

Coming back to life: From indicator to stereotype and a strange story of frequency¹

Jessi Elana Aaron

University of Florida

In the history of Spanish there are five forms, originally from the same lexical item, co-existing: *así*, *asín*, *ansí*, *asina*, and *ansina*, all meaning 'like that'. Standard Modern Spanish includes only one of these: *así*. This is not the case, however, in New Mexican Spanish. This corpus-based study examines the patterns of synchronic variation in New Mexican Spanish, as well as the near death and transformed rebirth of forms other than standard *así* in literature. Multivariate analysis suggests a decline in non-standard variants in New Mexico, associated with rural activities and objects, and with older, less-educated speakers. The synchronic idiosyncrasy of stereotypes is confirmed, while the quantitative diachronic patterns found may prove to be a regular pattern for developing stereotypes in literary texts: a slow decline in frequency followed by a sharp rise.

KEYWORDS: Linguistic stereotype, lexical variation, frequency, language change, Spanish, New Mexico

1. INTRODUCTION

If one were seeking to demonstrate the notion of variability to a new student, one might very well choose the apparently straightforward case of Spanish *así*. Here we have five forms, originally from the same lexical item, related historically, four of which still co-exist in a variety that is transmitted orally and that has not undergone a process of standardization. The forms I will examine here are: *así*, as in (1); *asín*, as in (2); *ansí*, as in (3); *asina*, as in (4); and *ansina* as in (5). Standard Modern Spanish includes only the first of these: *así*.

1. ... *ojalá fuera así* ... (Concolorcorvo, *El Lazarillo*, c.1775, Peru)
'I wish it were like that'
2. *¿Pos te crees tú que si no fuera asín, iba yo a aconsejarte que ...* (Reyes, *Para un vivo otro vivo*, 1888, Spain, CORDE)
'Well do you think that if it weren't like that, I would advise you to ...'
3. ... *que si ello no fuera ansí, nunca tal cosa dijeras*. (Cervantes Saavedra, *El rufián dichoso*, 1582, Spain, CORDE)
'if it weren't like that, you would never say such a thing'

4. *¡Ojalá fuera asina!* (Carrasquilla, *Hace tiempos*, 1951, Colombia, CORDE)
'I wish it were like that'
5. *...yo no pensaba que fuera ansina...* (NMCOS, 311)²
'I didn't think it was like that'

A clear case of variability in language, this alternation (henceforth *A-form variation* to avoid confusion and implicit preference for one variant) could appear haphazard. One of the main tenets of sociolinguistics, however, is that variability in language is characterized by structured heterogeneity (Weinreich, Labov and Herzog 1968), that is, that variation, when examined quantitatively, is rarely free or random (Stubbs 2002: 8); it is constrained by linguistic and social factors. These constraints are not rules, but rather probabilistic constraints on the optionality of forms within a stochastic system, with some constraints being linguistic in nature (e.g. Poplack and Tagliamonte 2001), some sociological (e.g. Wolfram 1993: 197), and others stylistic (e.g. Silva-Corvalán 1997: 115).

While much of the empirical evidence for the systematic nature of variation patterns has emerged from the study of phonology, it has become commonplace to study morpho-syntactic variation within such a framework as well, despite some doubts about the adequacy of the variationist model in handling morpho-syntactic variation (see Sankoff 1988). The main concern expressed regarding the use of quantitative variationist methods with morpho-syntactic variability is that an analyst, upon reflection, may be able to conjure up different meanings for very similar constructions, for example *I'll be happy* and *I'm gonna be happy*. If this is the case for morpho-syntactic variables, lexical variables may prove all the more troublesome. Indeed, it is not unpopular to assert the impossibility of true synonymy. (This does not preclude, however, the possibility of neutralization in context; see Sankoff 1988.) In the case of A-form variation, this notion leads us to seek out any possible meaning differences among these variants. Torres Cacoullos (2001) proposes that, in order for morpho-syntactic (and I add here, lexical) variants to be studied as a variable, it is not necessary for them to mean, *a priori*, exactly the same thing. Their common element is that they have occurred in overlapping contexts historically.

Given the evidence against the likelihood of 'free' variation, we might expect that speakers who use A-forms variably would show differentially structured patterns of use of these forms, and that these distribution patterns would reveal the divergent social histories of these variants, as reflected in descriptive and prescriptive literature and in literary representations. This expectation is strengthened in light of lexical models developed within recent decades, such as pattern grammar (Hunston and Francis 1999), which present an 'alternative to formalist conceptions of language as a series of syntactically generated slots which the speaker or writer fills with grammatically, semantically and pragmatically acceptable lexical items' (Groom 2005: 258). In Modern Standard Spanish, this structured heterogeneity in the case of *así* (example 1) is obvious: all other surviving A-forms seem to be afforded the status of what Labov (1972:

248) terms a linguistic 'stereotype'. However, this is not (yet) the case, it seems, in New Mexican Spanish, nor does it appear to have been the case in other varieties of Spanish before the 16th century.

This paper examines synchronic variation between A-forms in New Mexican Spanish, as well as the literary representation of these forms in Spanish-language literature from the year 1000 through 1974. Multivariate analyses show that non-standard variants, associated with older, less-educated speakers are in decline in New Mexico. The diachronic literary evidence shows a decline followed by a sharp increase in the literary representation of these same variants, which tells us more about the social status of variants than about actual use. A closer look reveals the relative frequencies to be altogether misleading: a recent rise in frequency is indicative of these forms' arrival, or literary/representational 're-birth', as linguistic stereotypes.

Abundant independent evidence for the status of 'stereotype' for non-standard A-forms can be found in their persistent appearance, particularly of *asina*, in word lists of non-standard forms. For instance, in her discussion of the goals of teaching Spanish to heritage speakers, Valdés-Fallis mentions that 'we can hope . . . that the student will remember *así* instead of *asina*, *deletrear* ['spell'] instead of *espeletear*, *faltado* ['missed'] instead of *mistido*, etc.', placing *asina* alongside salient English-origin borrowings (1973: 1041). Also, Bernal-Enríquez and Hernández-Chávez, in their description of New Mexican Spanish in comparison with other varieties of Spanish, list *asina* as a non-standard, rural form (2003: 100), and Cotton and Sharp characterize *asina* as a 'rustic usage' common in the Antilles and in Spain (2001: 205). López Morales mentions *asina* as a form, along with others, commonly used to characterize black speech, 'all serving the purpose of easy laughter' (1994: 14). Finally, Santa Ana and Parodi list both *asina* and *ansina* as 'stigmatized' in the Spanish variety spoken in Mexico's southern state of Michoacán. Distinguishing these 'stigmatized' forms from 'markers' (see Labov 1972), they explain: 'Native speakers are not aware of the fact that stigmatized forms are elements of the old koiné; they simply label them as rural or uneducated speech. In school these forms are censured' (1998: 36). These authors further explain:

The stigmatized words in Mexican Spanish are 16th century lexical items which were part of the vernacular Spanish of the first settlers of the Americas. Many, such as *haiga* and *asina/ansina*, are found throughout the non-metropolitan New World (e.g. Cárdenas 1967; Rona 1973: 319). These items were subsequently displaced by newer prestige items in Latin American cities, such as Mexico City. Because there was greater interaction across metropolitan areas in Latin America than between Mexico City and its provinces, these items are stigmatized throughout Latin America by metropolitan speakers. Because the provinces have not displaced them to the same degree, they are labeled as rural ways of speech by metropolitan speakers. (Santa Ana and Parodi 1998: 48, n14)

It is important to note, then, that the A-forms considered to be non-standard, like other non-standard variants, such as Present Subjunctive *haiga* 'there

is+SUBJ' vs. standard *haya*, or Preterit *truje* 'I bring+PRET' vs. standard *traje*, have remained in robust usage throughout much of the Spanish-speaking world on both sides of the Atlantic. Their apparent 'disappearance' from literary sources and their relegation to the category of 'archaic' has very little to do with their frequency in actual language use. Instead, what we have left to us is, on the diachronic side, their representation in comparatively heavily standardized texts, and on the synchronic side, their continued usage in some communities today. It is a through the examination of multiple sources, then, including (misrepresentational) literary occurrence, clues gleaned from prescriptive grammars, and everyday patterns of actual use, that we may come to a better understanding of the quantitative nature of the mismatch between literary representation and actual usage patterns of features acquiring stereotype status.

The basic notion of the linguistic stereotype is behind a large body of scholarly work, spanning discourse analysis, language ideology, language-based discrimination, and language and identity research (e.g. Preston 1989; Cameron 1995; Schilling-Estes 1998; Preston and Niedzielski 1999; Coupland 2001; Agha 2003; Wolfram 2003). The range of themes that converge around linguistic stereotypes is impressively vast, and I cannot possibly hope to do justice to this work here. Instead, I will briefly mention some of the research that impinges more directly on the involvement of linguistic stereotypes in language variation and change. Some of the most fascinating quantitative work regarding stereotypes, in my opinion, has been in the examination of language attitudes and social perception, particularly in relation to accent identification (see e.g. Giles 1970; Kalin 1982; Giles and Coupland 1991; Dixon, Mahoney and Cocks 2002; Coupland and Bishop 2007). Much of this synchronic work shows how speakers use indexical cues in language to assign values of both overt and covert prestige, as well as other locally significant meanings. What is known empirically about the diachronic development of these social indices, however, is somewhat more limited. In a recent qualitative study, in the only empirical diachronic study of stereotype development of which I am aware, Johnstone, Andrus, and Danielson draw on historical evidence, ethnography, discourse analysis, and sociolinguistic interviews to trace the development of certain dialect features of 'Pittsburguese' from indicators to markers to stereotypes (2006).

Despite the importance of the notion of stereotype in sociolinguistic literature, there are relatively few studies on the diachronic quantitative behavior of stereotypes (Kristiansen 2003: 78).³ Linguistic stereotypes tend to have an irregular synchronic stratification (Labov 1972: 248), and they have a propensity to be exaggerated (Allport 1954: 91) and imprecise (Honey 1997: 99) in popular belief.⁴ Nonetheless, more systematic empirical research of these outliers is still necessary:

It is beyond a doubt that a 'Theory of the Shibboleth' is an important desideratum, to which also linguistics can make its contribution. The theoretical implications and also the notation are to a great extent still unclear. Of special importance is in

my opinion an extensive empirical analysis of concrete examples and an extended definition of the scope of this phenomenon: which are the most important and common 'shibboleth speech events' and how may we describe them in adequate ways? (Kniffka 1991: 175; cited in Kristiansen 2003: 79)

While the current study corroborates Labov's early findings on the synchronic idiosyncrasy of stereotypes, the quantitative patterns shown here may be a first glimpse of what may prove to be a typical diachronic quantitative pattern for developing stereotypes: a slow decline in frequency, nearly to extinction, followed by a sharp and exaggerated rise. If this is indeed the case, we may be better able to identify erstwhile unknown linguistic stereotypes in other eras through quantitative diachronic study, thus finding diachronic regularity within linguistic features known best for their synchronic chaos.

2. PRESCRIPTIVE NORMS AND SOCIO-HISTORICAL REPRESENTATION

Before we examine the patterns of A-variation in New Mexican Spanish (section 3), it will be useful to uncover the prescriptive and descriptive histories of these forms in Standard Spanish (section 2.1). This will be followed by a corpus-based diachronic study, both quantitative (section 2.2) and qualitative (section 2.3).

2.1 Dictionaries and grammars

Despite the selective and skewed nature of many historical dictionaries and reference grammars, such texts, when used alongside sociolinguistic analysis, may offer clues about the existence and social status of certain linguistic features during a given time period (Poplack 2007). The earliest prescriptive mention of *así* and its variants I have found comes in the early-17th century. Correas, in his 1625 *Arte de la lengua española castellana*, comments on the variation found with three A-forms across Spain, mentioning *ansi*, *asi*, and *ansina*. Here, no form in particular appears to carry any particular social meaning, except for an association of *ansina* with Andalusia:

Causatives and rationals, which give reason and cause, are: *ca*, *que*, *porque*, *por ende*, *por tanto*, *pues*, *pues que*, *ansi que*, *asi que*, *ansina*, *onde*, *luego*, *por eso*, and *por tanto*, *es á saber*, *conviene á saber*.⁵ *Ca* is old, and it means 'then', 'because', 'thus', and due to its elegance Father Mariana and others use it. [...] *Ansina* is used in Andalusia, and at times *ansi*, *asi*, *ansina* threaten. (Correas 1625: paragraph 25)

By the time these forms appeared in the Real Academia dictionaries in 1726 (or 1770, in the case of *asina*), all of these except (now-)standard *así* were characterized as archaic forms.

If we look at the evolution of the dictionary definitions of *ansi/ansi* from 1726 through 2007, we see that, for nearly 300 years, this 'archaic' form 'of little

use' (*Diccionario de la Real Academia Española*, henceforth DRAE, 1726) has nonetheless persisted in Spain's foremost dictionary. While in 1726 it was only characterized as 'archaic' (DRAE 1726), by 1770 it was attributed to 'rustic people' (DRAE 1770). In 1884, these 'rustic' speakers are 'still' using this form (DRAE 1884), and in 1933 it is described as 'vulgar' (DRAE 1933), a characterization removed in 1936 and thereafter. Only in 2007 is this form described as being 'in disuse', though it is also described as a rural usage. The origins of *ansí* are unclear: in 1884 its etymology is listed as Latin *in sic* 'as such', but in 1956–2007, its etymology is listed as Latin *æque sic* 'just like this' or *ad sic* 'about like this'.

The form *ansina* has a similar prescriptive history to *ansí*, though even more strongly associated with lower social classes. In the first Academia dictionary in 1726, it is already described as 'low' and 'archaic' (DRAE 1726). In 1770 it is ascribed to 'rustic' speakers (DRAE 1770), and in 1822 it is said to be used 'only among rustic and vulgar people' (DRAE 1822). The only etymology offered first appears in 1914, locating *ansina*'s origins in *ansí*, an etymology listed through 2007. In 1933 *ansina* is defined as 'archaic and vulgar', then upgraded back to 'rustic' from 1936 to 1984. In 2007, it is no longer described as 'rustic' or 'vulgar', but simply as a 'rural' form 'in disuse'. Similarly, Lara Ramos, in his 2000 dictionary of Mexican Spanish, also defines *ansina* as a rural term, and offers the following examples:

Se parte, se le echa su dulce y entonces ya se come ansina en dulce 'You cut it, you put its sugar on it and then you eat it sweet like that', and *Llegó la yegua ansina de panzona* 'The mare arrived fat like this'. (Lara Ramos 2000)

Both *asín* and *asina* are less thoroughly described in the Academia dictionaries. *Asín* is described as an archaic form from 1726 through 1914, given no other qualification except a regional association with Aragon until 1869. This is the only A-form prescriptively associated with a certain region. From 1925 through 1970, *asín* is listed as having 'familiar usage', and from 1984 through 2007 it is listed as being a 'vulgar' usage. *Asina*, first appearing in the dictionary in 1770, follows the same trajectory as *asín*, from archaic to familiar to vulgar, during the same time periods, though it is never associated with a particular region. Lara Ramos (2000), who does not list *asín* in his dictionary, does include *asina* as a 'rural' use, and offers the following example:

Le aseguro que el verdadero México usted no lo conoce. –Pos francamente, si asina es la puntita, no quero ver lo que hay más pa dentro 'I assure you that you do not know the real Mexico. –Well frankly, if the tip is like this, I don't want to see what there is further in'.

Note that this example is replete with non-standard features, including phonological reduction (*pos* 'well' < *pues*; *pa* 'for' < *para*) and regularization (*quero* 'I want' < *quiero*), typical of 20th-century textual representations of A-forms (except *así*). I will return to this observation in section 2.3.

2.2 A millennium of A-form variation: 1000 – 1974

While the prescriptive history does offer some social and frequency information about these variants, more detail – and temporal depth – can be garnered through diachronic quantitative and qualitative analysis. I employed the *Corpus diacrónico del español* (CORDE), a nearly 300-million-word online corpus, compiled by the *Real Academia Española* ‘Royal Spanish Academy’ (RAE) and available on their website. According to the RAE, in CORDE ‘the intention is to collect all of the geographic, historical and genre varieties so that the entire collection will be sufficiently representative’ (RAE 2008a), and ‘today it is a necessary source for any diachronic study related to the Spanish language’ (RAE 2008c). The corpus is vastly collected from books (97%), with only three percent of the data coming from the press. In light of the historical diffusion of Spanish in the world, 74 percent of the data come from Spain, 25 percent from Spanish America, and one percent from other regions. Prose represents 85 percent of the data, and verse represents 15 percent. Prose includes lyrical prose, narrative, theater, science and technology, society, religion, press and publicity, history and documents, and legal prose. The verse data include lyric, narrative and dramatic verse (RAE 2008b). Over half of the data (56%) is non-fiction; the rest (44%) is made up of verse, fictional prose and theater (RAE 2008c). A fuller description is available (in Spanish) at the *Real Academia Española’s* website.

Of course, using corpus data to uncover quantitative evidence about social variables, from a sociolinguistic standpoint, is highly questionable. As Sigley quite justly notes, ‘the use of corpora has, to date, been extremely productive for variation linked to variety, situation, or linguistic function, but less so for social variation’ (2006: 220). There are many reasons to exercise great caution when studying written texts, since these are not a good representation of speech. Not only is writing generally more conservative than speech, it also relies much more heavily on codified and standardized forms, eliminating much of the variation one would find in everyday language use (Lippi-Green 1997: 18–25). The stylistic, syntactic and lexical differences between written genres are great, and the distance between different genres of writing can shift over time (e.g. Biber 1995; Biber et al. 2006). Since many linguists, including myself, rely on textual representations of language use in our quantitative work, often without the possibility of comparable speech samples, a more robust characterization of the quantitative behavior of the literary representation of emerging stereotypes may help us to understand diachronic changes in rates or constraints on variation in the final stages of this process.

CORDE, just like other corpora of this nature, cannot purport to offer us a portrait of actual language use. What is of interest here – because it is what is available to us – is precisely the changing textual *representation* of an emerging linguistic stereotype (for a discussion of reality and textual representation see Lipski 1995: 141). What is represented in literature is not usage, but representation itself; representation, especially of language and other complex

symbolic systems, such as culture, is always partial and rarely representational. This is especially the case if the social meaning attached to some variants changes over time: as a variant takes on evermore salient social meaning, its appearance in literature may stray ever farther from quantitative accuracy in terms of its actual use. The occurrence of certain variants in written discourse is a far cry from independent evidence of their use in actual speech; nonetheless, I agree with Biber that ‘it is possible to estimate the linguistic characteristics of speech in earlier periods from a consideration of speech-based registers’ (1995: 287). At the same time, while quantitatively these representations may be increasingly inaccurate, I assume, like Lipski in his diachronic study of literary ‘Africanized’ Spanish, that they are ‘qualitatively accurate’ representations, and that ‘comparative analysis of these texts will aid in determining the chronology’ (Lipski 1995: 134) of A-variant stereotypification.

In Table 1, we see the raw and relative frequencies of the A-forms in CORDE as a whole and in Mexico data within CORDE. In general, standard *así* is the most frequent A-form, making up 83 percent of the data (N = 241,294), 3 percent (N = 7039) of which can be attributed to Mexican data. *Ansí* is second-most frequent, representing 17 percent of the A-form data in CORDE (N = 48,186), with 2 percent (N = 861) of this occurring in Mexican data. The other three A-forms, *asín*, *asina* and *ansina*, make up less than one percent (N = 395,278 and 195 respectively) of the CORDE data. While *asín* does not occur in the Mexican data, one percent (N = 3) of *asina* occurrences are found in Mexican data. *Ansina* is the only form that may have a slight association with Mexico in particular, with seven percent (N = 14) of its occurrences found in Mexican texts. In total, then, this study analyzes the diachronic representation of 290,348 occurrences of A-forms by examining both relative frequency and context of use. Due to exceedingly low frequencies for some A-forms, the rest of this analysis will focus on CORDE as a whole, without reference to the region of the texts’ origin.

While it may be most intuitive to compare A-form frequencies to *each other*, this information is not particularly enlightening: of course *así* is generally the most frequent form. Relative frequencies alone only suggest what we already

Table 1: Raw and relative frequencies of A-forms in all regions and in Mexico, CORDE

A-form	N	% of data	N in Mexico	% type in Mexico
<i>Así/asi</i>	241,294	83	7039	3
<i>Asín/asin</i>	395	<1	0	–
<i>Ansí/ansi</i>	48,186	17	861	2
<i>Asina</i>	278	<1	3	1
<i>Ansina</i>	195	<1	14	7
Total	290,348	100	7917	3

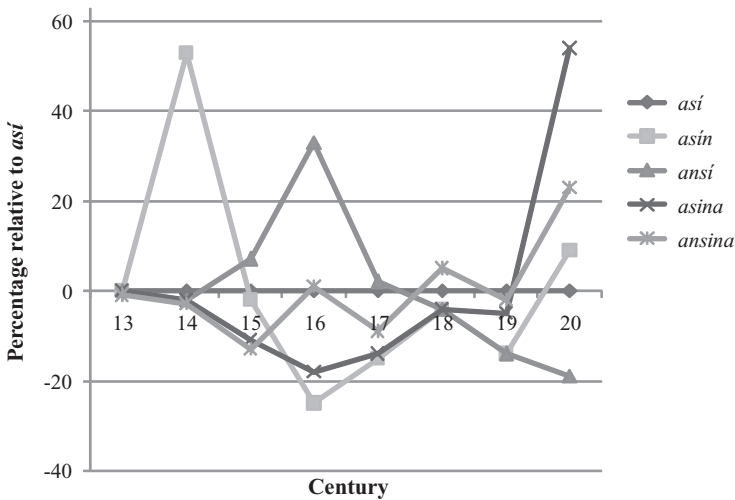


Figure 1: Distribution of A-forms in CORDE, 13th – 20th centuries, showing the percentage of each A-form that occurs in each century, normalized according to the form *así*

know. If we look instead, however, *within* the diachronic trajectory of frequency for each form across the centuries, comparing each form to itself, and using *así* as a (relatively) unmarked baseline, a story emerges. In Figure 1 I have normalized the data according to the distribution of *así*, which has been made a constant, at zero on the y-axis, in order to account for vastly differing corpus sizes in each century. This shows how other A-forms pattern diachronically in these data in comparison with ‘normative’ *así*. In other words, if we assume (though we cannot be overly cautious in doing so) that *así*, in a corpus balanced in size across the centuries (e.g. 10 million words per century), would occur with a similar raw frequency in each century, then we can artificially ‘balance’ the CORDE by setting *así*’s frequency at zero. Figure 1 shows us, then, when other A-forms, assuming normativity of *así*, are over- or under-represented in CORDE in each century, given the size of the corpus for each. We see that three forms, *asín*, *ansina*, and particularly *asina*, are over-represented in the 20th century. This may be due to many factors, including the birth of new literary genres, as well as these forms’ indexical utility as linguistic stereotypes.

Like a speaker who has access to linguistic knowledge about the standard and non-standard varieties of a language, the authors of the 20th-century texts, when representing non-standard speech, may have been using their

knowledge of the prestige dialect to avoid giving any vernacular form which is identical or similar to the standard, and so produce[d] stereotyped forms which are simply a collection of the ‘most different’ or ‘worst’ sentence types. (Labov 1972: 214)

Such exaggeration is evident both quantitatively (Figure 1) and qualitatively.

2.3 Words for the lower classes: Becoming marked variants

Qualitative evidence suggests that the status of all remaining A-forms except *así* changed sometime during the 18th or 19th centuries in real usage, and they re-emerged in literature as stereotypes in the literary movements of the 20th century. In examples (6) – (9), which date from the 13th through the 17th centuries, we see present-day non-standard A-variants occurring in otherwise unmarked, even formal, discourse.

6. ...*cornel asina sos consules contra los de carthago.* (Alfonso X, *General estoria*, Cuarta parte, c. 1280, CORDE)
'his consuls with him like that against those from Cartago'
7. *Luego que le di los dineros arrastró ansina la pierna; mas luego que se fue . . .* (Lope de Rueda, *Pasos*, 1545, Spain, CORDE)
'After I gave her the money she dragged her leg like that; but after she left'
8. *El dicho señor Inquisidor lo mandó ansí hacer . . .* (*Proceso de Pedro de Ocharte*, 1572-1574, Mexico, CORDE)
'The said Sir Inquisitor ordered it to be done like that'
9. *Daré mil quejas al aire. Y ansina diré a las selvas . . .* (Lope de Vega Carpio, *Romances, en Romancero general*, 1600, Spain, CORDE)
'I will give a thousand complaints in the air. And so I will say to the jungles'

Examples (6) – (9) are linguistically unremarkable; nothing except the occurrence of a present-day non-standard A-form suggests the use of a non-standard variety. This is not the case, however, in the 20th century. Here, present-day non-standard A-forms co-occur regularly with a multitude of non-standard features, manifested in phonology, morphology, lexical choice, and orthography (in bold in the following examples). In (10), we find *porra*, a colloquial, slang term for 'bother', *cargao* for *cargado* 'full', a common phonological reduction associated with casual speech, *mieu* for *miedo* 'fear', a similar though less common phonological reduction, and *naide* for *nadie* 'nobody', an archaic form associated with rural speech. In (11), phonological reduction of /d/ is found in *tós* for *todos* 'all' and in *mentás* for *mentadas* 'damned', phonological merger of liquids in *cuelpo* for *cuerpo* 'body', the archaic *mesmo* for *mismo* 'same', and the orthographically modified *prinsesas* for *princesas* 'princesses' (likely representing *seseo*).⁶ In (12), we find archaic morphology in *trujo* for *trajo* 'bring+PRET', and the regularization of the past participle of *escribir* 'write' in *escribida* for *escrita* 'write+PP'. While these phenomena have been documented and certainly do occur with some speakers, such density of occurrence is unlikely.

10. ...*de la **porra** el cachorrillo **cargao** hasta la boca, asina que no tengo **mieu** a **naide**.* (Hermilio Alcalde del Río, *Escenas cántabras (apuntes del natural)*, 1914, Spain, CORDE)

'full of kids **bothering** me up to my neck [lit. mouth], so I'm not afraid of anybody [lit. I have no **fear** of **nobody**].'

11. ...tós los resplandores de su **cuelpo** de estrella. Y asina mesmo como las **prinsesas mentás**, ella gobernaría... (Luis Palés Matos, *Poesía 1915-1956*, 1957, Mexico, CORDE)
'all of the splendors of her **body** as a star. And so just like damned princesses, she would govern...'
12. ...el nombre que **trujo** el muchacho para que ansina quede en la partida **escribida** en el libro. (Victor Cáceres Lara, *Humus (cuentos completos)*, 1952, Honduras, CORDE)
'the name the boy **brought** so that like that it will remain **written** in the book upon departure'

If dense co-occurrence of present-day non-standard A-forms with other non-standard features were not enough to alert the reader of these texts to the social identity of the represented speaker, these A-forms also tend to co-occur with overt references to the A-form-speakers' race or social class. In (13) and (14), social identities are double underlined, while non-standard features appear in bold.

13. ¡Después de mi Dios, él! Ni el **mesmo** Rey será asina. Lo dice este triste negro que lo conoce más... (Tomás Carasquilla, *La marquesa de Yolombó*, 1928, Colombia, CORDE)
'After my God, him! Not even the King **himself** will be as such. This sad black man who knows him more says this'
14. ...semos sino los negros y las pionadas. Pero no es asina, Amita, y perdone que yo se lo manifieste... (Tomás Carasquilla, *La marquesa de Yolombó*, 1928, Colombia, CORDE)
'we are but blacks and peons. But that isn't so, little Lady, and forgive me that I make this known to you'

These 20th-century examples are most likely related to stereotyped, exaggerated representations of actual speech of *bozales* (non-Spanish-speaking Africans), thus reflecting social, not linguistic reality (see López Morales 1994). A look at the distribution of A-forms according to genre in the 20th-century data in CORDE offers some further evidence that the A-forms *asín*, *asina* and *ansina* (but not *así* or *ansí*) served a particular literary function. In Table 2, we see that the former three forms occur nearly exclusively in two genres: lyrical verse and narrative prose, genres which make up only a small portion of *así* and *ansí* occurrences. The relative diachronic rise in these three forms' frequency, then, may point not to an increase in actual use, or even to the moment these forms became stereotypes, but rather to an increase in the production of literary genres in which authors aimed to reproduce non-standard speech, such as the Latin American literary movement known as *criollismo* (e.g. Alonso 1996).

Table 2: Rates of occurrence and distribution of A-forms across genres, 20th-century CORDE

Genre	<i>así</i>		<i>asín</i>		<i>ansí</i>		<i>asina</i>		<i>ansina</i>		Total N
	N	%	N	%	N	%	N	%	N	%	
Did. Prose	5,289	2	1	1	22	4	–	–	–	–	5,312
Dram. prose	1,494	<1	–	–	1	<1	3	1	1	1	1,499
Dram. verse	1	<1	–	–	1	<1	–	–	–	–	2
Hist. prose	5,504	2	–	–	224	36	8	4	2	2	5,738
Journ. prose	609	<1	–	–	2	<1	–	–	–	–	611
Legal prose	1,504	<1	–	–	–	–	–	–	–	–	1,504
Lyr. Verse	1,105	<1	37	30	–	–	18	8	2	2	1,162
Narr. Prose	18,768	6	86	69	27	4	184	86	88	95	19,153
Relig. prose	4,179	1	–	–	14	2	–	–	–	–	4,193
Sci. prose	12,496	4	–	–	292	47	–	–	–	–	12,788
Soc. prose	256,266	83	–	–	33	5	–	–	–	–	256,299
Total	307,215	100	124	100	616	100	213	100	93	100	308,261

In Section 3, we will see how these A-forms are used in naturally occurring speech in a variety of Spanish in which most of these A-forms are still in use: New Mexican Spanish.

3. NEW MEXICO: MODERN-DAY VARIATION

3.1 *Corpus and data*

The Traditional Spanish of New Mexico, a unique and endangered variety of Spanish (e.g. Espinosa 1911a; Bills 1997; Bills and Vigil 1999), has received its share of scholarly attention in Southwestern linguistics due to its unique lexical characteristics (e.g. Bills and Vigil 2000), phonology (Espinosa 1909; Torres Cacoullos and Ferreira 2000; Alba 2005), morphology (e.g. Espinosa 1911b) and characteristic code-mixing (Espinosa 1914, 1917; Torres Cacoullos and Aaron 2003; Aaron 2004; Clegg 2006). New Mexican Spanish is a heterogeneous variety spoken by a community that was in relative isolation from 1598 until the late-19th century, when ‘contact with English began in earnest’ (Bills and Vigil 1999: 52). In 1880, Anglo-Americans made up less than nine percent of the New Mexican population (Williams 1986: 126), but by the middle of the 20th century, only half of the population was Hispanic (Simmons 1977: 163). This isolation meant that the diachronic developments found in New Mexican Spanish did not parallel those found in neighboring and closely related Mexican Spanish, resulting in what has often been termed an ‘archaic’ dialect. Which features merit the label ‘archaism’, however, is relative:

...retentions from earlier periods are typically classified as 'archaisms' only when they remain in one variety of the language after they have disappeared in another variety that the classifier accepts as the 'correct' or 'standard' variety. Because of its history of relative seclusion, New Mexican Spanish is often characterized as highly 'archaic' in this restricted sense. (Bills and Vigil 1999: 49)

In the case of English, variation patterns in enclave communities have been hypothesized to provide synchronic evidence for the reconstruction of vernacular varieties, in particular African American Vernacular English and Appalachian Vernacular English, in the face of a lack of representative written documentation (e.g. Poplack and Sankoff 1987; Montgomery 1989; Poplack and Tagliamonte 1991, 2001; Poplack 1999; Mallinson and Wolfram 2002; Wolfram and Thomas 2002; Montgomery and Hall 2003). We must, of course, keep in mind that, just as diachronic literary representation does not show us an accurate picture of actual spoken language use in the past, enclave communities' language use patterns do not show us language relatively frozen in time, what Wolfram and Schilling-Estes refer to as 'the relic assumption' (2003: 208; see also Andersen 1988). Thus, though New Mexican Spanish may share some features with 16th-century Spanish that standard varieties no longer share, it is a misstep to take New Mexican Spanish as a living representation of the colonial era (Bills 1997: 169–170); indeed, it has its share of 'home-grown linguistic adaptations' (Bills and Vigil 1999: 53). Like in other varieties, variation patterns in enclave communities may demonstrate simultaneous retention and innovation (Wolfram and Schilling-Estes 2003: 209), and may even undergo an ebb and flow in usage patterns (Hickey 2002). Nonetheless, as Wolfram and Schilling-Estes note,

'Because enclave communities are by definition relatively self-contained, they allow us to better isolate and examine interrelations among the various social and linguistic factors that affect directionality in language change than is sometimes possible with larger, more diffuse societies. (2003: 209)

It is with this benefit in mind that I have chosen to examine New Mexican Spanish, a variety that Bills and Vigil characterize as 'profoundly "Mexican"' in nature (1999: 52), in hopes of finding clues to the path all A-forms except *así* may have taken as they were lost in standard varieties of Spanish.

One major challenge of studying the variety of Spanish spoken in New Mexico is that it is being displaced by Mexican Spanish from immigrant populations and by English (Bills 1997; Bills and Vigil 1999). In a major NEH-funded project aimed to document this variety, scholars and graduate students at the University of New Mexico carried out 350 interviews with Traditional Spanish speakers in New Mexico and Southern Colorado during the early 1990s, creating a database known as the New Mexico-Colorado Spanish Survey (NMCSS) (see Bills 1997; Bills and Vigil 1999). With the goal of obtaining both lexical information to document lexical variation throughout the region as well as natural speech samples that would document the cultural traditions and stories of the region,

each recording includes both item identification and interviews, the latter of which are comprised of long stretches of spontaneous speech. The corpus used for the present study included 17 of these interviews, which included 18 speakers in total: six men and 12 women ranging from 45 to 96 years of age at the time of the interview, from Bernalillo County (including Albuquerque), Española, Mora, Río Arriba, SW Colorado, Taos, and Tucumcari. Interviewer data were excluded.

In this variety of Spanish, four of the five lexemes of interest here (*así*, *ansí*, *asina* and *ansina*, but not *asín*) are in variation today, as seen in (15) – (17).⁷ In (15) and (16), the same speaker repeats the same linguistic construction, *y A he sido siempre yo* ‘and I have always been like that’, using *así* in (15) and *ansina* in (16). Similarly, in (17), the speaker explains how beans were sold before and how they are sold today. In both occurrences, the verb (*mercar* ‘sell’), the subject and object, and the word order (OV+A) are the same, yet the speaker uses *ansina* for the first occurrence and *así* for the second.⁸

15. ... *y así he sido siempre yo, yo puedo ayudar una gente le ayudo sí* (NMCOS, 311)
‘and I’ve always been like that, I can help a person, I do help them’
16. *Ellos sabían que tenían que respetarme, y ansina he sido siempre yo.* (NMCOS, 311)
‘They know that they had to respect me, and I’ve always been like that’
17. *la sal, el frijol lo mercaban ansina. Todavía el frijol sí todavía lo mercan así, pero to’ empaca’o ‘hora.* (NMCOS, 10)
‘salt, beans, they sold them like that. With beans, they still do sell them like that, but everything all packaged up now’

Given the history of New Mexican Spanish, which has been spoken in the region since 1540, and given further these variants’ presence in other varieties of Spanish in CORDE, it seems safe to assume that the persistence of these forms in New Mexico today is evidence that they were common in actual speech in the 16th century. New Mexican speech provides what the CORDE cannot: convincing evidence of variation in actual use in the 16th century. This finding coincides well with the story told by CORDE: it is precisely in the 16th century that all three non-standard forms still present in New Mexico had their first literary peaks (see Figure 1). *Asín*, the only form not present in the NMCOS data, had its literary peak earlier, in the 14th century.

Table 3 shows the relative frequencies in the data of the four A-forms that occur in the NMCOS sample. It is not, as one might expect, standard *así* that emerges as most frequent, at 21 percent (84/397), but rather *ansina*, the form most phonologically distinct from *así*, at nearly half (48%, 190/397). *Asina* trails behind *ansina* at 30 percent (120/397), while *ansí* is exceedingly uncommon at less than one percent (3/397).

Table 3: Raw and relative frequencies of *así*, *ansí*, *asina* and *ansina* in the NMCOS

A-form	N	% of data
<i>Así</i>	84	21
<i>Ansí</i>	3	<1
<i>Asina</i>	120	30
<i>Ansina</i>	190	48
Total	397	100

3.2 Methodology

The independent variables taken into account for this study included both social and linguistic factors. Social factors included: years of formal education; age at time of interview; and speaker sex. The linguistic factors included: Verb class; Noun class; and Construction type. The factor Verb class applied to adverbial uses of A-forms, and included:

- activity, e.g. *bailar* 'dance';
- stative, e.g. *estar* 'be';
- transitive motion, e.g. *sacar* 'take out';
- intransitive motion, e.g. *ir* 'go';
- psychological, e.g. *querer* 'want';
- exchange, e.g. *mercar* 'sell'; *hacer* 'do'; and
- other.

The factor Noun class applied to adjectival uses of A-forms, and included:

- abstract/intangible/mass, e.g. *remedios* 'medicine';
- tangible count, e.g. *jarrito* 'little jar';
- human/body part, e.g. *viejita* 'old lady';
- organic, not human, e.g. *mata* 'plant'; and
- *cosa* 'thing'.

The Construction type factor was aimed at discovering any form-specific tendencies to occur in certain constructions. The constructions examined, identified through qualitative examination of the extracted data, included:

- A + adj., (e.g. *un little strap asina chiquito* 'a little strap like that little', NMCOS, 214);
- A como, (e.g. *y luego ansina como ahora* 'and then like this like now', NMCOS, 10);
- A es 'that's how it is';
- A no más 'just like that', A nada más 'just like that', no más A 'just like that', nada más A 'just like that';

- adj. + A, (e.g. *unas cosas chaparras así* 'some **short** things like that', NMOCSS, 219);
- *cosas* A 'things like that', *una cosa* A 'a thing like that';
- *y* A 'and like that';
- other (including AA 'so-so', *algo* A 'something like that', *aun* A 'even so', *como* A 'like that', *tanto* A 'so much as that'); and
- none.

Two multivariate analyses using GoldVarb X (Sankoff, Tagliamonte and Smith 2005) were carried out, maintaining social factors and linguistic factors separate throughout the analysis. GoldVarb provides the *hierarchy of constraints* on variation, seen in the *probability weight* for each factor, and the *relative magnitude of effect* of each factor group, seen in the *range*. These analyses aimed to provide evidence to answer two questions. First, given that the *Real Academia Española* considered alternates of *así* to be archaic and rural even in the early 1700s, is there any indication of such traits in New Mexican Spanish, where *así* is one of three viable options? Second, is there any evidence that some A-forms will be lost in New Mexico?

For purposes of the statistical analysis, the data were divided into two groups: first, *así*; and second, all other variants. Such a division gives rise to a concern: it is most likely the case that speakers do not treat the three non-standard variants in exactly the same way. Ideally, all variants would be analyzed separately. Such a nuanced analysis, however, was complicated by the fact that the only variant shared by all speakers in this NMCOS sample was standard *así* (see Table 5). All other A-forms were used by only a subset of speakers, making a form-by-form analysis empirically deceptive, since the statistical results would end up comparing small groups of speakers, instead of comparing the patterns of use associated with forms. In linguistic terms, however, such a dichotomous grouping is not completely unjustified. The distribution patterns shown in 20th-century CORDE (Table 2) suggest that, at least in other varieties of Spanish, *asina* and *ansina* are functionally different from *así*. *Asín*, which aligns with *así* in Table 2, represents less than one percent of the NMCOS data (Table 3). Since A-form variation was hypothesized to be both socially and linguistically conditioned, working with the data from a subset of this sample seemed less felicitous than the simplistic division between standard and surviving non-standard forms. Such a division also makes sense socially: *así* may have acquired a special status in New Mexico due to its association with standard speech in other varieties of Spanish that speakers may encounter in their daily lives, such as Spanish on TV, in their classes in school, or in interaction with immigrants from other Spanish-speaking regions. This is particularly the case for younger speakers, who often participate in a process of dialect leveling – or 'the eradication of socially or locally marked variants (both within and between linguistic systems) in conditions of social or geographical mobility and resultant dialect contact' (L. Milroy 2002: 7) – through participation in high school or university

Spanish courses, where Traditional New Mexican Spanish is rarely taught. While dialect leveling in post-isolated communities is not obligatory (Schilling-Estes 2002), the hypothesis of leveling of this feature is supported by *así*'s high diffusion rate (see Table 5), and follows previous researchers' findings on lexical leveling in New Mexican Spanish (Bills and Vigil 1999: 56–57). With the data grouped in this way, we can see which linguistic and social factors condition the occurrence of the standard *así*, and with which contexts the prescriptively archaic forms tend to co-occur. As more of the NMCOS is transcribed and becomes available for study, a more complex and complete analysis of A-form variation in New Mexico may become feasible.

3.3 Results of variable rule analyses

Table 4 shows the results of the GoldVarb analysis for linguistic factors. The first column shows the number of occurrences of *así* in the specified linguistic context, the second shows the relative frequency of *así*, the third the probability of the occurrence of *así*, and the fourth the percentage of the data represented by that linguistic context. As shown, both Verb class and Noun class emerged as significant factors. Verb class, with a range of 51, has the highest magnitude of effect, and Noun class has a lower magnitude of effect, with a range of 39. Construction was not significant.

Examining the results for Verb class, we see that *así* is strongly disfavored with transitive motion verbs as in (18) and (19), a context that constitutes 12 percent

Table 4: Linguistic factors conditioning *así* occurrence vs. *ansí*, *asina* and *ansina* in NMCOS (input probability = .121)*

Factor group	N <i>así</i>	% <i>así</i>	Probability	% of data
Verb class				
Stative	22	28	.64	30
Activity/saying/doing	14	26	.62	20
Other	16	16	.48	37
Transitive motion	1	3	.13	12
Range			51	
Noun class				
Other	10	29	.71	46
Tangible object	3	8	.32	54
Range			39	
Special construction				
Other	24	30	[.56]	20
None	56	20	[.50]	72
A como/A+adj./adj.+A	4	13	[.34]	8
Total	84	21		100

Log likelihood = -194.750; $p \leq 0.01$; Total Chi-square = 11.1471; Chi-square/cell = 0.5067

*Square brackets indicate that this factor group did not achieve statistical significance.

of the data. In this context, *así* is used only one time. In contrast, *así* is favored with other verbs of activity, verbs of saying and doing (probability = .62), such as *curar* 'cure' in (20), as well as with stative verbs (probability = .64). In the case of noun class, with tangible count nouns, as in (21) – (22), with *bola* 'ball' and *granos* 'grains', *así* is strongly disfavored, at a probability of .32. All other noun classes favor *así*.

18. ...partir leña, todo eso lo hacía yo, enterrar zacate y **amontonar** zacate *ansina* con la horquilla ... (NMCOS, 10)
'chop firewood, I would do all that, plant grass and **pile** grass **up like that** with the rake'
19. Lo raisaban, ... lo lavaban y luego ... lo echaban moja'o en un saco y lo **ponían** *ansina* como, como ahora atrás de la estufa en el saco y luego lo tapaban ... (NMCOS, 10)
'they would let it rise, they would wash it and then they would put it wet in a sack and they would **put it like that** like, like now behind the stove in the sack and then they would cover it'
20. Oh, yeah, *así* solamente me **curaban** a mí ... nunca veía el doctor. (NMCOS, 102)
'Oh yeah, they only **treated** me **that way** ... I never saw the doctor'
21. Jaspe. Era una **bola** *ansina* como una piedrita y luego la quemaban echaban lumbre en el horno ... (NMCOS, 10)
'Jasper. It was a **ball** **like that** like a little rock and then they would burn it, they would light the stove'
22. Medidas de maíz, y otras cosas, **granos** *ansina*. Se fueron recontentos los indios. (NMCOS, 142)
'Measurements of corn, and other things, **grains** **like this**. The Indians left very happy.'

At first glance, these very specific contexts of disfavor, transitive motion verbs and tangible count nouns, appear somewhat arbitrary. With a bit of reflection, however, a meaningful pattern does seem to emerge. The A-forms other than *así* are used when describing actions that relate to a traditional, rural way of life. In a way, this may be akin to the 'extreme raising' of (*ay*) that Eckert found in the Detroit high school 'burnouts' discussion of themes related to 'trouble', an important and definitive feature of their everyday lives (1996: 58–59). In her study of variation and change in progressive constructions in Mexican Spanish, Torres Cacoullos (2001) found that the *andar* 'go around' + NDO progressive construction was associated more than *estar* 'be' + NDO with verbs denoting rural activities, and thus came to be associated with rural speech. This was a case in which the aspectual differences between the constructions led to differences in rural/urban distribution contexts, which became absorbed into the social meaning of the construction, thus following a path from lexical to grammatical to social meaning.

Table 5: Relative frequencies and diffusion of *así*, *ansí*, *asina* and *ansina* in NMCOS

A-form	% of data	# of speakers	Diffusion rate (%)
<i>Así</i>	21	18	100
<i>Ansí</i>	<1	2	11
<i>Asina</i>	30	14	78
<i>Ansina</i>	48	8	44
Total	100	18	100

With A-form variation, there is no known *a priori* semantic distinction. This distinction has been created through everyday use in the local context. As Eckert notes,

If variation carries textured local meaning, this meaning must be constructed in the course of local life – in the course of day-to-day practice. And if this local meaning links to broader social patterns, then the connection must be made in the course of this same practice. (1996: 57)

Here, it appears that linguistic distribution is a reflection of the rural lifestyle: the archaic variants co-occur with the discussion of ‘traditional’, rural activities. The linguistic constraints mirror the social stratification, inferred from clues in prescriptive literature and representations in modern-day written texts. In this case, then, we see what was once lexical variation become associated with social variation, which then conditioned the grammatical patterns associated with some variants. In this case, in which no semantic distinction was originally evident, variants followed a path from lexical to social to grammatical.

While the ‘social’ space of the path followed by A-variants – in which all A-forms but *así* come to be associated with rural speech – arguably could have skipped over the isolated New Mexican Spanish speech community, evidence in the social distribution of A-forms in New Mexican speech suggests that this is not the case. Table 5 shows the relative frequency and diffusion rate of each A-form present in the New Mexican data. Only *así*, with a diffusion rate of 100 percent (18 speakers), seems likely to continue in future generations. The most frequently occurring form, *ansina*, making up 48 percent of the data, is used by only 44 percent (N = 8) of speakers. *Asina*, with 30 percent relative frequency, is used by 78 percent of speakers; only 11 percent (N = 2) of speakers use *ansí*.

Variable rule analysis of the social factors conditioning *así* occurrence, shown in Table 6, tells a similar story: it is likely that all A-forms except *así* will be lost in the future. Age emerges as the strongest social factor, with a range of 36. Younger speakers (aged 45 – 82 at the time of interview) favor *así* at .68,

Table 6: Social factors conditioning *así* occurrence vs. *ansí*, *asina* and *ansina* in NMCOS (input probability = .178)*

Factor group	N <i>así</i>	% <i>así</i>	Probability	% of data
Age group				
45–82	63	34	.68	49
88–96	17	9	.32	51
Range			.36	
Years of formal education				
13+	14	61	.77	6
1–12	68	19	.48	94
Range			.29	
Sex				
Male	32	32	[.58]	25
Female	52	18	[.47]	75
Total	84	21		100

Log likelihood = −182.153; $p \leq 0.01$; Total Chi-square = 7.3604; Chi-square/cell = 0.9201
*Square brackets indicate that this factor group did not achieve statistical significance.

while older speakers (aged 88 – 96) disfavor *así* at .32. Speakers with some college education favor *así* as well, at .77, while those without disfavor *así* only slightly, at .48. Gender did not emerge as significant. Since *así* is favored by the youngest, most educated speakers, it does appear to be a prestige form in New Mexico as well. This social meaning is seen in the linguistic contexts that condition its use: *así* is not associated with the activities and objects of traditional rural life, and is thus disfavored with transitive motion verbs and tangible count nouns.

4. DISCUSSION AND CONCLUSIONS

This two-component study has offered a panchronic look at a unique phenomenon. In the case of New Mexico, archaic forms are associated with the rural way of life, as shown in their tendency to co-occur with transitive verbs of motion and with tangible count nouns. It appears that these archaic A-forms are disappearing from the New Mexican variety of Spanish, shown by their strong association with speakers over the age of 88.

In the diachronic analysis, it was found that comparison of prescriptive descriptions of A-forms with actual frequency counts supports previous findings that speakers of standard varieties overestimate the actual rate of non-standard variants in non-standard varieties (Sankoff 1988). In order to measure diffusion and relative frequencies, standard *así* was used as a quantitative baseline for

comparison over time. This way of measuring frequency can account for uneven distribution of texts and differing corpus sizes in different time periods, as well as offer a fresh perspective on the interpretation of relative frequencies. The use of an apparently stable standard variant as a baseline, especially if paired with contextual or prescriptive clues as in the case of A-variation, may allow at least tentative conclusions about the social meaning of other variants in societies long gone. Using this technique, a strange story of frequency was discovered: in one thousand years of literature, some 'archaic' forms were most frequent, relative to the standard form, in the 20th century. Given this finding, it was suggested that, during the path from standard/regional variant to non-standard/rural variant, a form may all but disappear from literary texts, only to reappear with greater frequency when its status as linguistic stereotype has been established. If found to be common, this pattern may alert us to the existence of linguistic stereotypes in times or places where we have little sociolinguistic access or awareness.

NOTES

1. I would like to thank the editors Dave Britain and Allan Bell, as well as two anonymous referees, for their helpful comments. Thanks also to Neddy Vigil and Garland Bills at the University of New Mexico, for the use of the NMCOS. Furthermore, I am grateful to Stephanie Knouse for assistance with data extraction, and to Esther L. Brown for her preliminary participation.
 2. In NMCOS examples, the number following the corpus name indicates the interview number.
 3. For qualitative studies of linguistic stereotypes, see, for example, Wells (1982), Hodge and Kress (1988), and Kristiansen (2003). For quantitative studies of stereotypes, numerous studies have analyzed such items, such as *h*-dropping (e.g. J. Milroy 1983; Coelho 1997) or negative concord in English (e.g. Kalle 2007). Such studies are rarer in the case of Spanish (though see Lipski 1995).
 4. As an example, Labov (1972: 248) offers the following: '... the evolution of the New York City vernacular has led to the raising of the vowel in *off*, *lost*, *shore*, *more*, etc. until it has merged with the vowel of *sure* and *moor*. This high vowel has been stigmatized, and is now being corrected irregularly by middle-class speakers. But the same vowel, raised simultaneously in the nucleus of *boy*, *toy*, etc., is never corrected.'
 5. Most of these are conjunctions, such as 'because', 'so', 'therefore', 'thus'.
 6. In many regions of Southern Spain and all of Latin America, the sounds represented by the letters *c* and *z* are pronounced the same as *s*, as a voiceless alveolar fricative. In Standard Madrid Castilian, the former are pronounced as a theta, the latter as a voiceless alveolar fricative.
 7. Recall that *asín* is not represented in Mexico data in CORDE, and it is possible that this variant never reached North America. See Table 1.
 8. While this does constitute a switch from a non-standard to a standard variant, this example would not constitute self-correction, since the change in tense offers new, not repeated, information.
-

REFERENCES

- Aaron, Jessi Elana. 2004. 'So respetamos un tradición del uno al otro': So and entonces in New Mexican bilingual discourse. *Spanish in Context* 1: 161–179.
- Agha, Asif. 2003. The social life of a cultural value. *Language and Communication* 23: 231–273.
- Alba, Matthew. 2005. Hiatus resolution between words in New Mexican Spanish: A usage-based account. Unpublished PhD dissertation, University of New Mexico, Albuquerque, New Mexico.
- Allport, Gordon W. 1954. *The Nature of Prejudice*. New York: Doubleday Books.
- Alonso, Carlos J. 1996. The *criollista* novel. In Roberto González Echevarría and Enrique Pupo-Walker (eds.) *The Cambridge History of Latin American Literature, II: The Twentieth Century*. Cambridge, U.K.: Cambridge University Press. 195–212.
- Andersen, Henning. 1988. Center and periphery: Adoption, diffusion, and spread. In Jacek Fisiak (ed.) *Historical Dialectology*. Berlin, Germany: Mouton de Gruyter. 39–83.
- Bernal-Enríquez, Ysaura and Eduardo Hernández-Chávez. 2003. La enseñanza del español en Nuevo México. In Ana Roca and María Cecilia Colombi (eds.) *Mi Lengua: Spanish as a Heritage Language in the United States*. Georgetown, Washington, D.C.: Georgetown University Press. 96–119.
- Biber, Douglas. 1995. *Dimensions of Register Variation*. Cambridge, U.K.: Cambridge University Press.
- Biber, Douglas, Mark Davies, James K. Jones and Nicole Tracy-Ventura. 2006. Spoken and written register variation in Spanish: A multi-dimensional analysis. *Corpora* 1: 1–37.
- Bills, Garland D. 1997. New Mexican Spanish: Demise of the earliest European variety in the United States. *American Speech* 72: 154–171.
- Bills, Garland D. and Neddy Vigil. 1999. From ashes to ashes: The historical basis for dialect variation in New Mexican Spanish. *Romance Philology* 53: 43–67.
- Bills, Garland D. and Neddy A. Vigil. 2000. The continuity of change: *Nahuatlismos* in New Mexican Spanish. In Ana Roca (ed.) *Research on Spanish in the United States: Linguistic Issues and Challenges*. Somerville, Massachusetts: Cascadilla Press. 137–153.
- Cameron, Deborah. 1995. *Verbal Hygiene*. London: Routledge.
- Cárdenas Negrete, Daniel. 1967. *El Español de Jalisco: Contribución a la Geografía Lingüística Hispanoamericana*. Madrid, Spain: Consejo Superior de Investigaciones Científicas.
- Clegg, Jens. 2006. Lone English-origin nouns in the Spanish of New Mexico: A variationist analysis of phonological and morphological adaptation. Unpublished PhD dissertation, University of New Mexico, Albuquerque, New Mexico.
- Coelho, Gail M. 1997. Anglo-Indian English: A nativized variety of Indian English. *Language in Society* 26: 561–589.
- CORDE (Real Academia Española: Banco de datos). *Corpus Diacrónico del Español*. Last accessed on 5 March 2007 at <http://www.rae.es>
- Correas, Gonzalo. 1625. *Arte de la Lengua Española Castellana*. Madrid, Spain. Accessed at <http://www.cervantesvirtual.com>
- Cotton, Eleanor Greet and John M. Sharp. 2001. *Spanish in the Americas*. Georgetown, Washington, D.C.: Georgetown University Press.
- Coupland, Nikolas. 2001. Dialect stylization in radio talk. *Language in Society* 30: 345–375.

- Coupland, Nikolas and Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics* 11: 74–93.
- Dixon, John A., Berenic Mahoney and Roger Cocks. 2002. Accents of guilt: Effects of regional accent, race, and crime type on attributions of guilt. *Journal of Language and Social Psychology* 21: 162–168.
- DRAE (Diccionario de la Real Academia Española). 1726. *Diccionario de la lengua castellana, en que se explica el verdadero sentido de las voces, su naturaleza y calidad, con las frases o modos de hablar, los proverbios o refranes, y otras cosas convenientes al uso de la lengua [...]. Compuesto por la Real Academia Española*. Vol. 1. *Que contiene las letras A.B.* Madrid. Imprenta de Francisco del Hierro. 1726. Reproduced from the original at the Biblioteca de la Real Academia Española. Last accessed on 2 November 2008 at <http://buscon.rae.es/ntlle/SrvltGUILoginNtll>
- DRAE (Diccionario de la Real Academia Española). 1770. *Diccionario de la lengua castellana compuesto por la Real Academia Española. Segunda impresión corregida y aumentada*. Vol. 1. A-B. Madrid. Joaquín Ibarra. Reproduced from the original at the Biblioteca de la Real Academia Española. Last accessed on 2 November 2008 at <http://buscon.rae.es/ntlle/SrvltGUILoginNtll>
- DRAE (Diccionario de la Real Academia Española). 1822. *Diccionario de la lengua castellana por la Real Academia Española*. 6th edition. Madrid: Imprenta Nacional. Reproduced from the original at the Biblioteca de la Real Academia Española. Last accessed on 2 November 2008 at <http://buscon.rae.es/ntlle/SrvltGUILoginNtll>
- DRAE (Diccionario de la Real Academia Española). 1884. *Diccionario de la lengua castellana por la Real Academia Española*. 12th edition. Madrid: Imprenta de D. Gregorio Hernando. Reproduced from the original at the Biblioteca de la Real Academia Española. Last accessed on 2 November 2008 at <http://buscon.rae.es/ntlle/SrvltGUILoginNtll>
- DRAE (Diccionario de la Real Academia Española). 1933. *Diccionario histórico de la lengua española*. Vol. I. A. Madrid: Imprenta de Librería y Casa Editorial Hernandon. Reproduced from the original at the Biblioteca de la Real Academia Española. Last accessed on 2 November 2008 at <http://buscon.rae.es/ntlle/SrvltGUILoginNtll>
- Eckert, Penelope. 1996. (ay) goes to the city: Exploring the expressive use of variation. In Gregory Guy (ed.) *Towards a Social Science of Language: Papers in Honor of William Labov, I: Variation and Change in Language and Society*. Amsterdam, The Netherlands: John Benjamins. 47–68.
- Espinosa, Aurelio M. 1909. Studies in New Mexican Spanish, Part I: Phonology. *Revue de Dialectologie Romane* 1: 157–239, 269–300. Also published in 1909, in *Bulletin University of New Mexico Language Series* 1: 1–161 (Part of unpublished PhD dissertation, University of Chicago, Chicago, Illinois, 1909).
- Espinosa, Aurelio M. 1911a. The Spanish language in New Mexico and Southern Colorado. *Historical Society of Mexico* 16: 1–37. Reprinted in 1980 in Carlos E. Cortés (ed.) *Spanish and Portuguese Languages in the United States*. New York: Arno Press.
- Espinosa, Aurelio M. 1911b. Studies in New Mexican Spanish, part II: Morphology. *Revue de Dialectologie Romane* 3: 251–286 (Part of unpublished PhD dissertation, University of Chicago, Chicago, Illinois, 1909).
- Espinosa, Aurelio M. 1914. Studies in New Mexican Spanish, part III: The English elements. *Revue de Dialectologie Romane* 6: 241–317 (Part of unpublished PhD dissertation, University of Chicago, Chicago, Illinois, 1909).

- Espinosa, Aurelio M. 1917. Speech mixture in New Mexico: The influence of the English language on New Mexican Spanish. In H. Morse Stephens and Herbert E. Bolton (eds.) *The Pacific Ocean in History: Papers and Addresses Held at the Panama-Pacific Historical Congress Held at San Francisco, Berkeley and Palo Alto, California, July 19–23, 1915*. New York: The Macmillan Co. 408–428. Reprinted in 1975, in Anthony F. Beltramo, Eduardo Hernández-Chávez and Andrew D. Cohen (eds.) *El Lenguaje de los Chicanos: Regional and Social Characteristics Used by Mexican Americans*. Georgetown, Washington, D.C.: Center for Applied Linguistics. 99–114.
- Giles, Howard. 1970. Evaluative reactions to accents. *Educational Review* 22: 211–227.
- Giles, Howard and Nikolas Coupland. 1991. *Language: Contexts and Consequences*. Milton Keynes, U.K.: Open University Press.
- Groom, Nicholas. 2005. Pattern and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes* 4: 257–277.
- Hickey, Raymond. 2002. Ebb and flow: A cautionary tale of language change. In Teresa Fanego, Belén Méndez-Naya and Elena Seoane (eds.) *Sounds, Words, Texts, Change: Selected Papers from the Eleventh International Conference on English Historical Linguistics*. Philadelphia, Pennsylvania: Benjamins. 105–128.
- Hodge, Robert and Gunter Kress. 1988. *Social Semiotics*. New York: Cornell University Press.
- Honey, John. 1997. Sociophonology. In Florian Coulmas (ed.) *The Handbook of Sociolinguistics*. Oxford, U.K.: Blackwell Publishers. 92–106.
- Hunston, Susan and Gill Francis. 1999. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam, The Netherlands: John Benjamins.
- Johnstone, Barbara, Jennifer Andrus and Andrew E. Danielson. 2006. Mobility, indexicality, and the enregisterment of 'Pittsburghese'. *Journal of English Linguistics* 34: 77–104.
- Kalin, Rudolf. 1982. The social significance of speech in medical, legal and occupational settings. In Ellen Bouchard Ryan and Howard Giles (eds.) *Attitudes towards Language Variation: Social and Applied Contexts*. London: Arnold. 148–163.
- Kallel, Amel. 2007. The loss of negative concord in Standard English: Internal factors. *Language Variation and Change* 19: 27–49.
- Kniffka, Hannes. 1991. Schibboleths: psychologische Bestandsaufnahme und Gesichtspunkte zu ihrer sociolinguistischen Analyse. *Deutsche Sprache* 19: 159–177.
- Kristiansen, Gitte. 2003. How to do things with allophones: Linguistic stereotypes as cognitive reference points in social cognition. In René Dirven, Roselyn M. Frank and Martin Pütz (eds.) *Cognitive Models in Language and Thought: Ideology, Metaphors, and Meanings*. Berlin, Germany: Mouton de Gruyter. 69–120.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia, Pennsylvania: University of Pennsylvania Press.
- Lara Ramos, Luis Fernando. 2000. *Diccionario del español usual en México*. Alicante, Spain: Biblioteca Virtual Miguel de Cervantes. Digital edition based on the Mexico edition. El Colegio de México, Centro de Estudios Lingüísticos y Literarios, 1996. Last accessed on 31 January 2008 at <http://www.cervantesvirtual.com/servlet/SirveObras/35716130101359941976613/index.htm>
- Lippi-Green, Rosina. 1997. *English with an Accent: Language, Ideology and Discrimination in the United States*. New York: Routledge.

- Lipski, John M. 1995. Literary 'Africanized' Spanish as a research tool: Dating consonant reduction. *Romance Philology* 49: 130–167.
- López Morales, Humberto. 1994. The history of literary language. In A. James Arnold (ed.) *A History of Literature in the Caribbean, I: Hispanic and Francophone Regions*. Amsterdam, The Netherlands: John Benjamins. 9–26.
- Mallinson, Christine and Walt Wolfram. 2002. Dialect accommodation in a bi-ethnic mountain enclave community: More evidence on the development of African American Vernacular English. *Language in Society* 31: 743–775.
- Milroy, James. 1983. On the sociolinguistic history of /h/-dropping in English. In Michael Davenport, Erik Hansen and Hans Nielsen (eds.) *Current Topics in English Historical Linguistics*. Odense, Denmark: Odense University Press.
- Milroy, Lesley. 2002. Introduction: Mobility, contact and language change – working with contemporary speech communities. *Journal of Sociolinguistics* 6: 3–15.
- Montgomery, Michael. 1989. Exploring the roots of Appalachian English. *English World-Wide* 10: 227–278.
- Montgomery, Michael and Joseph S. Hall. 2003. *Dictionary of Smoky Mountain English*. Knoxville, Tennessee: University of Tennessee Press.
- NMCOSS (New Mexico-Colorado Spanish Survey). Collection housed at the University of New Mexico, by the Department of Spanish and Portuguese, overseen by Neddy Vigil. See Bills and Vigil 1999.
- Poplack, Shana (ed.). 1999. *The English History of African American English*. Malden, Massachusetts: Blackwell.
- Poplack, Shana. 2007. Stability and change over four and a half centuries of real time. Paper presented at the 36th Conference on New Ways of Analyzing Variation, University of Pennsylvania, Pennsylvania, October.
- Poplack, Shana and David Sankoff. 1987. The Philadelphia story in the Spanish Caribbean. *American Speech* 62: 291–314.
- Poplack, Shana and Sali Tagliamonte. 1991. African American English in the diaspora: Evidence from old-line Nova Scotians. *Language Variation and Change* 3: 301–339.
- Poplack, Shana and Sali Tagliamonte. 2001. *African American English in the Diaspora*. Oxford, U.K.: Blackwell.
- Preston, Dennis. 1989. Standard English spoken here. Ulrich Ammon (ed.) *Status and Function of Languages and Language Varieties*. Berlin, Germany: Walter de Gruyter. 324–354.
- Preston, Dennis R. and Nancy Niedzielski. 1999. *Folk Linguistics*. Amsterdam, The Netherlands: John Benjamins.
- Real Academia Española. 2007. Diccionarios académicos, 1726–2007. Last accessed on 3 January 2007 at <http://www.rae.es>
- Real Academia Española. 2008a. CORDE: descripción. Last accessed on 2 November 2008 at <http://www.rae.es/rae/gestores/gespub000019.nsf/voTodosporId/B4E26FC2520104D8C125716400455C06?OpenDocument&i=1>
- Real Academia Española. 2008b. CORDE: Tamaño y criterios de selección. Last accessed on 2 November 2008 at [http://www.rae.es/rae/gestores/gespub000019.nsf/\(voAnexos\)/arch475E744872738671C125716500381CF8/\\$FILE/TamanoycriteriosCORDE.htm](http://www.rae.es/rae/gestores/gespub000019.nsf/(voAnexos)/arch475E744872738671C125716500381CF8/$FILE/TamanoycriteriosCORDE.htm)
- Real Academia Española. 2008c. Diseño del *Corpus Diacrónico del Español* CORDE. Last accessed on 2 November 2008 at http://corpus.rae.es/ayuda_c.htm

- Rona, José Pedro. 1973. Normas locales, regionales, nacionales y universales en la América Española. *Nueva Revista de Filología Hispánica* 22: 310–321.
- Sankoff, David. 1988. Sociolinguistics and syntactic variation. In Frederick J. Newmeyer (ed.) *Linguistics: The Cambridge Survey. IV: The Socio-Cultural Context*. Cambridge, U.K.: Cambridge University Press. 140–161.
- Sankoff, David, Sali Tagliamonte and Eric Smith. 2005. GoldVarb X: A multivariate analysis application. University of Toronto/University of Ottawa. Last accessed on 23 December 2008 at http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm
- Santa Ana, Otto and Claudia Parodi. 1998. Modeling the speech community: Configuration and variable types in the Mexican Spanish setting. *Language in Society* 27: 23–51.
- Schilling-Estes, Natalie. 1998. Investigating 'self-conscious' speech: The performance register in Ocracoke English. *Language in Society* 27: 53–83.
- Schilling-Estes, Natalie. 2002. On the nature of isolated and post-isolated dialects: Innovation, variation and differentiation. *Journal of Sociolinguistics* 6: 164–185.
- Sigley, Robert. 2006. Corpora in studies of variation. In Keith Brown (ed.) *Encyclopedia of Language and Linguistics* (2nd edition). Amsterdam, The Netherlands: Elsevier. 220–226.
- Silva-Corvalán, Carmen. 1997. Variación sintáctica en el discurso oral: Problemas metodológicos. In Francisco Moreno Fernández (ed.) *Trabajos de sociolingüística hispánica*. Alcalá, Spain: Universidad de Alcalá. 115–135.
- Simmons, Marc. 1977. *New Mexico: An Interpretive History*. Albuquerque, New Mexico: University of New Mexico Press.
- Stubbs, Michael. 2002. On text and corpus analysis: A reply to Borsley and Ingham. *Lingua* 112: 7–11.
- Torres Cacoullos, Rena. 2001. From lexical to grammatical to social meaning. *Language in Society* 30: 443–478.
- Torres Cacoullos, Rena and Jessi Elana Aaron. 2003. Bare English-origin nouns in Spanish: Rates, constraints and discourse functions. *Language Variation and Change* 15: 287–326.
- Torres Cacoullos, Rena and Fernanda Ferreira. 2000. Lexical frequency and voiced labiodental-bilabial variation in New Mexican Spanish. *Southwest Journal of Linguistics* 19: 1–17.
- Valdés-Fallis, Guadalupe. 1973. Spanish as a native language. *Hispania* 56: 1041–1043.
- Weinreich, Uriel, William Labov and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann and Yakov Malkiel (eds.) *Directions for Historical Linguistics*. Austin, Texas: University of Texas Press. 97–195.
- Wells, John C. 1982. *Accents of English I: An Introduction*. Cambridge, U.K.: Cambridge University Press.
- Williams, Jerry L. 1986. *New Mexico in Maps*. Albuquerque, New Mexico: University of New Mexico Press.
- Wolfram, Walt. 1993. Identifying and interpreting variables. In Dennis R. Preston (ed.) *American Dialect Research*. Amsterdam, The Netherlands: John Benjamins. 193–221.
- Wolfram, Walt. 2003. On the construction of vernacular dialect norms. In Christina Bratt Paulston and G. Richard Tucker (eds.) *Sociolinguistics: The Essential Readings*. Malden, Massachusetts: Blackwell. 251–271.

- Wolfram, Walt and Natalie Schilling-Estes. 2003. Language change in 'conservative' dialects: The case of past-tense *be* in Southern enclave communities. *American Speech* 78: 208–227.
- Wolfram, Walt and Erik R. Thomas. 2002. *The Development of African American English*. Malden, Massachusetts: Blackwell.
-

Address correspondence to:

Jessi Elana Aaron
Spanish and Portuguese Studies
University of Florida
P.O. Box 117405
Gainesville, FL 32611-7405
U.S.A.
jeaaron@ufl.edu