Focusing, implicational scaling, and the dialect status of New York Latino English[1]

Michael Newman

Queens College and the Graduate Center, The City University of New York

This study examines the status of New York Latino English (NYLE) as a focused systematic dialect versus an unsystematic menu of features, what Benor (2008, 2009, 2010) calls an 'ethnolinguistic repertoire.' Systematicity is assessed through implicational scaling, under which high degrees of scaling are assumed to represent high levels of systematicity. Data from 20 young Latino New Yorkers are examined with respect to five variables presenting variants of presumed substrate origin. Results are initially contradictory. Depending on the form of scaling used, NYLE can be seen as only marginally systematic or highly so. The paradox can be resolved in two findings. First, non-substrate factors obscure substrate effects at low frequencies, particularly in variables that tend to relatively low substrate-variant-usage rates. Second, individuals show different degrees of systematicity in variable usage. These findings are neutral about the eventual outcome of NYLE as fitting best a dialect or repertoire model. However, the results further imply that if systematicity is developing in NYLE, then it may follow a pattern of leaders and followers in its organization.

KEYWORDS: Focusing, New York, Latino English, Hispanic English, implicational scaling, ethnolinguistic repertoire

INTRODUCTION

Clyne (2008: 86) defines an *ethnolect* as a dialect of a minority ethnic community that emerges in the wake of shift to the majority language and succeeds the former language as linguistic index of that group's identity. Although this definition has informed the study of many varieties including Maori English (Holmes 1997), Cajun English (Dubois and Horvath 2003), and Latino Englishes, (Fought 2003; Wolfram, Carter and Moriello 2004; Mendoza-Denton 2008) among others, it presents a number of difficulties.

One is that the crucial role ascribed to language shift is hard to sustain. For instance, the origins and development of African American English (AAE) – though often presented as the pre-eminent American ethnolect – are famously far more complex than any simple language-shift scenario (e.g. Wolfram and Thomas 2002). Moreover, seeing ethnolects as originating in this way distorts the

nature of ethnolinguistic variation because it deracinates putative 'mainstream dialects.' As Fought (2006) and Eckert (2008) have argued with respect to European American English in the U.S., majority dialects can be as ethnically marked as any other variety in local communities.

Another problem involves how much differentiation there has to be between the speech of different ethnic communities for them to be considered different dialects. After all, as Eckert (2008: 27) points out, 'there is no obvious way to distinguish between a dialect with ethnic features and an ethnolect.' For example, in New York and Boston, Jewish and Italian Americans show differences between them, but only in rates of usage of widespread local variants (Laferriere 1979; Labov [1966]2006, 2008). Similarly, low-levels of inter-ethnic variation appear in new Northern European 'multi-ethnolects' that again show quantitative differences in usage between ethnic groups of features that are more widely shared (Chesire, Fox, Kerswill and Torgersen 2008; Svendesen and Røyneland 2008; Wiese 2009). It is not clear in either case that this level of difference is enough to index an ethnic identity.

Yet even if the speech forms of a community are robust and have been shown to serve as an ethnic index, Benor (2008, 2009, 2010) presents a third problem: lack of systematicity. Benor proposes that ethnic variation is better seen in terms of 'ethnolinguistic repertoires' than different ethnolects. The idea of repertoire, she argues, better captures how sometimes one feature and sometimes another is deployed by speakers in an indexing function. This kind of disorder violates the variationist axiom by which dialects are seen as systematic, but by the same token, in contemporary multicultural societies many ethnic groups are likely to be in intense contact. So to the extent that linguistic isolation is necessary for dialect development (Sapir 1921; Trudgill 1992), no dialectal system is likely to emerge. Thus, a repertoire outcome is certainly plausible in terms of variationist theory.

New York Latino English (NYLE) provides an excellent test case for the systematicity issue since it grows out of a well-established ethnic community in a sociolinguistically-complex urban environment. Furthermore, NYLE fits Clyne's language shift scenario and robustness is not at issue. Not only do New York Latinos use a series of characteristic variants, but their speech is highly recognizable to New Yorkers. The present study assesses just how systematic a sample of young-NYLE speakers are in their use of the features indexing Latino identity. In this way the study begins to explore whether an ethnolect or repertoire account better describes the speech of this community.

BACKGROUND

Focusing and new-dialect emergence

There is relatively little work examining the systematicity of new dialects, and the studies that exist are concentrated not on ethnolects but on koines that

emerge from dialect contact (Kerswill and Williams 2000; Kerswill and Trudgill 2005; Trudgill 2008; among others). In the model developed in those studies, full dialect status is realized only with *focusing*, defined as 'the acquisition of norms and stability' (Kerswill and Trudgill 2005: 199) following koineization. The authors borrow that concept from work on creoles (LePage and Tabouret-Keller 1985), and examine it as an intergenerational process. To give one example, Kerswill and Williams (2000) discuss the development of (ow) in the new town of Milton Keynes. They note that older migrants use variants at three levels from back to front, depending on their origins. By contrast, the range of variation among their offspring is considerably reduced and concentrated at the fronted end. The fact that analogous patterns repeat across a number of variables allows Kerswill and Williams to consider the dialect focused with that generation.

In the present study, however, Kerswill and Williams's operationalization of focusing needs to be modified. First, the study is based on ethnographic research with no comparisons over apparent time. Therefore, rather than study focusing as a process, the goal is to determine the relative status – from focused on one end to diffuse on the other – of participants' speech. On this point, most participants are the children or grandchildren of immigrants, equivalent to or past the focusing generation in Milton Keynes. Second, note that Kerswill and Williams' (ow) norm is unipolar, i.e., centered on the fronted end. However, rather than one norm, research on ethnic variation characteristically shows a bipolar normative tension between features associated with in-group and non-in-group identities (e.g. for Latino Englishes, Fought 2003; Slomanson and Newman 2004; Carter 2007; Mendoza-Denton 2008). Focusing, in such cases, needs to be understood as a simplification and ordering of variation so that it is consistent, systematic, and socially meaningful, not resolved on one potential pole or the other. In fact, focusing must allow for coherent bipolar variation to the extent that this concept is based on norms. After all, constructing styles and multifaceted identities involves shifting responses to competing normative pressures (Eckert 2000, 2008; Mendoza-Denton 2008).

The Latino community in New York

In the U.S., ethnolinguistic divergence is often associated with racial difference (Wolfram and Schilling-Estes 2005; Fought 2006). Exceptions exist, like Cajuns (Dubois and Horvath 2003) and orthodox Jews (Benor 2008, 2009). However, such groups present highly-marked cultural identities, which suggests that racial-dialect differentiation is a particular case of a larger association of linguistic differentiation with social distinctiveness. It is hard to avoid the conclusion that English in America is strongly racialized because race itself is so salient as a social divider.

The New York Latino community exemplifies the role of race in American English since, unlike earlier European immigrants to New York (Labov 2006), this group maintains prominent substrate features. The community was

established by Puerto Ricans who arrived in New York City after WWI. These migrants entered a community dominated by a traditional U.S. phenotype based Black-White racial binary. Nevertheless, they were not absorbed into either group, despite the fact that many Puerto Ricans are physically indistinguishable from either African Americans or European Americans (Urciuoli 1996; Zentella 1999). Later immigrants from the Dominican Republic, who present a similar range of phenotypes, repeated this pattern in areas where they were the first large group of Latino immigrants (Bailey 2000, 2001; Toribio 2000, 2003). In New York, Dominicans and other Latin Americans began to arrive in the late 1960s. Like the Puerto Ricans, they are usually considered part of a larger Latino community that is usually considered neither Black nor White.

This lack of assimilation to pre-existing racial categories is not always recognized by autochthonous residents. Bailey and Toribio discuss non-Latinos' efforts to shoehorn Dominicans with visible African ancestry into the Black category. Efforts to impose traditional U.S. racial identities are even institutionalized in the U.S. Census (Federal Register 1997). 'Hispanic' or 'Latino' is listed as a separate category labeled 'Ethnicity,' which is to be checked or not in addition to a racial one such as Black or White. However, race is often defined, contested, and redefined. As Urciuoli (1996) and Fought (2006) note, groups like the Irish and Jews, now 'obviously' White, were not always so considered (Ignatiev 1995; Brodkin 1999). Similarly, Ibrahim (2004) discusses how many African diaspora immigrants only really 'become Black' once in North America.

One reason Latin American immigrants 'become Latino' lies in their cultural construction of race, referred to by its Spanish name *raza*. According to Mendoza-Denton (2008), *raza* grows from the intricate caste distinctions imposed by Spanish colonial authorities. In British colonies, there was a greater tendency to discount mixed heritages and to impose more discrete categories. Urciuoli (1996: 20) explains the results as follows:

> Where [race] is a discrete, binary, naturalized classification, [*raza*] can mean a people, a lineage, a social group; it has finer shades and intermediate categories. Race is assessed as combinations of personal characteristics and social circumstances. Latin *raza* is about a complex identity in the New World.

Upon immigration to the U.S., these two constructions hybridize. The discrete nature of race is accepted, but phenotype is marginalized in favor of a single ascribed identity encompassing all Latinos. The presence of specific characteristics of Latino English thus marks a linguistic dissimilation that parallels Latinos' lack of racial assimilation into either White or Black identities. What is at issue in this study is whether such linguistic differentiation is systematic enough to merit the status of dialect in New York.

Systematicity in dialects and implicational relationships

Implicational scaling been suggested as a proxy for systematicity (Rickford 2001). Scaling involves two dimensions: one, usually represented horizontally in

a table, presents a set of implicational relationships between variables; whereas the other, represented vertically, presents implicational relationships among speakers. I refer to the degrees to which a corpus actually fulfills each set of implicational relations as *intervariable scalability* and *interspeaker scalability* respectively.

Take a hypothetical case of three variables with two variants apiece used by three speakers. For each variable one variant is in-group and the other is out-group, but the variables and speakers differ in terms of the proportions in which each variant appears. In 100 percent intervariable scalability, the three variables will appear ordered horizontally from least to most frequent in-group variant appearance, usually from left to right for the three speakers. For 100 percent interspeaker scalability, the speakers are ordered in a similar fashion from most frequent to least frequent in-group variant user, this time for all three variables from top to bottom. An important point is that all three speakers participate in the same ethnolectal system the same way that an acrolectal, a mesolectal, and a basolectal speaker all participate in a creole continuum; they are just at different points on that continuum. What marks the degree of participation in the system, by contrast, is scalability. With every exception to the expected ordering, the systematicity is reduced, until at 50 percent scalability the use of variants is random. After reviewing a number of studies, Rickford (2001) concludes that robust scaling requires a rate of 85 to 90 percent of relationships following the general implicational pattern on both dimensions.

The study of implicational relations has been mostly limited to creoles and L2 varieties (Rickford 1991, 2001; Bayley 1999). One interesting L2 study is Sharma's (2005) use of scaling to assess the systematicity of Indian English in the U.S. Among the variables Sharma examined were three phonological features for which she found intervariable scalability of about 90 percent. However, the interspeaker scalability showed an 'apparently individualistic' 53 percent (Sharma 2005: 211). On the other hand, Sharma found that participants employ variants strategically, implying their use as a repertoire along the lines later proposed by Benor (2008, 2009, 2010). Because it involves an immigrant-racial-minority community in the U.S., Sharma's study provides an interesting comparison with the New York Latino situation. In particular, since Indian English is an indigenized second language whereas the participants in this study are all native or very near native speakers, greater focusing may be expected in this study if systematicity is a natural outcome of ethnolect formation.

METHODS

Participants and interviews

Participants included 20 Latino New Yorkers and four non-Latino comparison speakers between 16 and 20 years old. Data from fourteen Latinos and three comparison speakers were gathered from semi-structured interviews conducted

as part of a larger two-year ethnography at 'The Urban Arts Academy' (UAA), the prime study site.[2] UAA is a small arts-oriented secondary school in Queens, New York, with students from grades 7–12. UAA interviewees were all in grades 11, 12 or were recent graduates, and ranged in age from 16 to 19. At the time of the study from 2000 to 2002, the school was about 50 percent Latino, and most students were either from working-class or low-income homes. Interviews were conducted at the school either alone or in pairs or, in one case, at a participant's home.

Due to the ethnography, participants' social characteristics were well attested. Peer culture was an important category at UAA, and it has been shown to be an important social variable in other sociolinguistic studies of adolescents (e.g. Eckert 2000; Fought 2003; Cutler 2008; Mendoza-Denton 2008). Most, if not all, peer cultural affiliations entail constellations of activities and behaviors that link them to macrosocial identities (Eckert 2000). To use an example from this study, being a skater has as its core activity performing acrobatic tricks on skateboards, which requires athleticism focused on virtuosity but invariably leads to sometimes serious injuries. It thus supports a form of masculinity (the overwhelming majority of skaters are males) focused on mastery, toughness, and individual accomplishment rather than aggression and teamwork. Less obviously, skating is often coupled with listening to English-language rock music. The music connects skating to a mostly European American youth identity since it is mainly European Americans who listen to and perform rock. Not surprisingly, most of the prominent skating celebrities, like most rockers, are White (e.g. Tony Hawke, Ryan Scheckler). One prominent exception, Paul Rodriguez, is often explicitly discussed in terms of his Latino identity, which does not happen with Hawke or Scheckler's White identities. The presence of Latinos like Rodriguez, however, is important too since it makes skating diverse, and it is not hard to find Latino skaters like the ones listed as skaters in Table 1. Thus, it is wrong to consider skating a White peer culture; it is rather a diverse one, though it has a European American center of cultural gravity. This diversity plus the orientation allows skating to function in an integrative way, drawing together a community of individuals with different roots around a common set of cultural referents. At UAA, two skaters were Latino, although the data of only one, Javier, were used. However, they were friends with several European American classmates and few if any other non-European Americans.

Skaters were a relatively small minority at UAA. Many more students were affiliated with Hip-Hop, which is equally diverse and integrative but oriented to African American culture. Cutler (2008, in press) notes how attraction to Hip-Hop can lead to modified usage of AAE features even in European immigrants. Hip-Hop has a strong ideological component that emphasizes identity as a personal quality as opposed to group loyalties (Newman 2005, 2007). In fact, Hip-Hop is a complex cultural movement, organized artistically around four 'elements,' which include besides MCing (i.e. rapping), turntablism, break dancing, and graffiti art. Four participants in the study listed in Table 1 – Perry,

Jesse, Lanny, and Jaime – contributed artistically to the culture as MCs, and three – César, Humberto, and Cristóbal – participated receptively as rap fans. Humberto was also a graffiti artist.

UAA also presented considerable gender diversity in the sense of having a number of openly-gay male students, and the level of acceptance can be appreciated in that one gay participant, Victor, was elected senior-class president. The term used for their peer cultural identity in the school was *Diva*. More than just gay, being a diva implied participation in the broader gay-male youth scene, as well as playful flouting of gender-role norms, and interests in fashion and musical artists such as Madonna. In New York, many working-class African Americans and Latino gay youth socialize with each other but not with many European Americans, Asian Americans, or middle-class gay youths of any race. However, these three divas had friendships that crossed all racial lines and included some straight boys and girls.

In addition, Victor along with one other participant, Dalia, was also involved in student government, i.e., planning events and coordinating student issues with teachers and the administration. Student government probably attracted the most racially-diverse group in the school, including European Americans, African Americans, and others. Student government members were also marked by particularly close relationships with teachers, although all of the students, with the exception of the skaters, had close relationships with at least one faculty member.

Some youths did not participate in any wider U.S. peer culture, and interviews with them revealed substantial intergenerational socializing involving families and family friends. Edwin, Horacio, and Wilson all fit this pattern, and all three had other commonalities too. They were members of all-Latino peer social networks, listened mostly to Latin music, occasionally wore beads with national flag colors, and paid attention to current events in those countries, although they had been born in the U.S. They were also the only ones who spoke any Spanish to peers, although usually mixed with English. Thus, the close family connections appeared to be part of a larger pattern of rootedness in the heritage their families brought from their home countries.

Orientation to academics was another important characteristic that differentiated students at UAA, and the degree of investment in educational practices was often revealed through interviews and from peers and teachers. The student-government members were all academically ambitious, but they were not the only ones; the most academically-successful student in the study was Jaime, a rap artist. Admission to college was a public event, and so it was easy to determine who had been accepted where. On the other hand, some participants dropped out before graduation, and that was also public knowledge. In addition, after the main part of the study was over, I returned to the school and was able to determine the subsequent trajectory of a number of students who had kept in touch with former teachers. Finally, the ethnic composition of participants' social networks was also of interest, and this

varied widely. All participants were asked to list their close friends, and that list supplemented observation to determine the composition of their social networks.

Unfortunately, the UAA sample contained only one skater and five college-bound students. Two other Latino skaters, Chuckie and Mariano, were therefore interviewed at a skate site in Manhattan, and Mariano also turned out to also be a rock musician. In addition, Johan, Colton, and Alberto were interviewed at two public New York colleges, and thus provided more speakers with upward social mobility, although none had middle-class origins. Alberto also sang rock music and Colton was a free runner (a form of street acrobat),[3] so they added additional European-American-oriented peer cultural affiliates.

This improved socio-economic and ethnicity-related stratification was done, however, at the cost of a balance of sexes as well as an uneven knowledge of, and familiarity with, participants. All but one of the participants analyzed here are males. Therefore, the effects of gender (beyond sexual orientation) will have to await further research.[4] The potentially confounding effects of interviewer familiarity on speech need to be acknowledged, although they are famously hard to eliminate in any interview context (see discussion of ethnography and observer's paradox in Mendoza-Denton 2008).

The four comparison participants were:

- Rashid: A low-income vernacular-speaking African American gang member, affiliated with Hip-Hop as a fan. He was known to, but not a close friend of, some participants at UAA. He did not graduate.
- Darryl: An upwardly-mobile Jamaican American who spoke a less vernacular, though still non-standard, AAE than Rashid's with peers and in the interviews. He was a close friend of Jaime's and had another, Lanny, in his social network. He went on to, and graduated from, a competitive state university.
- Justin L: An upwardly-mobile working-class Irish American, who spoke New York European American English (NYEAE). He was a friend of Darryl's but affiliated with Heavy Metal, as well as Rap. He also went on to college. He had an African American girlfriend during much of the study.
- Justin S: A low-income Italian American who spoke a more vernacular NYEAE than Justin L's. He did not attend UAA and was oriented more towards computer gaming and sports than other peer cultures. He anticipated a career in the police, but I lost contact with him soon after the interview.

Participants are listed by pseudonym in Table 1 along with social characteristics.

The interviews from UAA and the skate site were recorded with a Sony cassette recorder with an external Lavalier mic, and were digitalized using a Kay Labs CSL 4200. The later recordings were made with an Olympus DS 20, using the same microphone.

Table 1: New York Latino participants and demographic factors*

Participant	Site	Gen. or Im. age	National heritage	Academic orientation	Peer culture	Social networks
Alberto	Col.	3rd	Puerto Rican	In college	Rock music	Latinos/Whites
César	UAA	2nd	Ecuadorian	College bound	Rap fan	Mixed
Chuckie	Skate	3rd	Puerto Rican	In college	Skater	Latinos
Colton	Col.	3rd	Puerto Rican	In college	Free running	Latinos/Whites
Cristóbal	UAA	0.5 y/o	Ecuadorian	College bound	Rap fan	Latinos
Dalia	UAA	2nd	Dominican	College bound	Student gov't	Mixed
Edwin	UAA	2nd	Colombian	High school only	Family oriented	Latinos
Horacio	UAA	2nd	Dominican	Uncertain	Family oriented	Latinos
Humberto	UAA	2nd	P.R. and Salv.	High school only	Graffiti artist	Latinos
Jaime	UAA	2nd	Dom. and Ecuad.	College bound	Rap artist	Latinos/Blacks
Javier	UAA	10 y/o	Guatemalan	High school only	Skater	Latinos/Whites
Jesse	UAA	3rd/2nd	Puerto Rican	High school only	Rap artist	Latinos/Blacks
Johan	Col.	5 y/o	Venezuelan	In college	Computer geek	Mixed
John	UAA	3rd	Puerto Rican	High sch. drop out	Diva	Latinos/Whites
Jonny	UAA	3rd/2nd	Puerto Rican	High sch. drop out	Diva	Latinos/Whites
Lanny	UAA	2nd	Colombian	High sch. drop out	Rap artist	Latinos/Blacks
Mariano	Skate	11 y/o	Mexican	Uncertain	Skater/rock mus.	Latinos
John	UAA	2nd	Puerto Rican	High sch. drop out	Rap artist	Latinos/Blacks
Perry	UAA	3rd	Puerto Rican	College bound	Diva/stud. gov't	Mixed
Victor	UAA	3rd	Puerto Rican	High sch. drop out	Family oriented	Latinos

*Site = site of interview; Col. = one of two public New York colleges; UAA = arts-oriented secondary school in Queens; skate = skating site; Gen. = generation American; Im. age = age at which immigrated to the U.S.

Variables

The variables summarized in Table 2 were chosen to be those in which a Spanish substrate variant could be found in opposition to an endogenous one, obviating differences between AAE and EAE, on the one hand, and Spanish dialects, on the other, as much as possible. In fact, many forms that vary between Spanish dialects, such [s] ∼ [h] ∼ ø as coda /s/, [n] ∼ [ŋ] in coda nasals, or [j] ∼ [ʒ] ∼ [ʃ] ∼ [dʒ] for /j/, correspond to English phonemic distinctions. These variations did not appear in participants' English despite being robust in some of their Spanish. Similarly, the various realizations of Spanish /r/ showed no impact on respondents' English either, which followed New York patterns. It is known that ethnic indexes need not be substrate variants (e.g. Kerswill, Torgersen and Fox 2008; Wiese 2009), but substrate variants provide a connection to Spanish and so are the least ambiguous examples.

The first two variables, (b) and (d), involve the application of the Spanish spirantization rule, consisting of lenition of phonological voiced stops most consistently in intervocalic positions (Hualde 2006), the only context examined here. The third possible similar variable (g) was not examined because it is less frequent. Results were assessed first by ear, and, if there appeared to be any doubt, by visual examination of the spectrogram and waveform. Substrate variants were tokens realized with fricatives, approximants, or elisions. The (b) and (d) variables were considered separately because they showed a non-significant correlation of .22 (df = 22, p < .305) using Wessa's (2008) online Pearson's R calculator. The first 25 tokens were gathered from each participant, except in two cases in which that number was not reached for (b). Occasional ambiguous tokens were skipped.

A third variable (d/t) was coda postvocalic coronal stops, /d/ and /t/, which show a strong tendency to elision or less frequently in the case of /d/ spirantization in Spanish. Both the fricatives and the far more frequent elisions were counted as substrate variants, but glottal stops were not. The two phonemes were treated as a single variable because the total means were similar (39.8% and 42.9% for /d/ and /t/ respectively) and they correlated highly significantly at .64 (df = 22, p < .001) (Wessa 2008). Again the first 25 tokens were gathered from each participant, and judgments were made by ear. Tokens before stops were eliminated because of the difficulty in reliably discerning variants.

The fourth variable, onset (l), was examined in Slomanson and Newman (2004). In that study clear variants – characteristic of Spanish – were identified as the most salient marker of Latino English in New York. This finding was confirmed in informal discussions with New York Latinos about identifiable features of 'sounding Spanish.' Unlike the stop variables for which binary applications of variable rules are a natural measure, /l/ is more like a vowel in showing degrees of variability. A clear pronunciation is a more fronted one, and it varies analogously to the fronting of a vowel. Therefore, this variable was calculated based on the F2 of the midpoint. This strategy was complicated

Table 2: The five New York Latino English variables chosen to be examined (those in which a Spanish substrate variant could be found in opposition to an endogenous one)

	(b)	(d)	(d/t)	(l)	(nPVI)
Definition	/b/ between vowels within or between words	/d/ between vowels within or between words	/d/ or /t/ after vowel in syllable coda position	/l/ in onset position	Type of prosodic timing
Examples	*A boy* *Chubby*	*A dog* *Studio* *Bad ass*	*Bed##* *Bat##* *Badness* *Bought land*	*Like* *Play* *Slick*	–
Substrate variants	[β, ɓ]	[ð, Ø]	[Ø, ð]	[l]	Syllable timing
Endogenous variants	[b]	[d, ɾ]	[dˀ, d, tˀ, t, ʔ]	[ɫ]	Stress timing
Form of measurement	Proportion spirantized tokens	Proportion spirantized tokens	Proportion elided or fricative tokens	F2 adjusted for following V	Median normalized Pairwise Variability Index (nPVI)

because of the high degree of co-articulation of /l/ with surrounding sounds, and so it is relatively back before a back vowel, and front before a front one. To control for this effect to the extent possible, 40 separate measures were taken. Half were done before front vowels, /æ/, /ɛ/, /e/, /ɪ/, and /i/, and half before /aɪ/ (aided by the high frequency of *like*); the first 20 tokens in each condition were used. Evidence for the effectiveness of this strategy is the very strong correlation of 0.890 (df = 22, p < .001) between respondents' means for the two contexts (Wessa 2008). As expected, each respondent's mean pre-front vowel (l) was more fronted than their mean pre /aɪ/ (l). The value of (l) used was the mean of these forty measures, which will be referred to as the 'adjusted mean.'

The final variable involves prosody. Spanish has traditionally been described as 'syllable timed,' which has been hypothesized as corresponding to relatively homogeneous syllable lengths. English by contrast is said to be 'stress timed,' based on hypothesized foot homogeneity. One of the most productive metrics for speech timing was developed by Low and Grabe (1995) and is called the 'normalized pairwise variability index' (nPVI), which compares lengths of pairs of adjacent vowels. A lower nPVI corresponds to a greater degree of syllable timing whereas a higher one indicates stress timing. For example, Grabe and Low (2002) find a mean nPVI of about .30 for a sample of Spanish compared to .59 for British English. Thomas and Carter (2006) used nPVI to compare a number of varieties of English including Carolina and Texas Latino English, although they used medians rather than means so their results are not directly comparable to Grabe and Low's. In Thomas and Carter's data, all groups of speakers varied greatly in their nPVI, but the Latinos had consistently lower median nPVIs (below .425) than their non-Latino counterparts. The AAE and EAE speakers were not differentiated in this way.

In this study, nPVI was determined by a comparing 100 pairs of adjacent syllables taken from the first long uninterrupted stretches of talk. The value was calculated in Excel using Grabe and Low's formula as supplied by Erik Thomas. Medians were also used as indicator of central tendency study because 19 out of 20 participants did not present a Gaussian distribution (i.e. their nPVI data failed a D'Agostino and Pearson omnibus normality test performed with Graphpad Prism 5.0 for Mac), which would justify using means in that way.

FINDINGS

Implicational scales of spirant variant frequencies

It is possible to set up frequency-based implicational scales – the form considered most reliable by Rickford (2001) – only for variables with binary variants, in this case (d), (b), and (d/t). Table 3 shows the implicational scales of the Latinos for these three variables. Anomalies are indicated by triangles pointing towards the incongruent cell. The variable (d) shows the least proportion of substrate variants

Table 3: Proportional implicational scale for substrate stop variables of New York Latino participants*

Speaker	Variable (d)	(b)	(d/t)
1. Javier	.36	.40	.60
		▼	
2. César	.32	.60 ▶	.56
			▼
3. Horacio	.20	.33	.80
4. Johan	.20 ▶	.16	.68
		▼	
5. Wilson	.08	.32	.64
	▼		
6. Alberto	.16	.32	.48
		▼	
7. Cristóbal	.08	.40 ▶	.32
			▼
8. Colton	0	.24	.68
	▼		
9. Chuckie	.16 ▶	.12	.48
		▼	
10. Jaime	.16	.24	.36
			▼
11. Jonny	.08	.20	.48
			▼
12. Lanny	.00	.00	.68
		▼	
13. Perry	.00	.44 ▶	.32
	▼		▼
14. Dalia	.08	.28	.36
			▼
15. Jesse	.04	.08	.52
	▼		
16. Edwin	.12 ▶	.05	.48
		▼	
17. Victor	.00	.24	.32
	▼		▼
18. Mariano	.04	.16	.36
	▼		
19. Humberto	.08	.08	.16
		▼	
20. John	.00	.48 ▶	.04

*Number = proportional rate of substrate variant used by speaker for that variable. Triangles (▶ and ▼) point into cells out of the expected order – i.e. those in which the proportion of substrate variant use breaks the general implicational trend, either horizontally (intervariable) or vertically (interspeaker)

followed by (b), with most frequent number of such variants for (d/t). This order is realized in 33 out of 40 horizontally adjacent cells, for an intervariable scalability of 82.5 percent; which is strong but not quite as robust as Sharma's 90 percent. By contrast, the interspeaker scalability is fulfilled between only 37 of the 57 pairs of vertically adjacent cells or 64.9 percent; a proportion nevertheless higher than Sharma's near-chance rate. This measure supports a view of New York Latino English, like Sharma's L2 Indian English, as not quite focused. Still, it appears to be, at least with these three variables, somewhat more organized, at least along the interspeaker dimension.

Ternary implicational scales for all variables

To examine the relationship among all five variables requires establishing a common benchmark for determining presence or absence of substrate effects across the data set. This is because three variables are binary and two are continuous, and those two are measured in quite different ways. This effort was complicated by the fact that variants indistinguishable from Spanish substrate ones turned out to be sporadically produced by the non-Latinos. For example, in the case of (d), three non-Latinos have non-zero rates of apparent spirants, and one of these, Justin S's, is equal to, or higher than, all but two Latinos. Yet, there is good reason to believe that those variants are not of Spanish origin. First, they are all elisions, unlike those of the Latinos in which fricatives and approximants alternate with elided forms. Furthermore, all of Darryl's and Rashid's and one of Justin S's cases were found in the *everybody, anybody, somebody, nobody* set.[5] Justin S's other two elided (d)s were found between high front vowels (i.e. *video, Grady*), a pattern that was not noted in the Latino participants.

However, the presence of endogenous 'spirants' complicates an effort to determine whether the Latinos' elisions originated in endogenous or substrate processes. Nor is it self-evident at what rates they become indicators of Latino English. Only surpassing the non-Latinos' usage by a sufficient amount identifies a speaker as 'sounding Spanish' on that variable, but establishing that threshold is difficult. For the purposes of this study, a threshold was set in two different ways that take into account this uncertainty. The first step involved establishing each speaker's 95 percent adjusted Wald (Agresti and Coull 1998) confidence interval (CI) for that variable.[6] This measure allows the determination of significant differences among proportional means analogously to the way that ANOVAs can be used for continuous variables that are distributed normally. For the variable (l), ANOVAs were used to establish CIs because the distributions met that criterion.

A threshold for confidence in substrate influence was then set at the level at which the lowest point of a Latino speaker's 95 percent CI was above the mean non-Latinos' usage. In other words, if the Latino participant's usage rate was significantly higher than the non-Latino mean, that participant was deemed to be a substrate-variant user and marked by a (+) on the implicational table

(Table 4). The rationale for this choice is that significantly greater usage of the appropriate variant marks that usage as Spanish influenced. A benchmark for confidence in *lack of* substrate influence was also established at the mean of the comparison speaker with the highest rate for that particular variable. Latinos whose usage of a variant was at or below that rate were given a (−) on Table 4. The choice of that particular benchmark was that it marks usage equal to or below the rate of someone whose usage was not owed to Spanish influence. Some speakers' usage fell between the thresholds, and they were given a (0) on the implicational table (Table 4), indicative of the uncertainty of origins and meaning of their usage.

Evidently, these thresholds are stipulations, and the use of different comparison speakers may have resulted in somewhat different benchmarks. However, they provide consistent criteria that arise from the sample and respond to the differences between Latino and non-Latino realizations in that sample. As such they reflect variants that New York Latinos encounter as they interact within and across racial lines. They provide a workable standard with which to identify Latino English usage, although no definitive claims about what constitutes substrate influence are intended.

Figure 1 illustrates how the benchmarks work in the form of a means-and-whiskers diagram marking the 95 percent CIs for all speakers for (d), as determined by the adjusted Wald calculation. As with all subsequent means-and-whiskers diagrams, the Latinos' means are at the flat bisecting marks, and the AAE and NYEAE speakers' means are at the black and white circles respectively. The dotted line represents the upper benchmark: the non-Latino mean of 0.09. Note that only César and Javier's 95 percent CIs surpass
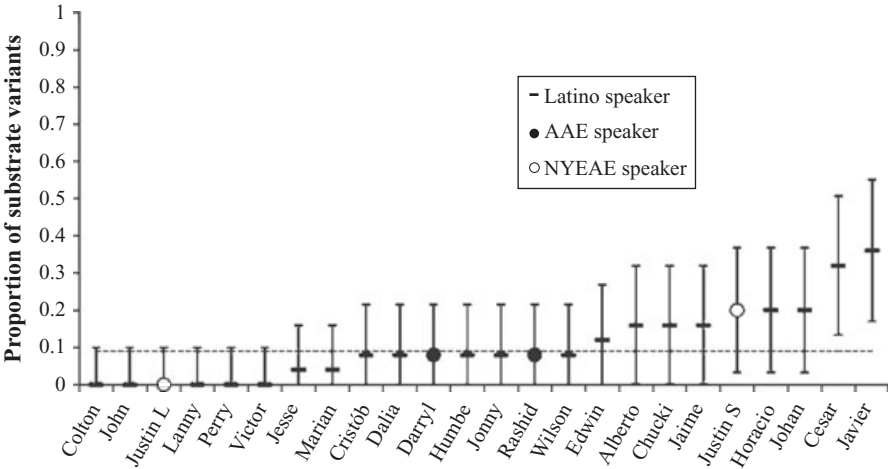

Figure 1: Mean use of substrate variants for (d), with adjusted Wald 95 percent confidence intervals. (Dotted line = mean of the four non-Latino speakers)
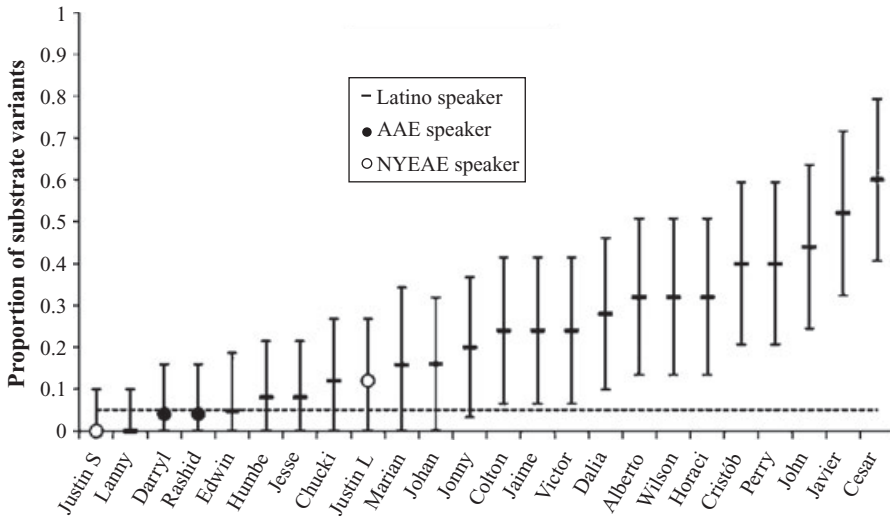
Figure 2: Mean use of substrate variants for (b), with adjusted Wald 95 percent confidence intervals. (Dotted line = mean of the four non-Latino speakers)

this threshold. The highest non-Latino's usage is 0.2, visible at Justin S's mean marker. The remaining Latino participants receive a (−), and no one receives a (0) (Table 4).

The results for (b) again present variants from the non-Latinos that, superficially at least, resemble substrate variants.[7] However, their variants are far fewer than with (d) and the Latinos' are much more frequent, as can be seen in Figure 2. Twelve of the 20 Latino speakers' CIs do not overlap with the mean rate of the non-Latinos (0.05), and are therefore are given a (+) (see Table 4). The highest non-Latino's mean is Justin L's at 0.12. Lanny – who categorically does not use the spirant – Edwin, Humberto, Jesse, and Chuckie are all at or below this rate and receive a (−). Mariano, Johan, and Jonny, who is actually at the cusp of a (+), are the uncertain cases marked (0).

The third variable (d/t) shows substantial elisions by two non-Latinos; Rashid, the vernacular AAE speaker has 0.24, whereas Darryl has 0.12. Note that again there are explanations for this use apart from of NYLE influence. For example, simple coda obstruent deletion has been observed for AAE (Wolfram 1994; Bailey and Thomas 1998; Rickford 1999). Justin S also has one such token. Eighteen of the twenty Latinos show significantly higher proportions than the non-Latino mean, and so receive (+) on the implicational table (Table 4). The others, John and Humberto, get a (−). The data are shown in Figure 3.

Figure 4 shows the results for (l) based on an ANOVA. On the implicational table (Table 4) 17 Latinos significantly surpass the non-Latino mean and get a (+). Darryl shows the clearest variants of the comparison speakers and so sets the upper benchmark for (−), below which are John, Alberto, and Lanny.
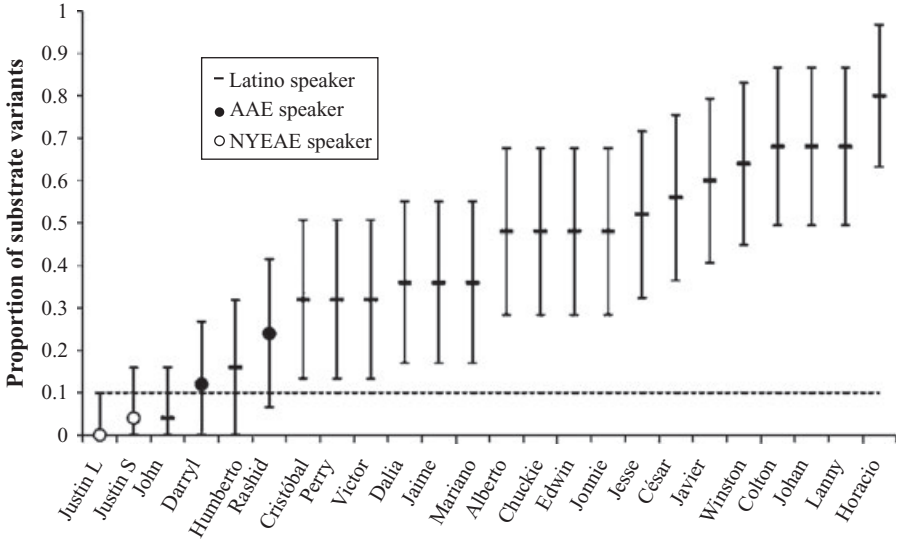
Figure 3: Mean use of substrate variants for (d/t), with adjusted Wald 95 percent confidence intervals. (Dotted line = mean of the four non-Latino speakers)

Rhythm, as prosodic, is of a different nature than the other variables, and it is measured in a different way, with nPVI. Although the distributions are continuous, recall that they were too skewed to use ANOVAs. Non-parametric tests like the Kruskal-Wallis test are normally used for skewed data, but this kind of test is too weak to produce results that are comparable to the other statistical tests. Fortunately, Thomas and Carter's (2006) data provide an alternative strategy for determining the (+), (0), and (−) assignments. Recall that all their Latinos from Texas and North Carolina Mexican American communities had median nPVIs below 0.425, whereas only one contemporary European American and no contemporary African American did. As can be appreciated in Figure 5, all the comparison speakers follow that pattern, having a median well above 0.425 (shown in reverse numerical order because substrate usage is indicated by the lower number). However, the New York Latinos depart from Thomas and Carter's results since most are also above this rate. Still, it can be reasonably supposed that these higher numbers result from a lack of substrate influence, and those participants were assigned a (−) on Table 4. Given the differences between various forms of Latino English in different parts of the U.S. (Fought 2006), there is no reason to expect that variables should behave similarly. Four participants – Cristóbal, Horacio, Javier, and Johan – are similar to Thomas and Carter's Latino English speakers. They were assigned a (+). The two that hover at that division, Wilson and César, get (0)s.

Table 4 provides the ternary substrate implicational scale for each participant by variable.
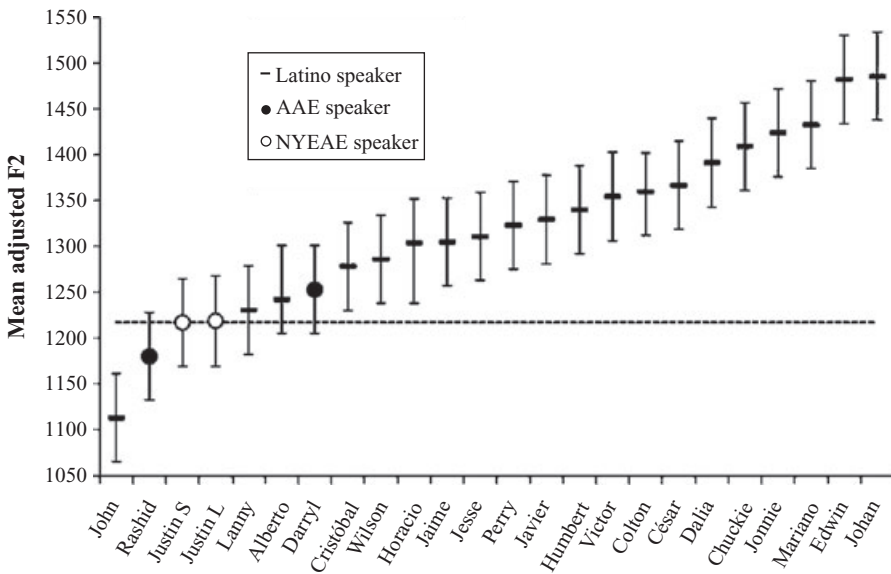
Figure 4: Mean adjusted F2 for (l), with 95 percent confidence intervals from ANOVAs. (Dotted line = mean of the four non-Latino speakers)

The horizontal (intervariable) scalability involves only five exceptions (out of 80) – marked by triangles as before – to the expected pattern for a scalability of 93.8 percent. The vertical (interspeaker) scalability is similarly robust also with five exceptions (out of 95) yielding a scalability of 94.7 percent. Thus, these ternary measures contradict the stop-variable findings and make New York Latino English look strikingly systematic.

RESOLVING THE CONTRADICTIONS

Scaling determines where a sociolinguistic data set lies on a range from most to least systematic. In ethnolects, the most systematic data consist of two monotonic clines on the interpersonal and intervariable dimensions, beginning at usage characteristic of the out-group(s) and ending at usage characteristic of the in-group. The least systematic data – in a pure repertoire – are randomly distributed on both dimensions. The contradictory results indicate that NYLE can be seen as more or less systematic depending on the form of scaling used. I suggest that examining the sources of this paradox reveals two places that scaling breaks down in the data set: one more linguistic; the other more social.

The more linguistic source of the inconsistency is that substrate influence is not the only source of differences in rates of variant usage. Recall that elided variants of each of the three stop variables appear among the comparison speakers, and that there appears to be evidence that this usage is endemic at low rates
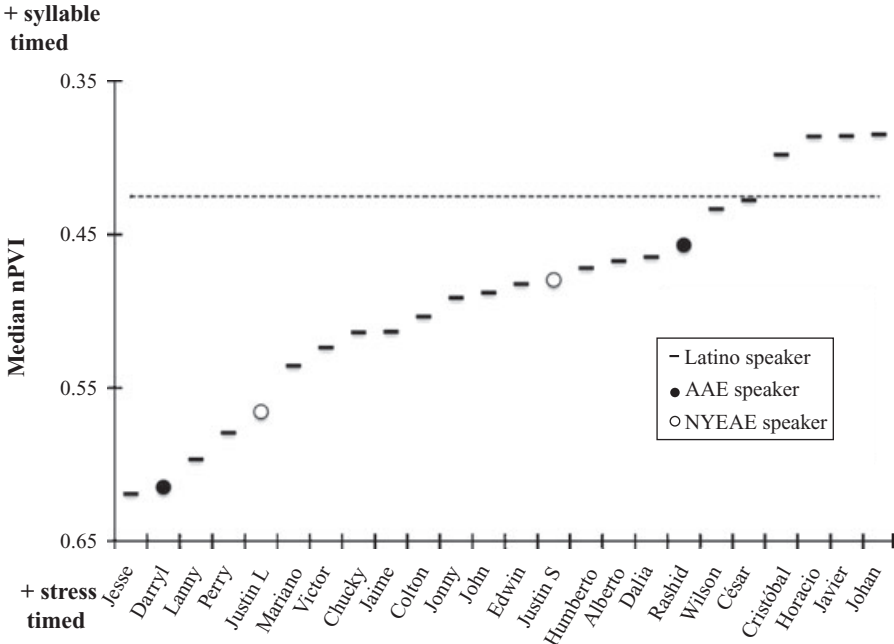
Figure 5: Type of prosodic timing (median normalized Pairwise Variability Index) for Latino and non-Latino speakers in New York. (Dotted line = boundary of substrate influence in Thomas and Carter 2006)

to the non-Latino groups. Also, there was more variation in (nPVI) on the non-substrate side of the threshold than on the substrate side, and even some variation in (l) was found among the non-Latinos. Therefore, a monotonic climb due to increasing substrate influence can only potentially begin to form above the point of influence of non-substrate factors. Assuming the lower threshold – indicated as (−) on Table 4 – as the starting point, the variation in spirant rates for 18 of the 20 Latinos for (d) should therefore be eliminated from the scales. The same is true for the five lowest Latinos for (b) and the lowest two for (d/t). This reasoning is reflected in Table 5, which reiterates Table 3, but with the proportions of all participants below the lower thresholds receiving zero. The result is a slight improvement of intervariable scalability to 94.7 percent and a substantial rise in interpersonal scalability to 84.2 percent, now barely below the 85–90 percent range given as robust by Rickford (2001). Note, that if the upper benchmark had been used, the result would have been exactly the same.

The more social source of difficulties for scaling is one of the most potentially interesting findings of the study. The results suggest that systematicity is present, but it is adopted by different speakers to different degrees. Figures 6–9 represent

Table 4: Ternary implicational scale for all five substrate variables for New York Latino participants*

Speaker	Variable (d)	(nPVI)	(b)	(l)	(d/t)
1. Javier	+	+	+	+	+
2. César	+ ▶	0 ▼	+	+	+
3. Horacio	−	+	+	+	+
4. Cristóbal	−	+	+	+	+
5. Johan	−	+ ▶	0 ▼	+	+
6. Wilson	−	0	+	+	+
7. Victor	−	−	+	+	+
8. Perry	−	−	+	+	+
9. Jaime	−	−	+	+	+
10. Dalia	−	−	+	+	+
11. Colton	−	−	+	+	+
12. Mariano	−	−	0	+	+
13. Jonny	−	−	0	+	+
14. Humberto	−	−	−	+ ▶	− ▼
15. Edwin	−	−	−	+	+
16. Jesse	−	−	−	+	+
17. Chuckie	−	−	− ▼	+	+
18. Alberto	−	−	+ ▶	−	+
19. John	−	−	+ ▶	−	− ▼
20. Lanny	−	−	−	−	+

*+ = substrate variant use above upper threshold; 0 = substrate variant use between upper and lower thresholds; − = substrate variant use below lower threshold. Triangles (▶ and ▼) point into cells out of the expected order – i.e. those whose proportion of substrate variant use break the general implicational trend, either horizontally (intervariable) or vertically (interspeaker)

Table 5: Modified proportional implicational scale for the three substrate stop variables for New York Latino participants*

Speaker	Variable		
	(d)	(b)	(d/t)
1. Javier	.36	.40	.60
		▼	
2. César	.32	.60 ►	.56
			▼
3. Horacio	0	.33	.80
4. Wilson	0	.32	.64
			▼
5. Colton	0	.24	.68
6. Johan	0	.16	.68
		▼	
7. Alberto	0	.32	.48
		▼	
8. Perry	0	.44 ►	.32
			▼
9. Jonny	0	.20	.48
10. Dalia	0	.28	.36
11. Jaime	0	.24	.36
12. Victor	0	.24	.32
			▼
13. Mariano	0	.16	.36
			▼
14. Lanny	0	0	.68
15. Jesse	0	0	.52
16. Chuckie	0	0	.48
17. Edwin	0	0 ►	.48
18. Cristóbal	0	0	.32
		▼	
19. John	0	.48	0
20. Humberto	0	0	0

*Number = proportion of substrate variant used by speaker for that variable but with all frequencies at or below upper threshold scored at 0. Triangles (► and ▼) point into cells out of the expected order – i.e. those whose proportion of substrate variant use break the general implicational trend, either horizontally (intervariable) or vertically (interspeaker).
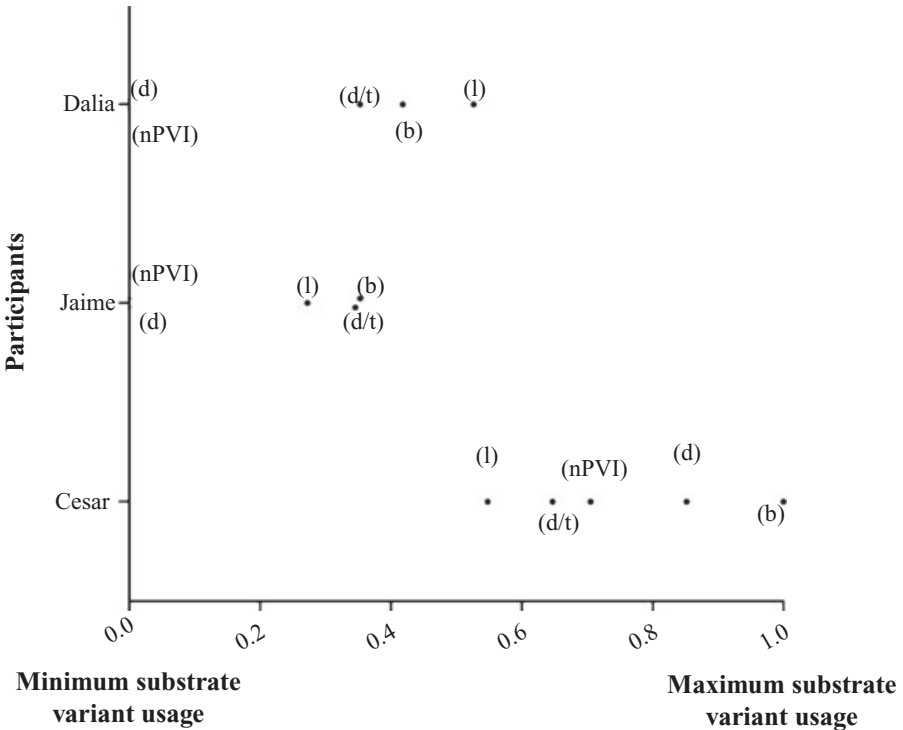
Figure 6: Relative use of Spanish substrate variants by participants with consistent usage

graphically the various patterns in speakers' usage from most to least substrate variants across variables holding intervariable differences as constant as possible. The score of 1 on each of these figures represents the maximum substrate use for that variable across all participants. The score of 0 is set at the non-Latino mean *for that variable,* but it includes all below that threshold as well. Therefore all Latinos who received a minus on Table 4 are given a 0 for the variable in question in Figures 6 to 9. The point of these figures is to equal out, to a certain extent, the differences between variables in substrate usage and make the consistency (or lack of it) across variables more easily visible. This visibility is unavoidably limited by a floor affecting those variables that have rarely used substrate variants, most strikingly (nPVI) and (d), and this effect has to be considered in interpreting the figures.

Figure 6 shows the most consistent set, César, Jaime, and Dalia. These are the only three who show rates of substrate-variant use across all five variables that would be expected given strict systematicity: César's use is uniformly high across all variables, and Jaime's is uniformly low, with (d) and (nPVI) at zero. Dalia is moderate across the three most common variables (l), (b) and (d/t), but again (d)
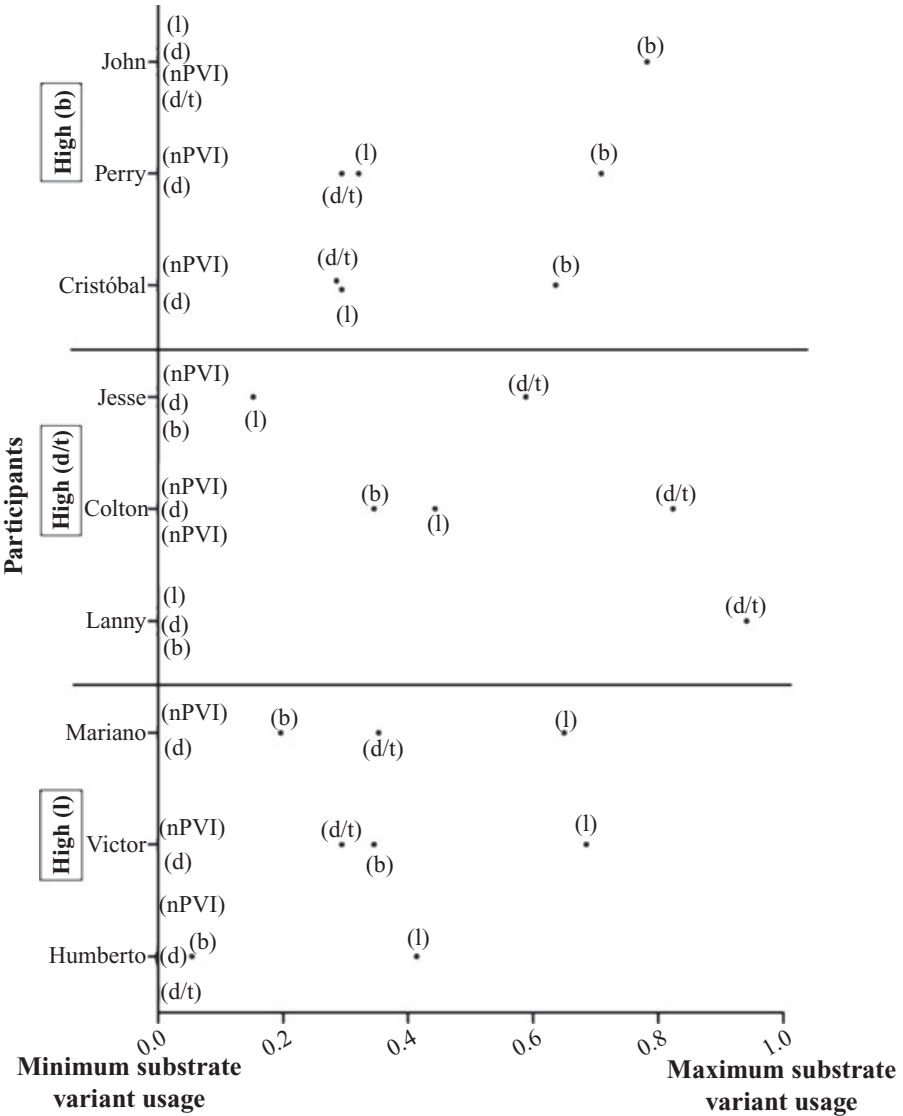
Figure 7: Relative use of Spanish substrate variants by participants with (b), (d/t), or (l) anomalously high.

and (nPVI) are, as expected, at the floor. Therefore, these three – although their actual profiles of use differ substantially – are alike in using NYLE resources in a systematic way.

Figure 7 shows nine speakers who are mainly systematic, but have one anomalously high variable. For Humberto, Victor, and Mariano, (l) is unusually
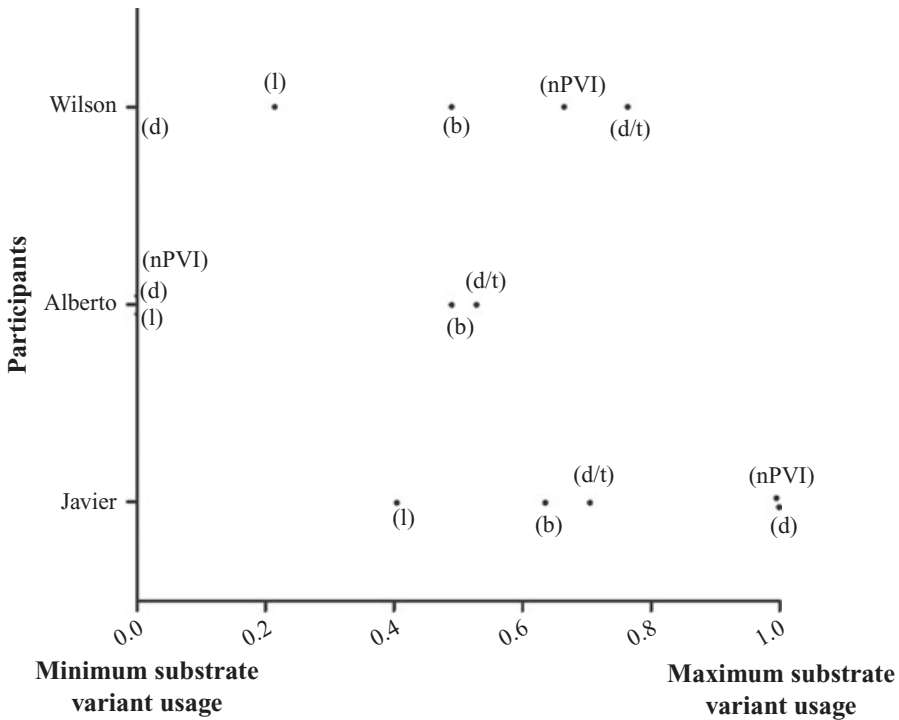
Figure 8: Relative use of Spanish substrate variants by participants with (l) anomalously low

high; for Lanny, Colton, and Jesse it is (d/t), and for Cristóbal, Perry, and John it is (b).

   Figure 8 provides a similar pattern, but the other way around. It shows Javier, Alberto, and Wilson, who show a consistent pattern except that (l) is well *lower* than expected; recall that 0s for (d) and (nPVI) are expected in moderate users of other variables.

   One tempting explanation for the odd-variant-out pattern would be that the switching among anomalous variables could be socially motivated. If this were the case, far from being unsystematic, a finer level of sociolinguistic analysis should be discoverable in Figures 7 and 8. Yet no social associations or interactions can be found there. On the contrary, in no case do the participants who behave alike on specific variables – higher, (l)s, (d/t)s, or (b)s or lower (l)s – form socially coherent groups. For example, the three with high clear (l)s are:

1. Humberto, a graffiti artist, who did not go to college, and had only Latino friends, and was second generation and bilingual. He was of mixed Salvardorian and Puerto Rican background.
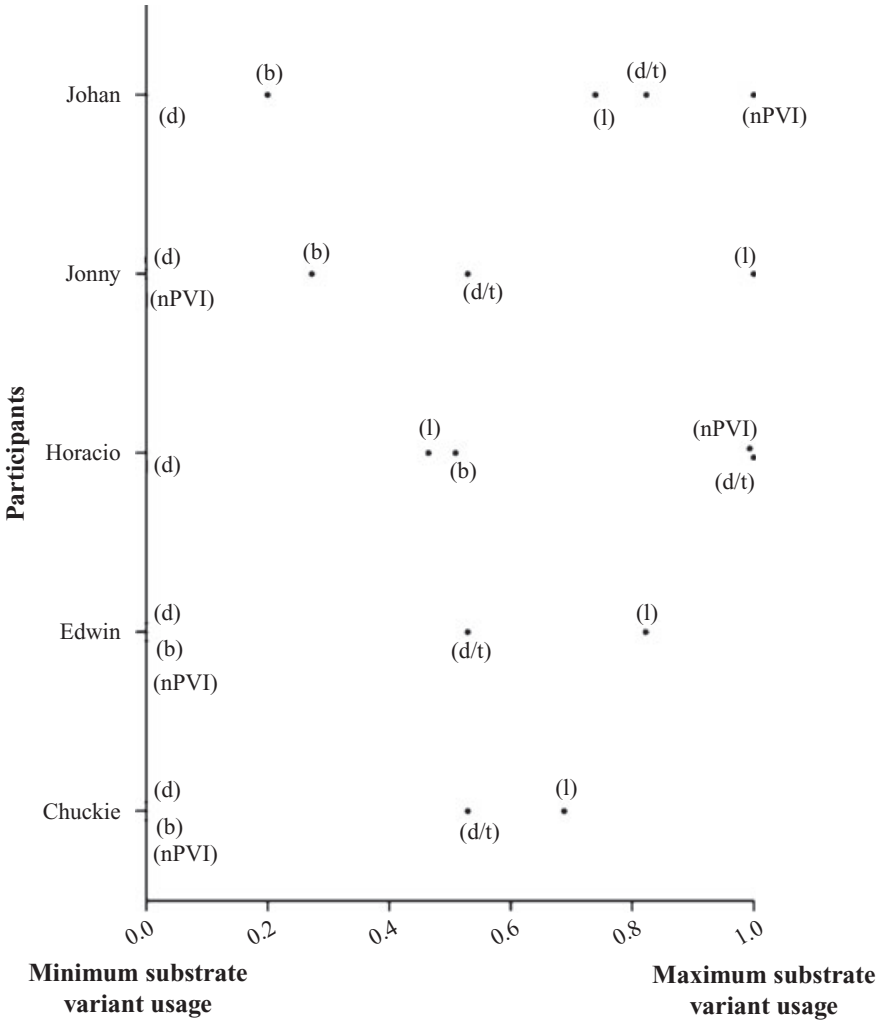
Figure 9: Relative use of Spanish substrate variants by participants with individualistic usage

2. Victor, a Diva who was also involved in student government and did go to college, and had a heterogeneous social network, was third generation and spoke little Spanish. He was of Puerto Rican background.
3. Mariano, a skater with Latino and European American friends, who did not have college plans, and came to the U.S. from Mexico at age 11.

Even social networks do not help: Humberto is in the high (1) group, but his friend, Cristóbel, is in the high (b) group. The high (d/t) group, and the low (l) group were similarly heterogeneous socially.

Figure 9 shows the remaining participants – Chuckie, Edwin, Horacio, Jonny, and Johan – who appear individualistic in their usage. Again, they form a heterogeneous group on all social variables examined.

Thus, we can break down speakers into three classifications with respect to their systematicity: systematic; partially systematic; and individualistic. The conclusion is that another source of lack of systematicity is its partial incorporation by some speakers and primarily individualistic use of variants by others.

SOCIAL IMPLICATIONS

The different levels of systematicity mean that the question of who is considered a *bona fide* speaker of NYLE depends on how the variety itself is conceived. If NLYE is understood as referring to a full dialect, only systematic users of variants – such as César, Dalia, and Jaime – or, arguably, partially-systematic users would qualify. However, if it is defined as a repertoire, then all the participants are NYLE speakers because all present high rates of usage of at least one NYLE variant. Being seen as an NYLE speaker could be important at UAA because there was a prominent norm of racial authenticity enforced on language (see also Cutler 2008, in press). An interesting example of the resulting linguistic policing was conveyed to me in an interview with John and a European American female friend:

Janet: This one girl she's like really, really tall, she came in one day. She's 12. She came in one day, a skirt up her ass, a shirt where she showed her breasts. I'm like – this was like when last year first began. . . . and I'm explaining to her, you don't come in to school dressed as something like that. That is inappropriate, and she's all, 'oh yeah what are you talking about?' this and that, and all up in my face, trying to talk slang. I'm like, 'you're White, deal with it; you're not Black, you're not Puerto Rican. Talk like a White girl.'

John: That's gotta be hard for her, cuz a White girl talks a certain way like Spanish (i.e. NYLE) or a ghetto accent, she gets called a poser because she's trying to be something she's not, but when she acts like a typical White girl, she gets made fun of.

For Janet, White speech is constructed as entailing lack of access to culturally-valuable transgressive forms, but John has a more nuanced view. He constructs Spanish-influenced English as tempting even for non-Latinos, which creates a double bind. Either this girl violates the authenticity norm or she accepts an undesirable 'typical White' identity. In fact, in another case, a mixed Egyptian and Polish student actually claimed a Spanish identity she did not have. She supported this with not just NYLE pronunciations, but an occasional Spanish word or two in conversations. By contrast, Latinos' use of forms associated with NYLE conforms to the authenticity norm.

On the other hand, 'too much' investment in racial identity was also rejected by many Latinos (and others) at UAA. A number of participants expressed as a matter of pride that at UAA, supposedly unlike other schools, race and national heritage were not social barriers or causes of tension. Sticking only with individuals of one's own ethnicity or national heritage was liable to be seen as cliquish, or at minimum lacking in sophistication – particularly for skaters and MCs (Newman 2005). As a review of Table 1 shows, multiracial social networks were common. Thus, it is not hard to see that participants were subject to normative tension between being too Latino and not Latino enough, and as the citation above shows, speech was a site where this tension played out.

However, limited systematicity has implications for efforts to study the resulting social meanings because it implies that examination of any single variable is unreliable as an index of social significance. One individual may index Latino identity exclusively with (l) whereas another may use primarily (d/t) or (b). Given the lack of evidence of secondary social-indexical roles or interactions, only if variables are considered together can such mappings be performed with confidence.

In fact, wholesale variation did map with the locally-salient peer-cultural categories in the form of a difference between the core members of integrative peer cultures and others. These core members include Lanny, Jesse, Perry, and Jaime, who were MCs, and Humberto, the graffiti artist, who were associated with the African-American-rooted Hip-Hop. They also included a set associated with European American cultures. Those include Javier, Mariano, and Chuckie, who were skaters; Mariano (again) and Alberto, who were rock singers; and Alejandro, who was a free runner. Finally, Victor and Dalia were involved in student government, which was highly heterogeneous. All these participants, except the skater Javier, show low overall use of substrate variants. Furthermore, there is plausible explanation for Javier's exceptional usage: he had immigrated at age 10 and may have residual second-language-learning effects.

In contrast to active artistry, receptive affiliation to Hip-Hop peer culture played out quite differently, though not entirely contrastively. César and Cristóbal – rap fans but not MCs, DJs, or graffiti artists – were among the highest substrate-variant users. Wilson and Horacio, who were family oriented had relatively high rates of substrate usage, but Edwin – also family oriented – did not, except for the highly salient (l). Similar variation could be found among the three Divas, who varied widely in their use of substrate variants; although Victor, who was also involved in student government, had mostly low rates. John and Jonny, although close friends, were quite different in their usage. Finally, one participant, the college student Johan, is exceptional in terms of his peer cultural orientations. He simultaneously belonged to a techno-geek group in college – which if anything has a White/Asian orientation – but was a passionate fan of *reguetón*, a Spanish-language musical genre. He reported having a highly racially-diverse group of friends in high school – including a Tibetan and an

African – all of whom, he insisted 'had better like' *reguetón*! Appropriately, he can be found right in the middle of the implicational scale.

Beyond this pattern of peer cultural influence, no other social factor seems to be associated with greater or lesser overall usage of substrate variants. Participants reporting greater or lesser Spanish ability and usage,[8] different national-heritage backgrounds, close cross-racial versus intra-racial friendship networks, and different immigration generations could be found on either end of the table or in the middle. Even though Javier's relatively-late arrival was suggested as an explanation for his exceptionally-high substrate usage, the same could not be said for Mariano, who arrived one year older yet used fewer substrate variants. Finally, academic orientation did not appear to play a role. College-bound and non-college-bound students were distributed in no discernable pattern.

CONCLUSION

LePage and Taboret-Keller (1985: 115) originally formulated focusing in terms of a metaphor, 'drawn . . . from cinema projection and focusing on a screen.'

> As the individual speaks he is seen as always using language with reference to the inner models of the universe he has constructed for himself; he projects in words images of that universe (or, of those universes) on to the social screen, and these images may be more or less sharply focused, or more or less diffuse, in relation to each other or in relation to those projected by others in their interaction with him.

In this case, we see that not all parts of the NYLE picture are equally in focus, although when we show that picture in a very zoomed-out fashion – to update the metaphor – as in the ternary division, it can appear so. Furthermore, the social transparency of even disordered substrate variants allows them to be used in the repertoire fashion described by Benor (2008, 2009, 2010). That said, the results suggest that there does seem to be an NYLE system, *but that this system is adopted to different degrees by different speakers.* Some adopt it wholesale; others take it up in good part by leaving one variable out; and yet others do not seem to adopt a system at all. If a repertoire is understood as including the possibility of individuals using a system or only parts of that system, then this study supports a repertoire account.

Ultimately, these findings neither foreclose the possibility of increased systematizing over time nor do they necessitate it. Deciding this question will require additional research in later years. The eventual outcome is worth exploring: is a natural tendency to systematicity in dialect development sufficiently strong to organize variation over time in the face of the complex sociolinguistic context that NYLE exists in? Although the data here cannot answer that question, they can, speculatively, suggest a scenario by which systematicity may plausibly spread in an ethnic community. Just as there are leaders and followers in the use of individual variables, so there may be leaders

and followers in focusing. The three most systematic speakers here fit the profile of leaders in language change outlined in Milroy (1985) and Labov (2001). As Mendoza-Denton (2008: 211–212) argues, 'Within communities of practice, iconic speakers are not necessarily the *sources* of innovation, but they are the ones who put together a style that is salient, identifiable and indeed recognizable and prone to imitation by others.' Jaime was a popular rap artist at the school, whose linguistic status can be seen – among other things – in his role as the one both boys and girls turned to write their valentine cards. He also stood out in becoming the school 'salutatorian' at graduation, a position given to the student with the second-highest average grades. César also had a strong personality and was successful academically. He won a prize for his work as a peer mediator and had friendships ranging from academically motivated students to thugs.[9] Dalia, who was involved in student government, also was a leader. Yet Victor, widely admired and elected class president in his senior year, showed the individualistic pattern, but as a gay male it is difficult to see him as a model for straight-identified boys or for girls. Jesse, the most successful rap artist, was in the odd-variable-out group, but again, certain personal qualities made him less of a potential model for others. He was quite soft spoken and less group-oriented in some ways. So these patterns are at best suggestive. It can only be hoped that future research on NYLE and other ethnic varieties can pinpoint precisely how systematicity develops, or fails to.

NOTES

1. This research was supported by the Research Institute for the Study of Language in Urban Society, the Sociological Initiatives Foundation, and PSC-CUNY.
2. The name is a pseudonym as are all participants' names.
3. See http://www.americanparkour.com/ for a page devoted to freerunning in the U.S. and its cousin parkour. Both were developed in Europe, which is still the center of these activities.
4. Girls in the original data sample are mostly members of an acting peer group or those involved in student government. Both cliques were also somewhat more upwardly mobile than most of the boys. Therefore, they would have confounded gender with other social factors.
5. Forms of –*body* were limited to three tokens per participant. Forms of *don't*, in which the *d* was frequently vocalized were not counted.
6. These were calculated using Graphpad's online calculator, accessed on 8 June 2009 at http://www.graphpad.com/quickcalcs/ConfInterval1.cfm
7. In this case, the few cases of weakened (b)s among the non-Latinos were not the result of any apparent conditioning; I suspect that they were caused by uncompleted lip gestures.
8. Mendoza-Denton (2008) comments on the differences between reported and observed Spanish usage among her participants. In this study, I observed Spanish usage only among the family-oriented participants, Wilson, Edwin, and Horacio. That said, reports were triangulated with observations from a bilingual teacher who was popular

enough with students to go out to lunch and spend time after school with them. Asked to list their three friends, a number of students listed him and occasionally one or two other teachers. Close relationships with teachers were a remarkable feature of the school, and being a former teacher there provided a counter-intuitive entree.

9. Peer pressure against academic success existed at UAA, but the notorious association of educational commitment with social failure so striking in Labov (1972) was absent. Neither success nor dropping out broke down along racial lines.

REFERENCES

Agresti, Alan and Brent Coull. 1998. Approximate is better than exact for interval estimation of binomial proportions. *The American Statistician* 52: 119–126.

Bailey, Benjamin. 2000. Language and negotiation of ethnic/racial identity among Dominican Americans. *Language in Society* 29: 555–582.

Bailey, Benjamin. 2001. Dominican American ethnic/racial identities and United States social categories. *International Migration Review* 35: 677–708.

Bailey, Guy and Erik Thomas. 1998. Some aspects of African American Vernacular English phonology. In Salikoko Mufwene, John Rickford, Guy Bailey and John Baugh (eds.) *African American English: Structure, History, and Use*. London: Routledge. 85–109.

Bayley, Robert. 1999. The primacy of aspect hypothesis revisited: Evidence from language shift. *Southwest Journal of Linguistics* 18: 1–22.

Benor, Sarah B. 2008. Towards a new understanding of Jewish language in the twenty-first century. *Religion Compass* 2: 1062–1080.

Benor, Sarah B. 2009. Do American Jews speak a 'Jewish language'? A model of Jewish linguistic distinctiveness. *The Jewish Quarterly Review* 99: 230–269.

Benor, Sarah B. 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics* 14: 159–183.

Brodkin, Karen. 1999. *How Jews Became White Folks and What That Says About Race in America*. Rutgers, New Jersey: Rutgers University Press.

Carter, Philip. 2007. Linguistic reinvention: Phonetic variation and the agentive speaker. *Penn Working Papers in Linguistics* 13: 1–14. Available at http://repository.upenn.edu/cgi/viewcontent.cgi?article=1000&context=pwpl

Cheshire, Jenny, Sue Fox, Paul Kerswill and Eivind Torgersen. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica* 22: 1–23.

Clyne, Michael. 2008. Lingua franca and ethnolects in Europe and beyond. *Sociolinguistica* 14: 83–89.

Cutler, Cecelia. 2008. Brooklyn style: Hip-hop markers and racial affiliation among European immigrants in New York City. *International Journal of Bilingualism* 12: 7–24.

Cutler, Cecelia. In press. White immigrant youth and hip-hop. *Journal of English Linguistics*.

Dubois, Sylvie and Barbara Horvath. 2003. Verbal morphology in Cajun Vernacular English. *Journal of English Linguistics* 31: 34–59.

Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford, U.K.: Blackwell.

Eckert, Penelope. 2008. Where do ethnolects stop? *International Journal of Bilingualism* 12: 25–42.

Federal Register. 1997. Standards for maintaining, collecting, and presenting Federal data on race and ethnicity. Last accessed 10 July 2009 at http://www.doi.gov/diversity/doc/racedata.htm

Fought, Carmen. 2003. *Chicano English in Context*. Basingstoke, U.K.: Palgrave.

Fought, Carmen. 2006. *Language and Ethnicity*. Cambridge, U.K.: Cambridge University Press.

Grabe, Esther and Ee Ling Low. 2002. Durational variability in speech and the rhythm class hypothesis. In Carlos Gussenhoven and Natasha Warner (eds.) *Papers in Laboratory Phonology*, Vol. 7. Berlin, Germany: Mouton de Gruyter. 377–401. Available at http://www.phon.ox.ac.uk/_esther/

Holmes, Janet. 1997. Maori and Pakeha English: Some New Zealand social dialect data. *Language in Society* 26: 65–101.

Hualde, José Ignacio. 2006. *The Sounds of Spanish*. Cambridge, U.K./New York: Cambridge University Press.

Ibrahim, Awad. 2004. One is not born Black: Becoming and the phenomenon(ology) of race. *Philosophical Studies in Education* 35: 89–97.

Ignatiev, Noel. 1995. *How the Irish became white*. New York: Routledge.

Kerswill, Paul and Ann Williams. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in Society* 29: 65–115.

Kerswill, Paul and Ann Williams. 2005. New towns and koineisation: Linguistic and social correlates. *Linguistics* 43: 1023–1048.

Kerswill, Paul and Peter Trudgill. 2005. The birth of new dialects. In Peter Auer, Frans Hinskens and Paul Kerswill (eds.) *Dialect Change: Convergence and Divergence in European Languages*. Cambridge, U.K.: Cambridge University Press. 196–220.

Kerswill, Paul, Eivind Nessa Torgersen and Susan Fox. 2008. Reversing 'drift': Innovation and diffusion in the London diphthong system. *Language Variation and Change* 20: 451–491.

Labov, William. [1966] 2006. *The Social Stratification of English in New York City* (2nd ed.). Cambridge, U.K./New York: Cambridge University Press.

Labov, William. 1972. *Language in the Inner City*. Philadelphia, Pennsylvania: University of Pennsylvania Press.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Malden, Massachusetts: Blackwell.

Labov, William. 2006. Unendangered dialects, endangered people. Natalie Schilling-Estes (ed.) *GURT'06*. Last accessed 21 May 2009 at www.ling.upenn.edu/~wlabov/Papers/UDEP.pdf

Labov, William. 2008. Mysteries of the substrate. In Miriam Meyerhoff and Naomi Nagy (eds.) *Social Lives in Language – Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff*. Amsterdam, The Netherlands/ Philadelphia, Pennsylvania: John Benjamins Publishing Company. 315–326.

Laferriere, Martha. 1979. Ethnicity in language variation and change. *Language* 55: 603–617.

LePage, Robert and Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge, U.K.: Cambridge University Press.

Low, Ee Ling and Esther Grabe. 1995. Prosodic patterns in Singapore English. In Kjell Elenius and Peter Branderud (eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Vol. III). Stockholm, Sweden: KTH and Stockholm University. 636–639.

Mendoza-Denton. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Malden, Massachusetts: Blackwell.

Milroy, Leslie. 1985. *Language and Social Networks*. Oxford, U.K.: Blackwell.

Newman, Michael. 2005. Rap as literacy: A genre analysis of Hip-Hop ciphers. *Text* 25: 399–436.

Newman, Michael. 2007. 'I don't want my ends to just meet; I want my ends overlappin': Personal aspiration and the rejection of Progressive Rap. *Language Identity and Education* 6: 131–145.

Rickford, John R. 1991. Variation theory: Implicational scaling and critical age limits in models of linguistic variation, acquisition, and change. In Thom Heubner and Carles Fersons (eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories*. Amsterdam, The Netherlands: Benjamins. 225–246.

Rickford, John R. 1999. *African American Vernacular English: Features, Evolution, Educational Implications*. Oxford, U.K.: Blackwell.

Rickford, John R. 2001. Implicational Scaling. In J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.) *The Handbook of Language Variation and Change*. Oxford, U.K.: Blackwell. 669–702.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt.

Sharma, Devyani. 2005. Dialect stabilization and speaker awareness in non-native varieties of English. *Journal of Sociolinguistics* 9: 194–224.

Slomanson, Peter and Michael Newman. 2004. Peer group identification and variation in New York Latino English laterals. *English Worldwide* 25: 199–216.

Svendsen, Bente and Unn Røyneland. 2008. Multiethnolectal facts and functions in Oslo, Norway. *International Journal of Bilingualism* 12: 63–83.

Thomas, Erik and Phillip M. Carter. 2006. Prosodic rhythm and African American English. *English World Wide* 27: 331–355.

Toribio, Almeida Jacqueline. 2000. Language variation and the linguistic enactment of identity among Dominicans. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 38: 1133–1159.

Toribio, Almeida Jacqueline. 2003. The social significance of language loyalty among Black and White Dominicans in New York. *The Bilingual Review/La Revista Bilingüe* 27: 3–11.

Trudgill, Peter. 1992. Dialect typology and social structure. In Ernst H. Jahr (ed.) *Language Contact and Language Change*. Berlin, Germany/New York: Mouton de Gruyter. 195–212.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society* 37: 241–280.

Urciuoli, Bonnie. 1996. *Exposing Prejudice: Puerto Rican Experiences of Language, Race, and Class*. Boulder, Colorado: Westview.

Wessa, Patrick. 2008. Pearson Correlation v1.0.3 in Free Statistics Software v1.1.23-r2. Office for Research Development and Education, URL. Accessed at http://www.wessa.net/rwasp_correlation.wasp/

Wiese, Heike. 2009. Grammatical innovation in multi-ethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119: 782–806.

Wolfram, Walt and Natalie Schilling-Estes. 2005. *American English: Dialects and Variation* (2nd edition). Hoboken, New Jersey: Wiley.

Wolfram, Walt, Phillip M. Carter and Rebecca Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8: 339–358.

Wolfram, Walt and Erik R. Thomas. 2002. *The Development of African American English*. Oxford, U.K./Malden, Massachusetts: Blackwell.

Wolfram, Walt. 1994. The phonology of a sociocultural variety: The case of African American Vernacular English. In John E. Bernthal and Nicholas W. Bankson (eds.) *Child Phonology: Characteristics, Assessment, and Intervention with Special Populations*. New York: Thieme. 227–244.

Zentella, Ana Celia. 1999. *Growing Up Bilingual*. London: Blackwell.

Address correspondence to:

Michael Newman
Queens College & the Graduate Center/CUNY and the Research Institute
for the Study of Language in Urban Society/CUNY
Department of Linguistics and Communication Disorders
Queens College
65-30 Kissena Boulevard
Flushing, NY 11367
U.S.A.

michael.newman@qc.cuny.edu