: Language and Cognition* 8(2):131–143.
- Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88:526–57.
- Geeslin, K.L., & Long, A.Y. (2014). *Sociolinguistics and second language acquisition: Learning to use language in context*. London: Routledge.
- Girard Lomheim, F. (2017). *Le pronom dans le français de Louisiane: Trois siècles de cheminement*. Paris: L’Harmattan.
- Gorman, K. & D.E. Johnson. (2013). Quantitative Analysis. In R. Bayley, R. Cameron, & C. Lucas (eds.), *The Oxford handbook of sociolinguistics*. Oxford: Oxford University Press. 214–40.
- Gudmestad, A., Edmonds, A., Donaldson, B., & Carmichael, K. (2018). On the role of the present indicative in variable future-time reference in Hexagonal French. *Canadian Journal of Linguistics*, 63 (1):42–69.
- Harrison, K.D. & G. Anderson. 2008. Tofa language change and terminal generation speakers. In K.D. Anderson, D.S. Rood, & A. Dwyer, *Lessons from documented endangered languages* [Typological Studies in Language 78]. Amsterdam: John Benjamins. 243–70.
- Johnson, D.E. (2009). Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3(1):359–83.
- Kasstan, J. R. (2019). Emergent sociolinguistic variation in severe language endangerment. *Language in Society* 48:1–36.
- King, R. (1989). On the social meaning of linguistic variability in language death situations: Variation in Newfoundland French. In N. C. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press. 139–48.

- Labov, William. (1966). Hypercorrection by the Lower Middle Class as a Factor in Linguistic Change. In W. Bright (Ed.), *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference*. The Hague: Mouton. 84–113.
- McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research* 52(5):661–70.
- Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164B:309–27.
- Nagy, N. (2017). Documenting variation in (endangered) heritage languages: How and why? *Language Documentation and Conservation* 13:33–64.
- Nagy, N., Iannozzi, M., & Heap, D. (2018). Faetar null subjects: A variationist study of a heritage language in contact. *International Journal of the Sociology of Language* 249:31–47.
- Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: Language contact, dialect leveling, and structural continuity*. Oxford: Oxford University Press.
- Pabst, K., Konnelly, F., Wilson, F., Meslin, S., & Nagy, N. (2020). Variation in subject doubling in homeland and heritage Faetar. *Toronto Working Papers in Linguistics*, 42:1–15.
- Palosaari, N. & L. Campbell. (2011). Structural aspects of language endangerment. In Austin, P.K. and Sallabank, J. (Eds.), *The Cambridge handbook of endangered languages*. Cambridge: Cambridge University Press. 100–19.
- Picone, M. D. (1997). Enclave dialect contraction: An external overview of Louisiana French. *American Speech* 72(2):117–53.
- Rottet, K. J. (1996). Language change and language death: Some changes in the pronominal system of declining Cajun French. *Plurilinguismes* 11:117–52.
- Rottet, K. J. (2001). *Language shift in the coastal marshes of Louisiana*. New York: Peter Lang.
- Rottet, K. J. (2005). Variation et étiollement en français cadien: Perspectives comparées. In A. Valdman, J. Auger, & D. Piston-Hatlen (Eds.), *Le Français en Amérique du Nord: État présent*. Saint-Nicolas, Québec, Canada: Les Presses de l'Université Laval. 243–60.
- Salmon, C. (2007). *Français acadien, français cadien: variation stylistique et maintenance de formes phonétiques dans le parler de quatre générations de femmes cadiennes*. Unpublished dissertation, Louisiana State University.
- Schmidt, A. (1985). *Young people's dyirbal: An example of language death from Australia*. Cambridge: Cambridge University Press.
- Szmrecsanyi, B., Biber, D., Egbert, J., & Franco, K. (2016). Toward more accountability: Modeling ternary genitive variation in Late Modern English. *Language Variation and Change*. 1:1–29.
- Walker, J. A. (2010). *Variation in linguistic systems*. New York: Routledge.
- Weinreich, U. (1974) [1953]. *Languages in contact: Findings and problems*. The Hague: Mouton.
- Wolfram, W. (2004). Language death and dying. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change*. Malden, MA: Blackwell. 764–87.

Appendix Participant Information

Alias	Age	Gender	Education	1 st language
Older Fluent Speakers				
OFS1	55	Female	High School	French
OFS2	59	Male	7th Grade	French
OFS3	65	Male	8th Grade	French
OFS4	60	Female	10th Grade	French
OFS5	54	Female	7th Grade	French
OFS6	73	Female	2nd Grade	French

(Continued)

Appendix (Continued.)

Alias	Age	Gender	Education	1 st language
OFS7	57	Male	6th Grade	French
OFS8	55	Female	8th Grade	French
OFS9	63	Male	5th Grade	French
OFS10	66	Female	8th Grade	French
OFS11	58	Male	6th Grade	French
OFS12	67	Male	2nd Grade	French
OFS13	51	Male	12th Grade	N/A
Younger Fluent Speakers				
YFS1	39	Male	High School	French
YFS2	37	Female	N/A	N/A
YFS3	35	Female	8th Grade	French
YFS4	44	Female	Middle School	French
YFS5	42	Male	2nd Grade	French, English
YFS6	30	Male	Some College	French
YFS7	34	Female	High School	French, English
YFS8	43	Male	8th Grade	French
Semi-speakers				
SS1	43	Female	Vocational School	English
SS2	28	Male	N/A	English
SS3	36	Female	N/A	French, English
SS4	41	Female	N/A	English
SS5	36	Male	N/A	English
SS6	33	Male	N/A	English
SS7	39	Male	N/A	English
SS8	38	Female	N/A	N/A

'N/A' means 'not reported'.

Cite this article: Gudmestad, A., & Carmichael, K. (2022). A variationist analysis of first-person-singular subject expression in Louisiana French. *Language Variation and Change* 34, 53–77. <https://doi.org/10.1017/S0954394521000235>